Pavol Zlatoš

LINEÁRNA ALGEBRA A GEOMETRIA

Cesta z troch rozmerov s presahmi do príbuzných odborov

Bratislava 2011

Obsah

Predhovor

Lineárna algebra je jazykom, slúžiacim na vyjadrenie geometrických ideí a vzťahov pôvodne vychádzajúcich z nášho geometrického názoru, ktoré však tento rámec – keďže sa v priebehu vývoja ukázal priúzky – čoraz výraznejšie prekračujú. Kým napr. naša geometrická predstavivosť je ohraničená dvoma rozmermi roviny resp. troma rozmermi priestoru, v lineárnej algebre sa zaoberáme priestormi ľubovoľnej konečnej dimenzie a okrajovo sa zatúlame aj do priestorov nekonečnorozmerných, hoci tie už skôr patria do oblasti pôsobnosti funkcionálnej analýzy. Aparát lineárnej algebry je tak akousi „slepeckou palicou“, ktorou si „oťukávame“ svet neprístupný nášmu pohľadu, zatiaľ čo pôvodný geometrický názor sa stáva zdrojom pojmov, intuitívnych vhľadov a metafor, ktoré do tohto sveta prenášame a s ich pomocou sa v ňom orientujeme.

V súčasnosti kurz lineárnej algebry tvorí spolu s kurzom matematickej analýzy dva základné piliere univerzitného matematického vzdelania. Lineárna algebra je aspoň v počiatočnej fáze tohto procesu tou pojmovo priezračnejšou a logicky jednoduchšou, no zároveň abstraktnejšou z oboch disciplín. Jej základné pojmy, ako *pole, vektorový priestor, báza, lineárne zobrazenie* a pod., sú definované ako abstraktné objekty resp. obory objektov, vyhovujúce istým podmienkam (axiómam). Všetky ich ďalšie vlastnosti sú z nich odvodené logickými úvahami, pri ktorých geometrický názor a intuícia hrajú nanajvýš pomocnú úlohu. Lineárnu algebru tak možno do veľkej miery budovať z minima základných pojmov a vopred stanovených predpokladov axiomatickou metódou, čo z jej výuky robí ideálny prostriedok kultivácie logického myslenia a estetického cítenia v matematike.

Dôležitú úlohu v matematike hrá voľba vhodného označenia. Kým výpočty s ťažkopádnou symbolikou sa môžu ľahko dostať do slepej uličky, šikovne zvolené označenie nám už samotnou svojou štruktúrou odhaľuje niektoré stránky označovaného predmetu, a tak nás vedie správnym smerom. Zároveň nám ponúka mnemotechnické pomôcky umožujúce predchádzať chybám, či ich aspoň ľahšie odhaliť. Práve algebra (a nielen tá lineárna) asi najviac zo všetkých matematických disciplín cieľavedome rozvíja cit pre primeranú symboliku.

Základné pojmy a metódy lineárnej algebry zasahujú do takmer všetkých oblastí modernej matematiky a jej aplikácií – nielen vo fyzike, technike, štatistike či informatike (kde to asi málokoho prekvapí), no taktiež v chémii, biológii, ekonómii a sociálnych vedách. S istou dávkou zjednodušenia možno povedať, že väčšina úspechov aplikácií matematiky napr. vo fyzike (no nielen v nej) sa zakladá na jednej pozoruhodnej vlastnosti mnohých prírodných procesov: závislosť ich prírastkov v malých časových úsekoch možno vzhľadom na dĺžku týchto úsekov považovať za lineárnu. Podobne, väčšinu geometrických objektov vyskytujúcich sa v technickej praxi možno v malých oblastiach veľmi dobre aproximovať pomocou lineárnych (t. j. priamočiarych resp. rovných plošných) útvarov. Takto chápaná myšlienka linearizácie je spolu s myšlienkou limitného prechodu (prípadne – v historicky pôvodnej podobe – s myšlienkou nekonečne malých veličín) základom diferenciálneho a integrálneho počtu. Na druhej strane, často sa môžeme stretnúť napr. s grafmi rôznych závislostí, pozostávajúcimi z empirických dát vynesených nad časovú škálu a pospájaných úsečkami do lomenej čiary, zobrazujúcimi napr. vývoj cien, priebeh priemerných denných teplôt, množstva exhalátov v ovzduší a pod. Pokiaľ však nie je k dispozícii hlbšia analýza zaručujúca, že v čiastkových intervaloch nemôže dôjsť k výraznejším výkyvom sledovaného ukazovateľa, treba byť k podobným „bulvárnym“ uplatneniam myšlienky linearizácie prinajmenšom skeptický.

Iným spôsobom vstupuje lineárna algebra (a funkcionálna analýza) napr. do kvantovej mechaniky. Stavom kvantovomechanického systému zodpovedajú vektory vo vhodnom (typicky nekonečnorozmernom) vektorovom priestore nad poľom komplexných čísel. Superpozíciou dvoch stavov je opäť stav systému, ktorému zodpovedá súčet vektorov zodpovedajúcich pôvodným stavom. „Celá fyzika“ takého systému sa potom do značnej miery redukuje na štúdium spektrálnych vlastností istých lineárnych operátorov na jeho stavovom priestore.

* * * * *

Zoči-voči množstvu učebníc lineárnej algebry vo svetovej matematickej literatúre sa autor len ťažko vyhne otázke, či má vôbec zmysel písať ďalšiu. Čiastočné oprávnenie by mu vari mohla poskytnúť situácia v slovenskej (a českej) matematickej produkcii. Ale ani tá – hoci má od ideálnej ďaleko (čo je však takmer tautológia) – nie je nijako kritická. Pokúsim sa teda tejto otázke čeliť tak, že odhalím osobné ciele a ambície, ktoré som svojou knihou sledoval. Nakoľko také zámery oprávňujú podobný podnik a do akej miery sa mi ich podarilo naplniť, to nech už posúdia jej čitatelia.

Predovšetkým som chcel napísať učebnicu, akú by som si želal mať k dispozícii, keď som ako študent navštevoval prednášky z lineárnej algebry a geometrie. Tým nevdojak prezrádzam, že s učebnicami, podľa ktorých sme vtedy postupovali, som príliš spokojný nebol. Neskôr, keď som už sám prednášal lineárnu algebru na Matematicko-fyzikálnej fakulte Univerzity Komenského v Bratislave (kurz som zdedil po kolegovi Jožovi Širáňovi, spolu s brilantnými stručnými poznámkami k prednáškam), som sa sám začal rozhliadať po vhodnom učebnom texte. Najviac ma zaujali *Linejnaja algebra i geometrija* od Kostrikina a Manina [22] a *Pěstujeme lineární algebru* od Motla a Zahradníka [29]. Zakrátko som si však uvedomil, že predbežná pripravenosť a všeobecná matematická kultúra, ktorú tieto knihy predpokladajú, sú na podstatne vyššej úrovni, než možno očakávať od študentov prvého ročníka. Ja sám ako prvák by som sa podľa nich učil len ťažko.

Rozhodol som sa teda zvoliť elementárnejší spôsob výkladu, ktorý – okrem gymnaziálnej látky a paralelne rozvíjaných poznatkov z matematickej analýzy – aspoň spočiatku nepredpokladá rozsiahlejšie vedomosti z iných matematických disciplín ani veľkú zbehlosť a všeobecný rozhľad. Mojim cieľom bolo napísať učebnicu, ktorá by jednak plne a logicky ucelene pokrývala materiál základného dvojsemestrálneho kurzu lineárnej algebry, jednak ponúkala dostatočne reprezentatívny výber tém tento materiál presahujúcich – či už smerom k iným matematickým disciplínam alebo aplikáciám mimo rámca „čistej“ matematiky. Ich výber bol popri vkuse autora ovplyvnený jednak spontánnym vývojom, kedy vznikajúca kniha akoby začala žiť svojím vlastným životom a občas sa vymkla spod kontroly autora, jednak faktom, že som túto látku dlhodobo prednášal študentom fyziky. Snažil som sa preto zaradiť doň témy (ako využitie metód lineárnej algebry v špeciálnej teórii relativity a v kvantovej mechanike), ktoré by ich mohli či mali zaujať. Napriek tomu som zvolil abstraktnejší prístup než je pre potreby fyzikov nevyhnutné. Z ich hľadiska by napr. úplne stačilo zaoberať sa vektorovými priestormi nad poľami reálnych a komplexných čísel. S ohľadom na potreby študentov matematiky a informatiky som však, pokiaľ to bolo možné, volil jednotný spôsob výkladu a pripustil ľubovoľné polia skalárov, vrátane tých konečných (hoci na ukážku ich aplikácií v kódovaní a kryptografii už nezostalo miesto). Na druhej strane som práve z pedagogických dôvodov odolal pokušeniu používať jazyk teórie kategórií aj za cenu, že niektoré partie (napr. otázky týkajúce sa duality) nebudú prezentované najelegantnejším možným spôsobom.

Napriek značnému rozsahu sa do knihy nedostali ani niektoré ďalšie témy, ktoré s látkou lineárnej algebry úzko súvisia. Napr. o projektívnych priestoroch alebo reprezentáciách grúp je tu len zopár letmých zmienok. Témy ako symplektické priestory, nezáporné a stochastické matice, markovovské reťazce, lineárne programovanie alebo numerické metódy lineárnej algebry, nie sú pokryté ani v náznaku. Čitateľovi (nielen) preto odporúčam, aby siahol po inej literatúre – či už kvôli nim ako aj pre porovnanie nášho prístupu k spracovanému materiálu s prístupom iných autorov. Nový pohľad totiž spravidla odkrýva súvislosti, ktoré by inak mohli zostať utajené.

Z moderne poňatých kníh, ktoré ma pri výbere a spracovaní materiálu najviac ovplyvnili, vrelo odporúčam do pozornosti tak prednášajúceho ako aj študentov už spomínané *Linejnaja algebra i geometrija* [22] a *Pěstujeme lineární algebru* [29]. Na druhej strane som hodne čerpal z klasických (hoci dnes už trochu zastaralých) učebníc Geľfanda [11] a Maľceva [27]. Pokiaľ sa čitateľ nebojí ruštiny, určite stojí za to sa do nich pozrieť.

Predbežnú predstavu o náplni knihy možno získať z jej pomerne podrobného obsahu. Už pri zbežnom pohľade naň (no vzhľadom na jej hrúbku aj bez toho) bude asi zrejmé, že látka pokrytá v knihe sa v dvojsemestrálnom kurze nedá nijako rozumne odprednášať. Akési nevyhnutné minimum tvorí časť I (t. j. kapitoly 0–10), ďalej kapitoly 11–14 časti II a kapitoly 18, 19 a prvé dva paragrafy kapitoly 20 z časti III. Pri bližšom pohľade vyjde najavo, že i z tohto minima možno ešte stále všeličo vynechať. Podľa úrovne predbežnej matematickej prípravy poslucháčov možno napr. celú kapitolu 0 prenechať na samostatné čítanie. Z viacerých kapitol časti I možno vypustiť spravidla posledné paragrafy – konkrétne 3.4, 4.6, 5.5–6, 6.5 a 7.6; z kapitol 8 a 9 o afinných priestoroch sa stačí obmedziť na úvodné paragrafy 8.1–2 a 9.1–3. Podobne v časti II možno paragraf 12.3 prenechať na prednášku z analýzy viac premenných a taktiež vynechať paragrafy 14.3-4. Zo spomínaných kapitol časti III možno ešte vypustiť paragraf 19.2.

Tento minimálny zoznam ponecháva dostatok priestoru na rozšírenie o niekoľko ďalších ucelených tém. Je už len na prednášajúcom, ktoré si vyberie. Pri plánovaní mu môže pomôcť *Schéma nadväznosti kapitol*. Ďalšie si študent môže podľa vlastného záujmu osvojiť samostatne. Po niektorých kapitolách môže siahnuť až neskôr, pri štúdiu iných matematických disciplín. Z toho dôvodu som sa usiloval, aby pestrý materiál mohol podľa knihy prednášať aj učiteľ, ktorý nie je špecialistom na danú problematiku, a zároveň sa z nej študent prvého ročníka mohol učiť aj bez sprievodnej prednášky. Celá prvá časť, úvodné kapitoly ďalších častí, ako aj úvodné paragrafy väčšiny kapitol sú preto (občas možno aj za cenu istej rozvláčnosti) napísané pomerne podrobne, vrátane rozboru technických detailov, s rozsiahlym vysvetľujúcim komentárom a dôrazom na motivačné a ilustračné príklady. Smerom ku koncu jednotlivých častí resp. kapitol (s výnimkou prvej časti) sa očakáva, že poslucháč alebo čitateľ sa v problematike začína postupne orientovať, a výklad sa stáva „strmším“. Jeden alebo dva záverečné paragrafy kapitoly sú občas venované doplňujúcim informáciám, prípadne matematickým alebo filozofickým „úletom“ od danej témy.

Atypicky neskoré zaradenie kapitol o grupách je komentované v poznámke v paragrafe 27.1. Prednášajúci, ktorý sa nechce vzdať možnosti používať pojmy *grupy*, *podgrupy*, *homomorfizmu* a pod. ako výhodné skratky, však môže aspoň úvodné paragrafy 1–3 kapitoly 27 zaradiť do programu kedykoľvek skôr (pri nevyhnutnej modifikácii niektorých dôkazov a príkladov). Napr. samotný pojem grupy možno zaviesť už v paragrafe 0.4.

Jednotlivé kapitoly sú sprevádzané pomerne veľkým počtom cvičení rôznej náročnosti. Bolo by naivné očakávať, že čitateľ ich všetky vyrieši. Bolo by však dobre, aby si ich všetky aspoň prečítal a vybral si z nich niekoľko, ktorými sa potom bude zaoberať podrobnejšie. Niektoré slúžia len na overenie porozumenia základným pojmom alebo precvičenie výpočtových postupov. Vo chvíli, keď je čitateľovi jasné, o čo ide alebo ako na to, môže ich kľudne preskočiť. V iných cvičeniach sa od čitateľa žiada, aby doplnil vynechané dôkazy niektorých tvrdení, alebo dokázal ďalšie. Aspoň občas by sa mal o to pokúsiť, u tých ťažších prípadne s využitím návodu, ktorým sú spravidla doplnené. Niektoré cvičenia prehlbujú alebo rozširujú preberanú látku – k tým možno pristupovať podľa individuálnej miery záujmu. Často sú však do cvičení bez predbežného varovania zaradené niektoré pojmy a výsledky, na ktoré sa budeme odvolávať v ďalšom výklade. Čitateľ, ktorý ich preskočil a bude mu chcieť porozumieť, by mal počítať s tým, že sa k nim bude musieť vracať (čo však nijako neprekáža). Často som do cvičení zapracúval otázky (a náznaky odpovedí), ktoré mi kedysi ešte ako študentovi napadali v podobných súvislostiach, no na prednáškach ani vo vtedy u nás dostupnej literatúre som na ne nenašiel odpoveď. Potešilo by ma, keby čitateľovi napadli ďalšie – i také, na ktoré táto kniha neprináša odpoveď, – a vyprovokovali ho k hľadaniu a premýšľaniu.

Istú časť cvičení, ako aj mnoho ďalších podobných, no rozsahovo a numericky podstatne náročnejších (teda reálnym aplikáciám bližších) úloh možno efektívne riešiť pomocou interaktívnych softwarových systémov ako Maple®, *Mathematica*® alebo MATLAB®. I keď sa v knihe nevenujeme žiadnemu z nich, určite čitateľovi odporúčam, aby sa aspoň s jedným z nich zoznámil a zvykol si ho používať. Ucelený prehľad týchto systémov ako aj balíkov podprogramov pre lineárnu algebru možno nájsť v *Handbook of Linear Algebra* [15]. Tento (do jednej ruky priťažký) *Handbook* (t. j. *Príručka*) však obsahuje omnoho viac: takmer úplný prehľad (pozor, nie výklad) základných aj pokročilých pojmov, výsledkov a aplikácií lineárnej algebry (vrátane tých najnovších), systematicky zoradený podľa jej jednotlivých oblastí a doplnený podrobnými odkazmi na rozsiahlu literatúru. Čitateľ, ktorý sa bude chcieť dozvedieť z lineárnej algebry čokoľvek, čo v našej knihe nenájde, urobí najlepšie, ak začne hľadať práve tu. Často dostupnejšou, no menej autoritatívnou alternatívou je Wikipedia.

* * * * *

Veľmi potrebnú spätnú väzbu pri písani knihy mi poskytli a viacerými radami a kritickými poznámkami, ktoré, ako verím, viedli k zlepšeniu jej kvality, prispeli kolegovia Vlado Balek, Laco Kvasz, Martin Mojžiš, Martin Mačaj, Martin Niepel, Dano Ševčovič a Jožo Širáň. Obrázky nakreslila a sadzbu v LaTeXu pripravila Mária Benešová. Pri zháňaní finančnej podpory na jej vydanie mi pomohli Andrej Ferko, a hlavne Barbora Kamrlová; v samom závere Peter Mederly. Všetkým im patrí moja vďaka. Za všetky chyby, nepresnosti, preklepy, prehrešky proti spisovnému jazyku či iné nedopatrenia, ktoré v knihe ešte zostali, – snáď je to zbytočné i dodávať – však nesie plnú a výhradnú zodpovednosť jej autor.

Napokon by som chcel poďakovať svojej manželke Jane a dcéram Hanke a Janke za morálnu podporu, ktorú mi poskytovali počas tých dlhých rokov práce na tejto knihe, a trpezlivosť, s akou znášali moju častú duševnú neprítomnosť.

Bratislava, júl 2011

Pavol Zlatoš

Schéma nadväznosti kapitol

Čitateľ by si mal byť vedomý, že uvedená schéma je značne neúplná a má len orientačný charakter. Predpokladané znalosti potrebné na štúdium nejakej kapitoly nemusia nutne zahŕňať celý obsah kapitol, ktoré jej v schéme predchádzajú. Naopak, môže sa stať, že niektoré časti danej kapitoly si vyžadujú istú znalosť (spravidla malých) častí kapitol, ktoré nie sú zaradené v schéme ako jej nevyhnutné predpoklady. Spoľahlivejšia schéma by musela miesto kapitol operovať s jednotlivými paragrafmi (a čiastočne aj s cvičeniami). Vzhľadom na ich počet (len paragrafov je takmer dvesto) by však pravdepodobne bola značne neprehľadná a sotva použiteľná.

$$\begin{tikzcd}
\text{Časť I:} & 0 \arrow[r] & 1 \arrow[r] & 2 \arrow[r] & 3 \arrow[r] & 4 \arrow[r] & 5 \arrow[r] & 6 \arrow[r] & 7 \arrow[r] \arrow[d] & 10 \\
& & & & & & & & 8 \arrow[r] & 9
\end{tikzcd}$$

$$\begin{tikzcd}
& 10 \arrow[d] & & & & 8 \arrow[d] & & \\
\text{Časť II:} & 11 \arrow[r] & 12 \arrow[r] & 13 \arrow[rr] & & 14 \arrow[rr] \arrow[dl] \arrow[dr] & & 17 \\
& & & & 15 & & 16 &
\end{tikzcd}$$

$$\begin{tikzcd}
& 10 \arrow[d] & & & & & 17 \arrow[d] & & \\
\text{Časť III:} & 18 \arrow[r] & 19 \arrow[r] & 20 \arrow[r] \arrow[d] & 22 \arrow[rr] & & 23 \arrow[rr] \arrow[dl] \arrow[dr] & & 26 \\
& & & 21 & & 24 & & 25 &
\end{tikzcd}$$

$$\begin{tikzcd}
& 7 \arrow[d] & & 15 \arrow[dr] & 16 \arrow[d] & 17 \arrow[dl] & & 22 \arrow[d] \\
\text{Časť IV:} & 27 \arrow[r] & 28 \arrow[rr] & & 29 \arrow[rr] & & 30 \arrow[r] & 31
\end{tikzcd}$$

$$\begin{tikzcd}
& 13 \arrow[dr] & & 30 \arrow[dl] & \\
\text{Časť V:} & & 32 \arrow[rr] & & 33
\end{tikzcd}$$

Časť I

Základné pojmy

0. Základné pojmy z logiky a teórie množín

V tejto kapitole zavedieme niektoré základné logické a množinové pojmy a dohodneme sa na štandardnej symbolike, ktorú budeme ďalej používať. Nebudeme však systematicky budovať *axiomatickú teóriu množín*, práve naopak, s množinami budeme narábať skôr intutívne. Čitateľ, ktorý základné množinové pojmy ovláda, môže túto kapitolu vynechať, prípadne ju len letmo prelistovať, aby sa oboznámil s našou terminológiou a symbolikou.

0.1 Logické spojky a kvantifikátory

Kvôli prehľadnosti budeme niektoré matematické tvrdenia zapisovať v symbolickej podobe ako *matematické formuly*. S príkladmi rôznych formúl sa ešte stretneme. V tejto chvíli sa zameriame len na spôsob, ako možno z daných tvrdení či formúl tvoriť nové pomocou *logických spojok* a *kvantifikátorov*.

Nech P, Q sú ľubovoľné tvrdenia.

(a) Tvrdenie, ktoré je pravdivé práve vtedy, keď tvrdenie P je nepravdivé, nazývame *negáciou* tvrdenia P, značíme ho $\neg P$ a čítame ho „nie P“, prípadne „non P“.

(b) Tvrdenie, ktoré je pravdivé práve vtedy, keď sú pravdivé obe tvrdenia P, Q, nazývame *konjunkciou* alebo *logickým súčinom* tvrdení P, Q, značíme ho $P \mathbin{\&} Q$ a čítame „P a zároveň Q“, krátko len „P a Q“, prípadne „P et Q“.

(c) Tvrdenie, ktoré je pravdivé práve vtedy, keď je pravdivé aspoň jedno z tvrdení P, Q, nazývame *alternatívou* alebo *disjunkciou* či *logickým súčtom* tvrdení P, Q, značíme ho $P \vee Q$, a čítame „P alebo Q“, prípadne „P vel Q“.

(d) Tvrdenie $\neg P \vee Q$ skrátene označujeme $P \Rightarrow Q$ a nazývame ho *implikáciou* tvrdení P, Q. Výraz $P \Rightarrow Q$ čítame „ak P, tak Q“ alebo „z P vyplýva Q“, prípadne „P implikuje Q“. Tvrdenie P nazývame *predpokladom* a tvrdenie Q *záverom* implikácie $P \Rightarrow Q$. Uvedomte si, že implikácia $P \Rightarrow Q$ je nepravdivá jedine v tom prípade, ak predpoklad P je pravdivý a záver Q je nepravdivý.

(e) Tvrdenie $(P \Rightarrow Q) \mathbin{\&} (Q \Rightarrow P)$ skrátene označujeme $P \Leftrightarrow Q$ a nazývame ho *ekvivalenciou* tvrdení P, Q. Výraz $P \Leftrightarrow Q$ čítame „P práve vtedy, keď Q“, prípadne „P je ekvivalentné s Q“. Zrejme ekvivalencia

$P \Leftrightarrow Q$ je pravdivá vtedy a len vtedy, keď tvrdenia P, Q sú zároveň obe pravdivé alebo zároveň obe nepravdivé.

Znaky $\neg$, &, $\vee$, $\Rightarrow$, $\Leftrightarrow$ nazývame *logickými spojkami*. V literatúre sa možno tiež stretnúť s označením P', $-P$ alebo $\sim P$ pre negáciu, $P \wedge Q$ pre konjunkciu, $P \to Q$ alebo $P \supset Q$ pre implikáciu a $P \leftrightarrow Q$ alebo $P \equiv Q$ pre ekvivalenciu.

Okrem tvrdení zapisujeme formulami aj vlastnosti objektov a vzťahy medzi nimi. Na tento účel používame formuly s *voľnými premennými*. Označujeme ich $P(x)$, $Q(x,y)$, $R(x_1,\dots,x_n)$ a pod. Dosadením konkrétnych objektov do formúl namiesto voľných premenných dostávame tvrdenia. Napríklad, ak $Q(x,y)$ je formula s voľnými premennými x, y a a, b sú nejaké objekty, tak $Q(a,b)$ je tvrdenie, ktoré je pravdivé práve vtedy, keď sa objekty a, b nachádzajú vo vzťahu označenom formulou Q.

Zrejme aj na formuly s voľnými premennými možno aplikovať logické spojky, ktoré si pritom zachovajú svoj doterajší význam. Popri logických spojkách možno z formúl tvoriť nové formuly či tvrdenia aj pomocou *kvantifikátorov*.

Nech $P(x)$ je ľubovoľná formula.

(a) Tvrdenie „existuje x také, že $P(x)$" skrátene zapisujeme $(\exists\, x)P(x)$.

(b) Tvrdenie „pre každé (pre všetky) x platí $P(x)$" skrátene zapisujeme $(\forall\, x)P(x)$.

Znaky $\exists$ resp. $\forall$ sú *kvantifikátory*; $\exists$ nazývame *existenčný* a $\forall$ *univerzálny* alebo tiež *všeobecný kvantifikátor*. Zrejme premenná x už nie je vo formulách $(\forall\, x)P(x)$ a $(\exists\, x)P(x)$ voľná ale *viazaná*; ak x je jediná voľná premenná vo formule $P(x)$, tak $(\forall\, x)P(x)$ a $(\exists\, x)P(x)$ sú tvrdenia. Oba uvedené kvantifikátory sú zviazané pravidlami negácie kvantifikovaných formúl:

$$\neg(\exists\, x)P(x) \Leftrightarrow (\forall\, x)\neg P(x),$$
$$\neg(\forall\, x)P(x) \Leftrightarrow (\exists\, x)\neg P(x).$$

Pomocou existenčného a univerzálneho kvantifikátora už vieme vyjadriť i *kvantifikátor jednoznačnej existencie*. Ak $P(x)$ je nejaká vlastnosť, tak tvrdenie „existuje práve jedno x také, že $P(x)$", t. j. tvrdenie

$$(\exists\, x)(P(x)\ \&\ (\forall\, y)(P(y) \Rightarrow y = x)),$$

skrátene zapisujeme v tvare $(\exists!\, x)P(x)$. Toto tvrdenie je zrejme ekvivalentné s tvrdením

$$(\exists\, x)(\forall\, y)(P(y) \Leftrightarrow y = x).$$

0.2 Množiny

Pod *množinou* rozumieme ľubovoľné jednoznačne vymedzené zoskupenie nejakých (často i značne rôznorodých) objektov – *prvkov množiny* – chápané ako jediný objekt. Množiny budeme väčšinou značiť veľkými latinskými písmenami, ich prvky malými písmenami.

Tvrdenie „objekt x je prvkom množiny X“, zapisujeme $x \in X$; hovoríme tiež, že x *patrí* do množiny X. Tvrdenie „objekt x nie je prvkom množiny X“, t. j. x nepatrí do množiny X, zapisujeme $x \notin X$.

Množina je jednoznačne zadaná zoskupením svojich prvkov. Preto dve množiny, nezávisle od spôsobu ich zadania, považujeme za totožné, ak majú tie isté prvky. Pre ľubovoľné množiny X, Y teda platí

$$X = Y \;\Leftrightarrow\; (\forall\, x)(x \in X \;\Leftrightarrow\; x \in Y).$$

Túto vlastnosť množín nazývame *extenzionalitou*.

Hovoríme, že množina X je *podmnožinou* množiny Y, označenie $X \subseteq Y$, ak každý prvok množiny X patrí aj do množiny Y, t. j.

$$X \subseteq Y \;\Leftrightarrow\; (\forall\, x)(x \in X \;\Rightarrow\; x \in Y).$$

Vzťah $\subseteq$ nazývame *vzťahom inklúzie* Extenzionalitu množín teraz možno skrátene vyjadriť v tvare konjunkcie dvoch inklúzií

$$X = Y \;\Leftrightarrow\; X \subseteq Y \;\&\; Y \subseteq X.$$

Kvantifikácie uvedené v predchádzajúcom paragrafe sa nazývajú *neohraničené*, lebo oblasť pôsobnosti kvantifikátorov v nich nebola nijako ohraničená. V matematike (i v bežnom živote) sa však častejšie vyskytujú *ohraničené kvantifikácie*, v ktorých je oblasť pôsobnosti príslušného kvantifikátora ohraničená nejakou množinou X. Ide o kvantifikácie tvaru $(\exists\, x \in X)$, $(\forall\, x \in X)$ a $(\exists!\, x \in X)$, ktoré čítame postupne „existuje x z množiny X“, „pre každé (pre všetky) x z množiny X“, resp. „existuje práve jedno (jediné) x z množiny X“. Tieto kvantifikácie možno vyjadriť pomocou neohraničených kvantifikácii nasledujúcim spôsobom: Ak $P(x)$ je ľubovoľná vlastnosť a X je množina, kladieme

$$\begin{aligned}
(\exists\, x \in X)P(x) \;&\Leftrightarrow\; (\exists\, x)(x \in X \;\&\; P(x)),\\
(\forall\, x \in X)P(x) \;&\Leftrightarrow\; (\forall\, x)(x \in X \;\Rightarrow\; P(x)),\\
(\exists!\, x \in X)P(x) \;&\Leftrightarrow\; \big(\exists\, x \in X\big)\big(P(x) \;\&\; (\forall\, y \in X)(P(y) \;\Rightarrow\; y = x)\big).
\end{aligned}$$

V poslednom prípade môžeme tiež použiť vyjadrenie

$$(\exists!\, x \in X)P(x) \;\Leftrightarrow\; (\exists\, x \in X)(\forall\, y \in X)(P(y) \;\Leftrightarrow\; y = x).$$

Množinu nazývame *konečnou*, ak ju možno zadať vymenovaním všetkých jej prvkov. Ak X je konečná množina a $x_1, x_2, \ldots, x_n$ sú všetky jej prvky, píšeme

$$X = \{x_1, x_2, \ldots, x_n\}.$$

Z extenzionality potom vyplýva, že nezáleží na poradí vymenovania prvkov množiny X. Taktiež sa môže stať, že X má menej než n prvkov – v takom prípade sa niektoré z prvkov $x_1, \ldots, x_n$ opakujú a v zápise množiny X môžeme (no nemusíme) opakujúce sa prvky až na jeden z nich vynechať. Napríklad $\{x, y\} = \{y, x\}$, a ak $x = y$, tak $\{x, y\} = \{x\} = \{y\}$. Okrem množín, ktoré majú nejaké prvky, zavádzame aj tzv. *prázdnu množinu* $\emptyset$, ktorá neobsahuje nijaký prvok. Z extenzionality vyplýva, že prázdna množina je touto podmienkou jednoznačne určená.

Popri konečných množinách však v matematike často pracujeme i s *nekonečnými* množinami, t. j. takými, ktoré nemožno zadať vymenovaním všetkých ich jednotlivých prvkov. Takéto množiny zvykneme zadávať nejakou *charakteristickou vlastnosťou.* Ak $P(x)$ je nejaká vlastnosť, píšeme

$$X = \{x;\ P(x)\},$$

čím myslíme, že pre ľubovoľné x platí $x \in X$ práve vtedy, keď x spĺňa $P(x)$. Z extenzionality vyplýva, že takto definovaná množina X je určená jednoznačne. Napríklad vlastnosťou „x je párne celé číslo“ je určená množina všetkých párnych celých čísel.

Poznamenajme, že z rovnosti $X = \{x;\ P(x)\}$ ešte nijako nevyplýva, že množina X je nekonečná – rovnako dobre môže byť aj konečná, dokonca prázdna.

Na tomto mieste je potrebné poznamenať, že uvedený princíp, ktorý nám umožňuje zadávať množiny akýmikoľvek vlastnosťami ich prvkov, vedie k logickým sporom, a je preto v uvedenej intuitívnej a neobmedzenej podobe nepoužiteľný. Keďže sa však nehodláme púšťať do jeho upresňovania, čo by si vyžiadalo vybudovať základy axiomatickej teórie množín, nezostáva nám než čitateľovi vopred zaručiť, že všetky prípady použitia tohto princípu, ktoré sa v tomto texte vyskytnú, budú plne legálne z hľadiska teórie množín, a požiadať ho o dôveru. Zatiaľ stačí, ak prezradíme, že všetky množiny netvoria množinu, t. j. neexistuje množina všetkých množín. To znamená, že vlastnosťou „x je množina“ nie je vymedzená nijaká množina.

Najčastejšie budeme spomínaný princíp používať na vymedzovanie podmnožín nejakej vopred danej *množiny* pomocou vlastností popísaných *matematickými formulami.* Ak M je množina a $P(x)$ je nejaká (matematická) vlastnosť, tak existuje množina X všetkých tých prvkov x množiny M, ktoré majú vlastnosť $P(x)$, t. j. množina

$$X = \{x \in M;\ P(x)\} = \{x;\ x \in M \ \&\ P(x)\}.$$

Nech X, Y sú ľubovoľné množiny. *Prienikom*, *zjednotením*, a *rozdielom* množín X, Y nazývame porade nasledujúce množiny:

$$\begin{aligned} X \cap Y &= \{x;\ x \in X \ \&\ x \in Y\}, \\ X \cup Y &= \{x;\ x \in X \ \vee\ x \in Y\}, \\ X \smallsetminus Y &= \{x;\ x \in X \ \&\ x \notin Y\}. \end{aligned}$$

Množiny X, Y nazývame *disjunktné*, ak $X \cap Y = \emptyset$. Čitateľovi prenechávame, aby si sám premyslel základné vlastnosti uvedených množinových operácií.

Pod usporiadanou dvojicou objektov x, y rozumieme objekt označovaný (x, y), taký, že pre všetky x, y, u, v platí:

$$(x, y) = (u, v) \ \Leftrightarrow\ (x = u \ \&\ y = v).$$

Uvedomme si, že nepotrebujeme vedieť, čo je „naozaj“ usporiadaná dvojica (x, y), dôležitá je len uvedená vlastnosť. Analogicky zavádzame pre ľubovoľné celé číslo $n \geq 2$ usporiadanú n-ticu $(x_1, \ldots, x_n)$ tak, že pre všetky $x_1, \ldots, x_n$, $y_1, \ldots, y_n$ platí

$$(x_1, \ldots, x_n) = (y_1, \ldots, y_n) \ \Leftrightarrow\ (x_1 = y_1 \ \&\ \ldots \ \&\ x_n = y_n).$$

Množiny

$$\begin{aligned} X \times Y &= \{(x, y);\ x \in X \ \&\ y \in Y\}, \\ X_1 \times \ldots \times X_n &= \{(x_1, \ldots, x_n);\ x_1 \in X_1 \ \&\ \ldots \ \&\ x_n \in X_n\} \end{aligned}$$

nazývame *karteziánskym súčinom* množín X, Y, resp. množín $X_1, \ldots, X_n$. V prípade, že $X_1 = \cdots = X_n = X$, píšeme

$$X_1 \times \ldots \times X_n = X^n.$$

Pre úplnosť ešte kladieme

$$X^1 = X, \qquad X^0 = \{\emptyset\}.$$

X^n nazývame *n-tou karteziánskou mocninou* množiny X. *Počet prvkov* konečnej množiny X budeme značiť $\# X$. Taktiež prázdna množina je konečná a platí $\# \emptyset = 0$. Pre nekonečnú množinu X píšeme $\# X = \infty$. Zrejme pre ľubovoľné konečné množiny X, Y platí

$$\begin{aligned} \# (X \cup Y) &= \# X + \# Y - \# (X \cap Y), \\ \# (X \times Y) &= \# X \cdot \# Y. \end{aligned}$$

Z poslednej rovnosti vyplýva, že

$$\#\big(X^n\big) = (\# X)^n$$

pre každé celé číslo $n \geq 0$ a konečnú množinu X.

0.3 Zobrazenia

Zobrazením alebo tiež *funkciou* z množiny X do množiny Y rozumieme ľubovoľný predpis, ktorý každému prvku x množiny X priradí jednoznačne určený prvok y množiny Y. Zápis $f\colon X \to Y$ označuje, že f je zobrazenie (funkcia) z X do Y. Ten jednoznačne určený prvok $y \in Y$, ktorý zobrazenie f priradí prvku $x \in X$, budeme značiť $f(x)$, prípadne len fx alebo f_x. Vo vzťahu $y = f(x)$ nazývame x *nezávisle premennou* alebo *argumentom* a y *závisle premennou* alebo *funkčnou hodnotou* funkcie f. Píšeme tiež $f\colon x \mapsto y$.

Dve zobrazenia $f, g\colon X \to Y$ *sa rovnajú*, ak pre každé $x \in X$ platí $f(x) = g(x)$.

Množinu všetkých zobrazení z množiny X do množiny Y budeme označovať Y^X; teda

$$Y^X = \{f;\ f\colon X \to Y\}.$$

Toto označenie je motivované vzorcom pre počet prvkov množiny Y^X. Pre konečné množiny X, Y totiž platí

$$\#\left(Y^X\right) = (\# Y)^{(\# X)}.$$

(Samostatne si rozmyslite prečo!)

Zobrazenie $f\colon X \to X$ sa nazýva *transformáciou* množiny X alebo tiež *unárnou* (t. j. jednomiestnou) *operáciou* na množine X.

Zobrazenie $f\colon X \to Y$ sa nazýva *prosté* alebo tiež *injektívne* či *injekcia*, ak rôznym prvkom $x_1, x_2 \in X$ priraďuje rôzne prvky $f(x_1), f(x_2) \in Y$, t. j. ak platí

$$(\forall\, x_1, x_2 \in X)(x_1 \neq x_2 \;\Rightarrow\; f(x_1) \neq f(x_2)).$$

Uvedenú podmienku možno ekvivalentne vyjadriť v tvare

$$(\forall\, x_1, x_2 \in X)(f(x_1) = f(x_2) \;\Rightarrow\; x_1 = x_2).$$

Zobrazenie $f\colon X \to Y$ sa nazýva zobrazenie *na množinu* Y alebo tiež *surjektívne* či *surjekcia*, ak na každý prvok množiny Y sa zobrazí nejaký prvok množiny X, t. j. ak platí

$$(\forall\, y \in Y)(\exists\, x \in X)(y = f(x)).$$

Hovoríme, že $f\colon X \to Y$ je *prosté zobrazenie* X *na* Y alebo tiež *bijektívne zobrazenie* či *bijekcia*, ak f je zároveň prosté a na, t. j. injektívne i surjektívne. Ešte inak to môžeme vyjadriť podmienkou

$$(\forall\, y \in Y)(\exists!\, x \in X)(y = f(x)).$$

Namiesto uvedených pojmov niekedy tiež hovoríme, že f je *vzájomne jednoznačné zobrazenie množiny* X *na množinu* Y.

Ak $f\colon X \to Y$ je bijekcia, tak existuje jednoznačne určené zobrazenie $g\colon Y \to X$, ktoré každému $y \in Y$ priradí ten jediný prvok $x \in X$, pre ktorý platí $y = f(x)$. Toto zobrazenie nazývame *inverzným zobrazením* k zobrazeniu f a označujeme ho f^{-1}. Zrejme $f^{-1}\colon Y \to X$ je tiež bijekcia a pre všetky $x \in X$, $y \in Y$ platí

$$f^{-1}\big(f(x)\big) = x, \qquad f\big(f^{-1}(y)\big) = y.$$

Nech $f\colon X \to Y$, $g\colon Y \to Z$ sú zobrazenia. *Kompozíciou* zobrazení f, g alebo aj *zloženým zobrazením* z f a g rozumieme zobrazenie označené ako $g \circ f\colon X \to Z$, dané pre každé $x \in X$ predpisom

$$(g \circ f)(x) = g(f(x)).$$

Zložené zobrazenie možno znázorniť pomocou tzv. *komutatívneho diagramu*

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "g \circ f"'] & Y \arrow[d, "g"] \\ & Z \end{tikzcd}$$

(Všimnite si, že zobrazenie $g \circ f$ zapisujeme „v obrátenom poradí" – najprv totiž na prvok x aplikujeme f a až potom g. Núti nás k tomu zaužívaná konvencia, podľa ktorej argument x píšeme napravo od funkcie f. Poznamenajme, že niektorí autori dávajú prednosť „prirodzenému poradiu" a kompozíciu zobrazení $f\colon X \to Y$, $g\colon Y \to Z$, zapisujú ako $f \circ g$. Kvôli tomu však opúšťajú spomínanú konvenciu a namiesto $f(x)$ píšu xf. V tomto duchu fungujú napr. niektoré kalkulačky.)

Skladanie zobrazení je *asociatívne* v nasledujúcom zmysle: ak $f\colon X \to Y$, $g\colon Y \to Z$ a $h\colon Z \to W$ sú zobrazenia, tak

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Ľahko totiž nahliadneme, že obe zobrazenia priradia prvku $x \in X$ prvok $h(g(f(x))) \in W$.

Na každej množine X máme definované *identické zobrazenie* $\mathrm{id}_X\colon X \to X$, nazývané tiež *identita na* X, také, že

$$\mathrm{id}_X(x) = x$$

pre každé $x \in X$. Zrejme id_X je bijekcia pre každé X, a pre ľubovoľné zobrazenie $f\colon X \to Y$ platí

$$f \circ \mathrm{id}_X = f = \mathrm{id}_Y \circ f.$$

Pre $f\colon X \to X$ kladieme $f^2 = f \circ f$, $f^3 = f \circ f \circ f$, atď. Kvôli úplnosti definujeme aj $f^1 = f$, $f^0 = \mathrm{id}_X$. Zobrazenie f^n nazývame n-tou *iteráciou zobrazenia* f.

Ak $f\colon X \to Y$ je bijekcia, tak k nej inverzné zobrazenie $f^{-1}\colon Y \to X$ teraz môžeme charakterizovať rovnosťami

$$f^{-1} \circ f = \mathrm{id}_X, \qquad f \circ f^{-1} = \mathrm{id}_Y .$$

Čitateľ sám ľahko nahliadne, že pre ľubovoľné zobrazenia $f\colon X \to Y$, $g\colon Y \to Z$ platí:

(a) Ak f, g sú injektívne, tak aj $g \circ f$ je injektívne.
(b) Ak f, g sú surjektívne, tak aj $g \circ f$ je surjektívne.
(c) Ak f, g sú bijektívne, tak aj $g \circ f$ je bijektívne.
(d) Ak $g \circ f$ je injektívne, tak aj f je injektívne.
(e) Ak $g \circ f$ je surjektívne, tak aj g je surjektívne.
(f) Ak $g \circ f$ je bijektívne, tak f je injektívne a g je surjektívne.

Podmienka (c) nás oprávňuje zaviesť pre bijekcie $f\colon X \to X$ aj záporné iterácie

$$f^{-n} = \left(f^{-1}\right)^n = \left(f^n\right)^{-1}.$$

Nech $f\colon X \to Y$ je nejaké zobrazenie a $A \subseteq X$. *Zúžením zobrazenia* f na množinu A nazývame zobrazenie $f \restriction A\colon A \to Y$ také, že

$$(f \restriction A)(x) = f(x)$$

pre každé $x \in A$. *Obrazom množiny* A v zobrazení f nazývame množinu

$$f(A) = \{f(x);\, x \in A\} \subseteq Y.$$

Špeciálne, množinu $f(X)$ nazývame *obrazom zobrazenia* f a značíme ju

$$\mathrm{Im}\, f = f(X) = \{f(x);\, x \in X\}.$$

Pre $f\colon X \to Y$ a $A \subseteq X$ platí $\mathrm{Im}(f \restriction A) = f(A)$; zrejme f je surjekcia práve vtedy, keď $\mathrm{Im}\, f = Y$.

Podobne, *vzorom množiny* $B \subseteq Y$ v zobrazení $f\colon X \to Y$ nazývame množinu

$$f^{-1}(B) = \{x \in X;\ f(x) \in B\} \subseteq X.$$

Pre ľubovoľné $A \subseteq X$, $B \subseteq Y$ možno jednoducho overiť inklúzie

$$A \subseteq f^{-1}\big(f(A)\big), \qquad f\big(f^{-1}(B)\big) \subseteq B.$$

0.4 Binárne operácie

Ak X, Y, Z sú množiny, tak zobrazenie $f\colon X \times Y \to Z$ nazývame *binárnou* (t. j. dvojmiestnou) *operáciou na množinách X, Y s hodnotami v množine Z*. Binárne operácie väčšinou označujeme znakmi umiestňovanými medzi hodnoty argumentov, ako napr $+$, $\cdot$, $\circ$, $*$ a pod. Hodnotu takej operácie na dvojici prvkov $x \in X$, $y \in Y$ potom označujeme $x + y$, $x \cdot y$ (prípadne len xy), $x \circ y$, $x * y$ a pod.

Podobným spôsobom možno zaviesť aj n-miestne operácie $X_1 \times \ldots \times X_n \to Y$, prípadne $X^n \to Y$, či $X^n \to X$ pre ľubovoľné celé číslo $n \geq 0$.

Najčastejšie budeme pracovať s binárnymi operáciami tvaru $f\colon X \times X \to X$, ktoré nazývame jednoducho *binárnymi operáciami na množine X*.

Binárna operácia $*$ na množine X sa nazýva *asociatívna*, ak pre všetky $x, y, z \in X$ platí

$$x * (y * z) = (x * y) * z.$$

Asociativita operácie nám dovoľuje vynechávať zátvorky a písať len $x * y * z$. Podobne si možno počínať i v prípade viacerých argumentov.

Binárna operácia $*$ na množine X sa nazýva *komutatívna*, ak pre všetky $x, y \in X$ platí

$$x * y = y * x.$$

Prvok $e \in X$ sa nazýva *neutrálny prvok* binárnej operácie $*$ na množine X, ak pre všetky $x \in X$ platí

$$x * e = e * x = x.$$

Zrejme neutrálny prvok operácie $*$ (ak existuje) je určený jednoznačne. Keby totiž $e_1, e_2 \in X$ boli dva neutrálne prvky, tak nevyhnutne

$$e_1 = e_1 * e_2 = e_2.$$

Ak binárna operácia $*$ na množine X má neutrálny prvok e a k danému prvku $x \in X$ existuje prvok $y \in X$ taký, že

$$x * y = y * x = e,$$

hovoríme, že y je *inverzný prvok* k prvku x. Ak $*$ je *asociatívna* binárna operácia na X, tak aj inverzný prvok k danému prvku $x \in X$ (pokiaľ existuje) je určený jednoznačne. Keby totiž y_1, y_2 boli dva inverzné prvky k x, tak

$$y_1 = y_1 * e = y_1 * (x * y_2) = (y_1 * x) * y_2 = e * y_2 = y_2.$$

Napríklad pre ľubovoľnú množinu X kompozícia $\circ$ je asociatívna binárna operácia na množine X^X všetkých transformácií množiny X. Zrejme ak $\# X \geq 2$, tak táto operácia nie je komutatívna. Identické zobrazenie $\mathrm{id}_X \in X^X$ je neutrálnym prvkom operácie $\circ$. K danému zobrazeniu $f \in X^X$ existuje inverzný prvok práve vtedy, keď f je bijekcia; v tom prípade je ním inverzné zobrazenie $f^{-1} \in X^X$.

Binárnu operáciu $*$ na konečnej množine X možno zadať pomocou tzv. *multiplikatívnej tabuľky*, ktorej stĺpce i riadky sú označené prvkami množiny X. Do poľa tabuľky ležiaceho v priesečníku x-tého riadku a y-tého stĺpca vpíšeme hodnotu $x * y$.

Napr. tabuľkami

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

sú zadané dve asociatívne a komutatívne operácie $+$ a $\cdot$ na množine $\{0, 1, 2, 3, 4\}$. Ďalej 0 je neutrálny prvok operácie $+$ a 1 je neutrálny prvok operácie $\cdot$. Navyše ku každému prvku a tejto množiny existuje inverzný prvok vzhľadom na operáciu $+$: k prvkom 0, 1, 2, 3, 4 sú to postupne prvky 0, 4, 3, 2, 1. Pokiaľ ide o operáciu $\cdot$, vidíme, že k prvku 0 neexistuje inverzný prvok; k prvkom 1, 2, 3 však inverzné prvky existujú: sú to postupne 1, 3, 2.

Komutativitu binárnej operácie možno ľahko nahliadnuť z jej multiplikatívnej tabuľky – prejaví sa symetriou tabuľky podľa hlavnej diagonály spájajúcej ľavý horný a pravý dolný roh. Taktiež neutrálny prvok možno odhaliť na prvý pohľad, lebo v jeho riadku i stĺpci sa zreprodukuje riadok resp. stĺpec zo záhlavia tabuľky. Ak už poznáme neutrálny prvok, možno overiť aj existenciu inverzného k danému – treba nájsť neutrálny prvok v riadku i v stĺpci daného prvku. Ak sa nám to podarí pre dvojicu polí tabuľky, ktoré ležia v stĺpci resp. riadku toho istého prvku, tak ide o hľadaný inverzný prvok. Asociatívnosť, žiaľ, tak jednoducho nahliadnuť nemožno.

0.5 Permutácie

Kým znalosť predchádzajúcich paragrafov je nevyhnutným predpokladom, aby čitateľ mohol začať so štúdiom kapitoly 1, tento paragraf budeme potrebovať až neskôr, keď začneme preberať determinanty.

Nech X je ľubovoľná množina. *Permutáciou* množiny X rozumieme ľubovoľné bijektívne zobrazenie $\sigma\colon X \to X$. Množinu všetkých permutácií množiny X značíme $\mathcal{S}(X)$. Ak X je konečná množina, tak počet prvkov množiny $\mathcal{S}(X)$ je daný známym vzťahom

$$\#\,\mathcal{S}(X) = (\#\,X)!\,,$$

kde $n! = 1 \cdot 2 \cdot \ldots \cdot n$ je *faktoriál* prirodzeného čísla n (pritom $0! = 1! = 1$).

Uvedomme si, že transformácia $f\colon X \to X$ *konečnej* množiny X je injektívna práve vtedy, keď je surjektívna. Jedna i druhá podmienka totiž hovorí, že množina $f(X) \subseteq X$ má rovnaký počet prvkov ako X. Teda už jedna z uvedených podmienok je postačujúca na to, aby f bola permutáciou konečnej množiny X.

Keďže zloženie $\sigma \circ \tau$ dvoch permutácií $\sigma, \tau \in \mathcal{S}(X)$ dáva opäť permutáciu množiny X, kompozícia $\circ$ je asociatívna binárna operácia na množine $\mathcal{S}(X)$. Ľahko sa možno presvedčiť, že – okrem prípadu, keď $\#\,X \le 2$, – táto operácia nie je komutatívna. Zrejme $\mathrm{id}_X \in \mathcal{S}(X)$ je neutrálny prvok tejto operácie a inverzným prvkom k permutácii $\sigma \in \mathcal{S}(X)$ je inverzná permutácia $\sigma^{-1} \in \mathcal{S}(X)$.

Pre $X = \{1, 2, \ldots, n\}$ namiesto $\mathcal{S}(X)$ píšeme $\mathcal{S}_n$. Permutáciu $\sigma \in \mathcal{S}_n$ zvyčajne zapisujeme v tvare

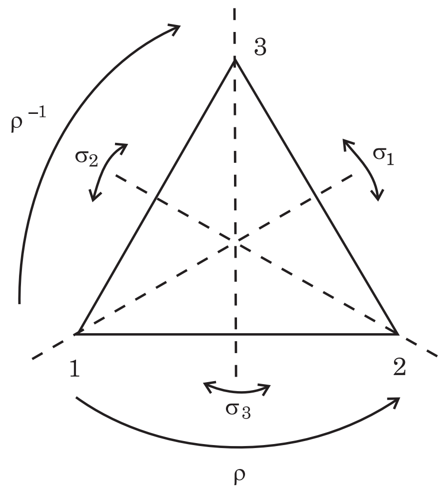
$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}.$$

Prvky množiny $\mathcal{S}_3$, t. j. permutácie množiny $\{1, 2, 3\}$, si môžeme predstaviť ako symetrie rovnostranného trojuholníka s vrcholmi označenými číslami 1, 2, 3.

Ak si identickú permutáciu tejto množiny označíme ako ι, otočenia okolo ťažiska trojuholníka proti smeru resp. v smere hodinových ručičiek o uhol $\pi/3$ ako ϱ resp. ϱ^{-1}, a osovú súmernosť podľa osi prechádzajúcej i-tým vrcholom a stredom protiľahlej strany ako σ_i, pre $i = 1, 2, 3$, tak množina permutácií $\mathcal{S}_3$ bude pozostávať z permutácií

$$\iota = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \qquad \varrho = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \varrho^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$



Obr. 0.1. Symetrie rovnostranného trojuholníka

Multiplikatívna tabuľka binárnej operácie $\circ$ na množine $\mathcal{S}_3$ vyzerá takto:

$\circ$	ι	ϱ	ϱ^{-1}	σ_1	σ_2	σ_3
ι	ι	ϱ	ϱ^{-1}	σ_1	σ_2	σ_3
ϱ	ϱ	ϱ^{-1}	ι	σ_3	σ_1	σ_2
ϱ^{-1}	ϱ^{-1}	ι	ϱ	σ_2	σ_3	σ_1
σ_1	σ_1	σ_2	σ_3	ι	ϱ	ϱ^{-1}
σ_2	σ_2	σ_3	σ_1	ϱ^{-1}	ι	ϱ
σ_3	σ_3	σ_1	σ_2	ϱ	ϱ^{-1}	ι

Permutáciu $\sigma \in \mathcal{S}(X)$ nazývame *transpozíciou*, ak existujú $x, y \in X$ také, že $x \neq y$, $\sigma(x) = y$, $\sigma(y) = x$ a $\sigma(z) = z$ pre každé $z \in X \setminus \{x, y\}$. Inak povedané, transpozícia je výmena dvoch prvkov množiny X.

Zrejme $\sigma_1, \sigma_2, \sigma_3 \in \mathcal{S}_3$ sú transpozície.

Z názoru je zrejmé (dôkaz je v cvičení 0.14), že každú permutáciu σ konečnej množiny X možno získať postupnými výmenami dvojíc prvkov, teda každá taká permutácia je kompozíciou transpozícií. Tento rozklad na transpozície nie je jednoznačný: napr. $\iota \in \mathcal{S}_3$ možno zapísať ako ι, t. j. kompozíciu 0 transpozícií, a taktiež ako $\iota = \sigma_1 \circ \sigma_1 = \sigma_2 \circ \sigma_2 = \sigma_3 \circ \sigma_3$, t. j. aspoň troma ďalšími spôsobmi ako kompozíciu dvoch transpozícií.

Dĺžkou permutácie σ konečnej množiny X nazveme najmenší počet transpozícií, na kompozíciu ktorých možno σ rozložiť, a označíme ju $|\sigma|$. Samotná dĺžka $|\sigma|$ nie je dôležitá, význam má len parita tohto čísla, t. j. vlastne výraz $\operatorname{sgn} \sigma = (-1)^{|\sigma|}$, ktorý nazývame *znakom*, prípadne *znamienkom permutácie* σ.

Permutácia σ konečnej množiny X sa nazýva *párna* resp. *nepárna*, ak číslo $|\sigma|$ je párne resp. nepárne, t. j. ak jej znak je 1 resp. -1.

Z nasledujúcej vety vyplýva, že pri určovaní znamienka permutácie σ môžeme použiť jej *ľubovoľný* rozklad na transpozície $\sigma = \tau_1 \circ \ldots \circ \tau_k$ a nemusíme sa starať o to, či tento rozklad je naozaj najkratší – pre ľubovoľný taký rozklad totiž platí

$$(-1)^{|\sigma|} = (-1)^k.$$

0.5.1. Veta. *Nech X je konečná množina. Potom pre ľubovoľné $\sigma, \tau \in \mathcal{S}(X)$ platí*

$$(-1)^{|\sigma\circ\tau|} = (-1)^{|\sigma|} \cdot (-1)^{|\tau|}.$$

Dôkaz. Zrejme stačí dokázať uvedenú rovnosť pre prípad, keď τ je transpozícia a $X = \{1, 2, \ldots, n\}$.

Pre každé $\sigma \in \mathcal{S}_n$ označme $p(\sigma)$ súčin všetkých rozdielov tvaru $\sigma(j)-\sigma(i)$, kde $1 \le i < j \le n$. Zrejme pre všetky $\sigma \in \mathcal{S}_n$ majú výrazy $p(\sigma)$ rovnakú absolútnu hodnotu a líšia sa nanajvýš znamienkom. Toto znamienko závisí od parity počtu záporných členov v súčine $p(\sigma)$. Člen $\sigma(j)-\sigma(i)$ je záporný práve vtedy, keď $i < j$ a $\sigma(i) > \sigma(j)$, – každú takú dvojicu (i, j) nazývame *inverziou* permutácie σ. Identita id_X má 0 inverzií a $p(\mathrm{id}_X) = 1^{n-1} \cdot 2^{n-2} \cdot \ldots \cdot (n-1)^1 = 1! \cdot 2! \cdot \ldots \cdot (n-1)! > 0$.

Stačí teda dokázať, že počet inverzií permutácií σ a $\sigma \circ \tau$ sa líši o nepárnu hodnotu. Nech $1 \le k < l \le n$ sú tie dva prvky, ktoré vymieňa transpozícia τ. Potom

$$\begin{aligned}
\sigma &= \begin{pmatrix} 1 & \ldots & k & \ldots & l & \ldots & n \\ \sigma(1) & \ldots & \sigma(k) & \ldots & \sigma(l) & \ldots & \sigma(n) \end{pmatrix}, \\
\sigma \circ \tau &= \begin{pmatrix} 1 & \ldots & k & \ldots & l & \ldots & n \\ \sigma(1) & \ldots & \sigma(l) & \ldots & \sigma(k) & \ldots & \sigma(n) \end{pmatrix}.
\end{aligned}$$

Inverzie (i, j) permutácie σ, v ktorých nevystupuje k ani l, sú tiež inverziami permutácie $\sigma \circ \tau$. Inverzie, v ktorých vystupujú prvky i, k, a i, l, kde $i \ne k, l$, alebo obe súčasne vzniknú alebo súčasne zaniknú v $\sigma \circ \tau$ oproti σ. Konečne, pokiaľ (k, l) nebola inverziou v σ, stane sa ňou v $\sigma \circ \tau$; pokiaľ ňou bola, táto inverzia v $\sigma \circ \tau$ zanikne. Teda celkový rozdiel počtu inverzií permutácií σ a $\sigma \circ \tau$ je nepárny.

0.6 Ekvivalencie a rozklady

Podobne ako predošlý, i tento paragraf môže čitateľ zatiaľ preskočiť. Jeho znalosť bude potrebná až neskôr, v súvislosti s niektorými otázkami teórie grúp. S pojmom ekvivalencie sa síce stretneme už predtým, dovtedy ho však nebudeme systematicky využívať.

Nech $\sim$ je nejaký dvojmiestny vzťah, do ktorého vstupujú prvky nejakého oboru objektov $\mathcal{M}$ (tento obor môže, ale nemusí byť množinou). Zápisom $x \sim y$ značíme, že prvky $x, y \in \mathcal{M}$ sa nachádzajú vo vzťahu $\sim$; ak sa $x, y \in \mathcal{M}$ nenachádzajú v tomto vzťahu, píšeme $x \not\sim y$.

Hovoríme, že vzťah $\sim$ je na obore $\mathcal{M}$

(a) *reflexívny*, ak pre všetky $x \in \mathcal{M}$ platí $x \sim x$;

(b) *symetrický*, ak pre všetky $x, y \in \mathcal{M}$ platí $x \sim y \;\Rightarrow\; y \sim x$;

(c) *tranzitívny*, ak pre všetky $x, y, z \in \mathcal{M}$ platí $x \sim y \;\&\; y \sim z \;\Rightarrow\; x \sim z$.

Vzťah $\sim$, ktorý je reflexívny, symetrický a tranzitívny na obore $\mathcal{M}$, nazývame *vzťahom ekvivalencie* alebo len krátko *ekvivalenciou* na obore $\mathcal{M}$. Ekvivalencie budeme väčšinou značiť znakmi $\sim$, $\approx$, $\equiv$ a pod.

Každý vzťah ekvivalencie na nejakej množine či obore objektov $\mathcal{M}$ predstavuje isté hľadisko, z ktorého považujeme niektoré prvky z $\mathcal{M}$ za rovnocenné, t. j. ekvivalentné, a iné nie. Napr. na množine všetkých hracích guličiek v danej jamke možno zaviesť vzťah ekvivalencie, v ktorom sa nachádzajú ľubovoľné dve guličky práve vtedy, keď majú rovnakú farbu. Vzťah, v ktorom sa nachádzajú dve takéto guličky práve vtedy, keď majú rovnakú hmotnosť, je iným príkladom ekvivalencie na tejto množine.

Jedným dychom však poznamenajme, že uvedené príklady neslobodno brať príliš vážne, lebo rovnocennosť sa v nich mieša s podobnosťou, – „naozajstné“ ekvivalencie predstavujú len v značne idealizovanom prípade. S reflexívnosťou a symetriou nie je problém, v reálnom živote však zvykne zlyhať tranzitívnosť. Môžeme sa napríklad zhodnúť, že guličky a, b majú rovnakú farbu, a takisto majú rovnakú farbu guličky b, c. No farba guličiek a, c sa nám už rovnakou zdať nemusí. Podobne môžeme v rámci presnosti našich váh dospieť k záveru, že guličky p, q ako aj guličky q, r majú rovnakú hmotnosť. Avšak hmotnosť guličiek p, r sa nám už vážením môže podariť rozlíšiť. Lepším príkladom ekvivalencie je tak vzťah na množine všetkých bankoviek danej meny, v ktorom sa nachádzajú dve bankovky práve vtedy, keď majú rovnakú nominálnu hodnotu.

Na rozdiel od reálneho života sa v matematike nemusíme trápiť podobnými ťažkosťami. Všetky ekvivalencie, s ktorými sa tu stretneme, budú mať v plnej miere všetky tri uvedené vlastnosti. Ešte jeden príklad za všetky: vzťahom

$$x \sim y \;\Leftrightarrow\; |x| = |y|$$

je definovaná ekvivalencia „mať rovnakú absolútnu hodnotu“ na množine $\mathbb{C}$ všetkých komplexných čísel.

Nech $\sim$ je ekvivalencia na množine X. Pre $x \in X$ označme

$$\tilde{x} = \{u \in X;\, u \sim x\}$$

množinu všetkých prvkov $u \in X$ ekvivalentných s x, ktorú nazývame *triedou* alebo *blokom ekvivalencie* prvku x. Zrejme pre ľubovoľné $x \in X$ platí $x \in \tilde{x}$. Ľahko tiež možno dokázať (skúste sami), že

$$x \sim y \;\Leftrightarrow\; \tilde{x} = \tilde{y} \;\Leftrightarrow\; x \in \tilde{y} \;\Leftrightarrow\; y \in \tilde{x} \;\Leftrightarrow\; \tilde{x} \cap \tilde{y} \neq \emptyset$$

pre všetky $x, y \in X$. Množinu

$$X/\!\sim\; = \{\tilde{x};\, x \in X\}$$

všetkých tried ekvivalencie prvkov množiny X nazývame *faktorovou množinou* množiny X podľa ekvivalencie $\sim$. (Podotýkame, že v zhode s paragrafom 0.2 sa každá trieda $\tilde{x}$ nachádza v množine $X/\!\sim$ iba raz, i keď prvkov $y \in X$, pre ktoré platí $\tilde{x} = \tilde{y}$, môže byť mnoho.)

Priradením $x \mapsto \tilde{x}$ je definované surjektívne zobrazenie $X \to X/\!\sim$, ktoré nazývame *prirodzenou* alebo tiež *kanonickou projekciou* množiny X na faktorovú množinu $X/\!\sim$.

Na faktorovú množinu $X/\!\sim$ sa možno dívať dvojakým spôsobom. Jednak ako na výsledok stotožnenia či zlepenia navzájom ekvivalentných prvkov množiny X; v takom prípade sa na bloky $\tilde{x}$ dívame predovšetkým ako na *prvky*, ktoré vznikli „stiahnutím“ celej triedy $\tilde{x}$ do jediného bodu, a vedome si nevšímame fakt, že sú to zároveň množiny. Použitím názvu „faktorová množina“ naznačujeme, že v danej chvíli dávame tomuto pohľadu prednosť. Na druhej strane sa na množinu $X/\!\sim$ možno dívať ako na *rozklad* množiny X na navzájom disjunktné neprázdne množiny $\tilde{x}$.

Rozkladom množiny X nazývame ľubovoľný systém (t. j. množinu) jej neprázdnych podmnožín $\mathcal{R}$ taký, že každý prvok množiny X padne do práve jednej množiny zo systému $\mathcal{R}$. Inými slovami, systém $\mathcal{R}$ neprázdnych podmnožín množiny X je jej rozkladom práve vtedy, keď spĺňa nasledujúce dve podmienky:

(1) zjednotením všetkých množín $A \in \mathcal{R}$ je celá množina X, t. j. $(\forall\, x \in X)(\exists\, A \in \mathcal{R})(x \in A)$;

(2) množiny z $\mathcal{R}$ sú navzájom disjunktné, t. j. $(\forall\, A, B \in \mathcal{R})(A \neq B \Rightarrow A \cap B = \emptyset)$.

Ľahko možno nahliadnuť, že faktorová množina $X/\!\sim$ množiny X podľa ekvivalencie $\sim$ je zároveň rozkladom množiny X, ktorý je tvorený triedami navzájom ekvivalentných prvkov. Taktiež naopak, každý rozklad $\mathcal{R}$ množiny X určuje predpisom

$$x \sim_{\mathcal{R}} y \;\Leftrightarrow\; (\exists\, A \in \mathcal{R})(x, y \in A)$$

ekvivalenciu na množine X. Inak povedané, prvky $x, y \in X$ sú vo vzťahu ekvivalencie určenej rozkladom $\mathcal{R}$ práve vtedy, keď sa nachádzajú v tej istej (jednoznačne určenej) množine z tohto rozkladu. Čitateľovi prenechávame, aby si samostane overil, že takto definovaný vzťah $\sim_{\mathcal{R}}$ je reflexívny, symetrický a tranzitívny, t. j. má všetky tri požadované vlastnosti ekvivalencie, ako aj rovnosť $X/\sim_{\mathcal{R}} = \mathcal{R}$, t. j. že rozklad (faktorová množina) určený ekvivalenciou $\sim_{\mathcal{R}}$ splýva s pôvodným rozkladom $\mathcal{R}$.

0.6.1. Príklad. Rozklad prislúchajúci k spomínanej ekvivalencii $x \sim y \Leftrightarrow |x| = |y|$ na množine $\mathbb{C}$ je vlastne rozkladom komplexnej roviny na navzájom sústredné kružnice so stredom v počiatku 0 a ľubovoľným polomerom $r \geq 0$ (kružnicu s nulovým polomerom prirodzene stotožňujeme s jej stredom).

0.7 O matematických dôkazoch

Matematika je veda vybudovaná prevažne (hoci nie výlučne) *deduktívne*. To znamená, že v tej-ktorej matematickej teórii vychádzame z určitých základných pojmov, ktoré považujeme za intuitívne jasné vďaka istým s nimi spojeným názorným predstavám. Ďalšie pojmy potom definujeme pomocou pojmov základných alebo skôr definovaných. Základné pojmy označujú základné objekty, ktoré tvoria predmet nášho štúdia, alebo určité základné vzťahy medzi nimi. Tieto objekty a vzťahy sú charakterizované istými východzími tvrdeniami, ktorým hovoríme *axiómy*. V najjednoduchších prípadoch je platnosť axióm jasná z názoru, ktorý stojí v pozadí príslušnej teórie. V zložitejších prípadoch však môžu názorné predstavy zlyhať – vtedy sa na axiómy dívame ako na *implicitné definície* základných pojmov. To znamená, že rezignujeme na otázku, čo „naozaj“ označujú základné pojmy. Môžu označovať čokoľvek, čo spĺňa dané axiómy – to je všetko, čo o nich predpokladáme. Zisk z takého prístupu spočíva v *univerzálnosti matematiky* – aj výsledky matematických teórií sa potom vzťahujú na veľmi rôznorodé oblasti reality. Totiž na tie, v ktorých možno interpretovať základné pojmy danej teórie tak, že sú pritom splnené jej axiómy.

Pri deduktívnej výstavbe nejakej teórie vyvodzujeme ďalšie poznatky z jej axióm logickými prostriedkami, t. j. dokazujeme ich. Týmto dokázaným poznatkom hovoríme *vety*, *tvrdenia*, *lemy* a *dôsledky*, čím naznačujeme rôzny stupeň dôležitosti, ktorý im pripisujeme. Názvom *veta* označujeme tie najdôležitejšie z nich, menej dôležité nazývame *tvrdeniami* a tvrdenia pomocného charakteru označujeme ako *lemy*. *Dôsledky*, ako už samotný názov napovedá, pripájame ako bezprostredné dôsledky niektorých viet, tvrdení či liem, pokiaľ ich význam nedosahuje úroveň viet. Poznamenajme, že toto rozdelenie má značne subjektívny charakter a vývoj ho často zvykne prekonať. Mnohé vety časom upadajú do zabudnutia, kým naopak mnohé lemy postupne nadobúdajú na význame.

Základným prostriedkom odvodzovania nových poznatkov v deduktívnej teórii je *dôkaz.* V tomto paragrafe sa veľmi stručne zoznámime s hlavnými typmi matematických dôkazov: s *priamym dôkazom*, s *nepriamym dôkazom* a s *dôkazom sporom.* Uvidíme, že toto rozdelenie tak trochu súvisí so stratégiou vedenia príslušného dôkazu. V nasledujúcom paragrafe sa ešte zoznámime s *dôkazom matematickou indukciou.*

Väčšina matematických tvrdení má tvar implikácie $P \Rightarrow Q$, t. j. tvrdí sa v nich, že z predpokladu P vyplýva záver Q. Pritom predpoklad P je často konjunkciou nejakých dielčích predpokladov, čiže má tvar $P_1 \mathbin{\&} \dots \mathbin{\&} P_n$. Na tomto mieste sa obmedzíme na niekoľko poznámok o dôkazoch tvrdení takéhoto tvaru.

0.7.1. Priamy dôkaz. Pri *priamom dôkaze* implikácie $P \Rightarrow Q$ dokazujeme (či sa aspoň pokúšame dokázať) záver Q z predpokladu P. Spočiatku sa snažíme dokázať priamo záver Q z daných axióm a už skôr dokázaných tvrdení. Postupujeme pri tom tak ďaleko, ako sa len dá, pričom jedným očkom stále poškuľujeme po predpoklade P, či dielčích predpokladoch $P_1, \dots, P_n$. Vo chvíli, keď už nevieme ako ďalej, siahneme po tom z dielčích predpokladov P_i, ktorý nám umožní pohnúť sa dopredu. Opäť postupujeme ďalej a vo vhodnej chvíli zasa použijeme niektorý dielčí predpoklad P_j (nie nevyhnutne rôzny od P_i). Ak sme úspešní, nakoniec sa nám podarí dospieť k záveru Q, čím dôkaz končí. Ak sme neúspešní, musíme to skúsiť inak, prípadne sa zamyslieť nad otázkou, či spomínaná implikácia vôbec platí.

Môže sa stať, že pri našom úspešnom dôkaze sme nepoužili všetky dielčie predpoklady $P_1, \dots, P_n$, ale povedzme prvý a posledný z nich sme nepotrebovali. To znamená, že miesto pôvodného tvrdenia $(P_1 \mathbin{\&} P_2 \mathbin{\&} \dots \mathbin{\&} P_{n-1} \mathbin{\&} P_n) \Rightarrow Q$ sme dokázali *silnejšie* tvrdenie $(P_2 \mathbin{\&} \dots \mathbin{\&} P_{n-1}) \Rightarrow Q$.

0.7.2. Nepriamy dôkaz. Pri *nepriamom dôkaze* implikácie $P \Rightarrow Q$ dokazujeme miesto nej logicky ekvivalentnú tzv. *transponovanú implikáciu* $\neg Q \Rightarrow \neg P$ práve opísanou metódou priameho dôkazu. Za tým účelom býva často užitočné (pokiaľ to ide) rozčleniť predpoklad $\neg Q$ na konjunkciu dielčích predpokladov $R_1 \mathbin{\&} \dots \mathbin{\&} R_m$. Ak pôvodný predpoklad P bol konjunkciou dielčích predpokladov $P_1 \mathbin{\&} \dots \mathbin{\&} P_n$, tak jeho negácia $\neg P$ je ekvivalentná s alternatívou $\neg P_1 \vee \dots \vee \neg P_n$. Potom transponovaná implikácia $\neg Q \Rightarrow \neg P$ je logicky ekvivalentná s ľubovoľnou z implikácií

$$(\neg Q \mathbin{\&} P_1 \mathbin{\&} \dots \mathbin{\&} P_{i-1} \mathbin{\&} P_{i+1} \mathbin{\&} \dots \mathbin{\&} P_n) \Rightarrow \neg P_i,$$

kde $1 \le i \le n$. Nový záver $\neg P_i$ sa, samozrejme, usilujeme vybrať čo najvýhodnejšie, na čo neexistuje jednoznačný recept, no časom sa nám azda podarí nadobudnúť cit, ktorým sa budeme môcť riadiť.

0.7.3. Dôkaz sporom. *Dôkaz sporom* do istej miery pripomína nepriamy dôkaz a často sa s ním zvykne zamieňať. Najmä začiatočník by mal k nemu siahnuť až vtedy, keď sa mu priamy ani nepriamy dôkaz nedarí, prípadne keď v ňom skrsne podozrenie, že dokazované tvrdenie neplatí. Namiesto dokazovanej implikácie $P \Rightarrow Q$ prijmeme predpoklad P & $\neg Q$, ktorý je logicky ekvivalentný s jej negáciou $\neg(P \Rightarrow Q)$. Tento predpoklad sa usilujeme *doviesť k sporu*, čím sa myslí nejaký logicky absurdný záver, ako napr. $x \neq x$, alebo spor s niektorým z pôvodných predpokladov P, $\neg Q$, prípadne spor s niektorou z axióm alebo s niektorým zo skôr dokázaných tvrdení.

Na rozdiel od priameho alebo nepriameho dôkazu, dôkaz sporom nemá vopred stanovený smer určený nejakým známym záverom – ten by sa mal objaviť až v jeho priebehu. Ak sa ani pokus doviesť k sporu predpoklad P & $\neg Q$ neskončí úspešne, je namieste pokúsiť sa ho dokázať, to znamená vyvrátiť pôvodnú hypotézu $P \Rightarrow Q$.

0.7.4. Dôkaz ekvivalencie. Niekedy sa nám môže podariť dokázať ekvivalenciu $P \Leftrightarrow Q$ postupnosťou logicky ekvivalentných krokov, no to je skôr výnimka než pravidlo. Vo všeobecnosti si jej dôkaz vyžaduje dokázať zvlášť každú z implikácií $P \Rightarrow Q$, $Q \Rightarrow P$. Pritom na každú z nich možno použiť ľubovoľnú z troch skôr spomínaných metód. Často sa jedna z uvedených implikácií dokazuje priamo a druhá nepriamo, teda dôkaz uvedenej ekvivalencie pozostáva napr. z priamych dôkazov implikácií $P \Rightarrow Q$ a $\neg P \Rightarrow \neg Q$.

V našom kurze sa neraz stretneme s vetami, v ktorých sa tvrdí ekvivalencia viacerých podmienok $P_1, \ldots, P_n$. V tom je zahrnutých $n(n-1)$ jednotlivých implikácií $P_i \Rightarrow P_j$ pre rôzne $i, j \leq n$. Dokazovať ich všetky by pre $n \geq 3$ bolo značne neefektívne a taktiež zbytočné. Stačí totiž dokázať n implikácií tvoriacich cyklus

$$P_{\sigma(1)} \Rightarrow P_{\sigma(2)} \Rightarrow \ldots \Rightarrow P_{\sigma(n-1)} \Rightarrow P_{\sigma(n)} \Rightarrow P_{\sigma(1)},$$

kde σ je ľubovoľná permutácia množiny indexov $\{1, \ldots, n\}$, ktorú si volíme tak, aby to bolo čo najvýhodnejšie. S príkladmi všetkých uvedených typov dôkazov sa budeme v našom kurze neustále stretávať.

0.8 Matematická indukcia a rekurzia

Množinu všetkých nezáporných celých čísel značíme $\mathbb{N} = \{0, 1, 2, \ldots, \}$ a nazývame ju tiež množinou všetkých *prirodzených čísel*.

0.8.1. Dôkaz matematickou indukciou. Platnosť nejakého tvrdenia $P(n)$ pre všetky prirodzené čísla, t. j. tvrdenie $(\forall\, n \in \mathbb{N})P(n)$ sa obvykle dokazuje *matematickou indukciou.* Dôkaz indukciou spočíva v dôkaze dvoch tvrdení: nato, aby sme dokázali, že každé prirodzené číslo n má vlastnosť P, stačí dokázať, že platí

1° $P(0)$, t. j. 0 má vlastnosť P;

2° $(\forall\, n \in \mathbb{N})(P(n) \Rightarrow P(n+1))$,

t. j. ak n je ľubovoľné prirodzené číslo, ktoré má vlastnosť P, tak aj číslo $n+1$ má vlastnosť P.

Štruktúru *dôkazu matematickou indukciou* tak možno zhrnúť do schémy

$$\big(P(0) \mathbin{\&} (\forall\, n \in \mathbb{N})(P(n) \Rightarrow P(n+1))\big) \Rightarrow (\forall\, n \in \mathbb{N})P(n).$$

V bode 2° sa vlastne tvrdí platnosť všetkých implikácií $P(0) \Rightarrow P(1)$, $P(1) \Rightarrow P(2)$, $P(2) \Rightarrow P(3)$, … . Z bodu 1° a prvej z nich vyplýva $P(1)$, z toho spolu s druhou implikáciou dostávame $P(2)$, z čoho pomocou tretej implikácie plynie $P(3)$, atď.

Princíp matematickej indukcie je logicky ekvivalentný so zdanlivo očividným *princípom dobrého usporiadania*, ktorý tvrdí, že každá neprázdna množina $A \subseteq \mathbb{N}$ má najmenší prvok. Keďže pre väčšinu študentov býva tento princíp ľahšie prijateľný než princíp indukcie, predvedieme ako možno princíp indukcie z neho dokázať. Dôkaz princípu dobrého usporiadania z princípu indukcie prenechávame na rozmyslenie čitateľovi.

Predpokladajme teda platnosť princípu dobrého usporiadania. Nech P je vlastnosť taká, že platí $P(0)$ a $(\forall\, n \in \mathbb{N})(P(n) \Rightarrow P(n+1))$. Označme $A = \{n \in \mathbb{N};\, \neg P(n)\}$. Ak neplatí $(\forall\, n \in \mathbb{N})P(n)$, tak $A \neq \emptyset$. Nech m je najmenší prvok množiny A. Potom zrejme $m \neq 0$ a $m-1 \notin A$, teda platí $P(m-1)$. No keďže $P(m-1) \Rightarrow P(m)$, platí $P(m)$, čiže $m \notin A$, čo je spor.

Z pedagogických dôvodov sa budeme (najmä spočiatku) pri dôkazoch indukciou odvolávať radšej na princíp dobrého usporiadania než na princíp indukcie, a tomu tiež podriadime redakciu dôkazu.

Poznámka. (a) Niekedy je potrebné miesto počiatočného tvrdenia 1° osobitne dokázať niekoľko prvých tvrdení $P(0), P(1), \dots, P(k)$ a potom prejsť k dôkazu modifikovaného tvrdenia 2°, totiž $(\forall\, n \geq k)(P(n) \Rightarrow P(n+1))$.

(b) Indukciou možno dokazovať aj tvrdenia tvaru $(\forall\, n \geq m)P(n)$, kde m je nejaké pevné prirodzené číslo. Stačí dokázať mierne upravené verzie tvrdení 1° a 2°: $P(m)$ a $(\forall\, n \geq m)(P(n) \Rightarrow P(n+1))$.

(c) Pri dôkaze indukciou možno bod 2° nahradiť tvrdením

$$(\forall\, n \in \mathbb{N})\big((P(0) \mathbin{\&} \dots \mathbin{\&} P(n)) \Rightarrow P(n+1)\big).$$

Inak povedané, pri dôkaze záveru $P(n+1)$ v bode 2° sa nemusíme opierať len o predpoklad $P(n)$, ale v prípade potreby môžeme ako predpoklady použiť všetky predchádzajúce tvrdenia $P(0), \ldots, P(n)$. Takýto dôkaz indukciou sa vlastne riadi schémou

$$(\forall\, n \in \mathbb{N})\big((\forall\, k < n)P(k) \;\Rightarrow\; P(n)\big) \;\Rightarrow\; (\forall\, n \in \mathbb{N})P(n).$$

Rozmyslite si, ako je predpoklad 1°, t. j. tvrdenie $P(0)$, už zahrnutý v predpoklade novej schémy pre $n = 0$, t. j. v tvrdení $(\forall\, k < 0)(P(k) \;\Rightarrow\; P(0))$. Ďalej si premyslite, ako možno transpozíciou uvedenej implikácie priamo dostať matematickú formuláciu princípu dobrého usporiadania.

0.8.2. Rekurzia. Princíp matematickej indukcie sa používa nielen na dôkazy tvrdení o prirodzených číslach. Možno ho použiť aj na konštrukciu rôznych, či už konečných alebo nekonečných postupností. V takom prípade miesto indukcie budeme radšej hovoriť o postupnosti definovanej či zostrojenej *rekurziou.*

Nech X je množina a F je zobrazenie, ktoré každej konečnej postupnosti, (usporiadanej n-tici) $(x_1, \ldots, x_n)$ prvkov z X (akejkoľvek dĺžky $n \in \mathbb{N}$) priradí nejaký prvok $F(x_1, \ldots, x_n) \in X$. Pomocou zobrazenia F možno zostrojiť nekonečnú postupnosť $(a_n)_{n=0}^{\infty}$ prvkov z X tak, že položíme

$$a_0 = F(\emptyset), \qquad a_{n+1} = F(a_0, a_1, \ldots, a_n)$$

pre každé $n \in \mathbb{N}$. V takom prípade, hovoríme, že postupnosť (a_n) je definovaná *rekurziou* pomocou zobrazenia F.

Druhú rovnosť možno samozrejme zapísať v tvare $a_n = F(a_0, a_1, \ldots, a_{n-1})$ pre $n > 0$. Taktiež možno definíciu rekurziou obmedziť len na nejaký počiatočný úsek $0, 1, \ldots, n$ množiny prirodzených čísel a dostať tak rekurziou konečnú postupnosť $(a_0, a_1, \ldots, a_n)$. Niekedy rekurziu začíname nie od nuly ale od jednotky, prípadne od ľubovoľného prirodzeného čísla k.

Prvým členom postupnosti (a_n) zostrojenej rekurziou pomocou zobrazenia F je prvok $a_0 = F(\emptyset) \in X$. Ďalšie členy potom vyzerajú takto: $a_1 = F(a_0)$, $a_2 = F(a_0, a_1)$, $a_3 = F(a_0, a_1, a_2)$, $\ldots$, $a_n = F(a_0, \ldots, a_{n-1})$, $a_{n+1} = F(a_0, \ldots, a_n)$, atď.

Najčastejšie sa stretáme s prípadom, keď sa pri rekurzívnej konštrukcii člena a_{n+1} nepoužíva celá predchádzajúca časť postupnosti $(a_0, \ldots, a_n)$ ale len jej posledný člen a_n. Napríklad v aritmetickej postupnosti reálnych čísel s počiatočným členom a_0 a diferenciou d platí $a_{n+1} = a_n + d$; podobne rekurentný vzťah pre geometrickú postupnosť reálnych čísel s počiatočným členom a_0 a kvocientom q má tvar $a_{n+1} = qa_n$.

Iným známym číselným príkladom je tzv. *Fibonacciho postupnosť* $(\phi_n)_{n=0}^{\infty}$, ktorej rekurzívna definícia

$$\phi_0 = \phi_1 = 1, \qquad \phi_{n+2} = \phi_n + \phi_{n+1}$$

používa dva predchádzajúce členy. Rozmyslite si, ako táto definícia zapadá do našej všeobecnej schémy.

0.8.3. Príklad. *Bellove čísla* sú definované rekurziou

$$B_0 = 1, \qquad B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k,$$

pri ktorej sa využívajú všetky predchádzajúce členy. Pre istotu pripomíname, že

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{k!}$$

označuje *binomický koeficient* alebo *kombinačné číslo* udávajúce počet všetkých k-prvkových podmnožín n-prvkovej množiny.

Vypočítame niekoľko počiatočných hodnôt Bellových čísel:

$$B_0 = 1, \qquad B_1 = \binom{0}{0} B_0 = 1, \qquad B_2 = \binom{1}{0} B_0 + \binom{1}{1} B_1 = 2,$$

$$B_3 = \binom{2}{0} B_0 + \binom{2}{1} B_1 + \binom{2}{2} B_2 = 5,$$

$$B_4 = \binom{3}{0} B_0 + \binom{3}{1} B_1 + \binom{3}{2} B_2 + \binom{3}{3} B_3 = 15,$$

$$B_5 = \binom{4}{0} B_0 + \binom{4}{1} B_1 + \binom{4}{2} B_2 + \binom{4}{3} B_3 + \binom{4}{4} B_4 = 52,$$

$$B_6 = \binom{5}{0} B_0 + \binom{5}{1} B_1 + \binom{5}{2} B_2 + \binom{5}{3} B_3 + \binom{5}{4} B_4 + \binom{5}{5} B_5 = 203, \ \dots$$

Matematickou indukciou teraz dokážeme, že počet všetkých rozkladov n-prvkovej množiny (teda aj ekvivalencií na n-prvkovej množine) je rovný číslu B_n. Zrejme na prázdnej možine existuje jediný rozklad $\mathcal{R} = \emptyset$. Predpokladajme teraz, že pre každé $k \leq n$ existuje práve B_k rozkladov k-prvkovej množiny. Všetky rozklady $(n+1)$-prvkovej množiny $\{0, 1, \dots, n\}$ možno získať nasledujúcim spôsobom:

(1) zvolíme si ľubovoľné $k \leq n$ a ľubovoľnú k-prvkovú podmnožinu A množiny $\{1, \dots, n\}$ – to pre dané k možno urobiť práve $\binom{n}{k}$ spôsobmi;
(2) vezmeme ľubovoľný rozklad $\mathcal{R}$ množiny A – ten podľa indukčného predpokladu možno vybrať práve B_k spôsobmi – a množinu $A' = \{0, 1, \dots, n\} \smallsetminus A$ pridáme k pôvodnému rozkladu $\mathcal{R}$.

Zrejme sme takto získali nejaký rozklad $\mathcal{R}_A = \mathcal{R} \cup \{A'\}$ $(n+1)$-prvkovej množiny $\{0, 1, \dots, n\}$, pričom každý rozklad $\mathcal{S}$ množiny $\{0, 1, \dots, n\}$ má tvar $\mathcal{S} = \mathcal{R}_A$ pre jednoznačne určenú dvojicu $(\mathcal{R}, A)$. Všetkých rozkladov $(n+1)$-prvkovej množiny teda je $\sum_{k=0}^{n} \binom{n}{k} B_k = B_{n+1}$.

Cvičenia

0.1. Pre ľubovoľné množiny X, Y sú nasledujúce štyri podmienky ekvivalentné:
(i) $X \subseteq Y$, (ii) $X \cap Y = X$, (iii) $X \cup Y = Y$, (iv) $X \smallsetminus Y = \emptyset$.
Dokážte.

0.2. Vzťah inklúzie je reflexívny a tranzitívny, t. j. pre ľubovoľné množiny X, Y, Z platí:
(a) $X \subseteq X$; (b) $X \subseteq Y$ & $Y \subseteq Z \Rightarrow X \subseteq Z$
Dokážte. Nájdite príklad dosvedčujúci, že nie je symetrický.

0.3. Nech $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ označuje množinu všetkých celých čísel. Pre každé $0 \neq n \in \mathbb{Z}$ označme $n\mathbb{Z} = \{nx;\ x \in \mathbb{Z}\}$ množinu všetkých celých čísel deliteľných číslom n. Dokážte, že pre ľubovoľné nenulové celé čísla m, n platí
(a) $m\mathbb{Z} \subseteq n\mathbb{Z} \Leftrightarrow m$ je násobkom n;
(b) $m\mathbb{Z} = n\mathbb{Z} \Leftrightarrow |m| = |n|$;
(c) $m\mathbb{Z} \cap n\mathbb{Z} = k\mathbb{Z}$, kde k je najmenší spoločný násobok čísel m a n;
(d) Môže pre niektoré m, n nastať rovnosť $m\mathbb{Z} \cap n\mathbb{Z} = \emptyset$, prípadne $m\mathbb{Z} \cap n\mathbb{Z} = \{0\}$? Zdôvodnite.

0.4. Pre ľubovoľné množiny X, Y, Z platí

$$\begin{aligned}
X \cap Y &= Y \cap X, & X \cup Y &= Y \cup X,\\
X \cap (Y \cap Z) &= (X \cap Y) \cap Z, & X \cup (Y \cup Z) &= (X \cup Y) \cup Z,\\
X \cap X &= X, & X \cup X &= X,\\
X \cap \emptyset &= \emptyset, & X \cup \emptyset &= X,\\
X \cap (Y \cup Z) &= (X \cap Y) \cup (X \cap Z), & X \cup (Y \cap Z) &= (X \cup Y) \cap (X \cup Z),
\end{aligned}$$

$$\begin{aligned}
X \smallsetminus X &= \emptyset, & X \smallsetminus \emptyset &= X,\\
X \cap (Y \smallsetminus X) &= \emptyset, & X \cup (Y \smallsetminus X) &= X \cup Y,\\
(X \cap Y) \smallsetminus Z &= (X \smallsetminus Z) \cap (Y \smallsetminus Z), & (X \cup Y) \smallsetminus Z &= (X \smallsetminus Z) \cup (Y \smallsetminus Z),\\
X \smallsetminus (Y \cap Z) &= (X \smallsetminus Y) \cup (X \smallsetminus Z), & X \smallsetminus (Y \cup Z) &= (X \smallsetminus Y) \cap (X \smallsetminus Z),\\
(X \smallsetminus Y) \smallsetminus Z &= X \smallsetminus (Y \cup Z), & X \smallsetminus (Y \smallsetminus Z) &= (X \smallsetminus Y) \cup (X \cap Z),\\
(X \smallsetminus Y) \cap Z &= (X \cap Z) \smallsetminus Y, & (X \smallsetminus Y) \cup Z &= (X \cup Z) \smallsetminus (Y \smallsetminus Z).
\end{aligned}$$

Dokážte. Pomenujte niektoré z uvedených vlastností binárnych operácií $\cap$, $\cup$ a $\smallsetminus$ (komutatívnosť, asociatívnosť, …).

0.5. Pre ľubovoľné množiny X, Y, Z platí

$$X \times (Y \cap Z) = (X \times Y) \cap (X \times Z), \quad X \times (Y \cup Z) = (X \times Y) \cup (X \times Z),$$
$$X \times (Y \smallsetminus Z) = (X \times Y) \smallsetminus (X \times Z), \qquad X \times \emptyset = \emptyset.$$

Dokážte. Napíšte analogické vzťahy distributívnosti karteziánskeho súčinu vzhľadom na operácie prieniku, zjednotenia a množinového rozdielu pri násobení sprava.

0.6. Nech X, Y, Z sú ľubovoľné množiny. Nájdite čo najprirodzenejšiu bijekciu a k nej inverzné zobrazenie medzi každou z nasledujúcich dvojíc množín (množina $\mathcal{P}(X)$ všetkých podmnožín množiny X z časti (f) sa nazýva *potenčná množina* množiny X):

(a) $X \times Y$, $Y \times X$; (b) $X \times (Y \times Z)$, $(X \times Y) \times Z$;
(c) $(X \times Y)^Z$, $X^Z \times Y^Z$; (d) $X^{Y \times Z}$, $\left(X^Y\right)^Z$;
(e) X^n, $X^{\{1,\dots,n\}}$; (f) $\{0,1\}^X$, $\mathcal{P}(X) = \{A;\, A \subseteq X\}$;
(g) $X^{Y \cup Z}$, $X^Y \times X^{Z \smallsetminus Y}$; (h) $X^Y \times X^Z$, $X^{(Y \times \{0\}) \cup (Z \times \{1\})}$.

0.7. Nech $\mathbb{R}$ označuje množinu všetkých reálnych čísel. Funkcie $f, g, h \colon \mathbb{R} \to \mathbb{R}$ sú dané predpismi $f(x) = (x+1)^2$, $g(x) = 1 - \sin 2x$, $h(x) = \mathrm{e}^{-x}$. Napíšte a zjednodušte predpisy pre zložené funkcie $f \circ f$, $f \circ g$, $g \circ f$, $f \circ h$, $h \circ f$, $g \circ h$, $h \circ g$, $f \circ g \circ h$, $h \circ g \circ f$, $g \circ f \circ h$ a $g \circ h \circ f$.

0.8. Nech $f \colon X \to Y$, $g \colon Y \to Z$ sú ľubovoľné zobrazenia. Dokážte postupne nasledujúce tvrdenia:
(a) Ak f, g sú injektívne, tak aj $g \circ f$ je injektívne.
(b) Ak f, g sú surjektívne, tak aj $g \circ f$ je surjektívne.
(c) Ak f, g sú bijektívne, tak aj $g \circ f$ je bijektívne.
(d) Ak $g \circ f$ je injektívne, tak aj f je injektívne.
(e) Ak $g \circ f$ je surjektívne, tak aj g je surjektívne.
(f) Ak $g \circ f$ je bijektívne, tak f je injektívne a g je surjektívne.
Na príklade ukážte, že v prípade (f) g nemusí byť injektívne ani f surjektívne. Čo z toho vyplýva pre prípady (d) a (e)?

0.9. Pre ľubovoľné zobrazenia $f \colon X \to Y$, $g \colon Y \to Z$ a množiny $A \subseteq X$, $B \subseteq Z$ dokážte rovnosti:
(a) $(g \circ f)(A) = g(f(A))$; (b) $(g \circ f)^{-1}(B) = f^{-1}(g^{-1}(B))$.

0.10. Nech $f \colon X \to Y$ je ľubovoľné zobrazenie. Potom pre všetky A, $B \subseteq X$, C, $D \subseteq Y$ platí

(a) $A \subseteq B \Rightarrow f(A) \subseteq f(B)$; (b) $C \subseteq D \Rightarrow f^{-1}(C) \subseteq f^{-1}(D)$;
(c) $f(A \cap B) \subseteq f(A) \cap f(B)$; (d) $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$;
(e) $f(A \cup B) = f(A) \cup f(B)$; (f) $f^{-1}(C \cup D) = f^{-1}(C) \cup f^{-1}(D)$;
(g) $f(A \smallsetminus B) \supseteq f(A) \smallsetminus f(B)$; (h) $f^{-1}(C \smallsetminus D) = f^{-1}(C) \smallsetminus f^{-1}(D)$;
(i) $f^{-1}(f(A)) \supseteq A$; (j) $f\left(f^{-1}(C)\right) \subseteq C$.

Dokážte. Na príkladoch sa presvedčte, že inklúzie v prípadoch (c), (g), (i) a (j) nemožno zameniť rovnosťami. Ukážte, že rovnosť pre všetky A, $B \subseteq X$ je v ľubovoľnom z prípadov (c), (g) a (i) ekvivalentná s injektívnoťou zobrazenia f. Podobne je rovnosť pre každé $C \subseteq Y$ v prípade (j) ekvivalentná so surjektívnosťou f.

0.11. Nech $f\colon X \to Y$, $g\colon Y \to Z$ sú bijektívne zobrazenia. Potom $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. Dokážte. Odvoďte z toho, že pre každú permutáciu $\sigma \in \mathcal{S}_n$ platí $|\sigma| = \left|\sigma^{-1}\right|$ a $\operatorname{sgn}\sigma = \operatorname{sgn}\sigma^{-1}$.

0.12. (a) Sformulujte pojmy ľavého a pravého neutrálneho prvku binárnej operácie $*$ na množine X.

(b) Nech $x \triangleright y = y$ pre všetky $x, y \in X$. Je táto binárna operácia na množine X asociatívna resp. komutatívna? Nájdite všetky ľavé a všetky pravé neutrálne prvky operácie $\triangleright$.

(c) Predpokladajme, že e je jediným neutrálnym prvkom (či už ľavým, pravým alebo obojstranným) binárnej operácie $*$ na množine X. Sformulujte pojmy ľavého a pravého inverzného prvku k danému prvku $x \in X$.

(d) Pomocou multiplikatívnej tabuľky nájdite príklad (neasociatívnej) binárnej operácie $*$ na množine X, ktorá má jediný (obojstranný) neutrálny prvok e, pričom niektorý prvok $a \in X$ má jediný ľavý inverzný a dva rôzne pravé inverzné prvky, a iný prvok $b \in X$ má jediný ľavý inverzný no žiadny pravý inverzný prvok. Aké situácie môžu ešte nastať? (Pozri cvičenie 0.18.)

0.13. Nech $\varrho = \begin{pmatrix} 1\,2\,3\,4 \\ 3\,4\,1\,2 \end{pmatrix}$, $\sigma = \begin{pmatrix} 1\,2\,3\,4 \\ 2\,3\,1\,4 \end{pmatrix}$ sú permutácie množiny $\{1, 2, 3, 4\}$. Vypočítajte permutácie σ^{-1}, $\varrho^{-1}\ \sigma \circ \varrho$, $\varrho \circ \sigma$, $\sigma \circ \varrho^{-1}$, $\varrho \circ \sigma^{-1}$ a $\varrho^{-1} \circ \sigma \circ \varrho$.

0.14. Nech $1 \le i \le n \in \mathbb{N}$. Permutácia $\alpha \in \mathcal{S}_n$ sa nazýva *i-cyklus*, ak na nejakej i-prvkovej množine $\{a_1, a_2, \ldots, a_i\} \subseteq \{1, 2, \ldots, n\}$ operuje cyklicky podľa schémy $\alpha\colon a_1 \mapsto a_2 \mapsto \ldots \mapsto a_i \mapsto a_1$ a na zvyšku množiny $\{1, \ldots, n\}$ identicky, t. j. $\alpha(a) = a$ pre $a \in \{1, \ldots, n\} \smallsetminus \{a_1, \ldots, a_i\}$. Skrátene píšeme $\alpha = (a_1, \ldots, a_i)$. Dva *cykly* $\alpha = (a_1, \ldots, a_i)$, $\beta = (b_1, \ldots, b_j)$ sa nazývajú *disjunktné*, ak $\{a_1, \ldots, a_i\} \cap \{b_1, \ldots, b_j\} = \emptyset$. Zrejme identickú permutáciu možno považovať za 1-cyklus a každá transpozícia je 2-cyklus. Dokážte postupne nasledujúce tvrdenia:

(a) Nech $\alpha, \beta \in \mathcal{S}_n$ sú dva cykly. Potom $\alpha \circ \beta = \beta \circ \alpha$ práve vtedy, keď cykly α, β sú disjunktné alebo aspoň jeden z nich je 1-cyklus.

(b) Každá permutácia $\sigma \in \mathcal{S}_n$, je kompozíciou disjunktných cyklov $\sigma = \alpha_1 \circ \ldots \circ \alpha_k$. Ak do tejto kompozície zahrnieme aj všetky 1-cykly, tak uvedený rozklad je určený jednoznačne až na poradie jednotlivých cyklov. (*Návod*: Vytvorte cyklus $(1, \sigma(1), \sigma^2(1), \ldots, \sigma^{i-1}(1))$. Na zvyšku množiny $\{1, \ldots, n\}$, ak nejaký zostal, postup opakujte.)

(c) Každý cyklus je kompozíciou transpozícií (napr. $(1, 2, \ldots, n) = (1, n) \circ \ldots \circ (1, 3) \circ (1, 2)$).

(d) Každá permutácia $\sigma \in \mathcal{S}_n$ je kompozíciou transpozícií.

(e) Dĺžka i-cyklu α je $|\alpha| = i - 1$ a jeho znak je $\operatorname{sgn}\alpha = (-1)^{i-1}$.

(f) Nech $\sigma = \alpha_1 \circ \ldots \circ \alpha_k$ je rozklad permutácie σ na disjunktné cykly, pričom α_j je i_j-cyklus. Potom dĺžka permutácie σ je $|\sigma| = i_1 + \ldots + i_k - k$ a jej znak je $\operatorname{sgn}\sigma = (-1)^{i_1 + \cdots + i_k - k}$. Ak uvedený rozklad zahŕňa aj všetky 1-cykly, tak $|\sigma| = n - k$ a $\operatorname{sgn}\sigma = (-1)^{n-k}$.

0.15. (a) Rozložte každú z permutácií zo zadania aj výsledkov v cvičení 0.11 na kompozíciu disjunktných cyklov (vrátane 1-cyklov) a pre každú z nich určte jej dĺžku a znamienko.

(b) Spočítajte pre každú z uvedených permutácií počet jej inverzií a porovnajte s jej dĺžkou. Presvedčte sa, že obe čísla sa nemusia rovnať, no vždy majú rovnakú paritu.

(c) Rozložte permutáciu $\tau = (1,3,5,7) \circ (1,2) \circ (2,4,6,8) \circ (4,5,6) \in \mathcal{S}_{10}$ na *disjunktné* cykly (vrátane 1-cyklov) a určte jej dĺžku a znamienko.

0.16. k:0.16 (a) Nech $n \in \mathbb{Z}$ (pozri cvičenie 0.3). Dokážte, že vzťahom $x \equiv_n y \Leftrightarrow x - y \in n\mathbb{Z}$ je definovaná ekvivalencia na množine $\mathbb{Z}$. Túto ekvivalenciu značíme tiež $x \equiv y \mod n$ a nazývame ju *kongruenciou modulo* n.

(b) Predpokladajme, že $n > 0$, a označme $\mathbb{Z}_n = \{0, 1, \dots, n-1\} = \{x \in \mathbb{Z};\, 0 \leq x < n\}$. Potom $(\forall\, x \in \mathbb{Z})(\exists!\, z \in \mathbb{Z}_n)(x \equiv_n z)$. Dokážte. Toto jednoznačne určené $z \in \mathbb{Z}_n$ nazývame *zvyškom po delení* čísla x číslom n.

(c) Pre ľubovoľné $n \in \mathbb{Z}$ platí $(\forall\, x, y, u, v \in \mathbb{Z})(x \equiv_n y \;\&\; u \equiv_n v \;\Rightarrow\; x + u \equiv_n y + v \;\&\; xu \equiv_n yv)$.

(d) Ako vyzerajú triedy rozkladu množiny $\mathbb{Z}$ podľa kongruencie $\equiv_n$?

0.17. Matematickou indukciou dokážte nasledujúce vzorce:

(a) $\sum_{i=1}^{n} i = 1 + 2 + \ldots + n = \frac{1}{2}n(n+1)$;

(b) $\sum_{j=1}^{n} j(j+1) = 1 \cdot 2 + 2 \cdot 3 + \ldots + n(n+1) = \frac{1}{3}n(n+1)(n+2)$;

(c) $\sum_{k=1}^{n} k^2 = 1^2 + 2^2 + \ldots + n^2 = \frac{1}{6}n(n+1)(2n+1)$;

(d) $\sum_{k=0}^{n} q^k = 1 + q + q^2 + \ldots + q^n = \frac{q^{n+1}-1}{q-1}$, $(q \neq 1)$;

(e) $\sum_{p=1}^{n} \sin px = \sin x + \sin 2x + \ldots + \sin nx = \frac{\sin(nx/2)\sin((n+1)x/2)}{\sin(x/2)}$, $(x \neq 2k\pi$, $k \in \mathbb{Z})$;

(f) $\sum_{p=0}^{n} \cos px = \cos 0 + \cos x + \ldots + \cos nx = \frac{\cos(nx/2)\sin((n+1)x/2)}{\sin(x/2)}$, $(x \neq 2k\pi$, $k \in \mathbb{Z})$;

(g) $\phi_n = \frac{\left(1+\sqrt{5}\right)^{n+1} - \left(1-\sqrt{5}\right)^{n+1}}{2^{n+1}\sqrt{5}}$, ($\phi_n$ sú Fibonacciho čísla).

0.18. (a) Priamym výpočtom overte, že binomické koeficienty $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ vyhovujú rovnosti $\binom{n}{k} = \binom{n}{n-k}$ a pravidlám *Pascalovho trojuholníka*, t. j. $\binom{n}{0} = \binom{n}{n} = 1$ a $\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$ pre $1 \leq k \leq n$.

(b) Označme $C(n,k)$ počet všetkých k-prvkových podmnožín n-prvkovej množiny. Kombinatorickou úvahou dokážte, že aj tzv. *kombinačné čísla* $C(n,k)$ vyhovujú pravidlám Pascalovho trojuholníka, t. j. $C(n,0) = C(n,n) = 1$ a $C(n+1,k) = C(n,k-1) + C(n,k)$ pre $1 \leq k \leq n$.

(c) Na základe (a) a (b) odvoďte známy vzorec $C(n,k) = \binom{n}{k}$. Kde treba pritom použiť matematickú indukciu?

(d) Pre $k > n$ dodefinujme $\binom{n}{k} = C(n,k) = 0$. Predpisom $n * k = C(n,k)$ je tak definovaná binárna operácia na množine $\mathbb{N}$. Je táto operácia komutatívna resp. asociatívna? Má nejaký ľavý resp. pravý neutrálny prvok? (Pozri cvičenie 0.12.) Ako je to s prípadnými ľavými či pravými inverznými prvkami?

(e) Riešte úlohu (d) aj pre binárnu operáciu $n \bullet k = C(n+k,k) = \frac{(n+k)!}{n!\,k!}$ na množine $\mathbb{N}$.

0.19. Matematickou indukciou dokážte princíp dobrého usporiadania množiny $\mathbb{N}$ (t. j.

každá neprázdna podmnožina A množiny $\mathbb{N}$ má najmenší prvok).
(*Návod*: Transponujte implikáciu $(\forall\, n \in \mathbb{N})\big((\forall\, k < n)P(k) \;\Rightarrow\; P(n)\big) \;\Rightarrow\; (\forall\, n \in \mathbb{N})P(n)$, a nahraďte vlastnosť $P(n)$ vlastnosťou $n \notin A$.)

1. Polia a vektorové priestory

V tejto kapitole zavedieme dva druhy algebraických štruktúr, ktoré budú hrať v celom ďalšom výklade kľúčovú úlohu, a dokážeme o nich niekoľko jednoduchých základných tvrdení. Ide štruktúry, ktoré zahŕňame pod pojem *poľa* a pojem *vektorového priestoru*.

Prvky poľa budeme nazývať *skaláry*, a niekedy len čísla. Fyzikálne ich možno interpretovať ako hodnoty fyzikálnych veličín, ktoré sú určené iba svojou veľkosťou a znamienkom. Prvky vektorového priestoru, t. j. *vektory*, zasa zodpovedajú fyzikálnym veličinám, ktoré sú okrem veľkosti určené tiež smerom a orientáciou.

1.1 Základné číselné obory

Predpokladáme, že čitateľ pozná základné číselné obory, ako sú *prirodzené čísla*, *celé čísla*, *racionálne čísla*, *reálne čísla* a *komplexné čísla*. Každý z týchto číselných oborov tvorí množinu. Dohodneme sa, že ich budeme označovať tzv. tučnými tabuľovými písmenami:

$\mathbb{N}$ – množina všetkých prirodzených čísel,
$\mathbb{Z}$ – množina všetkých celých čísel,
$\mathbb{Q}$ – množina všetkých racionálnych čísel,
$\mathbb{R}$ – množina všetkých reálnych čísel,
$\mathbb{C}$ – množina všetkých komplexných čísel.

Ešte poznamenajme, že i nulu považujeme za prirodzené číslo, t. j. $0 \in \mathbb{N}$. *Imaginárnu jednotku* (ktorá je prvkom $\mathbb{C} \setminus \mathbb{R}$) budeme značiť i.

Konštatovaním, že uvedené číselné obory tvoria množiny, sme však ich štruktúru zďaleka nevyčerpali. Omnoho dôležitejšie je, že na každej z týchto množín sú definované dve binárne operácie, *sčítanie* $+$ a *násobenie* $\cdot$. Pritom na každej z uvedených množín sú obe tieto operácie asociatívne a komutatívne. Navyše, násobenie je (z oboch strán) *distributívne* vzhľadom na sčítanie, t. j. pre všetky prvky x, y, z príslušnej množiny platí

$$x(y+z) = xy + xz, \qquad (x+y)z = xz + yz.$$

Číselný obor $\mathbb{N}$ je v porovnaní s obormi $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ a $\mathbb{C}$ akýsi „chudobnejší“ – kým rovnice tvaru $x + a = b$ majú v oboroch $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ riešenie $x = b - a$ pre ľubovoľné a, b, v $\mathbb{N}$ je takáto rovnica riešiteľná len ak $a \leq b$. Obory $\mathbb{Q}$, $\mathbb{R}$ a $\mathbb{C}$ sú však „bohatšie“ nielen v porovnaní s $\mathbb{N}$ no i so $\mathbb{Z}$ – rovnice tvaru

$ax = b$ majú v oboroch $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ riešenie pre ľubovoľné $a \neq 0$ a b, kým v $\mathbb{N}$ či $\mathbb{Z}$ sú riešiteľné len ak a je deliteľom b.

Nás budú zaujímať práve vlastnosti číselných oborov $\mathbb{Q}$, $\mathbb{R}$ a $\mathbb{C}$ s operáciami sčítania a násobenia. Pritom využijeme, že uvedené operácie na týchto oboroch majú rad spoločných vlastností, čo nám umožňuje skúmať ich do veľkej miery jednotným spôsobom a súčasne. To dosiahneme tým, že sformulujeme abstraktný pojem *poľa*, pod ktorý zahrnieme všetky spomínané prípady, ako i mnohé ďalšie, ktoré sa nám objavia až akosi dodatočne. Ako sme spomínali už v úvode, práve takýto prístup je charakteristický pre algebru, presnejšie, v ňom spočíva jej podstata.

1.2 Polia

Poľom nazývame množinu K s dvoma význačnými prvkami – *nulou* 0 a *jednotkou* 1 – a dvomi binárnymi operáciami na K – *sčítaním* $+$ a *násobením* $\cdot$ – takými, že platí

$$
\begin{array}{ll}
(\forall\, a,b \in K)(a+b=b+a), & (\forall\, a,b \in K)(a\cdot b=b\cdot a),\\
(\forall\, a,b,c \in K)(a+(b+c)=(a+b)+c), & (\forall\, a,b,c \in K)(a\cdot(b\cdot c)=(a\cdot b)\cdot c),\\
(\forall\, a \in K)(a+0=a), & (\forall\, a \in K)(1\cdot a=a),\\
(\forall\, a \in K)(\exists\, b \in K)(a+b=0), & (\forall\, a \in K\smallsetminus\{0\})(\exists\, b \in K)(a\cdot b=1),\\
(\forall\, a,b,c \in K)(a\cdot(b+c)=(a\cdot b)+(a\cdot c)), & 0\neq 1.
\end{array}
$$

Teda sčítanie a násobenie v poli sú komutatívne a asociatívne operácie a násobenie je distributívne vzhľadom na sčítanie. Ďalej 0 je neutrálny prvok sčítania a 1 je neutrálny prvok násobenia, pričom tieto dva prvky sú rôzne. Jednoducho možno nahliadnuť, že prvok $b \in K$ taký, že $a+b=0$, t. j. inverzný prvok vzhľadom na operáciu sčítania, je k danému prvku $a \in K$ určený jednoznačne (pozri paragraf 0.4). Tento jednoznačne určený prvok k danému a označujeme $-a$ a nazývame *opačný prvok* k a. Miesto $a+(-b)$ zvykneme písať len $a-b$. Takisto prvok $b \in K$ taký, že $a\cdot b=1$, je k danému $0 \neq a \in K$ určený jednoznačne – označujeme ho a^{-1} alebo $\frac{1}{a}$, prípadne $1/a$ a nazývame *inverzný prvok* k a alebo *prevrátená hodnota* prvku a. Miesto $a\cdot b^{-1}$ píšeme tiež $\frac{a}{b}$ alebo a/b.

Znak násobenia budeme väčšinou vynechávať a násobenie bude mať prednosť pred sčítaním, teda napr. miesto $(a\cdot b)+c$ budeme písať len $ab+c$. Asociatívnosť nám umožňuje vynechávať zátvorky a súčty či súčiny ľubovoľných konečných postupností prvkov poľa jednoznačne zapisovať v tvare $a_1+a_2+\ldots+a_n$ resp. $a_1\cdot a_2\cdot\ldots\cdot a_n$ prípadne len $a_1a_2\ldots a_n$; komutatívnosť nám navyše dovoľuje nestarať sa o poradie sčítancov resp. činiteľov. Kvôli úplnosti sa dohodneme, že pre $n=1$ sa oba uvedené výrazy rovnajú a_1; pre $n=0$ kladieme prázdny súčet rovný 0 a prázdny súčin rovný 1. Ak

$a_1 = \ldots = a_n = a$, tak miesto $a_1 + \ldots + a_n$ píšeme na a miesto $a_1 \ldots a_n$ len a^n.

Teraz si ukážeme, ako možno niektoré najzákladnejšie pravidlá počítania, na ktoré sme zvyknutí v číselných oboroch $\mathbb{Q}$, $\mathbb{R}$ a $\mathbb{C}$, odvodiť len z axióm poľa. Zhrnieme ich do nasledujúceho tvrdenia. Okrem iného z neho vyplýva, že k 0 nemôže v poli existovať inverzný prvok (podmienka (c)).

1.2.1. Tvrdenie. *Nech K je pole. Potom pre ľubovoľné $n \in \mathbb{N}$ a $a, b, c, b_1, \ldots, b_n \in K$ platí*
(a) $a + b = a + c \Rightarrow b = c$,
(b) $(ab = ac \ \& \ a \neq 0) \Rightarrow b = c$,
(c) $a0 = 0$,
(d) $ab = 0 \Rightarrow (a = 0 \vee b = 0)$,
(e) $-a = (-1)a$,
(f) $a(b - c) = ab - ac$,
(g) $a(b_1 + \ldots + b_n) = ab_1 + \ldots + ab_n$.

Dôkaz. (a), (b) Keďže obe podmienky možno dokázať v podstate rovnako, urobíme to len pre druhú z nich. Z $ab = ac$ vyplýva $a^{-1}ab = a^{-1}ac$. Ľavá strana sa rovná b a pravá c.

(c) $a0 + a0 = a(0 + 0) = a0 = a0 + 0$. Podľa (a) z toho vyplýva $a0 = 0$.

(d) Nech $ab = 0$. Potom podľa (c) $ab = 0 = a0$. Ak $a \neq 0$, tak podľa (b) z toho vyplýva $b = 0$.

(e) Vďaka jednoznačnosti opačného prvku k a stačí overiť, že $(-1)a + a = 0$. Jednoduchý výpočet dáva $(-1)a + a = (-1)a + 1a = (-1 + 1)a = 0a = 0$ podľa (c).

(f) Podľa (e) $a(b-c) = a(b+(-1)c) = ab+a(-1)c = ab+(-1)ac = ab-ac$.

(g) Rovnosť zrejme platí pre $n = 0, 1, 2$. Keby neplatila pre všetky prirodzené čísla, označme n najmenšie prirodzené číslo, pre ktoré existujú $a, b_1, \ldots, b_n \in K$ také, že uvedená rovnosť neplatí. Potom $n > 2$ a pre $n - 1$ rovnosť platí. Preto

$$a(b_1 + \ldots + b_{n-1} + b_n) = a(b_1 + \ldots + b_{n-1}) + ab_n = ab_1 + \ldots + ab_{n-1} + ab_n.$$

To je však spor.

Doplňme, že podmienky (a) a (b) sa nazývajú *zákony o krátení* pre sčítanie resp. násobenie v poli.

Podmienka (e) nám umožňuje zaviesť ľubovoľné celočíselné násobky prvkov z poľa. Pre $a \in K$, $n \in \mathbb{N}$ kladieme $(-n)a = -(na) = n(-a)$. Podobne možno pre nenulové prvky poľa zaviesť i ľubovoľné celočíselné mocniny. Pre $0 \neq a \in K$, $n \in \mathbb{N}$ kladieme $a^{-n} = (a^n)^{-1} = (a^{-1})^n$.

Čitateľovi prenechávame, aby si sám odvodil nasledujúce rovnosti známe z bežných číselných oborov:

$$
\begin{array}{ll}
0a = 0, \quad 1a = a, & a \in K, \\
n(a+b) = na + nb, & a, b \in K,\ n \in \mathbb{Z}, \\
(m+n)a = ma + na, & a \in K,\ m, n \in \mathbb{Z}, \\
(mn)a = m(na), & a \in K,\ m, n \in \mathbb{Z}, \\
(mn)(ab) = (ma)(nb), & a, b \in K,\ m, n \in \mathbb{Z}, \\
a^0 = 1, \quad a^1 = a, & a \in K, \\
(ab)^n = a^n b^n, & a, b \in K,\ n \in \mathbb{Z},\ n < 0 \Rightarrow a \neq 0 \neq b, \\
a^{m+n} = a^m a^n, & a \in K,\ m, n \in \mathbb{Z},\ (m < 0 \vee n < 0) \Rightarrow a \neq 0, \\
a^{mn} = (a^m)^n, & a \in K,\ m, n \in \mathbb{Z},\ (m < 0 \vee n < 0) \Rightarrow a \neq 0,
\end{array}
$$

Ešte podotýkame, že v rovnostiach v prvom a šiestom riadku označujú 0 a 1 na ľavých stranách prirodzené čísla, t. j. prvky množiny $\mathbb{N}$, kým 0 a 1 na pravých stranách v prvom riadku označujú prvky poľa K. Vzhľadom na to, že pre všetky tri príklady polí, s ktorými sme doteraz stretli, platí $\mathbb{N} \subseteq K$, môže sa nám toto rozlíšenie zdať nepodstatné. Vo všeobecnosti však uvedená inklúzia platiť nemusí.

Nech K je pole a $L \subseteq K$. Hovoríme, že L je *podpole* poľa K, ak $0, 1 \in L$ a pre všetky $a, b \in L$ platí $a + b \in L$, $ab \in L$, $-a \in L$ a, ak $a \neq 0$, tak aj $a^{-1} \in L$. Inak povedané, podpole poľa K je taká jeho podmnožina L, ktorá obsahuje nulu a jednotku a je uzavretá vzhľadom na sčítanie, násobenie, opačný a inverzný prvok. Zrejme každé podpole poľa K je s týmito operáciami zúženými z K na L i samo poľom. Hovoríme tiež, že pole K je *rozšírením* poľa L.

Zrejme pole $\mathbb{Q}$ je podpoľom poľa $\mathbb{R}$ i poľa $\mathbb{C}$; pole $\mathbb{C}$ je rozšírením poľa $\mathbb{Q}$ aj $\mathbb{R}$.

Charakteristikou poľa K, označenie $\operatorname{char} K$, nazývame najmenšie kladné celé číslo n také, že $n1 = 0$; ak také n neexistuje, t. j. $n1 \neq 0$ pre každé celé $n > 0$, hovoríme že K má charakteristiku ∞ (niektorí autori vtedy kladú $\operatorname{char} K = 0$).

Ak pole K je rozšírením poľa L, tak polia K a L majú tú istú jednotku, preto $\operatorname{char} K = \operatorname{char} L$.

Zrejme $\operatorname{char} \mathbb{Q} = \operatorname{char} \mathbb{R} = \operatorname{char} \mathbb{C} = \infty$.

1.2.2. Veta. *Nech K je pole. Potom* $\operatorname{char} K$ *je ∞ alebo prvočíslo.*

Dôkaz Keďže $0 \neq 1$, zrejme $\operatorname{char} K > 1$. Predpokladajme, že $\operatorname{char} K = n$ je zložené číslo. Potom existujú celé čísla $k, l > 1$ také, že $n = kl$. Keďže $k, l < n$, je $k1 \neq 0 \neq l1$. Na druhej strane $(k1)(l1) = (kl)(1 \cdot 1) = n1 = 0$. Podľa tvrdenia 1.2.1(d) z toho vyplýva $k1 = 0$ alebo $l1 = 0$, čo je spor.

1.3 Polia $\mathbb{Z}_p$

V tomto krátkom paragrafe si ukážeme príklady polí, ktorých charakteristika nie je ∞. Z toho dôvodu sa tieto polia výrazne odlišujú od našich dôverne známych číselných polí. Presnejšie, pre každé prvočíslo p zostrojíme isté konečné pole $\mathbb{Z}_p$, ktoré má p prvkov a charakteristiku p. Na druhej strane, spomínané číselné polia (ako vôbec všetky polia nekonečnej charakteristiky) sú nekonečné. Poznamenajme, že pre každé prvočíslo p a kladné celé číslo k existuje p^k-prvkové pole s charakteristikou p ako aj nekonečné polia charakteristiky p. Ich konštrukcia však presahuje rámec nášho úvodného kurzu.

Pre potreby matematickej analýzy, teda aj z hľadiska fyzikálnych aplikácií, sú najdôležitejšími poľami $\mathbb{R}$ a $\mathbb{C}$. Konečné polia však v súčasnosti zohrávajú dôležitú úlohu napr. v kódovaní a kryptografii.

Pre každé kladné celé číslo n označme

$$\mathbb{Z}_n = \{k \in \mathbb{N};\ k < n\} = \{0, 1, \ldots, n-1\}.$$

Množinu $\mathbb{Z}_n$ zo zrejmých dôvodov (pozri cvičenie 0.12) nazývame *množinou zvyškových tried modulo n*. Na tejto množine teraz zavedieme dve binárne operácie – sčítanie $\oplus$ a násobenie $\odot$ (toto trochu ťažkopádne označenie budeme používať len v tomto paragrafe, neskôr sa vrátime k obvyklým $+$ a $\cdot$; v definícii však treba odlíšiť sčítanie a násobenie v $\mathbb{Z}_n$ od príslušných operácií v $\mathbb{Z}$). Pre $a, b \in \mathbb{Z}_n$ kladieme

$$\begin{aligned} a \oplus b &= \text{zvyšok po delení } (a+b) : n, \\ a \odot b &= \text{zvyšok po delení } (ab) : n. \end{aligned}$$

Čitateľovi prenechávame na overenie (prípadne na uverenie), že $\oplus$ a $\odot$ sú asociatívne a komutatívne operácie na $\mathbb{Z}_n$ a násobenie je distributívne vzhľadom na sčítanie. Ďalej 0 je neutrálny prvok sčítania a, pre $n > 1$, je 1 neutrálny prvok násobenia. Navyše $\ominus a = n - a$ je opačný prvok k $a \in \mathbb{Z}_n \setminus \{0\}$; pre $a = 0$ je samozrejme $\ominus 0 = 0$.

1.3.1. Veta. *Množina $\mathbb{Z}_n$ s operáciami $\oplus$ a $\odot$ je pole práve vtedy, keď n je prvočíslo.*

Dôkaz. Zrejme n je najmenšie kladné celé číslo také, že

$$n1 = \underbrace{1 \oplus \ldots \oplus 1}_{n\text{-krát}} = 0.$$

Preto, ak $\mathbb{Z}_n$ je pole, tak $\operatorname{char} \mathbb{Z}_n = n$, a podľa vety 1.2.2 je n prvočíslo.

Dokážeme, že $\mathbb{Z}_p$ je pole pre každé prvočíslo p. Najprv overíme, že v $\mathbb{Z}_p$ platí zákon o krátení

$$(a \odot b = a \odot c \;\&\; a \neq 0) \;\Rightarrow\; b = c.$$

Rovnosť $a \odot b = a \odot c$ znamená, že číslo $ab - ac = a(b - c)$ je deliteľné číslom p. Keďže p je prvočíslo, musí byť aspoň jedno z čísel a, $b - c$ deliteľné číslom p. Nakoľko $0 < a < p$, môže to byť len $b - c$. Pre $b, c \in \mathbb{Z}_p$ to však znamená $b = c$.

Zostáva overiť existenciu inverzného prvku ku každému $0 \neq a \in \mathbb{Z}_p$. Uvažujme postupnosť mocnín $a^1 = a$, $a^2 = a \odot a$, $a^3 = a \odot a \odot a, \ldots$, atď. Keďže $a \neq 0$, z dokázaného krátenia vyplýva, že všetky jej členy sú nenulové. Pretože množina $\mathbb{Z}_p$ je konečná, nemôžu byť všetky členy uvedenej postupnosti rôzne. Musia preto existovať kladné celé čísla k, l také, že $a^k = a^{k+l} = a^k \odot a^l$. Potom platí $a^k \odot a^l = a^k \odot 1$, z čoho krátením dostávame $a^l = 1$. Keďže $a^l = a \odot a^{l-1}$, je $a^{-1} = a^{l-1}$ inverzný prvok k a.

Multiplikatívne tabuľky sčítania a násobenia v poli $\mathbb{Z}_5$ sme si ako príklady binárnych operácií uviedli v paragrafe 0.4.

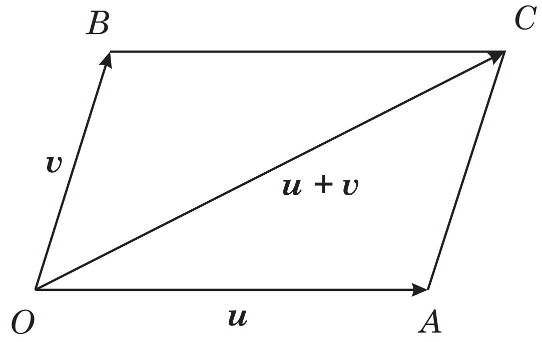
1.4 Vektory v rovine a v trojrozmernom priestore

Vektory v rovine či v priestore si predstavujeme ako orientované úsečky, t. j. úsečky, ktorých jeden krajný bod považujeme za počiatočný a druhý za koncový – ten je označený šípkou. Pritom dve rovnako dlhé, rovnobežné a súhlasne orientované úsečky predstavujú ten istý vektor – hovoríme, že sú umiestneniami toho istého vektora. Ak si teda zvolíme nejaký pevný bod O, tak všetky vektory v rovine či priestore môžeme jednoznačne reprezentovať ako orientované úsečky $\overrightarrow{OA}$ s počiatkom v O, pričom ich koncom môže byť ľubovoľný bod A roviny či priestoru, bod O nevynímajúc – orientovaná úsečka $\overrightarrow{OO}$ totiž predstavuje tzv. nulový vektor.

Vektory v rovine i v priestore možno sčítať pomocou tzv. *vektorového rovnobežníka*. Súčet vektorov $\boldsymbol{u} = \overrightarrow{OA}$, $\boldsymbol{v} = \overrightarrow{OB}$ je potom znázornený orientovanou uhlopriečkou $\boldsymbol{u} + \boldsymbol{v} = \overrightarrow{OC}$ rovnobežníka, ktorého dve priľahlé strany tvoria úsečky OA, OB.

Vektory možno taktiež násobiť ľubovoľnými skalármi, t. j. reálnymi číslami: ak $c \in \mathbb{R}$ a $\boldsymbol{v}$ je vektor, tak $c\boldsymbol{v}$ je vektor, t. j. orientovaná úsečka s počiatkom v O, ktorej dĺžka je $|c|$-násobkom dĺžky úsečky $\boldsymbol{v}$, leží na tej istej priamke ako $\boldsymbol{v}$ a je orientovaná súhlasne s $\boldsymbol{v}$, ak $c > 0$, resp. nesúhlasne s $\boldsymbol{v}$, ak $c < 0$, (ak $c = 0$ alebo $\boldsymbol{v}$ je nulový vektor, tak, samozrejme, aj $c\boldsymbol{v}$ je nulový vektor, takže nezáleží na jeho smere ani orientácii).

Ak si okrem počiatku O zvolíme v rovine či priestore ešte dve resp. tri súradné osi, t. j. navzájom kolmé priamky prechádzajúce počiatkom, a na každej z nich jeden bod v rovnakej jednotkovej vzdialenosti od počiatku



Obr. 1.1. Vektorový rovnobežník

(čím zadáme jednotku dĺžky a orientáciu osí), dostaneme pravouhlý súradnicový systém v rovine či v priestore. Každý bod roviny či priestoru je potom jednoznačne určený usporiadanou dvojicou, resp. trojicou svojich súradníc a tiež naopak, každá dvojica resp. trojica súradníc jednoznačne určuje nejaký bod roviny či priestoru. Taktiež každý vektor v rovine či v priestore je potom jednoznačne určený súradnicami svojho koncového bodu a tiež naopak ľubovoľná usporiadaná dvojica resp. trojica súradníc jednoznačne určuje nejaký vektor v rovine či priestore. Pri pevnom súradnicovom systéme tak možno množinu všetkých vektorov v rovine stotožniť s množinou $\mathbb{R}^2$ a množinu všetkých vektorov v priestore s množinou $\mathbb{R}^3$.

Ak (pri takomto stotožnení) $\boldsymbol{u} = (u_1, u_2) \in \mathbb{R}^2$, $\boldsymbol{v} = (v_1, v_2) \in \mathbb{R}^2$ sú dva vektory v rovine, tak ľahko nahliadneme, že pre ich súčet $\boldsymbol{u} + \boldsymbol{v}$, daný vektorovým rovnobežníkom, platí

$$\boldsymbol{u} + \boldsymbol{v} = (u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2).$$

Ak $c \in \mathbb{R}$, tak pre skalárny násobok $c\boldsymbol{u}$ dostávame

$$c\boldsymbol{u} = c(u_1, u_2) = (cu_1, cu_2).$$

Podobne možno reprezentovať aj operácie súčtu a skalárneho násobku vektorov v priestore príslušnými operáciami na množine $\mathbb{R}^3$ všetkých usporiadaných trojíc reálnych čísel.

Ešte si všimnime, že predpoklady kolmosti súradných osí a rovnosti jednotkových dĺžok v jednotlivých smeroch nehrali v našich úvahách nijakú úlohu. Stačí, aby systém súradných osí tvorili dve rôznobežné priamky (v rovine) resp. tri nekomplanárne priamky (v priestore) pretínajúce sa v počiatku O. Za jednotkové dĺžky v smeroch jednotlivých súradných osí možno zvoliť dĺžky ľubovoľných (nie nevyhnutne rovnako dlhých) úsečiek.

Operácie súčtu vektorov a násobenia vektora skalárom majú rad vlastností, ktoré nie sú viazané len na ich špecifickú geometrickú reprezentáciu v rovine či priestore. Napríklad, prostredníctvom súradnicovej reprezentácie vektorov by sme ich mohli priamočiaro zovšeobecniť na usporiadané n-tice skalárov z ľubovoľného poľa K pre akékoľvek $n \in \mathbb{N}$. Tým by sme dostali akési „n-rozmerné vektorové priestory nad poľom K“. V duchu algebry teraz zadefinujeme abstraktný pojem vektorového priestoru nad daným poľom, pričom budeme abstrahovať od akýchkoľvek súradníc aj „dimenzie“. Podstatné budú pre nás len algebraické vlastnosti operácií súčtu vektorov a skalárneho násobku vektora. K spomínaným príkladom sa však budeme sústavne vracať.

1.5 Vektorové priestory

Nech K je pole. *Vektorovým* alebo tiež *lineárnym priestorom* nad poľom K nazývame množinu V s význačným prvkom $\mathbf{0}$ a dvomi binárnymi operáciami – *sčítaním* $+ : V \times V \to V$ a *násobením* $\cdot : K \times V \to V$ – takými, že platí

$$
\begin{aligned}
&(\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V)(\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z})),\\
&(\forall \boldsymbol{x}, \boldsymbol{y} \in V)(\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}),\\
&(\forall \boldsymbol{x} \in V)(\boldsymbol{x} + \mathbf{0} = \boldsymbol{x}),\\
&(\forall \boldsymbol{x} \in V)(\exists \boldsymbol{y} \in V)(\boldsymbol{x} + \boldsymbol{y} = \mathbf{0}),\\
&(\forall a, b \in K)(\forall \boldsymbol{x} \in V)(a \cdot (b \cdot \boldsymbol{x}) = (ab) \cdot \boldsymbol{x}),\\
&(\forall \boldsymbol{x} \in V)(1 \cdot \boldsymbol{x} = \boldsymbol{x}),\\
&(\forall a \in K)(\forall \boldsymbol{x}, \boldsymbol{y} \in V)(a \cdot (\boldsymbol{x} + \boldsymbol{y}) = (a \cdot \boldsymbol{x}) + (a \cdot \boldsymbol{y})),\\
&(\forall a, b \in K)(\forall \boldsymbol{x} \in V)((a + b) \cdot \boldsymbol{x} = (a \cdot \boldsymbol{x}) + (b \cdot \boldsymbol{x})).
\end{aligned}
$$

Ako si čitateľ asi všimol, skaláry značíme „obyčajnými“ malými latinskými písmenami a vektory tučnými malými latinskými písmenami. Tejto implicitnej dohody sa budeme väčšinou držať, nie však za každú cenu. Kedykoľvek by nás obmedzovala, nebudeme váhať ju porušiť.

I keď sčítanie skalárov v poli a sčítanie vektorov značíme rovnakým znakom $+$, ide o rôzne operácie. Podobne násobenie v poli a násobenie vektora skalárom sú rôzne operácie, hoci obe značíme $\cdot$. Neskôr tento prístup dovedieme ešte ďalej, keď budeme rovnako značiť príslušné operácie a nuly v rôznych vektorových priestoroch. Rozlišovanie znakov pre nulu $0 \in K$ a $\mathbf{0} \in V$, hoci tieto prvky plnia rovnakú funkciu v K resp. vo V, je tak trochu proti duchu tohto prístupu. Ide vlastne o zbytočný luxus, ktorý je však v zhode s prijatou dohodou o značení skalárov a vektorov.

Z formálneho hľadiska pripomínajú axiómy vektorového priestoru axiómy poľa: sčítanie vektorov je opäť asociatívna a komutatívna binárna operácia na V s neutrálnym prvkom $\mathbf{0} \in V$, operácia násobenia vektora skalárom tiež spĺňa akúsi podmienku „asociatívnosti“, $1 \in K$ je jej „neutrálnym prvkom“

a platia dva „distributívne zákony". Je tu však jeden podstatný rozdiel – kým násobenie v poli K je binárnou operáciou na množine K, t. j. zobrazením $\cdot : K \times K \to K$, násobenie vo vektorovom priestore V nad poľom K nie je binárnou operáciou na V, ale binárnou operáciou $\cdot : K \times V \to V$. To nám však nebráni zaviesť obdobné dohody ako pre operácie v poli: i teraz bude mať násobenie prednosť pre sčítaním a znak násobenia budeme väčšinou vynechávať, t. j. písať napr. $a\boldsymbol{x} + \boldsymbol{y}$ miesto $(a \cdot \boldsymbol{x}) + \boldsymbol{y}$. Takisto budeme vynechávať zátvorky, ktorých umiestnenie neovplyvní výslednú hodnotu výrazov ako napr. v $ab\boldsymbol{x}$ alebo $a_1\boldsymbol{x}_1 + \ldots + a_n\boldsymbol{x}_n$. Posledný výraz budeme tiež značiť

$$\sum_{i=1}^{n} a_i\boldsymbol{x}_i$$

a nazývať *lineárnou kombináciou* vektorov $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ s koeficientmi $a_1, \ldots, a_n$. Špeciálne pre $n = 1$ to znamená $\sum_{i=1}^{1} a_i\boldsymbol{x}_i = a_1\boldsymbol{x}_1$; kvôli úplnosti pre $n = 0$ ešte kladieme prázdnu lineárnu kombináciu $\sum_{i=1}^{0} a_i\boldsymbol{x}_i$ rovnú $\mathbf{0}$.

Podobne ako v prípade polí, možno z axióm vektorových priestorov odvodiť niektoré základne pravidlá pre počítanie so skalármi a vektormi. Predovšetkým prvok $\boldsymbol{y} \in V$ taký, že $\boldsymbol{x} + \boldsymbol{y} = \mathbf{0}$, je k danému $\boldsymbol{x} \in V$ určený jednoznačne – značíme ho $-\boldsymbol{x}$ a nazývame *opačný vektor* k $\boldsymbol{x}$. Namiesto $\boldsymbol{x} + (-\boldsymbol{y})$ opäť píšeme len $\boldsymbol{x} - \boldsymbol{y}$. Tieto pravidlá zhrnieme v nasledujúcej analógii tvrdenia 1.2.1

1.5.1. Tvrdenie. *Nech V je vektorový priestor nad poľom K. Potom pre ľubovoľné $n \in \mathbb{N}$, $a, b, a_1, \ldots, a_n \in K$ a $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in V$ platí*

(a) $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z} \Rightarrow \boldsymbol{y} = \boldsymbol{z}$,

(b) $(a\boldsymbol{x} = a\boldsymbol{y} \ \&\ a \neq 0) \Rightarrow \boldsymbol{x} = \boldsymbol{y}$, $\qquad (a\boldsymbol{x} = b\boldsymbol{x} \ \&\ \boldsymbol{x} \neq \mathbf{0}) \Rightarrow a = b$,

(c) $a\mathbf{0} = \mathbf{0} = 0\boldsymbol{x}$,

(d) $a\boldsymbol{x} = \mathbf{0} \Rightarrow (a = 0 \vee \boldsymbol{x} = \mathbf{0})$,

(e) $-\boldsymbol{x} = (-1)\boldsymbol{x}$,

(f) $a(\boldsymbol{x} - \boldsymbol{y}) = a\boldsymbol{x} - a\boldsymbol{y}$, $\qquad (a - b)\boldsymbol{x} = a\boldsymbol{x} - b\boldsymbol{x}$,

(g) $a(\boldsymbol{x}_1 + \ldots + \boldsymbol{x}_n) = a\boldsymbol{x}_1 + \ldots + a\boldsymbol{x}_n$, $\qquad (a_1 + \ldots + a_n)\boldsymbol{x} = a_1\boldsymbol{x} + \ldots + a_n\boldsymbol{x}$.

Dôkaz Všetky podmienky, s výnimkou druhej implikácie v (b), možno dokázať celkom analogicky ako príslušné časti tvrdenia 1.2.1 Dokážeme aj túto. Nech $a\boldsymbol{x} = b\boldsymbol{x}$ a $\boldsymbol{x} \neq \mathbf{0}$. Potom $(a - b)\boldsymbol{x} = a\boldsymbol{x} - b\boldsymbol{x} = \mathbf{0}$. Podľa (d) z toho vyplýva $a - b = 0$, teda $a = b$.

Práve definované vektorové priestory by sme presnejšie mohli nazvať „ľavými" vektorovými priestormi, lebo v operácii skalárneho násobku píšeme skalár vľavo od vektora. Celkom obdobne by sme mohli definovať aj „pravé" vektorové priestory, v ktorých by sme operáciu skalárneho násobku chápali ako zobrazenie $V \times K \to V$ a zapisovali ju v tvare $\boldsymbol{x} \cdot a$ alebo len $\boldsymbol{x}a$ pre

$\boldsymbol{x} \in V$, $a \in K$. Vďaka komutatívnosti násobenia v poli K si však môžeme dovoliť chápať naše „ľavé“ vektorové priestory zároveň ako „pravé“. Pre všetky $a \in K$, $\boldsymbol{x} \in V$ jednoducho položíme $\boldsymbol{x}a = a\boldsymbol{x}$. Jediný problém – zabezpečiť pre všetky $a, b \in K$, $\boldsymbol{x} \in V$ rovnosť $(ab)\boldsymbol{x} = (ba)\boldsymbol{x}$, ktorá z takejto definície vyplýva výpočtom

$$(ab)\boldsymbol{x} = a(b\boldsymbol{x}) = a(\boldsymbol{x}b) = (\boldsymbol{x}b)a = \boldsymbol{x}(ba) = (ba)\boldsymbol{x},$$

– je vyriešený práve v dôsledku komutatívnosti násobenia v K. Teda, ak sa nám v operácii skalárneho násobku vyskytne skalár vpravo od vektora, nemusí nás to vyviesť z miery – kľudne ho môžeme prehodiť vľavo a ani o zátvorky sa nemusíme príliš starať.

1.6 Príklady vektorových priestorov

1.6.1. Rozšírenia polí. Zrejme každé pole K možno považovať za vektorový priestor nad sebou samým. Všeobecnejšie, ak pole L je rozšírením poľa K, tak L možno považovať za vektorový priestor nad poľom K (formálne stačí „zabudnúť“ násobenie niektorých dvojíc prvkov $a, b \in L$ a súčin ab pripustiť len pre $a \in K$, $b \in L$). Podobným spôsobom možno vektorový priestor V nad poľom L zúžením násobenia $L \times V \to V$ na násobenie $K \times V \to V$ prerobiť na vektorový priestor nad poľom K.

1.6.2. $\boldsymbol{n}$-rozmerné riadkové a stĺpcové vektory nad daným poľom. Pre ľubovoľné pole K a $n \in \mathbb{N}$ množina

$$K^n = \{(x_1, \dots, x_n);\ x_1, \dots, x_n \in K\}$$

všetkých usporiadaných n-tíc prvkov z K spolu s operáciami

$$\begin{aligned} \boldsymbol{x} + \boldsymbol{y} &= (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n), \\ c\boldsymbol{x} &= c(x_1, \dots, x_n) = (cx_1, \dots, cx_n), \end{aligned}$$

kde $\boldsymbol{x} = (x_1, \dots, x_n) \in K^n$, $\boldsymbol{y} = (y_1, \dots, y_n) \in K^n$ a $c \in K$, tvorí vektorový priestor nad poľom K. Zrejme usporiadaná n-tica $\mathbf{0}_n = (0, \dots, 0)$ hrá úlohu nuly v K^n. Ak bude potrebné rozlíšiť nulové vektory v priestoroch K^n pre rôzne prirodzené čísla n, budeme pre nulu v K^n používať označenie $\mathbf{0}_n$. Opačný prvok k $\boldsymbol{x} = (x_1, \dots, x_n) \in K^n$ je zrejme

$$-\boldsymbol{x} = -(x_1, \dots, x_n) = (-x_1, \dots, -x_n)$$

Hovoríme, že operácie na K^n sú definované *po zložkách.* Prvky tohto vektorového priestoru nazývame *n-rozmerné riadkové vektory* nad poľom K. Kvôli úplnosti ešte poznamenajme, že vektorový priestor K^0 pozostáva z jediného prvku $\emptyset$, predstavujúceho „usporiadanú nulaticu“, ktorá tak je nevyhnutne nulou v K^0.

Niekedy je (a väčšinou i bude) výhodnejšie pracovať s *n-rozmernými stĺpcovými vektormi* nad poľom K, t. j. s vektormi tvaru

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

kde $x_1, \ldots, x_n \in K$. Čitateľ si iste sám doplní definície príslušných operácií (opäť po zložkách) a ďalšie podrobnosti. Pokiaľ nebude hroziť nedorozumenie, budeme i tento priestor označovať K^n, prípadne len slovne naznačíme, či tým máme na mysli priestor n-rozmerných riadkových alebo stĺpcových vektorov. V súlade s tým $\mathbf{0}_n$ alebo len $\mathbf{0}$ môže označovať aj nulový vektor-stĺpec.

1.6.3. Polynómy nad daným poľom. Pod *polynómom* alebo tiež *mnohočlenom $f(x)$ stupňa n*, kde $n \in \mathbb{N}$, v premennej x nad poľom K rozumieme formálny výraz tvaru

$$f(x) = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + a_n x^n = \sum_{i=0}^{n} a_i x^i,$$

kde $a_0, a_1, \ldots, a_{n-1}, a_n \in K$ sú skaláry, nazývané *koeficienty* polynómu f, a $a_n \neq 0$; nulu $0 \in K$ považujeme za polynóm stupňa -1 a nenulové skaláry $a \in K$ za polynómy stupňa 0. Zrejme každý polynóm $f(x)$ definuje (rovnako značenú) funkciu $f\colon K \to K$ danú predpisom $c \mapsto f(c)$, t. j. dosadením konkrétnych hodnôt $c \in K$ za premennú x do polynómu $f(x)$. Množinu všetkých polynómov v premennej x nad K *stupňa nanajvýš n*, kde $-1 \leq n \in \mathbb{Z}$, budeme značiť $K^{(n)}[x]$; množinu *všetkých polynómov* v premennej x nad K značíme $K[x]$. Ľubovoľný polynóm $g(x) = \sum_{i=0}^{m} b_i x^i \in K[x]$ stupňa $m < n$ môžeme tiež písať v tvare

$$g(x) = b_0 + b_1 x + \ldots + b_m x^m + 0x^{m+1} + \ldots + 0x^n,$$

t. j. v tvare $g(x) = \sum_{i=0}^{n} b_i x^i$, kde $b_i = 0$ pre $m < i \leq n$. S použitím tejto konvencie možno definovať súčet $f(x) + g(x)$ polynómov $f(x) = \sum_{i=0}^{n} a_i x^i$, $g(x) = \sum_{i=0}^{m} b_i x^i$ z $K[x]$ predpisom

$$(f+g)(x) = f(x) + g(x) = \sum_{i=0}^{\max(m,n)} (a_i + b_i)x^i.$$

Ak navyše $c \in K$, kladieme

$$(cf)(x) = cf(x) \sum_{i=0}^{n} ca_i x^i.$$

Ľahko možno nahliadnuť, že s takto po zložkách definovanými operáciami súčtu a skalárneho násobku tvorí každá z množín polynómov $K^{(n)}[x]$, kde $-1 \leq n \in \mathbb{Z}$, ako i množina všetkých polynómov $K[x]$ vektorový priestor nad poľom K. Štruktúrou vektorového priestoru sa však algebra polynómov nevyčerpáva. Popri súčte a skalárnom násobku možno na $K[x]$ definovať aj súčin $f(x)\, g(x)$ uvedených polynómov $f(x)$, $g(x)$ predpisom

$$(fg)(x) = f(x)\, g(x) = \sum_{k=0}^{m+n} c_k x^k,$$

kde $c_k = \sum_{i=0}^{k} a_i b_{k-i}$.

1.6.4. Priame súčiny vektorových priestorov. Nech V_1 a V_2 sú vektorové priestory nad tým istým poľom K. *Priamym súčinom* (niekedy tiež *vonkajším priamym súčtom*) priestorov V_1, V_2 nazývame množinu $V_1 \times V_2$, t. j. karteziánsky súčin množín V_1, V_2, s operáciami súčtu vektorov a skalárneho násobku definovanými po zložkách. Teda pre $(\boldsymbol{u}_1, \boldsymbol{u}_2)$, $(\boldsymbol{v}_1, \boldsymbol{v}_2) \in V_1 \times V_2$, $c \in K$ kladieme

$$\begin{aligned}(\boldsymbol{u}_1, \boldsymbol{u}_2) + (\boldsymbol{v}_1, \boldsymbol{v}_2) &= (\boldsymbol{u}_1 + \boldsymbol{v}_1, \boldsymbol{u}_2 + \boldsymbol{v}_2),\\ c(\boldsymbol{u}_1, \boldsymbol{u}_2) &= (c\boldsymbol{u}_1, c\boldsymbol{u}_2).\end{aligned}$$

Zrejme $(\mathbf{0}, \mathbf{0})$ je nulou tohto vektorového priestoru a $-(\boldsymbol{u}_1, \boldsymbol{u}_2) = (-\boldsymbol{u}_1, -\boldsymbol{u}_2)$ je opačný prvok k $(\boldsymbol{u}_1, \boldsymbol{u}_2)$. Čitateľovi prenechávame, aby si overil, že priamy súčin $V_1 \times V_2$ s takto definovanými operáciami naozaj tvorí vektorový priestor nad poľom K, a taktiež, aby si premyslel, ako možno uvedenú konštrukciu zovšeobecniť na priamy súčin $V_1 \times \ldots \times V_n$ ľubovoľného konečného počtu vektorových priestorov $V_1, \ldots, V_n$ nad K. Ak $V = V_1 = \ldots = V_n$, tak píšeme $V_1 \times \ldots \times V_n = V^n$ a tento vektorový priestor nazývame n-tou *priamou mocninou* priestoru V. Pre $V = K$ uvedená konštrukcia dáva nám už známy vektorový priestor K^n z 1.6.2,

1.6.5. Vektorové priestory funkcií. Nech V je vektorový priestor nad poľom K a X je ľubovoľná množina. Pripomeňme, že V^X označuje množinu všetkých funkcií $f\colon X \to V$. Teraz ukážeme, ako možno z tejto množiny urobiť vektorový priestor nad poľom K. Operácie súčtu a skalárneho násobku budeme definovať opäť po zložkách. To znamená, že pre $f, g \in V^X$ a $c \in K$

budeme definovať funkcie $f+g \in V^X$ a $cf \in V^X$ tak, že pre každé $x \in X$ položíme

$$\begin{aligned}(f+g)(x) &= f(x)+g(x),\\ (cf)(x) &= cf(x).\end{aligned}$$

Znovu možno ľahko nahliadnuť, že V^X s takto definovanými operáciami tvorí vektorový priestor nad poľom K – nazývame ho *vektorovým priestorom všetkých funkcií* z X do V. Nulou vo V^X je funkcia $\mathbf{0}\colon X \to V$ identicky rovná prvku $\mathbf{0} \in V$; opačným prvkom k funkcii $f \in V^X$ je funkcia $-f \in V^X$ daná predpisom $x \mapsto -f(x)$ pre $x \in X$.

V špeciálnom prípade pre $V = K$ takto dostaneme vektorový priestor K^X všetkých funkcií z množiny X do poľa K. Ak K je pole všetkých reálnych prípadne komplexných čísel a X je napr. nejaký uzavretý interval $\langle a,b\rangle$ reálnych čísel, tak dostávame vektorové priestory funkcií $\mathbb{R}^{\langle a,b\rangle}$ resp. $\mathbb{C}^{\langle a,b\rangle}$, ktoré sa hojne vyskytujú v matematickej analýze.

Cvičenia

Cvičenia 1–4 sú opakovaním základných poznatkov o komplexných číslach.

1.1. Vypočítajte:
(a) $(5+3\mathrm{i})+(7-\mathrm{i})$, (b) $(11-10\mathrm{i})-(8-5\mathrm{i})$,
(c) $(-2+5\mathrm{i})\cdot(3+2\mathrm{i})$, (d) $(4-\mathrm{i})\cdot(2+9\mathrm{i})$,
(e) $(12+5\mathrm{i})^{-1}$, (f) $(7+\mathrm{i})/(3-4\mathrm{i})$.

1.2. (a) Pre komplexné číslo $x = a+b\mathrm{i}$, kde $a,b \in \mathbb{R}$, nazývame $a = \operatorname{Re} x$, $b = \operatorname{Im} x$ jeho *reálnou* resp. *imaginárnou časťou.* Teda $\operatorname{Re} x$ aj $\operatorname{Im} x$ sú *reálne čísla.* Dokážte vzorce:

$$\begin{aligned}\operatorname{Re}(x+y) &= \operatorname{Re} x + \operatorname{Re} y, & \operatorname{Re}(xy) &= \operatorname{Re} x \operatorname{Re} y - \operatorname{Im} x \operatorname{Im} y,\\ \operatorname{Im}(x+y) &= \operatorname{Im} x + \operatorname{Im} y, & \operatorname{Im}(xy) &= \operatorname{Re} x \operatorname{Im} y + \operatorname{Im} x \operatorname{Re} y.\end{aligned}$$

(b) Ak si v (reálnej) rovine zvolíme pravouhlý súradnicový systém, môžeme každé komplexné číslo $x = a+b\mathrm{i}$ reprezentovať bodom či vektorom so súradnicami (a,b). Ak prostredníctvom bijekcie $x \mapsto (\operatorname{Re} x, \operatorname{Im} x)$ stotožníme každé komplexné číslo s jeho obrazom a množinu $\mathbb{C}$ s rovinou (množinou $\mathbb{R}^2$), hovoríme o tzv. *Gaussovej rovine.* Znázornite čísla zo zadaní aj výsledkov cvičenia 1.1 v Gaussovej rovine.

1.3. *Absolútna hodnota* komplexného čísla $x = a+b\mathrm{i}$, kde $a,b \in \mathbb{R}$, je definovaná ako $|x| = \sqrt{a^2+b^2}$, t. j. ako vzdialenosť bodu x od počiatku v Gaussovej rovine. *Komplexne združené číslo* k číslu x je $\overline{x} = a-b\mathrm{i}$, t. j. číslo súmerne združené s x podľa reálnej osi.
(a) Nájdite absolútne hodnoty jednotlivých čísel zo zadaní aj výsledkov v cvičení 1.1.
(b) Dokážte nasledujúce vzťahy:

$$\begin{aligned}\operatorname{Re} x = \operatorname{Re}\overline{x} &= \frac{1}{2}(x+\overline{x}), & \operatorname{Im} x = -\operatorname{Im}\overline{x} &= \frac{1}{2\mathrm{i}}(x-\overline{x}),\\ \overline{\overline{x}} &= x, & xy^{-1} &= (x\overline{y})/|y|^2, \quad (y \neq 0),\\ \overline{x+y} &= \overline{x}+\overline{y}, & \overline{xy} &= \overline{x}\,\overline{y},\\ |x| &= |\overline{x}|, & |x|^2 &= x\overline{x},\\ |xy| &= |x|\,|y|, & |x+y| &\leq |x|+|y|.\end{aligned}$$

(c) V poslednom vzťahu nastane rovnosť práve vtedy, keď existuje nezáporné číslo $c \in \mathbb{R}$ také, že $x = cy$ alebo $y = cx$. Dokážte.

1.4. Každé komplexné číslo x možno vyjadriť v tzv. *goniometrickom tvare* $x = r(\cos\alpha + \mathrm{i}\sin\alpha)$, kde $r = |x|$ a α je uhol, ktorý (pre $x \neq 0$) zviera v Gaussovej rovine „vektor“ $\overrightarrow{01}$ s „vektorom“ $\overrightarrow{0x}$ (pre $x = 0$ vyhovuje ľubovoľné $\alpha \in \mathbb{R}$).
(a) Pre $x \neq 0$ vyjadrite $\cos\alpha$ a $\sin\alpha$ pomocou $\operatorname{Re} x$, $\operatorname{Im} x$ a $|x|$. Dokážte, že α je určené jednoznačne až na sčítanec $2k\pi$, kde $k \in \mathbb{Z}$.
(b) Pre $x = r(\cos\alpha + \mathrm{i}\sin\alpha)$, $y = s(\cos\beta + \mathrm{i}\sin\beta)$ platí $xy = rs(\cos(\alpha+\beta) + \mathrm{i}\sin(\alpha+\beta))$. Dokážte.
(c) Matematickou indukciou dokážte tzv. *Moivreovu vetu*: $(\cos\alpha+\mathrm{i}\sin\alpha)^n = \cos n\alpha + \mathrm{i}\sin n\alpha$, pre každé $n \in \mathbb{N}$. Rozšírte jej platnosť na všetky $n \in \mathbb{Z}$.
(d) Vyjadrite všetky čísla zo zadaní aj výsledkov v cvičení 1 v goniometrickom tvare.
(e) Pomocou Moivreovej vety vypočítajte $(\sqrt{3} + \mathrm{i})^{11}$, $(1 - \mathrm{i})^{-7}$.
(f) Na základe Moivreovej vety napíšte vzorec pre všetkých n riešení *binomickej rovnice* $x^n = c$, kde $c \in \mathbb{C}$. (*Návod*: Riešte najprv prípad $|c| = 1$.)
(g) Nájdite všetky riešenia binomických rovníc $x^3 = (\sqrt{3} - \mathrm{i})/2$, $y^4 = 1 + \mathrm{i}$ a $z^5 = -4 + 3\mathrm{i}$.

1.5. Podrobne dokážte vzťahy uvedené za dôkazom tvrdenia 1.2.1 Kde treba, použite matematickú indukciu.

1.6. V každom z nasledujúcich prípadov rozhodnite, či množina A je podpoľom poľa K. Svoje rozhodnutie zdôvodnite.
(a) $K = \mathbb{Q}$, $A = \mathbb{Z}$;
(b) $K = \mathbb{R}$, $A = \mathbb{Q}\big[\sqrt{2}\big] = \{a + b\sqrt{2};\ a, b \in \mathbb{Q}\}$;
(c) $K = \mathbb{R}$, $A = \langle -1, 1\rangle$;
(d) $K = \mathbb{C}$, $A = \mathbb{Z}[\mathrm{i}] = \{a + b\mathrm{i};\ a, b \in \mathbb{Z}\}$;
(e) $K = \mathbb{Z}_{11}$, $A = \mathbb{Z}_5$;
(f) $K = \mathbb{C}$, $A = \mathbb{Q}[\omega] = \{a+b\omega+c\omega^2;\ a, b, c \in \mathbb{Q}\}$, kde $\omega = (-1 + \mathrm{i}\sqrt{3})/2$.

1.7. Zostrojte multiplikatívne tabuľky sčítania a násobenia v $\mathbb{Z}_n$ pre $2 \leq n \leq 6$. Na ich základe zdôvodnite, prečo $\mathbb{Z}_4$ a $\mathbb{Z}_6$ nie sú polia.

1.8. Vynechajme z definície poľa podmienku $0 \neq 1$ a podmienku požadujúcu existenciu inverzného prvku vzhľadom na násobenie ku každému nenulovému prvku $a \in K$. Množina K s význačnými prvkami $0, 1 \in K$, vybavená binárnymi operáciami súčtu a súčinu, spĺňajúcimi zvyšné podmienky sa nazýva *komutatívny okruh s jednotkou.*[1] Komutatívny okruh s jednotkou sa nazýva *netriviálny*, ak v ňom predsa len platí $0 \neq 1$. Dokážte postupne nasledujúce tvrdenia:
(a) $\mathbb{Z}$ s obvyklými operáciami súčtu a súčinu je netriviálny komutatívny okruh s jednotkou.
(b) Pre každé $n \in \mathbb{N}$, $n \neq 0$, je $\mathbb{Z}_n$ so sčítaním a násobením modulo n komutatívny okruh s jednotkou. Tento okruh je netriviálny práve vtedy, keď $n \geq 2$.
(c) Komutatívny okruh s jednotkou je netriviálny práve vtedy, keď obsahuje aspoň dva rôzne prvky.

1.9. (a) V ľubovoľnom komutatívnom okruhu s jednotkou K zadefinujte výrazy tvaru na pre ľubovoľné $n \in \mathbb{Z}$, $a \in K$ rovnako ako v poli. Taktiež zadefinujte výrazy tvaru a^n pre $n \in \mathbb{N}$, $a \in K$. Dokážte pre ne analogické tvrdenia, ako platia v poli. Čo je prekážkou definície a^n pre všetky $n \in \mathbb{Z}$?
(b) Zadefinujte *charakteristiku* ľubovoľného *komutatívneho okruhu s jednotkou* rov-

[1] Občas sa v literatúre takáto štruktúra nazýva len komutatívny okruh.

nakým spôsobom ako v prípade poľa.
(c) Dokážte, že pre komutatívny okruh s jednotkou K platí $\operatorname{char} K = 1$ práve vtedy, keď K je triviálny.
(d) Pre každé $n \in \mathbb{N}$, $n \neq 0$, platí $\operatorname{char} \mathbb{Z}_n = n$.
(e) Pre každé prvočíslo p zostrojte príklad komutatívneho okruhu s jednotkou, ktorý má charakteristiku p, no nie je poľom. (*Návod*: Pozri cvičenie 1.13.)

1.10. (a) Matematickou indukciou dokážte platnosť binomickej vety *v ľubovoľnom komutatívnom okruhu s jednotkou* K (teda aj v ľubovoľnom poli). To znamená, že pre všetky $n \in \mathbb{N}$, $a, b \in K$ platí

$$(a+b)^n = a^n + \binom{n}{1} a^{n-1} b + \ldots + \binom{n}{n-1} a b^{n-1} + b^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

(b) Predpokladajme, že charakteristikou komutatívneho okruhu s jednotkou K je prvočíslo p. Nech $m \in \mathbb{Z}$ je násobkom p. Dokážte, že pre každé $c \in K$ platí $mc = 0$.
(c) Na základe (a) a (b) dokážte, že v komutatívnom okruhu s jednotkou prvočíselnej charakteristiky p platí pre exponent $n = p$ nasledujúci „populárny“ variant binomickej vety:

$$(a+b)^p = a^p + b^p.$$

1.11. Doplňte vynechané časti dôkazu tvrdenia 1.5.1

1.12. V každom z príkladov 1.6.1–5, podrobne overte, že uvedená množina s príslušnými operáciami tvorí vektorový priestor.

1.13. Rovnako ako v príklade 1.6.3, zadefinujte pre ľubovoľný komutatívny okruh s jednotkou K množinu $K[x]$ všetkých polynómov v premennej x s koeficientmi z K a na nej operácie súčtu a súčinu. Dokážte, že $K[x]$ s takto definovanými operáciami je opäť komutatívny okruh s jednotkou a platí $\operatorname{char} K[x] = \operatorname{char} K$.

1.14. Na množine $\mathbb{R}^+$ všetkých kladných reálnych čísel definujme nové „sčítanie“ $\oplus$ ako násobenie, t. j. $x \oplus y = xy$. Ďalej definujme novú operáciu „skalárneho násobku“ $\odot\colon \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}^+$ ako umocňovanie, t. j. predpisom $a \odot x = x^a$. Dokážte, že množina $\mathbb{R}^+$ s uvedenými operáciami tvorí vektorový priestor nad poľom $\mathbb{R}$. Čo je nulový vektor $\mathbf{0} \in \mathbb{R}^+$? Ako vyzerá opačný vektor $\ominus x$ k vektoru $x \in \mathbb{R}^+$? Vyjadrite pomocou pôvodných operácií násobenia a umocňovania lineárnu kombináciu $(a_1 \odot x_1) \oplus \ldots \oplus (a_n \odot x_n)$, kde $a_1, \ldots, a_n \in \mathbb{R}$, $x_1, \ldots, x_n \in \mathbb{R}^+$.

2. Základy maticového počtu

V tejto kapitole sa zoznámime s *maticami*, t. j. obdĺžnikovými tabuľkami, pomocou ktorých budeme kódovať najrôznejšie dôležité údaje o vektorových priestoroch, a naučíme sa s nimi zaobchádzať. Niektoré operácie s maticami budú zatiaľ nemotivované, ich význam vyjde najavo až neskôr. Od čitateľa tak žiadame istú dávku trpezlivosti, podobnú tej, akú musí prejaviť prváčik na základnej škole, ktorý tiež musí najprv zvládnuť jednotlivé písmenká, potom sa naučiť, ako sa z nich skladajú slová, a až potom môže začať čítať zmysluplné texty. Tento vklad sa nám zúročí neskôr, keď nám umožní hladko napredovať a nezdržiavať sa pri nepodstatných otázkach.

Pri prvom čítaní možno vynechať odstavce venované blokovým maticiam a maticiam nad vektorovými priestormi. Celkom postačí nalistovať si príslušnú časť až vo chvíli, keď sa s blokovými maticami stretneme v ďalších kapitolách.

2.1 Matice nad danou množinou

2.1.1. Typy matíc. Nech X je ľubovoľná množina a $m, n \in \mathbb{N}$. *Maticou typu* $m \times n$, alebo tiež $m \times n$*-rozmernou maticou* nad množinou X rozumieme obdĺžnikovú tabuľku

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

pozostávajúcu z prvkov množiny X. Skrátene tiež píšeme $\boldsymbol{A} = (a_{ij})_{m \times n}$, alebo len $\boldsymbol{A} = (a_{ij})$. Prvky $a_{ij} \in X$, kde $1 \le i \le m$, $1 \le j \le n$, sa nazývajú *prvkami matice* $\boldsymbol{A}$. Prvok a_{ij} nachádzajúci sa v i-tom riadku a j-tom stĺpci matice $\boldsymbol{A}$ nazývame tiež *prvok v mieste* (i, j), prípadne (i, j)*-ty prvok* matice $\boldsymbol{A}$. Množinu všetkých $m \times n$-rozmerných matíc nad množinou X značíme $X^{m \times n}$. Ak $m = n$, hovoríme o *štvorcových maticiach rádu* n nad množinou X.

Poznamenajme, že v prípade, keď niektoré z čísel m, n je 0, množina $X^{m \times n}$ pozostáva z jedinej a to *prázdnej* matice $\emptyset$. Neskôr sa ukáže rozumné stotožniť túto maticu s tzv. nulovou maticou. Aby sme sa vyhli trivialitám, budeme sa vždy baviť len o maticiach kladných rozmerov $m \times n$, čitateľ by si však mal aspoň občas uvedomiť, že väčšina našich úvah si zachováva platnosť aj v prípade, keď $m = 0$ alebo $n = 0$.

Dve matice nad množinou X považujeme za *navzájom rovné* alebo *totožné*, ak majú rovnaké rozmery a rovnaké prvky na príslušných miestach. To znamená, že pre matice $\boldsymbol{A} = (a_{ij})_{m\times n}$, $\boldsymbol{B} = (b_{ij})_{p\times q}$ nad X kladieme $\boldsymbol{A} = \boldsymbol{B}$ práve vtedy, keď $m = p$, $n = q$ a pre všetky $i = 1, \dots, m$, $j = 1, \dots, n$ platí $a_{ij} = b_{ij}$.

Množina matíc typu $1 \times n$ nad X splýva s množinou X^n, ak usporiadané n-tice prvkov z X zapisujeme do riadku. Podobne, ak usporiadné m-tice prvkov z X zapisujeme do stĺpca, tak množina matíc typu $m \times 1$ nad X splýva s množinou X^m. Pokiaľ bude z kontextu jasné, či ide o riadky alebo stĺpce, prípadne, ak na tom nebude záležať, budeme písať jednoducho X^n, X^m a pod. Podrobnejšie označenie $X^{1\times n}$, $X^{m\times 1}$ a pod. budeme používať, len ak bude treba rozlíšiť riadky a stĺpce.

Podobne by sme mohli zaviesť nielen dvoj- ale aj troj- či viacindexové matice $\boldsymbol{A} = (a_{ijk})_{m\times n\times r}$, $\boldsymbol{B} = \big(b_{i_1\dots i_p}\big)_{n_1\times\dots\times n_p}$, a pod. Pri ich prehľadnom zápise by sme však už nevystačili s dvojrozmernými tabuľkami. Napr. trojrozmerné matice by bolo teba zapisovať ako v priestore uložené vrstvy dojrozmerných matíc, čo presahuje obmedzené možnosti dvojrozmernej papierovej stránky prípadne obrazovky, ani nehovoriac o „viacrozmerných maticiach“.

2.1.2. Riadky a stĺpce matice. Nech $\boldsymbol{A} = (a_{ij}) \in X^{m\times n}$. Usporiadanú n-ticu

$$\boldsymbol{r}_i(\boldsymbol{A}) = (a_{i1}, a_{i2}, \dots, a_{in}) \in X^{1\times n},$$

kde $1 \le i \le m$, nazývame *i-tym riadkom* matice $\boldsymbol{A}$. Podobne, usporiadnú m-ticu

$$\boldsymbol{s}_j(\boldsymbol{A}) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix},$$

kde $1 \le j \le n$, nazývame *j-tym stĺpcom* matice $\boldsymbol{A}$. Maticu $\boldsymbol{A}$ tak možno stotožniť so stĺpcom jej riadkov ako aj s riadkom jej stĺpcov, t. j.

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} \boldsymbol{r}_1(\boldsymbol{A}) \\ \boldsymbol{r}_2(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{r}_m(\boldsymbol{A}) \end{pmatrix} = (\boldsymbol{s}_1(\boldsymbol{A}), \boldsymbol{s}_2(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A})).$$

2.1.3. Transponovaná matica. Maticu, ktorú získame z matice $\boldsymbol{A} = (a_{ij})_{m\times n}$ zámenou jej riadkov a stĺpcov, nazývame *transponovanou maticou* k matici

$\boldsymbol{A}$ a značíme ju $\boldsymbol{A}^\mathsf{T}$. Teda trochu podrobnejšie

$$\boldsymbol{A}^\mathsf{T} = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}.$$

To znamená, že $\boldsymbol{A}^\mathsf{T} \in X^{n\times m}$ a prvok v mieste (i,j) matice $\boldsymbol{A}^\mathsf{T}$ je a_{ji}.

Zrejme pre ľubovoľnú maticu $\boldsymbol{A} \in X^{m\times n}$ platí

$$(\boldsymbol{A}^\mathsf{T})^\mathsf{T} = \boldsymbol{A}.$$

Transpozíciou matíc-riadkov z $X^{1\times n}$ dostaneme matice-stĺpce z $X^{n\times 1}$ a transpozíciou matíc-stĺpcov z $X^{m\times 1}$ matice-riadky z $X^{1\times m}$. Na základe tejto poznámky možno nahliadnuť, že pre ľubovoľnú maticu $\boldsymbol{A} \in X^{m\times n}$ a $1 \le i \le m$, $1 \le j \le n$ platí

$$\boldsymbol{s}_i\big(\boldsymbol{A}^\mathsf{T}\big) = \boldsymbol{r}_i(\boldsymbol{A})^\mathsf{T}, \qquad \boldsymbol{r}_j\big(\boldsymbol{A}^\mathsf{T}\big) = \boldsymbol{s}_j(\boldsymbol{A})^\mathsf{T}.$$

Štvorcová matica $\boldsymbol{A} \in X^{n\times n}$ sa nazýva *symetrická*, ak $\boldsymbol{A} = \boldsymbol{A}^\mathsf{T}$, t. j. ak $a_{ij} = a_{ji}$ pre všetky $i, j = 1, \dots, n$. Postupnosť prvkov $(a_{11}, a_{22}, \dots, a_{nn})$ nazývame *diagonálou* štvorcovej matice $\boldsymbol{A}$. Transponovanú maticu k štvorcovej matici $\boldsymbol{A}$ zrejme získame „osovou súmernosťou“ jej prvkov podľa diagonály.

2.1.4. Blokové matice. Niekedy bude užitočné spojiť dve matice $\boldsymbol{A} \in X^{m\times n_1}$, $\boldsymbol{B} \in X^{m\times n_2}$ s rovnakým počtom riadkov do jednej matice tak, že príslušné tabuľky jednoducho napíšeme vedľa seba. Výsledná matica je typu $m \times (n_1 + n_2)$ a značíme ju $(\boldsymbol{A}, \boldsymbol{B})$, prípadne $(\boldsymbol{A} \,|\, \boldsymbol{B})$.

Podobne možno spojiť dve matice $\boldsymbol{A} \in X^{m_1\times n}$, $\boldsymbol{B} \in X^{m_2\times n}$ s rovnakým počtom stĺpcov do jednej matice tak, že príslušné tabuľky napíšeme pod seba. Výsledná matica je typu $(m_1 + m_2) \times n$ a značíme ju $\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix}$, prípadne $\begin{pmatrix} \boldsymbol{A} \\ \hline \boldsymbol{B} \end{pmatrix}$.

Práve popísané konštrukcie sú príkladmi tzv. *blokových matíc*. Pôvodné matice, z ktorých takto vytvárame blokovú maticu, potom nazývame jej *blokmi*. Takisto môžeme vedľa seba resp. pod seba zoradiť väčší počet blokov, nie len dva.

Naopak, niekedy sa môže ukázať účelné vyznačiť v danej matici nejaké menšie obdĺžnikove časti ako jej bloky. Vtedy hovoríme o tzv. *blokovom tvare* danej matice. Príkladom toho bol zápis matice $\boldsymbol{A} \in X^{m\times n}$ ako riadku jej stĺpcov, prípadne ako stĺpca jej riadkov.

Uvedené dve schémy vytvárania blokových matíc „vedľa seba“ a „pod seba“ možno tiež kombinovať. Napr. z matíc $\boldsymbol{A}_{11} \in X^{m_1\times n_1}$, $\boldsymbol{A}_{12} \in X^{m_1\times n_2}$, $\boldsymbol{A}_{21} \in X^{m_2\times n_1}$, $\boldsymbol{A}_{22} \in X^{m_2\times n_2}$ možno vytvoriť blokovú maticu

$$\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix}$$

typu $(m_1 + m_2) \times (n_1 + n_2)$.

Túto konštrukciu možno zrejmým spôsobom zovšeobecniť i na väčšie systémy matíc. Voľne povedané, blokové matice sú vlastne matice, ktorých prvkami sú opäť matice, pričom všetky matice v tom istom riadku blokovej matice majú rovnaký počet riadkov a všetky matice v tom istom stĺpci blokovej matice majú rovnaký počet stĺpcov. Takto chápanú blokovú maticu možno zapísať v tvare

$$\boldsymbol{A} = (\boldsymbol{A}_{ij})_{k\times l} = \begin{pmatrix} \boldsymbol{A}_{11} & \dots & \boldsymbol{A}_{1l} \\ \vdots & \ddots & \vdots \\ \boldsymbol{A}_{k1} & \dots & \boldsymbol{A}_{kl} \end{pmatrix},$$

pričom jednotlivé bloky $\boldsymbol{A}_{ij}$ sú matice nad X rozmerov $m_i \times n_j$, kde $(m_1, \dots, m_k)$, $(n_1, \dots, n_l)$ sú nejaké konečné postupnosti prirodzených čísel. Maticu nad množinou X z tejto „matice matíc“ dostaneme tak, že si v $\boldsymbol{A}$ odmyslíme vnútorné zátvorky oddeľujúce jej jednotlivé bloky $\boldsymbol{A}_{ij}$.

2.2 Matice nad daným poľom

Na množine X, nad ktorou sme vytvárali príslušné matice, sme zatiaľ nepredpokladali nijakú ďalšiu štruktúru. Jednako na množinách matíc $X^{m\times n}$ sa nám pomerne bohatá štruktúra prirodzene vynorila. Všetky doposiaľ zavedené maticové operácie a vlastnosti však mali výlučne *pozičný charakter* – zakladali sa na reprezentácii každej matice ako príslušnej obdĺžnikovej tabuľky. Ďalšie maticové operácie a vlastnosti, ktoré hodláme zaviesť a neskôr využívať, už budú podmienené prítomnosťou istej štruktúry na množine X.

Najdôležitejší a, až na pár výnimiek, vlastne jediný druh matíc, ktorými sa budeme v tomto kurze zaoberať, tvoria matice nad nejakým poľom. Teda v celom paragrafe K označuje pevne zvolené, inak však ľubovoľné pole. V súlade s predošlým paragrafom $K^{m\times n}$, kde $m, n \in \mathbb{N}$, označuje množinu všetkých matíc typu $m \times n$ nad poľom K.

2.2.1. Vektorový priestor matíc. Pre pevné $m, n \in K$ budeme na množine matíc $K^{m\times n}$ definovať po zložkách operácie súčtu a skalárneho násobku. Teda pre matice $\boldsymbol{A} = (a_{ij})_{m\times n}$, $\boldsymbol{B} = (b_{ij})_{m\times n}$ nad K a $c \in K$ položíme

$$\boldsymbol{A} + \boldsymbol{B} = (a_{ij} + b_{ij})_{m\times n},$$
$$c\boldsymbol{A} = (ca_{ij})_{m\times n}.$$

Podotýkame, že súčet matíc $\boldsymbol{A} + \boldsymbol{B}$ je definovaný len pre matice $\boldsymbol{A}$, $\boldsymbol{B}$ rovnakého typu a samotná matica $\boldsymbol{A}+\boldsymbol{B}$ je toho istého typu ako $\boldsymbol{A}$ a $\boldsymbol{B}$. Neutrálnym prvkom operácie sčítania na $K^{m\times n}$ je matica typu $m \times n$, ktorej všetky prvky sú nulové; nazývame ju *nulová matica* typu $m \times n$ a označujeme ju $\mathbf{0}_{m,n}$, prípadne len $\mathbf{0}$, keď jej rozmer je jasný z kontextu alebo na ňom nezáleží. Opačným prvkom k matici $A = (a_{ij})_{m\times n}$ je zrejme matica $-\boldsymbol{A} = (-a_{ij})_{m\times n}$.

Čitateľ si iste sám ľahko overí, že matice ľubovoľného pevného typu $m\times n$ nad poľom K s takto definovanými operáciami súčtu a skalárneho násobku tvoria *vektorový priestor* nad poľom K. Odteraz teda $K^{m\times n}$ už označuje nielen množinu takýchto matíc, ale príslušný vektorový priestor. Nám už známe vektorové priestory $K^{1\times n}$ a $K^{m\times 1}$ riadkových resp. stĺpcových vektorov sú zrejme špeciálnymi prípadmi vektorových priestorov matíc.

2.2.2. Násobenie matíc. Okrem štruktúry vektorového priestoru na množine matíc pevného typu $m \times n$ budeme definovať aj operáciu násobenia matíc, ktorá spája matice rôznych, „vhodne do seba zapadajúcich“ rozmerov.

Pod vplyvom doterajšieho výkladu čitateľ po takomto nadpise asi očakáva, že i súčin matíc budeme definovať na množine $K^{m\times n}$ po zložkách. Hoci by to, samozrejme, bolo možné a na prvý pohľad sa to zdá prirodzené, násobenie matíc budeme definovať diametrálne odlišným spôsobom, ktorý sa nám zatiaľ môže zdať čudný a neprirodzený. Dôvody pre takúto definíciu budú postupne vychádzať najavo a jej prednosti budeme mať mnohokrát možnosť oceniť.

Najprv sa naučíme násobiť niektoré dvojice vektorov. Pod *súčinom* $\boldsymbol{x}\cdot\boldsymbol{y}$ *riadkového vektora* $\boldsymbol{x} = (x_1, \dots, x_n) \in K^{1\times n}$ *a stĺpcového vektora* $\boldsymbol{y} = (y_1, \dots, y_n)^{\mathsf{T}} \in K^{n\times 1}$ rozumieme skalár

$$\boldsymbol{x}\cdot\boldsymbol{y} = (x_1, \dots, x_n)\cdot\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1y_1 + \ldots + x_ny_n = \sum_{i=1}^{n} x_iy_i\,.$$

Teda, až na „nepochopiteľné“ miešanie riadkových a stĺpcových vektorov, ide o bežný „skalárny súčin“ vektorov $\boldsymbol{x}, \boldsymbol{y} \in K^n$.

Pre takto definovaný súčin vektorov sú tiež splnené dobre známe vlastnosti „skalárneho súčinu“. Ľahko možno nahliadnuť, prípadne priamym výpočtom overiť, že pre všetky $n \in \mathbb{N}$, $c \in K$ a $\boldsymbol{x}, \boldsymbol{x}' \in K^{1\times n}$, $\boldsymbol{y}, \boldsymbol{y}' \in K^{n\times 1}$ platí

$$\begin{aligned}
\boldsymbol{x}\cdot(\boldsymbol{y}+\boldsymbol{y}') &= \boldsymbol{x}\cdot\boldsymbol{y} + \boldsymbol{x}\cdot\boldsymbol{y}',\\
(\boldsymbol{x}+\boldsymbol{x}')\cdot\boldsymbol{y} &= \boldsymbol{x}\cdot\boldsymbol{y} + \boldsymbol{x}'\cdot\boldsymbol{y},\\
\boldsymbol{x}\cdot c\boldsymbol{y} &= c(\boldsymbol{x}\cdot\boldsymbol{y}) = c\boldsymbol{x}\cdot\boldsymbol{y},\\
\boldsymbol{x}\cdot\boldsymbol{y} &= \boldsymbol{y}^{\mathsf{T}}\cdot\boldsymbol{x}^{\mathsf{T}}.
\end{aligned}$$

Hovoríme, že násobenie riadkových a stĺpcových vektorov je *distributívne* (z oboch strán) vzhľadom na sčítanie a *komutuje*, t. j. je zameniteľné s operáciou skalárneho násobku. Poslednú rovnosť možno chápať ako svojho druhu

„komutatívnosť“ tohto súčinu; vďačíme za ňu komutatívnosti násobenia v poli K.

Nech $m, n, p \in \mathbb{N}$ a $\boldsymbol{A} = (a_{ij})_{m\times n}$, $\boldsymbol{B} = (b_{jk})_{n\times p}$. Pod *súčinom matíc* $\boldsymbol{A}$, $\boldsymbol{B}$ rozumieme maticu

$$\boldsymbol{A}\cdot\boldsymbol{B} = (\boldsymbol{r}_i(\boldsymbol{A})\cdot\boldsymbol{s}_k(\boldsymbol{B}))_{m\times p}.$$

Všimnime si, že súčin matíc $\boldsymbol{A}$, $\boldsymbol{B}$ je definovaný, len ak sa počet stĺpcov matice $\boldsymbol{A}$ rovná počtu riadkov matice $\boldsymbol{B}$, t. j. práve vtedy, keď riadky matice $\boldsymbol{A}$ a stĺpce matice $\boldsymbol{B}$ majú rovnaký rozmer. Ďalej, súčin matíc typov $m\times n$ a $n\times p$ je matica typu $m\times p$, čo si možno ľahko zapamätať v symbolickom tvare

$$[m\times n]\cdot[n\times p] = [m\times p],$$

pripomínajúcom rozmerové vzťahy vo fyzike. Špeciálne, súčin dvoch štvorcových matíc typu $n\times n$ je opäť matica typu $n\times n$. Konečne, prvok na mieste (i,k) matice $\boldsymbol{A}\cdot\boldsymbol{B}$ dostaneme ako súčin i-teho riadku matice $\boldsymbol{A}$ a k-teho stĺpca matice $\boldsymbol{B}$, teda ako výraz

$$\boldsymbol{r}_i(\boldsymbol{A})\cdot\boldsymbol{s}_k(\boldsymbol{B}) = (a_{i1},\dots,a_{in})\cdot\begin{pmatrix} b_{1k}\\ \vdots\\ b_{nk}\end{pmatrix} = a_{i1}b_{1k}+\ldots+a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk}\,.$$

Na základe toho možno ľahko nahliadnuť (prípadne priamym výpočtom overiť) nasledujúce rovnosti

$$\boldsymbol{r}_i(\boldsymbol{A}\cdot\boldsymbol{B}) = \boldsymbol{r}_i(\boldsymbol{A})\cdot\boldsymbol{B}, \qquad \boldsymbol{s}_k(\boldsymbol{A}\cdot\boldsymbol{B}) = \boldsymbol{A}\cdot\boldsymbol{s}_k(\boldsymbol{B}).$$

Násobenie matíc je (z oboch strán) *distributívne* vzhľadom na sčítanie. To znamená že pre ľubovoľné $m, n \in \mathbb{N}$ a matice $\boldsymbol{A}$, $\boldsymbol{A}' \in K^{m\times n}$, $\boldsymbol{B}$, $\boldsymbol{B}' \in K^{n\times p}$ platí

$$\begin{aligned}\boldsymbol{A}\cdot(\boldsymbol{B}+\boldsymbol{B}') &= \boldsymbol{A}\cdot\boldsymbol{B}+\boldsymbol{A}\cdot\boldsymbol{B}',\\ (\boldsymbol{A}+\boldsymbol{A}')\cdot\boldsymbol{B} &= \boldsymbol{A}\cdot\boldsymbol{B}+\boldsymbol{A}'\cdot\boldsymbol{B}.\end{aligned}$$

Vďaka distributívnosti súčinu vektorov voči ich súčtu je totiž jasné, že (i,k)-ty prvok matice $\boldsymbol{A}\cdot(\boldsymbol{B}+\boldsymbol{B}')$ je

$$\boldsymbol{r}_i(\boldsymbol{A})\cdot\boldsymbol{s}_k(\boldsymbol{B}+\boldsymbol{B}') = \boldsymbol{r}_i(\boldsymbol{A})\cdot(\boldsymbol{s}_k(\boldsymbol{B})+\boldsymbol{s}_k(\boldsymbol{B}')) = \boldsymbol{r}_i(\boldsymbol{A})\cdot\boldsymbol{s}_k(\boldsymbol{B})+\boldsymbol{r}_i(\boldsymbol{A})\cdot\boldsymbol{s}_k(\boldsymbol{B}'),$$

teda sa rovná (i,k)-temu prvku matice $\boldsymbol{A}\cdot\boldsymbol{B}+\boldsymbol{A}\cdot\boldsymbol{B}'$. Rovnako pre druhú rovnosť.

Podobne, s využitím zameniteľnosti súčinu vektorov a skalárneho násobku možno dokázať, že pre ľubovoľný skalár $c\in K$ a všetky matice $\boldsymbol{A}\in K^{m\times n}$, $\boldsymbol{B}\in K^{n\times p}$ platí

$$\boldsymbol{A}\cdot c\boldsymbol{B} = c(\boldsymbol{A}\cdot\boldsymbol{B}) = c\boldsymbol{A}\cdot\boldsymbol{B}.$$

Hovoríme, že násobenie matíc *komutuje*, t. j. je zameniteľné s operáciou skalárneho násobku.

Násobenie matíc je tiež *asociatívne* v nasledujúcom zmysle: súčin matíc $\boldsymbol{A}\cdot(\boldsymbol{B}\cdot\boldsymbol{C})$ je definovaný práve vtedy, keď je definovaný súčin $(\boldsymbol{A}\cdot\boldsymbol{B})\cdot\boldsymbol{C}$, a v takom prípade sa obe matice rovnajú. Teda podrobnejšie, pre $m,n,p,q\in\mathbb{N}$ a $\boldsymbol{A}\in K^{m\times n}$, $\boldsymbol{B}\in K^{n\times p}$, $\boldsymbol{C}\in K^{p\times q}$ platí

$$\boldsymbol{A}\cdot(\boldsymbol{B}\cdot\boldsymbol{C})=(\boldsymbol{A}\cdot\boldsymbol{B})\cdot\boldsymbol{C}.$$

Na dôkaz toho si stačí uvedomiť, že pre ľubovoľné vektory $\boldsymbol{x}=(x_1,\dots,x_n)\in K^{1\times n}$, $\boldsymbol{y}=(y_1,\dots,y_p)^\mathsf{T}\in K^{p\times 1}$ platí

$$\begin{aligned}\boldsymbol{x}\cdot(\boldsymbol{B}\cdot\boldsymbol{y})&=(x_1,\dots,x_n)\cdot\begin{pmatrix}\sum_{k=1}^p b_{1k}y_k\\ \vdots\\ \sum_{k=1}^p b_{nk}y_k\end{pmatrix}=\sum_{j=1}^n x_j\Big(\sum_{k=1}^p b_{jk}y_k\Big)\\ &=\sum_{k=1}^p\Big(\sum_{j=1}^n x_jb_{jk}\Big)y_k=\Big(\sum_{j=1}^n x_jb_{j1},\dots,\sum_{j=1}^n x_jb_{jp}\Big)\cdot\begin{pmatrix}y_1\\ \vdots\\ y_p\end{pmatrix}=(\boldsymbol{x}\cdot\boldsymbol{B})\cdot\boldsymbol{y}.\end{aligned}$$

Potom pre $1\le i\le m$, $1\le l\le q$, prvok na mieste (i,l) matice $\boldsymbol{A}\cdot(\boldsymbol{B}\cdot\boldsymbol{C})$ je

$$\begin{aligned}\boldsymbol{r}_i(\boldsymbol{A})\cdot\boldsymbol{s}_l(\boldsymbol{B}\cdot\boldsymbol{C})&=\boldsymbol{r}_i(\boldsymbol{A})\cdot(\boldsymbol{B}\cdot\boldsymbol{s}_l(\boldsymbol{C}))\\ &=(\boldsymbol{r}_i(\boldsymbol{A})\cdot\boldsymbol{B})\cdot\boldsymbol{s}_l(\boldsymbol{C})=\boldsymbol{r}_i(\boldsymbol{A}\cdot\boldsymbol{B})\cdot\boldsymbol{s}_l(\boldsymbol{C}),\end{aligned}$$

teda sa rovná (i,l)-tému prvku matice $(\boldsymbol{A}\cdot\boldsymbol{B})\cdot\boldsymbol{C}$. Štvorcovú maticu rádu n, ktorá má všetky prvky na diagonále rovné 1 a mimo diagonály 0, označujeme $\boldsymbol{I}_n$ a nazývame *jednotkovou maticou* rádu n. S použitím tzv. *Kroneckerovho symbolu*

$$\delta_{ij}=\begin{cases}1, & \text{ak } i=j,\\ 0, & \text{ak } i\neq j,\end{cases}$$

môžeme písať

$$\boldsymbol{I}_n=(\delta_{ij})_{n\times n}=\begin{pmatrix}1&0&\dots&0&0\\ 0&1&\dots&0&0\\ \vdots&\vdots&\ddots&\vdots&\vdots\\ 0&0&\dots&1&0\\ 0&0&\dots&0&1\end{pmatrix}.$$

Jednotkové matice hrajú úlohu neutrálnych prvkov pre násobenie matíc. Presnejšie, pre ľubovoľnú maticu $\boldsymbol{A}\in K^{m\times n}$ platí

$$\boldsymbol{I}_m\cdot\boldsymbol{A}=\boldsymbol{A}=\boldsymbol{A}\cdot\boldsymbol{I}_n.$$

Rozmyslite si prečo.

Špeciálne, množina $K^{n\times n}$ všetkých štvorcových matíc rádu n je tak popri štruktúre vektorového priestoru navyše vybavená asociatívnou operáciou násobenia, ktorá je (z oboch strán) distributívna vzhľadom na sčítanie matíc, komutuje s operáciou skalárneho násobku a jednotková matica $\boldsymbol{I}_n$ je jej neutrálny prvok. To nám, podobne ako pre prvky poľa K, umožňuje zaviesť i *mocniny štvorcových matíc.* Pre $\boldsymbol{A} \in K^{n\times n}$, kladieme

$$\boldsymbol{A}^0 = \boldsymbol{I}_n \quad \text{a} \quad \boldsymbol{A}^k = \underbrace{\boldsymbol{A} \cdot \ldots \cdot \boldsymbol{A}}_{k\text{-krát}}$$

ak $0 < k \in \mathbb{N}$; teda $\boldsymbol{A}^1 = \boldsymbol{A}$, $\boldsymbol{A}^2 = \boldsymbol{A} \cdot \boldsymbol{A}$, $\boldsymbol{A}^3 = \boldsymbol{A} \cdot \boldsymbol{A} \cdot \boldsymbol{A}$, atď.

Na druhej strane si treba uvedomiť, že pre $n > 1$ – napriek komutatívnosti násobenia v poli K – násobenie matíc z pozičných dôvodov *nie je komutatívne* na $K^{n\times n}$. Napríklad

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1+1 \\ 1 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1+1 \end{pmatrix}.$$

(Uvedomte si, že na to, aby oba súčiny $\boldsymbol{A} \cdot \boldsymbol{B}$, $\boldsymbol{B} \cdot \boldsymbol{A}$ boli definované a mali rovnaké rozmery, teda, aby vôbec malo zmysel uvažovať o komutatívnosti súčinu, $\boldsymbol{A}$, $\boldsymbol{B}$ musia byť štvorcové matice rovnakého typu.)

Napriek tomu komutatívnosť násobenia v poli K má za dôsledok, že pre všetky m, n, p a matice $\boldsymbol{A} \in K^{m\times n}$, $\boldsymbol{B} \in K^{n\times p}$ platí rovnosť

$$(\boldsymbol{A} \cdot \boldsymbol{B})^\mathsf{T} = \boldsymbol{B}^\mathsf{T} \cdot \boldsymbol{A}^\mathsf{T}.$$

Naozaj, $(\boldsymbol{A} \cdot \boldsymbol{B})^\mathsf{T}$ aj $\boldsymbol{B}^\mathsf{T} \cdot \boldsymbol{A}^\mathsf{T}$ sú matice typu $p \times m$ a pre $1 \leq i \leq m$, $1 \leq k \leq p$, (k, i)-ty prvok matice $(\boldsymbol{A} \cdot \boldsymbol{B})^\mathsf{T}$ je (i, k)-ty prvok matice $\boldsymbol{A} \cdot \boldsymbol{B}$, t. j.

$$\boldsymbol{r}_i(\boldsymbol{A}) \cdot \boldsymbol{s}_k(\boldsymbol{B}) = \boldsymbol{s}_k(\boldsymbol{B})^\mathsf{T} \cdot \boldsymbol{r}_i(\boldsymbol{A})^\mathsf{T} = \boldsymbol{r}_k(\boldsymbol{B}^\mathsf{T}) \cdot \boldsymbol{s}_i(\boldsymbol{A}^\mathsf{T}),$$

čo je (k, i)-ty prvok matice $\boldsymbol{B}^\mathsf{T} \cdot \boldsymbol{A}^\mathsf{T}$. Pritom sme využili už spomínanú „komutatívnosť“ $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{y}^\mathsf{T} \cdot \boldsymbol{x}^\mathsf{T}$ súčinu vektorov.

Na margo poslednej rovnosti ešte podotknime, že pre $\boldsymbol{x} \in K^{1\times n}$, $\boldsymbol{y} \in K^{m\times 1}$ je taktiež definovaný súčin $\boldsymbol{y} \cdot \boldsymbol{x}$. Nie je to však skalár, ale matica typu $m \times n$:

$$\boldsymbol{y} \cdot \boldsymbol{x} = (y_i x_j)_{m\times n} = \begin{pmatrix} y_1 x_1 & \ldots & y_1 x_n \\ \vdots & \ddots & \vdots \\ y_m x_1 & \ldots & y_m x_n \end{pmatrix}.$$

Teda, okrem prípadu $m = n = 1$, rovnosť $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{y} \cdot \boldsymbol{x}$ nemôže nastať už z rozmerových dôvodov.

2.2.3. Operácie s blokovými maticami. Operácie maticového súčtu a skalárneho násobku, vďaka tomu, že boli definované po zložkách, možno na blokových maticiach rozložiť na jednotlivé bloky. Ak $\boldsymbol{A} = (\boldsymbol{A}_{ij})_{k\times l}$, $\boldsymbol{B} = (\boldsymbol{B}_{ij})_{k\times l}$ sú blokové matice nad poľom K, pričom zodpovedajúce si bloky $\boldsymbol{A}_{ij}$, $\boldsymbol{B}_{ij}$ majú rovnaký typ $m_i \times n_j$, tak ich súčet je opäť bloková matica

$$\boldsymbol{A} + \boldsymbol{B} = (\boldsymbol{A}_{ij} + \boldsymbol{B}_{ij})_{k\times l}$$

s blokmi rovnakých typov. S operáciou skalárneho násobku je to ešte jednoduchšie, lebo sa nemusíme starať o zhodnosť rozmerov jednotlivých blokov. Pre $c \in K$ jednoducho dostávame

$$c\boldsymbol{A} = (c\boldsymbol{A}_{ij})_{k\times l}.$$

Bloková štruktúra sa prenáša aj na súčin matíc za podmienky, že stĺpce prvej matice sú v rovnakom poradí rozdelené na rovnaký počet rovnako veľkých skupín, povedzme $n_1 + n_2 + \ldots + n_\nu$, ako riadky druhej matice. Teda ak $\boldsymbol{A} = (\boldsymbol{A}_{ij})_{\mu\times\nu}$, $\boldsymbol{B} = (\boldsymbol{B}_{jk})_{\nu\times\vartheta}$ sú blokové matice nad K, pričom blok $\boldsymbol{A}_{ij}$ je typu $m_i \times n_j$ a blok $\boldsymbol{B}_{jk}$ typu $n_j \times p_k$, tak aj ich súčin je bloková matica tvaru $\boldsymbol{A} \cdot \boldsymbol{B} = (\boldsymbol{C}_{ik})_{\mu\times\vartheta}$, kde blok

$$\boldsymbol{C}_{ik} = \boldsymbol{A}_{i1} \cdot \boldsymbol{B}_{1k} + \boldsymbol{A}_{i2} \cdot \boldsymbol{B}_{2k} + \ldots + \boldsymbol{A}_{i\nu} \cdot \boldsymbol{B}_{\nu k}$$

je typu $m_i \times p_k$. Inak povedané, blokové matice násobíme tak ako „obyčajné“ matice, len s tým rozdielom, že súčet resp. súčin v poli K nahradíme súčtom resp. súčinom matíc. Vo výsledku, ak chceme, si nakoniec môžeme odmyslieť zátvorky oddeľujúce jednotlivé bloky a matica, ktorú takto dostaneme, sa rovná matici, ktorú by sme dostali, keby sme „normálne“ vynásobili „odblokované“ matice $\boldsymbol{A}$ a $\boldsymbol{B}$.

Jednotkové matice $\boldsymbol{I}_n$ sú príkladom tzv. diagonálnych matíc. Štvorcovú maticu $\boldsymbol{A} = (a_{ij})_{n\times n}$ nazývame *diagonálnou*, ak $a_{ij} = 0$ pre všetky $i \neq j$, t. j. ak všetky jej prvky mimo diagonály sú nuly.

Diagonálnu maticu, ktorá má na diagonále postupne prvky $d_1, d_2, \ldots, d_n \in K$ značíme $\operatorname{diag}(d_1, d_2, \ldots, d_n)$. Teda napr.

$$\boldsymbol{I}_n = \operatorname{diag}(\underbrace{1, \ldots, 1}_{n\text{-krát}}).$$

Podobne možno definovať aj tzv. blokovo diagonálne matice. Ak $\boldsymbol{A}_1$, $\boldsymbol{A}_2, \ldots, \boldsymbol{A}_k$ sú štvorcové matice rádov $n_1, n_2, \ldots, n_k$, tak *blokovo diagonálnou maticou* s blokmi $\boldsymbol{A}_1$, $\boldsymbol{A}_2, \ldots, \boldsymbol{A}_k$ nazývame štvorcovú blokovú maticu

$$\operatorname{diag}(\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_k) = \begin{pmatrix} \boldsymbol{A}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{A}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \boldsymbol{A}_k \end{pmatrix},$$

kde $\mathbf{0}$ nachádzajúca sa na mieste (i, j) označuje nulovú maticu $\mathbf{0}_{n_i n_j}$.

Pred chvíľou uvedené pravidlo o súčine blokových matíc sa redukuje na obzvlášť jednoduchý tvar pre blokovo diagonálne matice – ich násobenie totiž funguje *diagonálne po zložkách.* Ak $\boldsymbol{A} = \mathrm{diag}(\boldsymbol{A}_1, \dots, \boldsymbol{A}_k)$, $\boldsymbol{B} = \mathrm{diag}(\boldsymbol{B}_1, \dots, \boldsymbol{B}_k)$ sú blokovo diagonálne matice, pričom zodpovedajúce si bloky $\boldsymbol{A}_i$, $\boldsymbol{B}_i$ sú štvorcové matice rovnakého rádu n_i, tak aj ich súčin je blokovo diagonálna matica tvaru

$$\boldsymbol{A} \cdot \boldsymbol{B} = \mathrm{diag}(\boldsymbol{A}_1 \cdot \boldsymbol{B}_1, \dots, \boldsymbol{A}_k \cdot \boldsymbol{B}_k)$$

so štvorcovými blokmi rádov $n_1, \dots, n_k$. Špeciálne, pre „obyčajné“ diagonálne matice platí

$$\mathrm{diag}(a_1, \dots, a_n) \cdot \mathrm{diag}(b_1, \dots, b_n) = \mathrm{diag}(a_1 b_1, \dots, a_n b_n).$$

Formuláciu analogických pravidiel pre súčet a skalárny násobok (blokovo) diagonálnych matíc prenechávame čitateľovi.

2.3 Matice nad vektorovým priestorom

Matice typu $m \times n$ nad poľom K sú špeciálnym druhom blokových matíc. Maticu $\boldsymbol{A} = (a_{ij}) \in K^{m \times n}$ môžeme považovať jednak za blokovú maticu s blokmi a_{ij} typu 1×1, jednak, ako sme už neraz naznačili, môžeme sa na ňu dívať ako na riadok jej stĺpcov resp. ako na stĺpec jej riadkov. V takom prípade $\boldsymbol{A}$ chápeme ako maticu typu $m \times 1$ nad vektorovým priestorom $K^{1 \times n}$, resp. ako maticu typu $1 \times n$ nad vektorovým priestorom $K^{m \times 1}$. Konkrétna podoba týchto vektorových priestorov je však teraz pre nás nepodstatná – pre ľubovoľné $m, n \in \mathbb{N}$ a ľubovoľný (abstraktný) vektorový priestor V máme totiž definovanú množinu $V^{m \times n}$ všetkých matíc nad množinou V.

Na množine $V^{m \times n}$ možno zaviesť operácie súčtu a skalárneho násobku po zložkách. $V^{m \times n}$ s týmito operáciami opäť tvorí vektorový priestor nad poľom K. Čitateľovi prenechávame, aby si sám doplnil a premyslel potrebné detaily.

My sa sústredíme na zovšeobecnenie operácie skalárneho násobku $K \times V \to V$ na operácie súčinu medzi maticami vhodných typov nad K a nad V. Pre matice $\boldsymbol{A} = (a_{ij}) \in K^{m \times n}$, $\boldsymbol{\alpha} = (\boldsymbol{u}_{jk}) \in V^{n \times p}$ kladieme $\boldsymbol{A} \cdot \boldsymbol{\alpha} = (\boldsymbol{v}_{ik}) \in V^{m \times p}$, kde

$$\boldsymbol{v}_{ik} = \sum_{j=1}^{n} a_{ij} \boldsymbol{u}_{jk} \,.$$

Teda súčin $\boldsymbol{A} \cdot \boldsymbol{\alpha}$ definujeme z formálneho hľadiska rovnako ako súčin matíc nad poľom K, len s tým rozdielom že operácia súčtu v K je nahradená operáciou súčtu vo V a operácia súčinu v K operáciou skalárneho násobku $K \times V \to V$.

Celkom obdobne ako v odstavci 2.2.2 aj pre násobenie matíc nad V maticami nad K možno overiť distributívnosť (z oboch strán) vzhľadom na sčítanie, zameniteľnosť s operáciou skalárneho násobku, asociatívnosť a postavenie jednotkových matíc ako neutrálnych prvkov. To znamená, že pre všetky $l, m, n, p \in \mathbb{N}$, $c \in K$, $\boldsymbol{A}$, $\boldsymbol{B} \in K^{m\times n}$, $\boldsymbol{C} \in K^{l\times m}$ $\boldsymbol{\alpha}$, $\boldsymbol{\beta} \in V^{n\times p}$ platí:

$$\begin{gathered}
\boldsymbol{A}\cdot(\boldsymbol{\alpha}+\boldsymbol{\beta}) = \boldsymbol{A}\cdot\boldsymbol{\alpha}+\boldsymbol{A}\cdot\boldsymbol{\beta},\\
(\boldsymbol{A}+\boldsymbol{B})\cdot\boldsymbol{\alpha} = \boldsymbol{A}\cdot\boldsymbol{\alpha}+\boldsymbol{B}\cdot\boldsymbol{\alpha},\\
\boldsymbol{A}\cdot(c\boldsymbol{\alpha}) = c(\boldsymbol{A}\cdot\boldsymbol{\alpha}) = (c\boldsymbol{A})\cdot\boldsymbol{\alpha},\\
\boldsymbol{C}\cdot(\boldsymbol{A}\cdot\boldsymbol{\alpha}) = (\boldsymbol{C}\cdot\boldsymbol{A})\cdot\boldsymbol{\alpha},\\
\boldsymbol{I}_n\cdot\boldsymbol{\alpha} = \boldsymbol{\alpha}.
\end{gathered}$$

Vzhľadom na našu dohodu, podľa ktorej $\boldsymbol{x}c = c\boldsymbol{x}$ pre $c \in K$, $\boldsymbol{x} \in V$, môžeme definovať aj súčin matíc $\boldsymbol{\beta} = (\boldsymbol{v}_{ij}) \in V^{m\times n}$, $\boldsymbol{B} = (\boldsymbol{b}_{jk}) \in K^{n\times p}$ v obrátenom poradí ako maticu $\boldsymbol{\beta}\cdot\boldsymbol{B} = (\boldsymbol{w}_{ik}) \in V^{m\times p}$ takú, že

$$\boldsymbol{w}_{ik} = \sum_{j=1}^{n} \boldsymbol{v}_{ij} b_{jk} = \sum_{j=1}^{n} b_{jk}\boldsymbol{v}_{ij}\,.$$

S využitím poslednej definície možno pre $\boldsymbol{A} \in K^{m\times n}$, $\boldsymbol{\alpha} \in V^{n\times p}$, $\boldsymbol{\beta} \in V^{m\times n}$, $\boldsymbol{B} \in K^{n\times p}$ dokázať tiež rovnosti

$$(\boldsymbol{A}\cdot\boldsymbol{\alpha})^{\mathsf{T}} = \boldsymbol{\alpha}^{\mathsf{T}}\cdot\boldsymbol{A}^{\mathsf{T}}, \qquad (\boldsymbol{\beta}\cdot\boldsymbol{B})^{\mathsf{T}} = \boldsymbol{B}^{\mathsf{T}}\cdot\boldsymbol{\beta}^{\mathsf{T}}.$$

Aplikáciou týchto vzťahov na predchádzjúci zoznam rovností (no taktiež priamo) možno aj pre súčin matíc tvaru $\boldsymbol{\beta}\cdot\boldsymbol{B}$, kde $\boldsymbol{\beta} \in V^{m\times n}$, $\boldsymbol{B} \in K^{n\times p}$, overiť jeho distributívnosť (z oboch strán) vzhľadom na sčítanie, zameniteľnosť s operáciou skalárneho násobku, asociatívnosť a postavenie jednotkových matíc ako neutrálnych prvkov. To znamená, že pre všetky $m, n, p, q \in \mathbb{N}$, $c \in K$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta} \in K^{m\times n}$, $\boldsymbol{A}$, $\boldsymbol{B} \in V^{n\times p}$, $\boldsymbol{C} \in K^{p\times q}$ platí:

$$\begin{gathered}
(\boldsymbol{\alpha}+\boldsymbol{\beta})\cdot\boldsymbol{A} = \boldsymbol{\alpha}\cdot\boldsymbol{A}+\boldsymbol{\beta}\cdot\boldsymbol{A},\\
\boldsymbol{\alpha}\cdot(\boldsymbol{A}+\boldsymbol{B}) = \boldsymbol{\alpha}\cdot\boldsymbol{A}+\boldsymbol{\alpha}\cdot\boldsymbol{B},\\
\boldsymbol{\alpha}\cdot(c\boldsymbol{A}) = c(\boldsymbol{\alpha}\cdot\boldsymbol{A}) = (c\boldsymbol{\alpha})\cdot\boldsymbol{A},\\
\boldsymbol{\alpha}\cdot(\boldsymbol{A}\cdot\boldsymbol{C}) = (\boldsymbol{\alpha}\cdot A)\cdot\boldsymbol{C},\\
\boldsymbol{\alpha}\cdot\boldsymbol{I}_n = \boldsymbol{\alpha}.
\end{gathered}$$

Taktiež vzťahy pre riadky a stĺpce súčinu z odseku 2.2.2 zostávajú zachované pre oba typu súčinov matíc nad K a V, t. j.

$$\begin{aligned}
\boldsymbol{r}_i(\boldsymbol{A}\cdot\boldsymbol{\alpha}) &= \boldsymbol{r}_i(\boldsymbol{A})\cdot\boldsymbol{\alpha}, & \boldsymbol{s}_k(\boldsymbol{A}\cdot\boldsymbol{\alpha}) &= \boldsymbol{A}\cdot\boldsymbol{s}_k(\boldsymbol{\alpha})\\
\boldsymbol{r}_i(\boldsymbol{\beta}\cdot\boldsymbol{B}) &= \boldsymbol{r}_i(\boldsymbol{\beta})\cdot\boldsymbol{B}, & \boldsymbol{s}_k(\boldsymbol{\beta}\cdot\boldsymbol{B}) &= \boldsymbol{\beta}\cdot\boldsymbol{s}_k(\boldsymbol{B})
\end{aligned}$$

pre všetky $\boldsymbol{A} \in K^{m\times n}$, $\boldsymbol{\alpha} \in V^{n\times p}$ $\boldsymbol{\beta} \in V^{m\times n}$, $\boldsymbol{B} \in K^{n\times p}$.

Napokon si ešte uvedomme, že definície súčinov $\boldsymbol{A}\cdot\boldsymbol{\alpha}$, $\boldsymbol{\beta}\cdot\boldsymbol{B}$ sú v zhode s pôvodným násobením matíc. Ak totiž maticu $\boldsymbol{A} \in K^{m\times n}$ chápame ako riadok, t. j. ako maticu typu $1\times n$ nad priestorom stĺpcových vektorov K^m, tak pre $\boldsymbol{B} \in K^{n\times p}$ splýva matica $(\boldsymbol{s}_1(\boldsymbol{A}),\dots,\boldsymbol{s}_n(\boldsymbol{A}))\cdot\boldsymbol{B}$ vypočítaná podľa „novej" definície s blokovým tvarom $(\boldsymbol{A}\cdot\boldsymbol{s}_1(\boldsymbol{B}),\dots,\boldsymbol{A}\cdot\boldsymbol{s}_p(\boldsymbol{B}))$ matice $\boldsymbol{A}\cdot\boldsymbol{B}$. Podobne, ak $\boldsymbol{B}$ chápeme ako stĺpec, t. j. ako maticu typu $n\times 1$ nad priestorom riadkových vektorov K^p, tak

$$\boldsymbol{A}\cdot\begin{pmatrix}\boldsymbol{r}_1(\boldsymbol{B})\\ \vdots\\ \boldsymbol{r}_n(\boldsymbol{B})\end{pmatrix} = \begin{pmatrix}\boldsymbol{r}_1(\boldsymbol{A})\cdot\boldsymbol{B}\\ \vdots\\ \boldsymbol{r}_m(\boldsymbol{A})\cdot\boldsymbol{B}\end{pmatrix} = \boldsymbol{A}\cdot\boldsymbol{B}.$$

(Doplňte si vynechané podrobnosti – pozri cvičenie 2.6.)

Špeciálne, lineárnu kombináciu $a_1\boldsymbol{x}_1 + \dots a_n\boldsymbol{x}_n$ vektorov $\boldsymbol{x}_1,\dots,\boldsymbol{x}_n \in V$ s koeficientmi $a_1,\dots,a_n \in K$ môžeme s využitím vektorových matíc zapísať v tvare súčinov

$$a_1\boldsymbol{x}_1 + \dots + a_n\boldsymbol{x}_n = (a_1,\dots,a_n)\cdot\begin{pmatrix}\boldsymbol{x}_1\\ \vdots\\ \boldsymbol{x}_n\end{pmatrix} = (\boldsymbol{x}_1,\dots,\boldsymbol{x}_n)\cdot\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}.$$

Cvičenia

2.1. Nech $\boldsymbol{A} = \begin{pmatrix}1 & -1 & 1\\ 0 & 5 & 2\\ 1 & -4 & 0\end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix}0 & 1 & 2\\ -2 & 9 & 3\\ 10 & 6 & 0\end{pmatrix}$, $\boldsymbol{C} = \begin{pmatrix}1 & 2 & 3\\ 1 & 4 & 9\\ 1 & 2 & 4\end{pmatrix}$ sú matice nad $\mathbb{R}$. Vypočítajte matice $\boldsymbol{A}+2\boldsymbol{B}$, $\boldsymbol{A}-\boldsymbol{B}^{\mathsf{T}}-3\boldsymbol{C}$, $\boldsymbol{A}\cdot\boldsymbol{B}$, $\boldsymbol{B}\cdot\boldsymbol{A}$, $\boldsymbol{A}\cdot(\boldsymbol{B}+\boldsymbol{C})$, $(3\boldsymbol{A}^{\mathsf{T}}+\boldsymbol{B})\cdot\boldsymbol{C}$, $\boldsymbol{B}\cdot\boldsymbol{C}^2$, $\boldsymbol{C}^2\cdot\boldsymbol{B}$, $\boldsymbol{C}\cdot\boldsymbol{B}\cdot\boldsymbol{C}$, $\boldsymbol{A}\cdot\boldsymbol{C}-\boldsymbol{C}\cdot\boldsymbol{A}$, $\boldsymbol{A}\cdot\boldsymbol{B}\cdot\boldsymbol{C}$ a $\boldsymbol{C}^{\mathsf{T}}\cdot\boldsymbol{A}\cdot\boldsymbol{C}$.

2.2. Vypočítajte súčin $\boldsymbol{A}\cdot\boldsymbol{B}$ komplexných matíc $\boldsymbol{A} = \begin{pmatrix}1+\mathrm{i} & -2 & -\mathrm{i}\\ 1-\mathrm{i} & -\mathrm{i} & 2+3\mathrm{i}\end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix}3\mathrm{i} & 2+\mathrm{i}\\ -4 & 1-2\mathrm{i}\\ 0 & 4\mathrm{i}\end{pmatrix}$.

2.3. Nájdite matice $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{Q}^{2\times 2}$ také, že $\boldsymbol{A}\cdot\boldsymbol{B} = \boldsymbol{0} \neq \boldsymbol{B}\cdot\boldsymbol{A}$.

2.4. Uvažujte matice $\boldsymbol{A} = \begin{pmatrix}1 & 3\\ 3 & 2\end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix}0 & 1\\ 4 & 1\end{pmatrix}$, $\boldsymbol{C} = \begin{pmatrix}4 & 3\\ 4 & 0\end{pmatrix}$ nad poľom
(a) $\mathbb{Z}_5$, (b) $\mathbb{Z}_7$, (c) $\mathbb{Z}_{11}$, (d) $\mathbb{Q}$.
V každom z uvedených prípadov vypočítajte maticu $\boldsymbol{A}\cdot(\boldsymbol{B}+\boldsymbol{C})$. Skúste riešiť úlohy (a)–(d) v optimálnom poradí.

2.5. Sú dané reálne blokové matice $\boldsymbol{A} = \begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \boldsymbol{B}_{13}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \boldsymbol{B}_{23}\end{pmatrix}$, kde

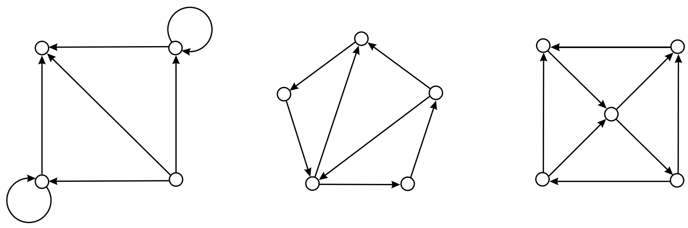
$\boldsymbol{A}_{11} = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$, $\boldsymbol{A}_{12} = \begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}$, $\boldsymbol{A}_{21} = \begin{pmatrix}2 & 1\\ 0 & 1\end{pmatrix}$, $\boldsymbol{A}_{22} = \begin{pmatrix}2 & 2\\ 0 & 1\end{pmatrix}$,
$\boldsymbol{B}_{11} = \begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$, $\boldsymbol{B}_{12} = \boldsymbol{B}_{13} = \begin{pmatrix}1 & 2\\ 1 & -1\end{pmatrix}$, $\boldsymbol{B}_{21} = \begin{pmatrix}2 & 1\\ -1 & 1\end{pmatrix}$, $\boldsymbol{B}_{22} = \boldsymbol{B}_{23} = \begin{pmatrix}2 & 2\\ -1 & -1\end{pmatrix}$.

Vynásobte $\boldsymbol{A}$ a $\boldsymbol{B}$ ako blokové matice aj ako matice, v ktorých ste zabudli na rozdelenie do blokov, a oba výsledky porovnajte.

2.6. Maticu $\boldsymbol{A} = (a_{ij})_{m\times n}$ nad ľubovoľným poľom K uvažujte ako riadok jej stĺpcov, t. j. ako blokovú maticu $\boldsymbol{A} = (\boldsymbol{u}_1,\dots,\boldsymbol{u}_n)$, kde $\boldsymbol{u}_j = \boldsymbol{s}_j(\boldsymbol{A})$ pre $j \le n$. Nech $\boldsymbol{c} =$

$(c_1, \ldots, c_n)^\mathsf{T} \in K^n$ je stĺpcový vektor. Ukážte, že lineárna kombinácia $c_1\boldsymbol{u}_1 + \ldots + c_n\boldsymbol{u}_n$ splýva s „obyčajným“ maticovým súčinom $\boldsymbol{A}\cdot\boldsymbol{c}$. Vysvetlite tento fakt pomocou násobenia blokových matíc.

2.7. Nech X je konečná množina. Ľubovoľná množina $H \subseteq X^2$ určuje *orientovaný graf* (X, H) s množinou *vrcholov* X a s množinou *orientovaných hrán* H: vrcholy (t. j. prvky množiny X) si znázorníme krúžkami v rovine a z vrcholu x vedieme orientovanú hranu (t. j. šípku) do vrcholu y práve vtedy, keď $(x, y) \in H$ (pozri obr. 2.1). Konečnú postupnosť $(z_0, z_1, \ldots, z_k)$ prvkov množiny X takú, že pre každé $1 \le i \le k$ platí $(z_{i-1}, z_i) \in H$, nazývame *cestou* dĺžky k *v orientovanom grafe* (X, H). Predpokladajme, že $X = \{x_1, \ldots, x_n\}$ má práve n prvkov. Maticu $\boldsymbol{H} = (h_{ij}) \in \mathbb{R}^{n\times n}$ takú, že $h_{ij} = 1$, ak $(x_i, x_j) \in H$, a $h_{ij} = 0$, ak $(x_i, x_j) \notin H$, nazývame *incidenčnou maticou orientovaného grafu* (X, H). Prvky k-tej mocniny incidenčnej matice $\boldsymbol{H}$ označme $h_{ij}^{(k)}$, t. j. $\boldsymbol{H}^k = \big(h_{ij}^{(k)}\big)$. Potom číslo $h_{ij}^{(k)}$ udáva počet ciest dĺžky k z vrcholu x_i do vrcholu x_j v orientovanom grafe (X, H). Dokážte (napr. matematickou indukciou).



Obr. 2.1. Príklady orientovaných grafov

2.8. Očíslujte vrcholy orientovaných grafov z obrázku 2.1). Pre každý graf napíšte jeho incidenčnú maticu a pre každú dvojicu (x_i, x_j) jeho vrcholov určte počet ciest dĺžky 2, 3, 4 a 5 z x_i do x_j.

2.9. Nech K je komutatívny okruh s jednotkou (pozri cvičenie 1.8). Presvedčte sa, že pre ľubovoľné $m, n \in \mathbb{N}$ možno na množine matíc $K^{m\times n}$ definovať operácie súčtu $\boldsymbol{A} + \boldsymbol{B}$ a skalárneho násobku $c\boldsymbol{A}$ rovnako ako v odstavci 2.2.1. Taktiež možno pre $\boldsymbol{A} \in K^{m\times n}$, $\boldsymbol{B} \in K^{n\times p}$ definovať súčin $\boldsymbol{A}\cdot\boldsymbol{B} \in K^{m\times p}$ rovnako ako v odstavci 2.2.2. Ukážte, že všetky vlastnosti maticových operácií uvedené v kapitole 2 zostávajú v platnosti aj v tomto všeobecnejšom prípade.

2.10. Vynechajme z definície poľa, popri nerovnosti $0 \ne 1$ a požiadavke existencie inverzného prvku ku každému nenulovému $a \in K$, aj podmienku komutatívnosti násobenia, namiesto ktorej pridajme ešte jeden distributívny zákon $(\forall\, a, b, c \in K)((a + b)c = ac + bc)$. Množina K s význačnými prvkami 0 a 1, vybavená operáciami sčítania a násobenia, ktoré vyhovujú uvedeným podmienkam, sa nazýva *okruh s jednotkou.*[1]

(a) Okruh s jednotkou K sa nazýva *netriviálny*, ak v ňom platí $0 \ne 1$. Dokážte, že okruh s jednotkou K je netriviálny práve vtedy, keď obsahuje aspoň dva rôzne prvky.

(b) Nech K je okruh s jednotkou. Presvedčte sa, že pre matice nad K možno zaviesť operácie súčtu, skalárneho násobku a súčinu rovnako ako pre matice nad poľom. Ukážte, že všetky vlastnosti týchto operácií uvedené v kapitole 2, s výnimkou

[1] Občas sa v literatúre takáto štruktúra nazýva len okruh.

rovností $(\boldsymbol{x} \cdot \boldsymbol{y})^\mathsf{T} = \boldsymbol{y}^\mathsf{T} \cdot \boldsymbol{x}^\mathsf{T}$, $(\boldsymbol{A} \cdot \boldsymbol{B})^\mathsf{T} = \boldsymbol{B}^\mathsf{T} \cdot \boldsymbol{A}^\mathsf{T}$ a možnosti zapisovať „lineárne kombinácie" v tvare $c_1\boldsymbol{u}_1 + \ldots c_n\boldsymbol{u}_n = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) \cdot (c_1, \ldots, c_n)^\mathsf{T}$, zostávajú v platnosti.

2.11. Nech K je okruh s jednotkou a $n \in \mathbb{N}$. Dokážte postupne nasledujúce tvrdenia:
(a) Množina $K^{n\times n}$ so sčítaním a násobením matíc tvorí okruh s jednotkou.

(b) Okruh s jednotkou $K^{n\times n}$ je triviálny práve vtedy, keď K je triviálny alebo $n = 0$.

(c) Okruh s jednotkou $K^{n\times n}$ je komutatívny práve vtedy, keď $n = 0$, alebo $n = 1$ a K je komutatívny.

2.12. (a) Rovnako ako v prípade poľa zadefinujte *charakteristiku* ľubovoľného *okruhu s jednotkou.*

(b) Dokážte, že okruh s jednotkou K je triviálny práve vtedy, keď $\operatorname{char} K = 1$.

(c) Nech K je ľubovoľný okruh a $1 \leq n \in \mathbb{N}$. Potom $\operatorname{char} K^{n\times n} = \operatorname{char} K$. Dokážte.

(d) Pre každé $2 \leq m \in \mathbb{N} \cup \{\infty\}$ uveďte príklad *nekomutatívneho* okruhu s jednotkou charakteristiky m.

2.13. Nech K je okruh s jednotkou.
(a) Pre všetky $n \in \mathbb{N}$ zadefinujte mocniny a^n prvku $a \in K$ rovnako ako v poli a dokážte, že pre $m, n \in \mathbb{N}$ platí $a^m a^n = a^{m+n}$, $(a^m)^n = a^{mn}$. Čo bráni definícii mocnín a^n pre záporné exponenty $n \in \mathbb{Z}$?

(b) Prvok $a \in K$ sa nazýva *invertovateľný*, ak k nemu existuje (obostranný, teda nutne jediný) inverzný prvok a^{-1} vzhľadom na násobenie. Pre invertovateľné prvky $a \in K$ rozšírte definíciu mocnín a^n na všetky $n \in \mathbb{Z}$ a dokážte rovnosti z (a) pre ľubovoľné $m, n \in \mathbb{Z}$.

(c) Nech $a, b \in K$. Čo je prekážkou všeobecnej platnosti rovnosti $(ab)^n = a^n b^n$ pre $n \geq 2$? Dokážte, že ak a, b *komutujú*, t. j. $ab = ba$, tak uvedená rovnosť platí pre všetky $n \in \mathbb{N}$.

(d) Pre $K = \mathbb{R}^{2\times 2}$ nájdite príklad matíc $\boldsymbol{A}, \boldsymbol{B} \in K$ takých, že $(\boldsymbol{A} \cdot \boldsymbol{B})^2 \neq \boldsymbol{A}^2 \cdot \boldsymbol{B}^2$.

(e) Nech $a, b \in K$ sú invertovateľné. Dokážte, že potom aj prvok ab je invertovateľný a platí $(ab)^{-1} = b^{-1}a^{-1}$. Čo je prekážkou všeobecnej platnosti rovnosti $(ab)^{-n} = b^{-n}a^{-n}$ pre $n \geq 2$?

2.14. (a) Rovnako ako v príklade 1.6.3 a v cvičení 1.12 zadefinujte množinu $K[x]$ všetkých polynómov v premennej x nad ľubovoľným okruhom s jednotkou K a na nej operácie súčtu a súčinu. Dokážte, že $K[x]$ s takto definovanými operáciami je opäť okruh s jednotkou a platí $\operatorname{char} K[x] = \operatorname{char} K$.

(b) Dokážte, že $K[x]$ je komutatívny práve vtedy, keď K je komutatívny.

(c) Dokážte, že polynóm $f(x)$ je invertovateľný prvok okruhu $K[x]$ práve vtedy, keď $f(x) = a$ je konštantný polynóm, pričom a je invertovateľný prvok okruhu K.

3. Sústavy lineárnych rovníc

V tejto kapitole sa predbežne zoznámime so sústavami lineárnych rovníc nad všeobecným poľom K a naučíme sa ich riešiť. Využijeme pri tom zápis sústavy pomocou istej matice. Štruktúrne vlastnosti množiny všetkých riešení danej sústavy a ich dôsledky preštudujeme a využijeme až neskôr, keď sa bližšie oboznámime so štruktúrou vektorových priestorov.

3.1 Maticový zápis sústavy lineárnych rovníc

Pod *lineárnou rovnicou* o n neznámych $x_1, \dots, x_n$ nad poľom K rozumieme formulu tvaru

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b,$$

kde $a_1, a_2, \dots, a_n, b \in K$, v premenných $x_1, x_2, \dots, x_n$. *Sústavou m lineárnych rovníc* o n neznámych $x_1, x_2, \dots, x_n$ nad poľom K rozumieme konjunkciu formúl tvaru

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{21}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \multicolumn{9}{c}{\dotfill} \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m, \end{array}$$

kde a_{ij}, b_i, pre $1 \le i \le m$, $1 \le j \le n$, sú skaláry z poľa K. Maticu $\boldsymbol{A} = (a_{ij}) \in K^{m\times n}$ nazývame *maticou sústavy*, stĺpcový vektor $\boldsymbol{b} = (b_1, \dots, b_m)^{\mathsf{T}} \in K^m$ nazývame jej *pravou stranou.* Konečne *rozšírenou maticou sústavy* nazývame blokovú maticu $(\boldsymbol{A} \,|\, \boldsymbol{b}) \in K^{m\times(n+1)}$. Sústava sa nazýva *homogénna*, ak $\boldsymbol{b} = \mathbf{0}$; v opačnom prípade sa nazýva *nehomogénna.*

Uvedenú sústavu možno stručne a úsporne zapísať v *maticovom tvare*

$$\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b},$$

prípadne, ak ide o homogénnu sústavu, v tvare

$$\boldsymbol{A} \cdot \boldsymbol{x} = \mathbf{0}.$$

Riešením sústavy $\boldsymbol{A}\cdot\boldsymbol{x} = \boldsymbol{b}$ nazývame ľubovoľný vektor-stĺpec $\boldsymbol{x} = (x_1, x_2, \dots, \dots, x_n)^{\mathsf{T}} \in K^n$, ktorého zložky vyhovujú každej z rovníc tejto sústavy, t. j. platí preň $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$. *Vyriešiť sústavu* znamená nájsť *všetky* jej riešenia, t. j. popísať *množinu* všetkých jej riešení.

Dve *sústavy* $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ a $\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}$, kde $\boldsymbol{A}, \boldsymbol{B} \in K^{m\times n}$, $\boldsymbol{b}, \boldsymbol{c} \in K^{m\times 1}$, sa nazývajú *ekvivalentné*, ak majú rovnakú množinu riešení, t. j. ak pre všetky $\boldsymbol{x} \in K^n$ platí $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ práve vtedy, keď $\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}$.

Skôr než prikročíme k otázke riešenia sústav lineárnych rovníc, považujeme za potrebné upozorniť čitateľa na dve veci.

(a) Podčiarkujeme, že riešením sústavy rozumieme vždy *vektor* $\boldsymbol{x}$ a nie jeho zložky. Tak napríklad sústava

$$\begin{aligned} 2x + 3y &= 12 \\ 3x - 2y &= 5 \end{aligned}$$

nad poľom $\mathbb{R}$ má, ako ľahko nahliadneme, jediné riešenie $\binom{x}{y} = \binom{3}{2}$ a nie dve riešenia $x = 3$, $y = 2$. Keď si toto poriadne uvedomíme, môžeme (a samozrejme aj budeme) sa naďalej vyjadrovať obvyklým spôsobom. Budeme teda hovoriť, že sústava má *jediné* riešenie $x = 3$, $y = 2$.

(b) Všimnite si, že počet rovníc sústavy a počet neznámych sa nemusia rovnať. V obvyklom prípade, keď rovníc je rovnaký počet ako neznámych, očakávame, že sústava bude mať jediné riešenie. Keď je rovníc menej než neznámych, môžeme očakávať, že sústava bude mať viacero (prípadne i nekonečne mnoho) riešení. Naopak, keď je rovníc viac ako neznámych, môže sa stať, že sústava nebude mať nijaké riešenie. Napriek tomu, že tieto očakávania vyjadrujú niečo ako „prevládajúci trend“, ľahko možno nájsť príklady, keď sa nemusia splniť. Zatiaľ len poznamenajme, že homogénna sústava $\boldsymbol{A}\cdot\boldsymbol{x} = \boldsymbol{0}$ má (bez ohľadu na počet neznámych a počet rovníc) vždy aspoň jedno riešenie – je ním nulový vektor $\boldsymbol{x} = \boldsymbol{0}$.

Nie je dôležité, akými znakmi sú označené neznáme v sústave $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$. Na jej riešenie nemá nijaký vplyv, či si vektor neznámych označíme $\boldsymbol{x} = (x_1, \dots, x_n)^\mathsf{T}$ alebo $\boldsymbol{y} = (y_1, \dots, y_n)^\mathsf{T}$ alebo nejako inak. To znamená, že celá informácia o tejto sústave, potrebná na nájdenie všetkých jej riešení, je obsiahnutá v rozšírenej matici sústavy $(\boldsymbol{A} \,|\, \boldsymbol{b})$, prípadne, ak ide o homogénnu sústavu, len v matici sústavy $\boldsymbol{A}$. Preto i metóda riešenia sústav lineárnych rovníc, s ktorou sa teraz zoznámime, bude založená len na úprave tejto matice.

Stručne povedané, rozšírenú maticu $(\boldsymbol{A} \,|\, \boldsymbol{b})$ sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ budeme upravovať tak, aby sme dostali vhodnú maticu $(\boldsymbol{B} \,|\, \boldsymbol{c})$, zodpovedajúcu novej sústave $\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}$, ktorá spĺňa nasledujúce dve podmienky:

(a) Je ekvivalentná s pôvodnou sústavou $\boldsymbol{A}\cdot\boldsymbol{x} = \boldsymbol{b}$, t. j. má rovnakú množinu riešení.

(b) Všetky jej riešenia možno priamo vyčítať z jej rozšírenej matice $(\boldsymbol{B} \,|\, \boldsymbol{c})$.

V takom prípade hovoríme, že sústava $\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}$ je *vyriešená*.

3.2 Redukovaný stupňovitý tvar matice

Našou prvou úlohou teda bude vyjasniť si, ako by mala vyzerať matica $(\boldsymbol{B} \,|\, \boldsymbol{c})$, aby sme príslušnú sústavu $\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}$ mohli považovať za vyriešenú. Za tým účelom teraz zavedieme niekoľko pojmov.

Hovoríme, že prvok a_{ij} matice $\boldsymbol{A} \in K^{m\times n}$ je *vedúci prvok* i*-teho riadku* matice $\boldsymbol{A}$, ak $a_{ij} \neq 0$, a $j = 1$ alebo $a_{il} = 0$ pre všetky $1 \leq l < j$. Inak povedané, vedúci prvok nenulového riadku je prvý nenulový prvok tohto riadku. Nulový riadok nemá vedúci prvok.

Hovoríme, že matica $\boldsymbol{A} = (a_{ij}) \in K^{m\times n}$ je v *redukovanom stupňovitom tvare*, ak spĺňa nasledujúce štyri podmienky:

(a) Ak $\boldsymbol{r}_i(\boldsymbol{A}) \neq \boldsymbol{0}$ a $\boldsymbol{r}_k(\boldsymbol{A}) = \boldsymbol{0}$, tak $i < k$; t. j. každý nenulový riadok matice $\boldsymbol{A}$ leží nad každým jej nulovým riadkom.

(b) Ak a_{ij}, a_{kl} sú vedúce prvky i-teho resp. k-teho riadku a $i < k$, tak aj $j < l$;

t. j. vedúci prvok vyššieho riadku leží viac vľavo než vedúci prvok nižšieho riadku.

(c) Ak a_{ij} je vedúci prvok i-teho riadku, tak $a_{ij} = 1$; t. j. vedúci prvok každého nenulového riadku je 1.

(d) Ak a_{ij} je vedúci prvok i-teho riadku, tak $a_{kj} = 0$ pre každé $k \neq i$; t. j. v stĺpci, v ktorom sa nachádza vedúci prvok nejakého riadku, sú všetky ostatné prvky rovné 0.

Pokiaľ matica $\boldsymbol{A}$ spĺňa len podmienky (a), (b), hovoríme, že je v *stupňovitom tvare*. Používa sa tiež názov (redukovaný) *schodovitý tvar*.

Napríklad z uvedených matíc nad poľom $\mathbb{R}$

$$\begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 2 & 3 & 0 & 1 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

ani jedna z matíc vľavo nie je v stupňovitom tvar; obe matice v strednom stĺpci sú v stupňovitom tvare, nie však v redukovanom stupňovitom tvare; konečne, obe matice vpravo sú v redukovanom stupňovitom tvare. (V každom jednotlivom prípade si podrobne premyslite prečo.)

Taktiež každá jednotková matica $\boldsymbol{I}_n$, ako aj všetky nulové matice $\boldsymbol{0}_{mn}$ sú v redukovanom stupňovitom tvare.

Uvedomme si teraz, akej sústave lineárnych rovníc zodpovedá rozšírená matica v redukovanom stupňovitom tvare. Napríklad

$$(\boldsymbol{B} \,|\, \boldsymbol{c}) = \left(\begin{array}{ccccc|c} 1 & 0 & -2 & 0 & 0 & 3 \\ 0 & 1 & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{array}\right)$$

je matica v redukovanom stupňovitom tvare nad $\mathbb{R}$. Táto matica zodpovedá sústave

$$\begin{aligned} x_1 \qquad - 2x_3 \qquad &= 3 \\ x_2 + 6x_3 \qquad &= 0 \\ x_4 &= 1 \end{aligned}$$

v neznámych x_1, x_2, x_3, x_4, x_5. Hneď vidíme, že táto sústava má nekonečne mnoho riešení. Každej voľbe *parametrov* $s, t \in \mathbb{R}$ totiž zodpovedá jedno riešenie

$$\begin{aligned} x_1 &= 3 + 2s \\ x_2 &= \quad - 6s \\ x_3 &= \qquad s \\ x_4 &= 1 \\ x_5 &= \qquad\qquad t. \end{aligned}$$

Asi sa zhodneme na tom, že preznačenie neznámych za parametre $x_3 = s$, $x_5 = t$ a ich presun na pravú stranu, je úprava natoľko bezprostredná, že sústavu prislúchajúcu k matici $(\boldsymbol{B} \,|\, \boldsymbol{c})$ už možno považovať za vyriešenú. Na napísanie jej riešenia nemusíme písať príslušnú sústavu, môžeme ho napísať priamo na základe matice $(\boldsymbol{B} \,|\, \boldsymbol{c})$.

Dohodneme sa teda, že sústavu lineárnych rovníc $\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}$ nad poľom K budeme nazývať *vyriešenou sústavou*, ak jej rozšírená matica $(\boldsymbol{B} \,|\, \boldsymbol{c})$ je v redukovanom stupňovitom tvare. V prípade homogénnej sústavy sa, samozrejme, stačí obmedziť na maticu $\boldsymbol{B}$.

Teraz si predvedieme, ako možno k danej blokovej matici $(\boldsymbol{B} \,|\, \boldsymbol{c})$ v redukovanom stupňovitom tvare, nájsť všetky riešenia sústavy $\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}$.

Najprv si ujasníme, kedy je taká sústava *riešiteľná*, t. j. má aspoň jedno riešenie. Odpoveď na túto otázku je jednoduchá: Sústava $\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}$ má riešenie práve vtedy, keď sa v matici $(\boldsymbol{B} \,|\, \boldsymbol{c})$ nenachádza riadok tvaru

$$(\underbrace{0, \dots, 0}_{n\text{-krát}} \,|\, 1).$$

Taký riadok totiž zodpovedá rovnici $0 = 1$, ktorá očividne nemá riešenie. To, že neprítomnosť takého riadku je i postačujúcou podmienkou riešiteľnosti sústavy, vyplýva z nasledujúceho postupu, ako toto riešenie nájsť.

Ak sa v j-tom stĺpci matice $\boldsymbol{B}$ nenachádza vedúci prvok žiadneho riadku, tak si neznámu x_j zvolíme za parameter; ak sa v j-tom stĺpci nachádza vedúci prvok nejakého riadku, tak neznámu x_j si vyjadríme pomocou parametrov tak, že stĺpce matice $\boldsymbol{B}$ prislúchajúce týmto parametrom „prehodíme s opačným znamienkom na druhú stranu“.

Presnejšiu formuláciu celého postupu vo všeobecnej podobe si odpustíme. Názornejšie bude osvetliť ho na ešte jednom príklade.

$$(\boldsymbol{B}\,|\,\boldsymbol{c}) = \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 2/3 & -1/2 & 5 \\ 0 & 1 & 0 & 3/4 & 0 & 2 \\ 0 & 0 & 1 & -4 & -2/5 & -2 \end{array}\right)$$

je reálna matica v redukovanom stupňovitom tvare. Vidíme, že sa v nej nenachádza riadok tvaru $(0,0,0,0\,|\,1)$, teda sústava $\boldsymbol{B}\cdot\boldsymbol{x} = \boldsymbol{c}$ by mala mať riešenie. Vedúce prvky riadkov matice $\boldsymbol{B}$ sa nachádzajú v stĺpcoch 1, 2 a 3. Za parametre si teda zvolíme neznáme x_4 a x_5. Riešením sústavy je každý vektor $(x_1,x_2,x_3,x_4,x_5)^{\mathsf{T}} \in \mathbb{R}$ tvaru

$$\begin{aligned} x_1 &= 5 - \frac{2}{3}s + \frac{1}{2}t \\ x_2 &= 2 - \frac{3}{4}s \\ x_3 &= -2 + 4s + \frac{2}{5}t \\ x_4 &= \qquad\quad s \\ x_5 &= \qquad\qquad\quad t, \end{aligned}$$

kde parametre $s,t \in \mathbb{R}$ môžu nadobúdať ľubovoľné hodnoty. Pre estétov ešte poznamenajme, že zlomkov pri parametroch sa možno jednoducho zbaviť. Je totiž jedno, či si parametrické premenné zvolíme v tvare $x_4 = s$, $x_5 = t$ alebo v tvare $x_4 = 12s$, $x_5 = 10t$, kde $s,t \in \mathbb{R}$. Pri takejto voľbe parametrov dostaneme všetky riešenia sústavy v tvare bez zlomkov

$$\begin{aligned} x_1 &= 5 - 8s + 5t \\ x_2 &= 2 - 9s \\ x_3 &= -2 + 36s + 4t \\ x_4 &= \qquad 12s \\ x_5 &= \qquad\qquad 10t. \end{aligned}$$

3.3 Gaussova-Jordanova eliminačná metóda

Zatiaľ sme si ujasnili, na *aký tvar* treba upraviť rozšírenú maticu $(\boldsymbol{A}\,|\,\boldsymbol{b})$ sústavy $\boldsymbol{A}\cdot\boldsymbol{x} = \boldsymbol{b}$, aby sme získali s ňou ekvivalentnú *vyriešenú* sústavu $\boldsymbol{B}\cdot\boldsymbol{x} = \boldsymbol{c}$: rozšírená matica $(\boldsymbol{B}\,|\,\boldsymbol{c})$ novej sústavy musí byť v *redukovanom stupňovitom tvare*. Teraz si ukážeme, *ako* to možno urobiť. Maticu $(\boldsymbol{A}\,|\,\boldsymbol{b})$ budeme na redukovaný stupňovitý tvar upravovať pomocou tzv. elementárnych riadkových operácií.

Elementárnou riadkovou operáciou, skrátene tiež ERO, na matici $\boldsymbol{A} \in K^{m\times n}$ rozumieme

I. výmenu dvoch riadkov matice $\boldsymbol{A}$;
II. vynásobenie niektorého riadku matice $\boldsymbol{A}$ *nenulovým* skalárom z poľa K;
III. pripočítanie skalárneho násobku niektorého riadku matice $\boldsymbol{A}$ k jej inému riadku.

Matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{m\times n}$ sa nazývajú *riadkovo ekvivalentné*, označenie $\boldsymbol{A} \sim \boldsymbol{B}$, ak jednu z nich možno upraviť na druhú konečným počtom elementárnych riadkových operácií. Riešenie sústavy lineárnych rovníc úpravou jej rozšírenej matice pomocou ERO na riadkovo ekvivalentnú maticu v redukovanom stupňovitom tvare sa nazýva *Gaussova-Jordanova eliminácia.*

Prenechávame čitateľovi, aby si sám sformuloval analogické pojmy *elementárnych stĺpcových operácií* (ESO) a *stĺpcovej ekvivalencie* matíc, označenie $\boldsymbol{A} \wr \boldsymbol{B}$. Ich význam vyjde najavo až v neskorších kapitolách.

Výmenou i-teho a k-teho riadku v matici

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{r}_1(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{r}_i(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{r}_k(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{r}_m(\boldsymbol{A}) \end{pmatrix} \qquad \text{dostaneme maticu} \qquad \begin{pmatrix} \boldsymbol{r}_1(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{r}_k(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{r}_i(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{r}_m(\boldsymbol{A}) \end{pmatrix}.$$

Opätovnou výmenou i-teho a k-teho riadku v tejto matici získame zasa maticu $\boldsymbol{A}$.

Vynásobením i-teho riadku matice $\boldsymbol{A}$ skalárom $c \neq 0$ dostaneme maticu

$$\begin{pmatrix} \boldsymbol{r}_1(\boldsymbol{A}) \\ \vdots \\ c\boldsymbol{r}_i(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{r}_k(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{r}_m(\boldsymbol{A}) \end{pmatrix}.$$

Vynásobením i-teho riadku tejto matice skalárom $c^{-1} \neq 0$ získame opäť maticu $\boldsymbol{A}$.

Konečne, pripočítaním c-násobku i-teho riadku matice $\boldsymbol{A}$ k jej k-temu riadku z nej dostaneme maticu

$$\begin{pmatrix} \boldsymbol{r}_1(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{r}_i(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{r}_k(\boldsymbol{A}) + c\boldsymbol{r}_i(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{r}_m(\boldsymbol{A}) \end{pmatrix}.$$

Všimnite si, že i-ty riadok pri tom zostáva nezmenený. Maticu $\boldsymbol{A}$ z tejto matice získame pripočítaním $(-c)$-násobku jej i-teho riadku ku k-temu riadku.

Ak $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ je sústava s rozšírenou maticou $(\boldsymbol{A} \,|\, \boldsymbol{b})$ a bloková matica $(\boldsymbol{A}' \,|\, \boldsymbol{b}')$ vznikne z $(\boldsymbol{A} \,|\, \boldsymbol{b})$ vykonaním jednej (nezáleží ktorej) ERO, tak sústava $\boldsymbol{A}' \cdot \boldsymbol{x} = \boldsymbol{b}'$ je ekvivalentná s pôvodnou sústavou $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$. Elementárne riadkové operácie na matici $(\boldsymbol{A} \,|\, \boldsymbol{b})$ totiž zodpovedajú postupne zámene poradia dvoch rovníc sústavy, vynásobením niektorej rovnice nenulovým skalárom a pripočítaním nejakého násobku jednej rovnice k inej rovnici (presnejšie nahradením dojice rovníc $\boldsymbol{r}_i(\boldsymbol{A}) \cdot \boldsymbol{x} = b_i$, $\boldsymbol{r}_k(\boldsymbol{A}) \cdot \boldsymbol{x} = b_k$ dvojicou rovníc $\boldsymbol{r}_i(\boldsymbol{A}) \cdot \boldsymbol{x} = b_i$, $(\boldsymbol{r}_k(\boldsymbol{A}) + c\boldsymbol{r}_i(\boldsymbol{A})) \cdot \boldsymbol{x} = b_k + cb_i$). Z pred chvíľou vykonaných úvah vyplýva, že ide o ekvivalentné úpravy, ktorými sa množina riešení sústavy nezmení – od novej sústavy $\boldsymbol{A}' \cdot \boldsymbol{x} = \boldsymbol{b}'$ sa možno vhodnou ERO vykonanou na jej rozšírenej matici opäť vrátiť k pôvodnej sústave $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$.

3.3.1. Tvrdenie. *Nech K je pole, $\boldsymbol{A}, \boldsymbol{B} \in K^{m \times n}$, $\boldsymbol{b}, \boldsymbol{c} \in K^m$. Ak sú blokové matice $(\boldsymbol{A} \,|\, \boldsymbol{b})$, $(\boldsymbol{B} \,|\, \boldsymbol{c})$ riadkovo ekvivalentné, tak i sústavy lineárnych rovníc $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}$ sú ekvivalentné.*

Dôkaz. Podľa predpokladu existuje postupnosť blokových matíc $(\boldsymbol{A}\,|\,\boldsymbol{b}) = (\boldsymbol{C}_0\,|\,\boldsymbol{d}_0)$, $(\boldsymbol{C}_1\,|\,\boldsymbol{d}_1)$, $\dots, (\boldsymbol{C}_p\,|\,\boldsymbol{d}_p) = (\boldsymbol{B}\,|\,\boldsymbol{c})$ typu $m \times (n+1)$ takých, že pre každé $l < p$ matica $(\boldsymbol{C}_{l+1}\,|\,\boldsymbol{d}_{l+1})$ vznikne vykonaním jedinej ERO z matice $(\boldsymbol{C}_l\,|\,\boldsymbol{d}_l)$. Potom všetky sústavy $\boldsymbol{C}_l \cdot \boldsymbol{x} = \boldsymbol{d}_l$ majú tú istú množinu riešení, t. j. sú ekvivalentné.

3.3.2. Veta. *Každá matica nad poľom K je riadkovo ekvivalentná s jednoznačne určenou maticou v redukovanom stupňovitom tvare.*

Dôkaz existencie a jednoznačnosti spomínanej matice odložíme do cvičení 3.12 a 3.13 (pozri tiež cvičenie 5.13). Zatiaľ sa radšej len na konkrétnych príkladoch naučíme, ako možno k danej matici $\boldsymbol{A}$ nájsť s ňou riadkovo ekvivalentnú maticu v redukovanom stupňovitom tvare. Takýto prístup má navyše tú výhodu, že z radu možných postupov, medzi ktorými si možno pružne voliť podľa okolností, nám nesugeruje jedinú stratégiu, na ktorú by sme sa nevyhnutne museli obmedziť pri všeobecnom dôkaze. Ambicióznejší čitateľ v uvedených príkladoch ľahko i sám zahliadne myšlienku všeobecného dôkazu, ktorú potom bude môcť uplatniť v cvičení 3.12. Napokon, aby sme sa nebavili iba o maticiach, začneme zakaždým s nejakou sústavou lineárnych rovníc. Tým sa zároveň naučíme *riešiť* ľubovoľnú sústavu lineárnych rovníc nad daným poľom Gaussovou-Jordanovou elimináciou, prípadne rozpoznať, že daná sústava nemá riešenie.

3.3.3. Príklad. Je daná sústava

$$\begin{aligned} 2x_1 + 3x_2 \phantom{{}+4x_3} - x_4 &= 1\\ 3x_1 + 2x_2 + 4x_3 - 2x_4 &= 0\\ x_1 - x_2 + 4x_3 - x_4 &= 2 \end{aligned}$$

troch rovníc o štyroch neznámych nad poľom $\mathbb{R}$. Jej rozšírená matica je

$$\left(\begin{array}{rrrr|r} 2 & 3 & 0 & -1 & 1\\ 3 & 2 & 4 & -2 & 0\\ 1 & -1 & 4 & -1 & 2 \end{array}\right).$$

Pri jej úprave na redukovaný stupňovitý tvar (podobne ako i v ďalších príkladoch) budeme vynechávať niektoré medzikroky a zaznamenáme len niektoré výsledky viacerých vykonaných ERO. Posledný riadok matice dáme na prvé miesto, potom jeho (-2)-násobok pripočítame k pôvodnému prvému riadku, ktorý posunieme na druhé miesto, a (-3)-násobok toho istého riadku pripočítame k pôvodnému druhému riadku, ktorý posunieme na tretie miesto. Dostaneme tak maticu

$$\left(\begin{array}{rrrr|r} 1 & -1 & 4 & -1 & 2\\ 0 & 5 & -8 & 1 & -3\\ 0 & 5 & -8 & 1 & -6 \end{array}\right).$$

Pripočítaním (-1)-násobku druhého riadku k tretiemu dostaneme maticu

$$\left(\begin{array}{cccc|c} 1 & -1 & 4 & -1 & 2 \\ 0 & 5 & -8 & 1 & -3 \\ 0 & 0 & 0 & 0 & -3 \end{array}\right).$$

Už z tohto tvaru vidíme, že sústava zodpovedajúca poslednej matici nemá riešenie – obsahuje totiž rovnicu $0 = -3$. Teda ani pôvodná sústava (hoci neznámych je v nej viac než rovníc) nemá riešenie. Z cvičných dôvodov však dokončíme úpravu na redukovaný stupňovitý tvar, ktorý dostaneme vynásobením tretieho riadku skalárom $-1/3$, pripočítaním (-2)-násobku resp. 3-násobku tohto nového riadku k prvému resp. druhému riadku a, konečne, vynásobením druhého riadku skalárom $1/5$:

$$\left(\begin{array}{cccc|c} 1 & 0 & 12/5 & -4/5 & 0 \\ 0 & 1 & -8/5 & 1/5 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right).$$

Čitateľ by si mal všimnúť, že po nastavení vedúceho prvku niektorého riadku na hodnotu 1 okamžite pristupujeme k nulovaniu zvyšných prvkov stĺpca, v ktorom leží tento vedúci prvok.

3.3.4. Príklad. Riešme sústavu

$$\begin{array}{rcrcrcl} x & + & 2\mathrm{i}y & & & = & 5+4\mathrm{i} \\ & & (3-\mathrm{i})y & + & (6-2\mathrm{i})z & = & 10 \\ 2x & & & - & z & = & 5+3\mathrm{i} \\ x & + & y & + & z & = & 5+2\mathrm{i} \end{array}$$

štyroch rovníc o troch neznámych nad poľom $\mathbb{C}$. Jej rozšírená matica je

$$\left(\begin{array}{ccc|c} 1 & 2\mathrm{i} & 0 & 5+4\mathrm{i} \\ 0 & 3-\mathrm{i} & 6-2\mathrm{i} & 10 \\ 2 & 0 & -1 & 5+3\mathrm{i} \\ 1 & 1 & 1 & 5+2\mathrm{i} \end{array}\right).$$

Prehoďme jej posledný riadok na prvé miesto a zvyšné riadky posuňme o jedno miesto nadol. V takto získanej matici pripočítajme (-1)-násobok prvého riadku k druhému riadku a (-2)-násobok prvého riadku k štvrtému riadku. Konečne vynásobme tretí riadok skalárom $(3+\mathrm{i})/10$. Dostaneme maticu

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 5+2\mathrm{i} \\ 0 & -1+2\mathrm{i} & -1 & 2\mathrm{i} \\ 0 & 1 & 2 & 3+\mathrm{i} \\ 0 & -2 & -3 & -5-\mathrm{i} \end{array}\right).$$

(-1)-násobok tretieho riadku pripočítame k prvému riadku, jeho $(1-2\mathrm{i})$-násobok k druhému a 2-násobok k štvrtému. Nakoniec výmenou druhého a tretieho riadku dostaneme maticu

$$\left(\begin{array}{ccc|c} 1 & 0 & -1 & 2+\mathrm{i} \\ 0 & 1 & 2 & 3+\mathrm{i} \\ 0 & 0 & 1-4\mathrm{i} & 5-3\mathrm{i} \\ 0 & 0 & 1 & 1+\mathrm{i} \end{array}\right).$$

Pripočítajme posledný riadok k prvému, (-2)-násobok posledného riadku k druhému a jeho $(-1+4\mathrm{i})$-násobok k tretiemu. Zostáva vymeniť tretí a štvrtý riadok – výsledná matica je už v redukovanom stupňovitom tvare

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 3+2\mathrm{i} \\ 0 & 1 & 0 & 1-\mathrm{i} \\ 0 & 0 & 1 & 1+\mathrm{i} \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Vidíme, že pôvodná sústava (hoci obsahuje viac rovníc než neznámych) má jediné riešenie $x = 3+2\mathrm{i}$, $y = 1-\mathrm{i}$, $z = 1+\mathrm{i}$, teda presnejšie vektor $(3+2\mathrm{i}, 1-\mathrm{i}, 1+\mathrm{i})^{\mathsf{T}} \in \mathbb{C}^3$.

3.3.5. Príklad. Uvažujme sústavu

$$\begin{aligned} x_1 + x_2 + 2x_3 + 3x_4 &= 0 \\ 2x_1 \qquad + 4x_3 \qquad &= 0 \\ x_1 + 2x_2 + x_3 + 3x_4 &= 0 \\ 3x_3 + 4x_4 &= 0 \end{aligned}$$

štyroch rovníc o štyroch neznámych nad poľom $\mathbb{Z}_5$. Keďže ide o homogénnu sústavu (ktorej ľavá strana je nulový stĺpcový vektor, teda sa nemení pri žiadnej ERO), stačí upravovať jej (nerozšírenú) maticu

$$\begin{pmatrix} 1 & 1 & 2 & 3 \\ 2 & 0 & 4 & 0 \\ 1 & 2 & 1 & 3 \\ 0 & 0 & 3 & 4 \end{pmatrix}.$$

(-2)-násobok, t. j. 3-násobok prvého riadku pripočítame k druhému riadku a jeho
(-1)-násobok, t. j. 4-násobok pripočítame k tretiemu riadku. Dostaneme tak

maticu

$$\begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 3 & 0 & 4 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 3 & 4 \end{pmatrix}.$$

(-1)-násobok, t. j. 4-násobok tretieho riadku pripočítame k prvému riadku a jeho
(-3)-násobok, t. j. 2-násobok pripočítame k druhému riadku. Konečne výmenou druhého a tretieho riadku dostaneme maticu

$$\begin{pmatrix} 1 & 0 & 3 & 3 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 3 & 4 \end{pmatrix}.$$

Tretí riadok odpočítame od prvého aj od štvrtého riadku. Ďalej ho vynásobíme skalárom $3^{-1} = 2$. Napokon jeho (-4)-násobok, t. j. priamo tento nový tretí riadok pripočítame k druhému riadku. Výsledná matica je už v redukovanom stupňovitom tvare

$$\begin{pmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Premennú x_4 si zvolíme za parameter. Všetky riešenia sústavy majú potom tvar $x_1 = t$, $x_2 = 2t$, $x_3 = 2t$, $x_4 = t$, kde $t \in \mathbb{Z}_5$. Vidíme teda, že pôvodná sústava (hoci počet jej rovníc je rovnaký ako počet neznámych) má viac než jedno riešenie; nie je ich však nekonečne veľa ale len 5. Práve toľko je totiž možných volieb parametra t, t. j. prvkov poľa $\mathbb{Z}_5$.

Zaznamenajme ešte jeden očakávaný dôsledok tvrdenia 3.3.1, vety 3.3.2 a spôsobu, ako napísať riešenie sústavy s (rozšírenou) maticou v redukovanom stupňovitom tvare, uvedeného v paragrafe 3.2.

3.3.6. Tvrdenie. *Nech $\boldsymbol{A} \in K^{m\times n}$, $\boldsymbol{b} \in K^m$ a $m < n$, t. j. sústavy $\boldsymbol{A}\cdot\boldsymbol{x} = \mathbf{0}$, $\boldsymbol{A}\cdot\boldsymbol{x} = \boldsymbol{b}$ obsahujú menej rovníc než neznámych. Potom*

(a) homogénna sústava $\boldsymbol{A}\cdot\boldsymbol{x} = \mathbf{0}$ má popri riešení $\boldsymbol{x}_0 = \mathbf{0}$ aspoň jedno riešenie $\boldsymbol{x} \neq \mathbf{0}$;

(b) ak existuje aspoň jedno riešenie sústavy $\boldsymbol{A}\cdot\boldsymbol{x} = \boldsymbol{b}$, tak táto sústava má viac než jedno riešenie.

Dôkaz. (a) Upravme maticu sústavy $\boldsymbol{A}$ na redukovaný stupňovitý tvar $\boldsymbol{B}$. Uvedomme si, že matice $\boldsymbol{A}$ aj $\boldsymbol{B}$ majú m riadkov a n stĺpcov. Riadky matice

$\boldsymbol{B}$ majú nanajvýš m vedúcich prvkov. Keďže $m < n$, aspoň v jednom stĺpci matice $\boldsymbol{B}$ neleží vedúci prvok žiadneho riadku. Nech je to napr. j-ty stĺpec. Potom voľbe parametra $x_j = t \in K$, $t \neq 0$ zodpovedá aspoň jedno nenulové riešenie sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{0}$.

(b) prenechávame ako cvičenie čitateľovi.

3.4 Gaussova eliminačná metóda

Hlavne z historických dôvodov ešte stručne spomenieme metódu riešenia sústav lineárnych rovníc tzv. *Gaussovou elimináciou.* Pri riešení touto metódou upravíme rozšírenú maticu sústavy len na stupňovitý (teda nie nevyhnutne redukovaný stupňovitý) tvar. Už z tohto tvaru možno ľahko spoznať, či sústava má nejaké riešenie (príslušná matica nesmie obsahovať riadok tvaru $(0, \dots, 0 \,|\, d)$, kde $0 \neq d \in K$). V tom prípade možno všetky riešenia sústavy získať voľbou parametrov (opäť si za ne volíme neznáme x_j také, že j-tom stĺpci sa nevyskytuje vedúci prvok žiadneho riadku) a spätným dosadzovaním, t. j. *elimináciou* neznámych pomocou parametrov.

3.4.1. Príklad. Predpokladajme, že rozšírenú maticu nejakej sústavy nad $\mathbb{R}$ sme už pomocou ERO upravili na stupňovitý tvar

$$\left(\begin{array}{cccccc|c} 0 & 2 & 3 & 0 & -1 & 4 & 1 \\ 0 & 0 & 0 & -2 & 5 & 4 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 & 4 \end{array}\right).$$

Táto matica zodpovedá sústave

$$\begin{aligned} 2x_2 + 3x_3 \qquad - \ \ x_5 + 4x_6 &= 1 \\ -\, 2x_4 + 5x_5 + 4x_6 &= 0 \\ 3x_5 + \ \ x_6 &= 4. \end{aligned}$$

Za parametre si zvolíme premenné x_1, x_3 a x_6. Spätným dosadzovaním postupne dostaneme všetky riešenia v parametrickom tvare

$$\begin{aligned} x_6 &= t \\ x_5 &= \frac{1}{3}(4 - x_6) = \frac{4}{3} - \frac{1}{3}t \\ x_4 &= \frac{1}{2}(5x_5 + 4x_6) = \frac{10}{3} - \frac{7}{6}t \\ x_3 &= s \\ x_2 &= \frac{1}{2}(1 - 3x_3 + x_5 - 4x_6) = \frac{7}{6} - \frac{3}{2}s - \frac{13}{6}t \\ x_1 &= r, \end{aligned}$$

kde $r, s, t \in \mathbb{R}$. Prípadne, po trochu „šikovnejšej“ voľbe parametrov, v tvare $x_6 = 6t$, $x_5 = \frac{4}{3} - 2t$, $x_4 = \frac{10}{3} - 7t$, $x_3 = 2s$, $x_2 = \frac{7}{6} - 3s + 13t$, $x_1 = r$.

Pozorný čitateľ si iste všimol, že spätné dosadzovanie možno nahradiť ďalšou úpravou rozšírenej matice sústavy pomocou ERO na *redukovaný* stupňovitý tvar. Stačí totiž vynásobiť nenulové riadky prevrátenými hodnotami ich vedúcich prvkov a pripočítaním vhodných násobkov týchto riadkov vynulovať zvyšné nenulové prvky v stĺpcoch obsahujúcich vedúce prvky jednotlivých riadkov.

I tak však môže byť Gaussova eliminačná metóda v niektorých prípadoch užitočná – najmä keď nám nejde ani tak o explicitný tvar riešení, ako skôr o samotnú otázku riešiteľnosti sústavy, prípadne o počet parametrov, ktoré sa v nich vyskytujú. Všetko to možno totiž spoznať už na základe nejakej matice v stupňovitom tvare, riadkovo ekvivalentnej s pôvodnou rozšírenou maticou sústavy. V takom prípade si teda môžeme odpustiť nielen ďalšiu úpravu na redukovaný stupňovitý tvar, ale aj spätné dosadzovanie.

Cvičenia

3.1. Podrobne prepočítajte sústavy lineárnych rovníc z príkladov 3.3.3 – 3.3.5

3.2. Gaussovou-Jordanovou elimináciou riešte sústavu lineárnych rovníc $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ nad poľom $\mathbb{R}$ pre matice:

(a) $\boldsymbol{A} = \begin{pmatrix} 2 & 1 & 2 & 0 & -6 \\ 1 & 0 & 0 & 2 & -5 \\ 1 & 1 & 2 & 0 & -3 \\ 0 & 1 & 2 & 0 & 1 \end{pmatrix}$, $\boldsymbol{b} = \begin{pmatrix} 4 \\ 1 \\ 1 \\ 2 \end{pmatrix}$; (b) $\boldsymbol{A} = \begin{pmatrix} 3 & -6 & 1 & 10 \\ 1 & -2 & 1 & 4 \\ 2 & -4 & 3 & 9 \\ 1 & -2 & 0 & 3 \end{pmatrix}$, $\boldsymbol{b} = \begin{pmatrix} 6 \\ 3 \\ 2 \\ 2 \end{pmatrix}$.

3.3. Gaussovou-Jordanovou elimináciou riešte sústavu lineárnych rovníc $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ nad poľom $\mathbb{C}$ pre matice:

(a) $\boldsymbol{A} = \begin{pmatrix} 1+\mathrm{i} & 0 & \mathrm{i}-1 \\ 2-\mathrm{i} & 1 & 1+\mathrm{i} \end{pmatrix}$, $\boldsymbol{b} = \begin{pmatrix} 1-2\mathrm{i} \\ 0 \end{pmatrix}$; (b) $\boldsymbol{A} = \begin{pmatrix} 1+2\mathrm{i} & 2-\mathrm{i} & 1-\mathrm{i} \\ 5\mathrm{i} & 5 & 3-\mathrm{i} \end{pmatrix}$, $\boldsymbol{b} = \begin{pmatrix} 1 \\ 2+\mathrm{i} \end{pmatrix}$.

3.4. Gaussovou-Jordanovou elimináciou riešte sústavu $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ lineárnych rovníc nad poľom $\mathbb{Z}_{11}$ pre matice:

(a) $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 1 \\ 7 & 3 & 5 & 3 \end{pmatrix}$, $\boldsymbol{b} = \begin{pmatrix} 1 \\ 5 \\ 1 \end{pmatrix}$; (b) $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 4 & 5 \\ 3 & 2 & 0 & 1 \\ 1 & 9 & 1 & 7 \end{pmatrix}$, $\boldsymbol{b} = \boldsymbol{0}$.

Pre každú sústavu určte počet jej riešení.

3.5. Riešte sústavy z cvičenia 3.4 nad poľom $\mathbb{Z}_{13}$ a opäť určte počet riešení každej z nich.

3.6. V paragrafe 3.2 definovaný (redukovaný) stupňovitý tvar matice by sme mohli presnejšie nazvať *riadkovým* (redukovaným) stupňovitým tvarom. Sformulujte definíciu *stĺpcového* (redukovaného) stupňovitého tvaru matice. Podrobne definujte elementárne stĺpcové operácie (ESO) typov I, II a III.

3.7. Nech $\boldsymbol{A} = (a_{ij})_{m\times n}$, $\boldsymbol{B} = (\boldsymbol{b}_{ik})_{m\times p}$ sú matice nad poľom K. Označme $\boldsymbol{b}_k = \boldsymbol{s}_k(\boldsymbol{B})$ k-ty stĺpec matice $\boldsymbol{B}$. Uvažujme maticovú rovnicu $\boldsymbol{A} \cdot \boldsymbol{X} = \boldsymbol{B}$ s neznámou maticou $\boldsymbol{X} = (x_{jk})_{n\times p}$. Dokážte postupne nasledujúce tvrdenia:
(a) Matica $\boldsymbol{X} \in K^{n\times p}$ je riešením maticovej rovnice $\boldsymbol{A} \cdot \boldsymbol{X} = \boldsymbol{B}$ práve vtedy, keď pre každé $k \le p$ je jej k-ty stĺpec $\boldsymbol{x}_k = \boldsymbol{s}_k(\boldsymbol{X})$ riešením sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}_k$.
(b) Maticová rovnica $\boldsymbol{A} \cdot \boldsymbol{X} = \boldsymbol{B}$ má riešenie práve vtedy, keď každá zo sústav

$\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}_k$ $(k = 1, \dots, p)$ má riešenie.
Na základe (a) a (b) navrhnite metódu, ako možno úpravou vhodnej blokovej matice pomocou ERO riešiť naraz viacero sústav $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}_k$ s rovnakou ľavou stranou a rôznymi pravými stranami $\boldsymbol{b}_1, \dots, \boldsymbol{b}_p$.

3.8. Nech $\boldsymbol{A} = (a_{ij})_{m\times n}$, $\boldsymbol{C} = (c_{kj})_{q\times n}$ sú matice nad poľom K. S hromadným riešením akých sústav lineárnych rovníc súvisí riešenie maticovej rovnice $\boldsymbol{Y} \cdot \boldsymbol{A} = \boldsymbol{C}$ s neznámou maticou $\boldsymbol{Y} = (y_{ki})_{q\times m}$? Navrhnite metódu založenú na úprave vhodnej blokovej matice pomocou ESO. Ako sa možno vyhnúť ESO a nahradiť ich ERO?

3.9. Riešte maticové rovnice $\boldsymbol{A}\cdot\boldsymbol{X} = \boldsymbol{B}$ a $\boldsymbol{Y}\cdot\boldsymbol{A} = \boldsymbol{C}$ pre matice $\boldsymbol{A} = \left(\begin{smallmatrix} 3 & 2 & 4 \\ 2 & 1 & 3 \end{smallmatrix}\right)$, $\boldsymbol{B} = \left(\begin{smallmatrix} 1 & 3 & -1 \\ 1 & 2 & 4 \end{smallmatrix}\right)$, $\boldsymbol{C} = \left(\begin{smallmatrix} 12 & 7 & 17 \\ 3 & 3 & 7 \end{smallmatrix}\right)$ nad poľom $\mathbb{Q}$. Určte najprv rozmery matíc $\boldsymbol{X}$ a $\boldsymbol{Y}$. Aké sústavy lineárnych rovníc ste takto vyriešili? Napíšte riešenie každej z nich (ak existuje).

3.10. Nad poľom $\mathbb{R}$ riešte sústavy lineárnych rovníc v neznámych x, y, z a urobte diskusiu počtu riešení vzhľadom na parametre $a \in \mathbb{R}$ resp. $c, d \in \mathbb{R}$:

(a) $x + y + (2a^2 - 1)z = a^2 + a + 1,$
$x + y + (a^2 + a - 1)z = 1,$
$x + a^2 z = a^2 + a;$

(b) $cx + y + (c+1)dz = c + 2d + 1,$
$cx + cdz = d + 1,$
$cy + 2cdz = 2c^2 + cd - 1.$

3.11. Nech K je pole a $m, n \in \mathbb{N}$. Dokážte, že vzťah $\boldsymbol{A} \sim \boldsymbol{B}$ riadkovej ekvivalencie na množine $K^{m\times n}$ spĺňa podmienky $\boldsymbol{A} \sim \boldsymbol{A}$, $\boldsymbol{A} \sim \boldsymbol{B} \Rightarrow \boldsymbol{B} \sim \boldsymbol{A}$ a $\boldsymbol{A} \sim \boldsymbol{B} \mathbin{\&} \boldsymbol{B} \sim \boldsymbol{C} \Rightarrow \boldsymbol{A} \sim \boldsymbol{C}$ pre ľubovoľné $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in K^{m\times n}$. Inak povedané, tento vzťah je reflexívny, symetrický a tranzitívny, teda je naozaj reláciou ekvivalencie na množine $\boldsymbol{K}^{m\times n}$ (pozri paragraf 0.6). Sformulujte analogický výsledok pre vzťah stĺpcovej ekvivalencie $\boldsymbol{A} \wr \boldsymbol{B}$.

3.12. Dokážte vetu 3.3.2 matematickou indukciou podľa počtu riadkov matice. (*Návod*: Ukážte, že matica s jediným riadkom je riadkovo ekvivalentná s maticou v redukovanom stupňovitom tvare. Predpokladajte, že matica $\boldsymbol{A}$ vznikla z matice v redukovanom stupňovitom tvare pridaním jedného riadku. Ukážte, že aj $\boldsymbol{A}$ je riadkovo ekvivalentná s maticou v redukovanom stupňovitom tvare.)

3.13. Dokážte jednoznačnosť redukovaného stupňovitého tvaru matice. Presnejšie, dokážte, že pre matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{m\times n}$ v redukovanom stupňovitom tvare platí $\boldsymbol{A} \sim \boldsymbol{B} \Rightarrow \boldsymbol{A} = \boldsymbol{B}$. (*Návod*: Ak $\boldsymbol{A}$, $\boldsymbol{B}$ sú v redukovanom stupňovitom tvare a $\boldsymbol{A} \neq \boldsymbol{B}$, ukážte, že homogénne sústavy lineárnych rovníc nemajú rovnaké riešenia; to je však spor s $\boldsymbol{A} \sim \boldsymbol{B}$.)

3.14. Dokážte zosilnenie tvrdenia 3.3.1 do podoby ekvivalencie, t. j. pre ľubovoľné $\boldsymbol{A}$, $\boldsymbol{B} \in K^{m\times n}$, $\boldsymbol{b}, \boldsymbol{c} \in K^m$ platí: sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ a $\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}$ sú ekvivalentné práve vtedy, keď $(\boldsymbol{A} \,|\, \boldsymbol{b}) \sim (\boldsymbol{B} \,|\, \boldsymbol{c})$.

3.15. Dokážte tvrdenie 3.3.6(b).

4. Lineárne podpriestory a lineárna nezávislosť

V tejto kapitole sa opäť vrátime k štúdiu abstraktných vektorových priestorov nad všeobecným poľom. K bude v celej kapitole označovať nejaké pevné, inak ľubovoľné pole a V bude nejaký pevne zvolený vektorový priestor nad K. Čitateľ sa však nedopustí nijakej chyby, ak si pod všeobecným poľom K bude predstavovať pole $\mathbb{R}$ všetkých reálnych čísel. Zakaždým, keď sa budeme odvolávať na geometrický názor, bude to dokonca užitočné. Na druhej strane by však nemal spúšťať zo zreteľa, že naše úvahy majú podstatne širšiu platnosť – okrem vektorových priestorov nad $\mathbb{R}$ sa z nám známych príkladov vzťahujú tak na vektorové priestory nad poľom $\mathbb{C}$ všetkých komplexných čísel, poľom $\mathbb{Q}$ všetkých racionálnych čísel ako i na vektorové priestory nad konečnými poľami $\mathbb{Z}_p$.

4.1 Lineárne podpriestory vektorového priestoru

Množina $S \subseteq V$ sa nazýva *lineárny podpriestor* vektorového priestoru V, ak $S \neq \emptyset$ a pre všetky skaláry $a \in K$ a vektory $\boldsymbol{x}, \boldsymbol{y} \in S$ platí $a\boldsymbol{x} \in S$ a $\boldsymbol{x}+\boldsymbol{y} \in S$. Inak povedané, neprázdna podmnožina $S \subseteq V$ je lineárny podpriestor práve vtedy, keď je uzavretá na operácie skalárneho násobku a súčtu vektorov.

Nasledujúce tvrdenie je bezprostredným dôsledkom práve vyslovenej definície.

4.1.1. Tvrdenie. *Nech S je lineárny podpriestor vektorového priestoru V. Potom $\mathbf{0} \in S$ a S s operáciami súčtu vektorov a skalárneho násobku zúženými z V na S tvorí vektorový priestor nad poľom K.*

V každom vektorovom priestore V sú $\{\mathbf{0}\}$ a V lineárne podpriestory (v prípade, keď $V = \{\mathbf{0}\}$, dokonca splývajú, inak ide o dva rôzne podpriestory) – $\{\mathbf{0}\}$ nazývame *triválny* alebo tiež *nulový* a V *nevlastný* alebo tiež *plný* lineárny podpriestor. Teda pre *vlastný netriviálny* lineárny podpriestor $S \subseteq V$ platí $\{\mathbf{0}\} \neq S \neq V$.

Napr. vo vektorovom priestore $\mathbb{R}^3$ netriviálne vlastné podpriestory sú práve všetky priamky a roviny prechádzajúce počiatkom $\mathbf{0}$.

Nasledujúce tvrdenie charakterizuje lineárne podpriestory ako množiny uzavreté na lineárne kombinácie.

4.1.2. Tvrdenie. *Pre ľubovoľnú podmnožinu S vektorového priestoru V nasledujúce podmienky sú ekvivalentné:*

(i) *S je lineárny podpriestor vo V;*
(ii) *$S \neq \emptyset$ a pre všetky skaláry $a, b \in K$ a vektory $\boldsymbol{x}, \boldsymbol{y} \in S$ platí $a\boldsymbol{x} + b\boldsymbol{y} \in S$;*
(iii) *pre každé $n \in \mathbb{N}$ a pre všetky skaláry $a_1, \ldots, a_n \in K$ a vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in S$ platí*

$$a_1\boldsymbol{x}_1 + \ldots + a_n\boldsymbol{x}_n \in S.$$

Dôkaz. Postupne dokážeme implikácie (i) $\Rightarrow$ (ii), (ii) $\Rightarrow$ (iii) a (iii) $\Rightarrow$ (i).

(i) $\Rightarrow$ (ii): Ak S je lineárny podpriestor, tak $S \neq \emptyset$. Nech $a, b \in K$, $\boldsymbol{x}, \boldsymbol{y} \in S$. Keďže S je uzavreté na skalárne násobky, platí $a\boldsymbol{x}, b\boldsymbol{y} \in S$. Z uzavretosti S na súčet vyplýva $a\boldsymbol{x} + b\boldsymbol{y} \in S$.

(ii) $\Rightarrow$ (iii): Nech platí (ii). Keďže $S \neq \emptyset$, existuje $\boldsymbol{s} \in S$. Potom $\mathbf{0} = 0\boldsymbol{s} + 0\boldsymbol{s} \in S$ a tiež $a\boldsymbol{x} = a\boldsymbol{x} + 0\boldsymbol{s} \in S$ pre každé $a \in K$, $\boldsymbol{x} \in S$. Teda podmienka z (iii) je splnená pre $n = 0$ (lebo prázdna lineárna kombinácia je $\mathbf{0}$) a $n = 1$; podľa (ii) je splnená tiež pre $n = 2$. Keby nebola splnená pre všetky $n \in \mathbb{N}$, označíme n najmenšie prirodzené číslo s touto vlastnosťou. Potom $n > 2$ a pre všetky $k < n$ podmienka z (iii) platí. Nech $a_1, \ldots, a_n \in K$, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in S$ sú také, že $a_1\boldsymbol{x}_1 + \ldots + a_n\boldsymbol{x}_n \notin S$. Avšak

$$a_1\boldsymbol{x}_1 + \ldots + a_n\boldsymbol{x}_n = (a_1\boldsymbol{x}_1 + \ldots + a_{n-1}\boldsymbol{x}_{n-1}) + a_n\boldsymbol{x}_n \in S,$$

keďže pre prirodzené čísla $n - 1$ a 2 podmienka z (iii) platí. To je spor.

(iii) $\Rightarrow$ (i): Z platnosti (iii) pre $n = 0$ vyplýva, že $\mathbf{0} \in S$ (prázdna lineárna kombinácia je totiž $\mathbf{0}$). Teda $S \neq \emptyset$. Voľbou $n = 1$ dostávame uzavretosť S na skalárne násobky. Uzavretosť S na súčet vyplýva z voľby $n = 2$, $a_1 = a_2 = 1$.

4.1.3. Príklad. Keďže s príkladmi lineárnych podpriestorov vektorových priestorov K^n sa ešte stretneme pri mnohých príležitostiach, uvedieme tu niekoľko „exotickejších" príkladov. Napospol pôjde o podpriestory priestorov V^X všetkých funkcií z ľubovoľnej množiny X do nejakého vektorového priestoru V nad poľom K; špeciálne môže ísť o prípad $V = K$ (pozri príklad 1.6.5).

(a) Označme $V^{(X)}$ množinu všetkých funkcií $f\colon X \to V$ takých, že množina $\{x \in X;\ f(x) \neq \mathbf{0}\}$ je konečná. Pre ľubovoľnú lineárnu kombináciu funkcií $f, g \in V^X$ platí

$$\{x \in X;\ af(x) + bg(x) \neq \mathbf{0}\} \subseteq \{x \in X;\ f(x) \neq \mathbf{0}\} \cup \{x \in X;\ g(x) \neq \mathbf{0}\}.$$

Z toho vyplýva, že $V^{(X)}$ je lineárny podpriestor vektorového priestoru V^X. Ak X je konečná, tak $V^{(X)} = V^X$; ak X je nekonečná, tak $V^{(X)}$ je netriválny vlastný podpriestor v V^X.

(b) Nech $X \subseteq \mathbb{R}$ je ľubovoľná množina reálnych čísel. Potom $\mathcal{C}(X, \mathbb{R})$, alebo len stručne $\mathcal{C}(X)$ označuje množinu všetkých spojitých funkcií $f\colon X \to \mathbb{R}$. Keďže lineárne kombinácie spojitých funkcií sú zrejme opäť spojité funkcie,

$\mathcal{C}(X)$ je lineárny podpriestor v $\mathbb{R}^X$.

(c) Ak X je nejaký (ohraničený alebo neohraničený) interval reálnych čísel, tak $\mathcal{D}(X)$ označuje množinu všetkých funkcií $f\colon X \to \mathbb{R}$, ktoré majú v každom bode $x \in X$ konečnú deriváciu (v prípadných krajných bodoch intervalu X sa žiada existencia konečnej derivácie zľava alebo sprava). Keďže každá diferencovateľná funkcia je spojitá na svojom definičnom obore a lineárna kombinácia diferencovateľných funkcií je opäť diferencovateľná, $\mathcal{D}(X)$ je lineárny podpriestor vektorového priestoru $\mathcal{C}(X)$.

4.2 Lineárny obal množiny vektorov

Množinu všetkých lineárnych kombinácií vektorov z podmnožiny X vektorového priestoru V nazývame *lineárnym obalom* množiny X a označujeme ju $[X]$. Teda

$$[X] = \{a_1\boldsymbol{x}_1 + \ldots + a_n\boldsymbol{x}_n;\ n \in \mathbb{N}\ \&\ a_1, \ldots, a_n \in K\ \&\ \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in X\}.$$

Ak $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ je konečná množina, tak miesto $[\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}]$ píšeme len $[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]$. Zrejme tento zápis má zmysel aj pre ľubovoľnú usporiadanú n-ticu (nie nutne rôznych) vektorov $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$, a platí

$$[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n] = \{a_1\boldsymbol{x}_1 + \ldots + a_n\boldsymbol{x}_n;\ a_1, \ldots, a_n \in K\}.$$

4.2.1. Tvrdenie. *Nech X je podmnožina vektorového priestoru V. Potom lineárny obal $[X]$ množiny X je najmenší lineárny podpriestor vektorového priestoru V taký, že $X \subseteq [X]$.*

Dôkaz. Musíme dokázať dve veci:
(a) $[X]$ je lineárny podpriestor vo V;
(b) pre každý lineárny podpriestor $S \subseteq V$ platí $X \subseteq S \Rightarrow [X] \subseteq S$.

(a) Zrejme $[X]$ obsahuje $\mathbf{0}$ ako prázdnu lineárnu kombináciu, teda $[X] \neq \emptyset$. Nech $c, d \in K$ a $\boldsymbol{u} = a_1\boldsymbol{x}_1 + \ldots + a_n\boldsymbol{x}_n$, $\boldsymbol{v} = b_1\boldsymbol{y}_1 + \ldots + b_m\boldsymbol{y}_m$ sú prvky z $[X]$, pričom $a_i, b_j \in K$, $\boldsymbol{x}_i, \boldsymbol{y}_j \in X$. Potom

$$c\boldsymbol{u} + d\boldsymbol{v} = ca_1\boldsymbol{x}_1 + \ldots + ca_n\boldsymbol{x}_n + db_1\boldsymbol{y}_1 + \ldots + db_m\boldsymbol{y}_m \in [X],$$

keďže je to opäť lineárna kombinácia vektorov z X. Podľa podmienky (ii) tvrdenia 4.1.2 je $[X]$ lineárny podpriestor vo V.

(b) Nech $S \subseteq V$ je lineárny podpriestor taký, že $X \subseteq S$. Potom podľa podmienky (iii) tvrdenia 4.1.2 všetky lineárne kombinácie vektorov z S, a tým skôr vektorov z X, patria do S. Teda $[X] \subseteq S$.

Dokázané tvrdenie nás oprávňuje nazývať lineárny obal $[X]$ množiny $X \subseteq V$ tiež lineárnym podpriestorom *generovaným* množinou X. Ak $[X] = S$, hovoríme, že X *generuje* lineárny podpriestor S, prípadne že X je *generujúca množina* alebo tiež *množina generátorov* lineárneho podpriestoru $S \subseteq V$. Ak $S = V$, t. j. ak $[X] = V$, hovoríme krátko o *generujúcej množine*. Používa sa tiež názov *vytvárajúca množina.*

Kvôli prehľadnosti ešte zhrnieme základné vlastnosti operácie lineárneho obalu $X \mapsto [X]$.

4.2.2. Tvrdenie. *Pre ľubovoľné podmnožiny X, Y vektorového priestoru V a $\boldsymbol{v} \in V$ platí:*
(a) $[\emptyset] = [\mathbf{0}] = \{\mathbf{0}\}$*;*
(b) $X \subseteq [X]$*;*
(c) $X \subseteq Y \Rightarrow [X] \subseteq [Y]$*;*
(d) X je lineárny podpriestor vo V práve vtedy, keď $X = [X]$;
(e) $[[X]] = [X]$*;*
(f) $\boldsymbol{v} \in [X] \Leftrightarrow [X \cup \{\boldsymbol{v}\}] = [X]$.

Dôkaz. (a), (b) a (c) sú triviálne, (d) priamo vyplýva z tvrdenia 4.2.1 a (e) je bezprostredným dôsledkom (d).

(f) Nech $\boldsymbol{v} \in [X]$. S použitím (b), (c) a (e) dostávame

$$[X \cup \{\boldsymbol{v}\}] \subseteq [[X] \cup \{\boldsymbol{v}\}] = [[X]] = [X].$$

Teda $[X \cup \{\boldsymbol{v}\}] = [X]$. Keďže $\boldsymbol{v} \in [X \cup \{\boldsymbol{v}\}]$, obrátená implikácia je triviálna.

4.3 Prienik a súčet lineárnych podpriestorov

Nech X, Y sú ľubovoľné podmnožiny vektorového priestoru V. Potom množinu

$$X + Y = \{\boldsymbol{x} + \boldsymbol{y};\, \boldsymbol{x} \in X \;\&\; \boldsymbol{y} \in Y\}$$

nazývame *súčtom* množín X, Y.

4.3.1. Tvrdenie. *Nech S, T sú lineárne podpriestory vektorového priestoru V. Potom aj $S \cap T$ a $S + T$ sú lineárne podpriestory vo V. Navyše platí*

$$S + T = [S \cup T],$$

t. j. $S + T$ je najmenší lineárny podpriestor vo V, ktorý obsahuje S aj T.

Dôkaz. Zrejme $\mathbf{0} \in S \cap T$. Z toho, že S aj T sú uzavreté na lineárne kombinácie, vyplýva, že aj $S \cap T$ má túto vlastnosť.

Dokážeme, že aj $S+T$ je uzavreté na lineárne kombinácie. Nech $a_1, a_2 \in K$ a $\boldsymbol{u}_1 = \boldsymbol{x}_1 + \boldsymbol{y}_1$, $\boldsymbol{u}_2 = \boldsymbol{x}_2 + \boldsymbol{y}_2$ sú vektory z $S+T$, pričom $\boldsymbol{x}_1, \boldsymbol{x}_2 \in S$, $\boldsymbol{y}_1, \boldsymbol{y}_2 \in T$. Potom

$$\begin{aligned} a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2 &= a_1(\boldsymbol{x}_1 + \boldsymbol{y}_1) + a_2(\boldsymbol{x}_2 + \boldsymbol{y}_2) \\ &= (a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2) + (a_1\boldsymbol{y}_1 + a_2\boldsymbol{y}_2) \in S + T, \end{aligned}$$

lebo S, T sú lineárne podpriestory, teda $a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 \in S$ a $a_1\boldsymbol{y}_1 + a_2\boldsymbol{y}_2 \in T$.

Dokážeme poslednú rovnosť. Inklúzie $S \cup T \subseteq S + T \subseteq [S \cup T]$ sú zrejmé. Keďže $S+T$ je lineárny podpriestor vo V a $[S \cup T]$ je najmenší lineárny podpriestor vo V, ktorý obsahuje množinu $S \cup T$, platí tiež $[S \cup T] \subseteq S + T$.

Na druhej strane čitateľ iste ľahko nájde príklady na to, že zjednotenie dvoch lineárnych podpriestorov S, T vektorového priestoru V nemusí byť lineárnym podpriestorom. Presnejšie, $S \cup T$ je lineárny podpriestor vo V práve vtedy, keď $S \subseteq T$ alebo $T \subseteq S$. Porozmýšľajte prečo.

Každý prvok $\boldsymbol{z} \in S + T$ súčtu lineárnych podpriestorov $S, T \subseteq V$ možno vyjadriť v tvare $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$ pre nejaké $\boldsymbol{x} \in S$, $\boldsymbol{y} \in T$. Vo všeobecnosti to však možno urobiť viacerými spôsobmi. Súčet lineárnych podpriestorov S, T vektorového priestoru V nazývame *priamym* alebo tiež *direktným súčtom*, ak každé $\boldsymbol{z} \in S + T$ možno *jednoznačne* vyjadriť v tvare $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$, kde $\boldsymbol{x} \in S$, $\boldsymbol{y} \in T$; takýto súčet zvykneme tiež označovať $S \oplus T$.

4.3.2. Tvrdenie. *Nech S, T sú lineárne podpriestory vektorového priestoru V. Potom nasledujúce podmienky sú ekvivalentné:*
(i) $S + T = S \oplus T$, *t. j. súčet $S + T$ je direktný;*
(ii) $S \cap T = \{\mathbf{0}\}$.

Dôkaz. (i) $\Rightarrow$ (ii): Nech $\boldsymbol{z} \in S \cap T$. Potom $\boldsymbol{z}$ možno vyjadriť v tvare $\boldsymbol{z} = \boldsymbol{z} + \mathbf{0}$, kde $\boldsymbol{z} \in S$, $\mathbf{0} \in T$, ako aj v tvare $\boldsymbol{z} = \mathbf{0} + \boldsymbol{z}$, kde $\mathbf{0} \in S$, $\boldsymbol{z} \in T$. Z predpokladanej jednoznačnosti vyplýva $\boldsymbol{z} = \mathbf{0}$. Teda $S \cap T = \{\mathbf{0}\}$.

(ii) $\Rightarrow$ (i): Nech $S \cap T = \{\mathbf{0}\}$. Predpokladajme, že vektor $\boldsymbol{z} \in S + T$ možno vyjadriť v tvaroch $\boldsymbol{z} = \boldsymbol{x}_1 + \boldsymbol{y}_1 = \boldsymbol{x}_2 + \boldsymbol{y}_2$, kde $\boldsymbol{x}_1, \boldsymbol{x}_2 \in S$, $\boldsymbol{y}_1, \boldsymbol{y}_2 \in T$. Potom $\boldsymbol{x}_1 - \boldsymbol{x}_2 = \boldsymbol{y}_2 - \boldsymbol{y}_1$. Keďže $\boldsymbol{x}_1 - \boldsymbol{x}_2 \in S$, $\boldsymbol{y}_2 - \boldsymbol{y}_1 \in T$, uvedená spoločná hodnota patrí do $S \cap T$. Preto $\boldsymbol{x}_1 - \boldsymbol{x}_2 = \boldsymbol{y}_2 - \boldsymbol{y}_1 = \mathbf{0}$, t. j. $\boldsymbol{x}_1 = \boldsymbol{x}_2$, $\boldsymbol{y}_1 = \boldsymbol{y}_2$. To dokazuje požadovanú jednoznačnosť.

Uvedenú definíciu možno zrejmým spôsobom zovšeobecniť na priamy súčet ľubovoľného konečného počtu lineárnych podpriestorov. Zodpovedajúce zovšeobecnenie podmienky (ii) z práve dokázaného tvrdenia však už celkom priamočiare nie je. Podrobnosti nájde čitateľ v cvičení 4.8.

4.4 Lineárna nezávislosť

Nech $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n \in V$. Hovoríme, že usporiadaná n-tica vektorov $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ je *lineárne závislá*, ak existujú skaláry $c_1, \dots, c_n \in K$ také, že $(c_1, \dots, c_n) \neq \mathbf{0}$

a $c_1\boldsymbol{u}_1 + \ldots + c_n\boldsymbol{u}_n = \mathbf{0}$. V opačnom prípade hovoríme, že usporiadaná n-tica vektorov $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ je *lineárne nezávislá.* Pre $n = 0$ kvôli úplnosti dodávame, že usporiadanú 0-ticu (t. j. prázdnu postupnosť) vektorov považujeme za lineárne nezávislú.

Miesto „lineárne (ne)závislá usporiadaná n-tica vektorov $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$“ budeme často hovoriť len o lineárne (ne)závislých vektoroch $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$.

Rozmeňme si teraz „na drobné“, čo znamená ono „v opačnom prípade“ v definícii lineárnej nezávislosti. Podľa tejto definície vektory $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ sú lineárne nezávislé práve vtedy, keď

$$(\forall\, c_1, \ldots, c_n \in K)(c_1\boldsymbol{u}_1 + \ldots + c_n\boldsymbol{u}_n = \mathbf{0} \Rightarrow c_1 = \ldots = c_n = 0).$$

Vidíme, že logická štruktúra pojmu lineárnej nezávislosti je trochu zložitejšia, než sme boli doteraz zvyknutí. Keďže ide o kľúčový pojem, je potrebné sa pri ňom na chvíľu pristaviť.

Uvedomme si, že pre n-ticu skalárov $(c_1, \ldots, c_n) = \mathbf{0}$ platí $c_1\boldsymbol{u}_1 + \ldots + c_n\boldsymbol{u}_n = \mathbf{0}$ *pre ľubovoľnú* n-ticu vektorov $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$, bez ohľadu na to, či je lineárne závislá alebo nezávislá. Avšak pre niektoré n-tice vektorov $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ môžeme ako výsledok lineárnej kombinácie $c_1\boldsymbol{u}_1 + \ldots + c_n\boldsymbol{u}_n$ dostať $\mathbf{0}$ aj pomocou *inej* n-tice skalárov $(c_1, \ldots, c_n)$ než len $\mathbf{0} = (0, \ldots, 0)$ – takéto usporiadané n-tice $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ nazývame *lineárne závislé.* Pre niektoré usporiadané n-tice vektorov $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ je voľba $(c_1, \ldots, c_n) = \mathbf{0}$ *jediná možnosť* ako lineárnou kombináciou $c_1\boldsymbol{u}_1 + \ldots + c_n\boldsymbol{u}_n$ získať výsledok $\mathbf{0}$ – takéto usporiadané n-tice nazývame *lineárne nezávislé.*

Na precvičenie práve definovaných pojmov čitateľovi odporúčame, aby si dokázal štyri jednoduché no užitočné pozorovania:

(a) jediný vektor $\boldsymbol{u}$ je lineárne nezávislý práve vtedy, keď $\boldsymbol{u} \neq \mathbf{0}$;

(b) vektory $\boldsymbol{u}, \boldsymbol{v}$ sú lineárne závislé práve vtedy, keď jeden z nich je násobkom druhého;

(c) ak niektorý z vektorov $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ je $\mathbf{0}$, tak tieto vektory sú lineárne závislé;

(d) ak sa niektoré dva z vektorov $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ rovnajú alebo niektorý z nich je násobkom iného, tak tieto vektory sú lineárne závislé.

Inak povedané, len usporiadaná n-tica nenulových a navzájom rôznych vektorov, z ktorých žiaden nie je násobkom druhého, môže (no stále ešte nemusí) byť lineárne nezávislá.

Nasledujúce tvrdenie asi vysvetľuje názov „lineárna závislosť“ lepšie než samotná definícia.

4.4.1. Tvrdenie. *Pre ľubovoľné $n \in \mathbb{N}$ a $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in V$ nasledujúce podmienky sú ekvivalentné:*

(i) vektory $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ sú lineárne závislé;

(ii) *niektorý z vektorov $\boldsymbol{u}_k$, $k \le n$, je lineárnou kombináciou predchádzajúcich;*

(ii') *niektorý z vektorov $\boldsymbol{u}_k$, $k \le n$, je lineárnou kombináciou nasledujúcich;*

(iii) *niektorý z vektorov $\boldsymbol{u}_k$, $k \le n$, je lineárnou kombináciou ostatných.*

Dôkaz. Dokážeme implikácie (i) ⇒ (ii) ⇒ (iii) a (iii) ⇒ (i). Rovnako by bolo možné dokázať aj implikácie (i) ⇒ (ii') ⇒ (iii).

(i) ⇒ (ii): Nech $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ sú lineárne závislé vektory a $c_1, \dots, c_n$ sú skaláry, nie všetky rovné 0, také, že $c_1\boldsymbol{u}_1 + \dots + c_n\boldsymbol{u}_n = \mathbf{0}$. Nech k je najväčší z indexov $1, \dots, n$ taký, že $c_k \neq 0$. Potom $c_i = 0$ pre $k < i \le n$, teda $c_1\boldsymbol{u}_1 + \dots + c_k\boldsymbol{u}_k = \sum_{i=1}^n c_i\boldsymbol{u}_i = \mathbf{0}$. Z toho dostávame

$$\boldsymbol{u}_k = c_k^{-1}(c_1\boldsymbol{u}_1 + \dots + c_{k-1}\boldsymbol{u}_{k-1}),$$

t. j. $\boldsymbol{u}_k$ je lineárnou kombináciou predchádzajúcich vektorov.

(ii) ⇒ (iii) platí triválne.

(iii) ⇒ (i): Ak $\boldsymbol{u}_k = \sum_{\substack{i=1\\ i\neq k}}^n c_i\boldsymbol{u}_i$ je lineárnou kombináciou ostatných vektorov, položme $c_k = -1$. Potom pre n-ticu skalárov $(c_1, \dots, c_n) \neq \mathbf{0}$ platí $c_1\boldsymbol{u}_1 + \dots + c_n\boldsymbol{u}_n = \mathbf{0}$, teda vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ sú lineárne závislé.

Poznámka. Všimnite si, že dôkaz implikácie (i) ⇒ (ii) pokrýva aj prípad $k = 1$. Vtedy $c_1 \neq 0$ a $c_1\boldsymbol{u}_1 = \mathbf{0}$, preto tiež $\boldsymbol{u}_1 = \mathbf{0}$. Teda $\boldsymbol{u}_1$ je naozaj lineárnou kombináciou predchádzajúcich (t. j. prázdnej postupnosti) vektorov.

Každý vektor $\boldsymbol{x}$ z lineárneho obalu $[\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]$ možno vyjadriť v tvare

$$\boldsymbol{x} = c_1\boldsymbol{u}_1 + \dots + c_n\boldsymbol{u}_n$$

pre nejakú n-ticu skalárov $(c_1, \dots, c_n)$. Nasledujúce tvrdenie ukazuje, že lineárna nezávislosť vektorov $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ je ekvivalentná s jednoznačnosťou tohto vyjadrenia.

4.4.2. Veta. *Vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ sú lineárne nezávislé práve vtedy, keď každý vektor $\boldsymbol{x} \in [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]$ možno vyjadriť v tvare $\boldsymbol{x} = c_1\boldsymbol{u}_1 + \dots + c_n\boldsymbol{u}_n$ pre jedinú usporiadanú n-ticu $(c_1, \dots, c_n) \in K^n$.*

Dôkaz. Nech $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ sú lineárne nezávislé vektory. Predpokladajme, že vektor $\boldsymbol{x} \in [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]$ možno vyjadriť v tvaroch

$$\boldsymbol{x} = c_1\boldsymbol{u}_1 + \dots + c_n\boldsymbol{u}_n = d_1\boldsymbol{u}_1 + \dots + d_n\boldsymbol{u}_n,$$

kde $(c_1, \dots, c_n), (d_1, \dots, d_n) \in K^n$. Potom

$$(c_1 - d_1)\boldsymbol{u}_1 + \dots + (c_n - d_n)\boldsymbol{u}_n = \mathbf{0}.$$

Z lineárnej nezávislosti vektorov $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ vyplýva $c_1 - d_1 = \ldots = c_n - d_n = 0$, čiže $(c_1, \dots, c_n) = (d_1, \dots, d_n)$. Teda vyjadrenie vektora $\boldsymbol{x}$ v tvare lineárnej kombinácie vektorov $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ je jednoznačné.

Predpokladajme teraz, že každý vektor $\boldsymbol{x} \in [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]$ má jednoznačné vyjadrenie v tvare lineárnej kombinácie vektorov $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$. Špeciálne to platí aj pre vektor $\boldsymbol{x} = \mathbf{0}$, ktorý má vyjadrenie $\mathbf{0} = 0\boldsymbol{u}_1 + \ldots + 0\boldsymbol{u}_n$. Z jednoznačnosti tohto vyjadrenia vyplýva

$$c_1\boldsymbol{u}_1 + \ldots + c_n\boldsymbol{u}_n = \mathbf{0} \Rightarrow c_1 = \ldots = c_n = 0$$

pre ľubovoľnú n-ticu skalárov $(c_1, \dots, c_n)$. Teda vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ sú lineárne nezávislé.

Nasledujúce tvrdenie dáva do súvislosti lineárnu (ne)závislosť s lineárnym obalom.

4.4.3. Tvrdenie. *Nech $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v} \in V$ pričom vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ sú lineárne nezávislé. Potom nasledujúce podmienky sú ekvivalentné:*
(i) $\boldsymbol{v} \in [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]$*;*
(ii) vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}$ sú lineárne závislé;
(iii) $[\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}] = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]$.

Dôkaz. (i) ⇒ (ii): Ak $\boldsymbol{v} \in [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]$ tak vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}$ sú lineárne závislé podľa tvrdenia 4.4.1

(ii) ⇒ (iii): Nech vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}$ sú lineárne závislé. Potom niektorý z nich je lineárnou kombináciou predchádzajúcich. Keďže vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ sú lineárne nezávislé, môže to byť len vektor $\boldsymbol{v}$. Teda $\boldsymbol{v} \in [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]$.

(iii) ⇒ (i) je obsiahnuté v bode (f) tvrdenia 4.2.2

4.4.4. Veta. *Nech $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}_1, \dots, \boldsymbol{v}_m \in V$, pričom vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ sú lineárne nezávislé. Potom z množiny $\{1, \dots, m\}$ možno vybrať indexy $i_1 < \ldots < i_k$ tak, že vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}_{i_1}, \dots, \boldsymbol{v}_{i_k}$ sú lineárne nezávislé a generujú rovnaký podpriestor ako vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}_1, \dots, \boldsymbol{v}_m$.*

Dôkaz. Označme $X = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$. Vektory $\boldsymbol{v}_{i_1}, \dots, \boldsymbol{v}_{i_k}$ vyberieme z množiny X nasledujúcim spôsobom. Ak $X \subseteq [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]$, položme $k = 0$, t. j. nevyberieme žiaden z nich. V opačnom prípade nech $\boldsymbol{v}_{i_1}$ je prvý z vektorov množiny X, ktorý neleží v podpriestore $[\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]$. Ak $X \subseteq [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}_{i_1}]$, tak $k = 1$ a $\boldsymbol{v}_{i_1}$ je jediný vybraný vektor. Podľa predchádzajúceho tvrdenia sú vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}_{i_1}$ lineárne nezávislé. Ak $X \nsubseteq [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}_{i_1}]$, označíme $\boldsymbol{v}_{i_2}$ prvý vektor množiny X, ktorý neleží v $[\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}_{i_1}]$ (zrejme $i_1 < i_2$ a $\boldsymbol{v}_{i_1} \neq \boldsymbol{v}_{i_2}$). Vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}_{i_1}, \boldsymbol{v}_{i_2}$ sú podľa tvrdenia 4.4.3 opäť lineárne nezávislé. Podľa potreby pokračujeme rovnakým spôsobom, až kým pre takto získané lineárne nezávislé vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}_{i_1}, \dots, \boldsymbol{v}_{i_k}$ neplatí inklúzia $X \subseteq [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{v}_{i_1}, \dots, \boldsymbol{v}_{i_k}]$, kedy sa zastavíme. (V krajnom prípade dostaneme $k = m$, t. j. vyberieme všetky vektory z množiny X.) Z uvedenej inklúzie okamžite vyplýva rovnosť

$$[\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m] = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n, \boldsymbol{v}_{i_1}, \ldots, \boldsymbol{v}_{i_k}],$$

keďže každý z generátorov podpriestoru na ľavej strane je prvkom podpriestoru na pravej strane.

4.5 Lineárny obal a lineárna nezávislosť v priestoroch K^m

V tomto paragrafe si ukážeme, ako možno na základe našich doterajších znalostí o sústavách lineárnych rovníc tou istou metódou úpravy matíc pomocou ERO na (redukovaný) stupňovitý tvar riešiť pre vektory z priestoru K^m nasledujúce tri otázky:

(1) rozhodnúť pre dané vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{y} \in K^m$ či $\boldsymbol{y}$ patrí alebo nepatrí do lineárneho obalu $[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]$;

(2) rozhodnúť pre dané vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in K^m$ či sú lineárne závislé alebo nezávislé;

(3) vybrať z vektorov $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in K^m$ lineárne nezávislé vektory $\boldsymbol{x}_{j_1}, \ldots, \boldsymbol{x}_{j_k}$ ($j_1 < \ldots < j_k$) tak, aby vektory $\boldsymbol{x}_{j_1}, \ldots, \boldsymbol{x}_{j_k}$ generovali vo V ten istý lineárny podpriestor ako vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$.

Hoci všetky tri otázky možno riešiť naraz jednotným spôsobom, z metodických dôvodov začneme jednoduchšími otázkami (1) a (2), a až potom pristúpime k trochu zložitejšej otázke (3). Navyše pri tom zavedieme označenie, ktorého sa budeme držať v celom paragrafe.

Nech $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{y} \in K^m$ sú stĺpcové vektory, pričom

$$\boldsymbol{x}_j = \begin{pmatrix} x_{1j} \\ \vdots \\ x_{mj} \end{pmatrix}, \qquad \boldsymbol{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Označme $\boldsymbol{X} = (x_{ij}) \in K^{m\times n}$ maticu so stĺpcami $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, a $(\boldsymbol{X} \,|\, \boldsymbol{y}) \in K^{m\times(n+1)}$ blokovú maticu zloženú z matice $\boldsymbol{X}$ a vektora $\boldsymbol{y}$. Potom pre $\boldsymbol{c} = (c_1, \ldots, c_n)^\mathsf{T} \in K^n$ platí:

(1) $c_1\boldsymbol{x}_1 + \ldots + c_n\boldsymbol{x}_n = \boldsymbol{y} \;\Leftrightarrow\; \boldsymbol{X} \cdot \boldsymbol{c} = \boldsymbol{y}$;

(2) $c_1\boldsymbol{x}_1 + \ldots + c_n\boldsymbol{x}_n = \boldsymbol{0} \;\Leftrightarrow\; \boldsymbol{X} \cdot \boldsymbol{c} = \boldsymbol{0}$.

Inak povedané: (1) $\boldsymbol{y} \in [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]$ práve vtedy, keď sústava $\boldsymbol{X} \cdot \boldsymbol{c} = \boldsymbol{y}$ s rozšírenou maticou $(\boldsymbol{X} \,|\, \boldsymbol{y})$ má aspoň jedno riešenie; (2) vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ sú lineárne nezávislé práve vtedy, keď homogénna sústava $\boldsymbol{X} \cdot \boldsymbol{c} = \boldsymbol{0}$ má jediné riešenie $\boldsymbol{c} = \boldsymbol{0}$; ak táto sústava má aj nejaké nenulové riešenie, tak vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ sú lineárne závislé. (Nedajte sa popliesť atypickým označením: x_{ij} sú teraz koeficienty sústavy, y_i sú zložky pravej strany a c_j sú neznáme.)

Otázku (1) už vieme riešiť. Stačí pomocou ERO upraviť maticu $(\boldsymbol{X}\,|\,\boldsymbol{y})$ na stupňovitý tvar. Ak výsledná matica obsahuje riadok tvaru $(0,\dots,0\,|\,z)$, kde $z\neq 0$, tak sústava $\boldsymbol{X}\cdot\boldsymbol{c}=\boldsymbol{y}$ nemá riešenie a $\boldsymbol{y}\notin[\boldsymbol{x}_1,\dots,\boldsymbol{x}_n]$. Ak sa taký riadok vo výslednej matici nenachádza, tak sústava má aspoň jedno riešenie a $\boldsymbol{y}\in[\boldsymbol{x}_1,\dots,\boldsymbol{x}_n]$.

Podobne je to s otázkou (2). Opäť stačí pomocou ERO upraviť maticu $\boldsymbol{X}$ na stupňovitý tvar a pozrieť sa, či v každom stĺpci leží vedúci prvok nejakého riadku. Ak je to tak, niet čo voliť za parametre, $\boldsymbol{c}=\boldsymbol{0}$ je jediným riešením sústavy $\boldsymbol{X}\cdot\boldsymbol{c}=\boldsymbol{0}$ a vektory $\boldsymbol{x}_1,\dots,\boldsymbol{x}_n$ sú lineárne nezávislé. V opačnom prípade máme možnosť voľby aspoň jedného parametra, sústava má aj nejaké nenulové riešenie a vektory $\boldsymbol{x}_1,\dots,\boldsymbol{x}_n$ sú lineárne závislé.

Ešte si všimnime úzku súvislosť oboch otázok. Vedúcim prvkom riadku $(0,\dots,0\,|\,z)$, kde $z\neq 0$, je práve v $(n+1)$-om stĺpci ležiaci prvok z. Teda matica v stupňovitom tvare riadkovo ekvivalentná s $(\boldsymbol{X}\,|\,\boldsymbol{y})$ neobsahuje taký riadok práve vtedy, keď v jej poslednom stĺpci neleží vedúci prvok žiadneho riadku.

4.5.1. Príklad. Uvažujme stĺpcové vektory $\boldsymbol{x}_1=(1,1,-1,-1)^{\mathsf T}$, $\boldsymbol{x}_2=(0,1,0,1)^{\mathsf T}$, $\boldsymbol{x}_3=(3,1,-3,-5)^{\mathsf T}$, $\boldsymbol{x}_4=(0,0,1,2)^{\mathsf T}$, $\boldsymbol{y}=(3,5,-2,1)^{\mathsf T}$, $\boldsymbol{z}=(1,1,1,1)^{\mathsf T}$ v priestore $\mathbb{R}^4$. Máme rozhodnúť, či vektory $\boldsymbol{y}$, $\boldsymbol{z}$ patria do lineárneho obalu $[\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3,\boldsymbol{x}_4]$. Označme si nasledujúce matice

$$(\boldsymbol{X}\,|\,\boldsymbol{y})=\left(\begin{array}{rrrr|r}1&0&3&0&3\\1&1&1&0&5\\-1&0&-3&1&-2\\-1&1&-5&2&1\end{array}\right),\qquad(\boldsymbol{X}\,|\,\boldsymbol{z})=\left(\begin{array}{rrrr|r}1&0&3&0&1\\1&1&1&0&1\\-1&0&-3&1&1\\-1&1&-5&2&1\end{array}\right).$$

Matice $(\boldsymbol{X}\,|\,\boldsymbol{y})$, $(\boldsymbol{X}\,|\,\boldsymbol{z})$ sú riadkovo ekvivalentné s maticami

$$\left(\begin{array}{rrrr|r}1&0&3&0&3\\0&1&-2&0&2\\0&0&0&1&1\\0&0&0&0&0\end{array}\right)\quad\text{resp.}\quad\left(\begin{array}{rrrr|r}1&0&3&0&1\\0&1&-2&0&0\\0&0&0&1&2\\0&0&0&0&-2\end{array}\right).$$

Okamžite vidíme, že platí $\boldsymbol{y}\in[\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3,\boldsymbol{x}_4]$ a $\boldsymbol{z}\notin[\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3,\boldsymbol{x}_4]$.

4.5.2. Príklad. Zistíme, či stĺpce reálnej matice

$$\boldsymbol{X}=\begin{pmatrix}2&0&1&3\\2&1&2&3\\0&2&3&2\\1&2&4&2\end{pmatrix}$$

sú lineárne závislé alebo nezávislé. Táto matica je riadkovo ekvivalentná s maticou

$$\begin{pmatrix} 1 & 2 & 4 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

Vidíme, že stĺpce matice $\boldsymbol{X}$ sú lineárne nezávislé. Na druhej strane, ako matica nad poľom $\mathbb{Z}_5$ je $\boldsymbol{X}$ riadkovo ekvivalentná s maticou

$$\begin{pmatrix} 1 & 2 & 4 & 2 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Teda stĺpce matice $\boldsymbol{X}$, chápané ako vektory z vektorového priestoru $\mathbb{Z}_5^4$, sú lineárne závislé.

Kľúčom k odpovedi na otázku (3) je nasledujúce tvrdenie.

4.5.3. Tvrdenie. *Nech $\boldsymbol{X}, \boldsymbol{Y} \in K^{m\times n}$ sú riadkovo ekvivalentné matice, pričom matica $\boldsymbol{Y}$ je v stupňovitom tvare. Pre $1 \le j \le n$ označme $\boldsymbol{x}_j = \boldsymbol{s}_j(\boldsymbol{X})$ j-ty stĺpec matice $\boldsymbol{X}$. Nech $j_1 < \ldots < j_k$ sú indexy všetkých stĺpcov matice $\boldsymbol{Y}$, v ktorých ležia vedúce prvky jej riadkov. Potom platí:*

(a) *vektory $\boldsymbol{x}_{j_1}, \ldots, \boldsymbol{x}_{j_k}$ sú lineárne nezávislé;*

(b) *ak v j-tom stĺpci matice $\boldsymbol{Y}$ neleží vedúci prvok žiadneho jej riadku (t. j. $1 \le j \le n$ a $j \ne j_1, \ldots, j_k$), tak vektor $\boldsymbol{x}_j$ je lineárnou kombináciou vektorov $\boldsymbol{x}_{j_1}, \ldots, \boldsymbol{x}_{j_l}$, kde $l \le k$ je najväčší index, pre ktorý platí $j_l < j$;*

(c) *$[\boldsymbol{x}_{j_1}, \ldots, \boldsymbol{x}_{j_k}] = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]$.*

Dôkaz. (a) Označme $\boldsymbol{X}', \boldsymbol{Y}'$ matice, ktoré pozostávajú len zo stĺpcov s indexmi $j_1, \ldots, j_k$ matíc $\boldsymbol{X}$ resp. $\boldsymbol{Y}$ (ostatné stĺpce vynecháme). Potom postupnosťou tých istých ERO, ktorými sme $\boldsymbol{X}$ upravili na $\boldsymbol{Y}$, dostaneme z $\boldsymbol{X}'$ maticu $\boldsymbol{Y}'$, teda $\boldsymbol{X}' \sim \boldsymbol{Y}'$. Matica $\boldsymbol{Y}'$ je však v stupňovitom tvare a má v každom stĺpci vedúci prvok nejakého svojho riadku. Preto homogénna sústava $\boldsymbol{X}' \cdot \boldsymbol{d} = \boldsymbol{0}$ má jediné riešenie $\boldsymbol{d} = \boldsymbol{0} \in K^k$, čo znamená, že stĺpce matice $\boldsymbol{X}'$, t. j. vektory $\boldsymbol{x}_{j_1}, \ldots, \boldsymbol{x}_{j_k}$, sú lineárne nezávislé.

(b) Bez ujmy na všeobecnosti môžeme predpokladať, že matica $\boldsymbol{Y}$ je dokonca v redukovanom stupňovitom tvare. Poloha vedúcich prvkov riadkov v jednotlivých stĺpcoch bude stále rovnaká. Nech $j \ne j_1, \ldots, j_k$. Pri voľbe parametra $c_j = 1$ a voľbou 0 za hodnotu všetkých ostatných parametrov (ak nejaké zostali) dostaneme jedno riešenie $\boldsymbol{c} = (c_1, \ldots, c_n)^\mathsf{T} \ne \boldsymbol{0}$ sústavy $\boldsymbol{X} \cdot \boldsymbol{c} = \boldsymbol{0}$. Nech $l \le k$ je najväčší index taký, že $j_l < j$. Pre naše riešenie $\boldsymbol{c}$ navyše platí $c_p = 0$, ak $j \le p \le n$. Ak je totiž c_p parameter, tak je to dôsledok našej voľby, a vo vyjadrení neznámych c_{j_h} pre $l < h \leq k$ sa (jediný nenulový) parameter c_j nevyskytuje. Označme $\boldsymbol{X}''$ maticu, ktorá pozostáva len zo stĺpcov matice $\boldsymbol{X}$ s indexmi $j_1, \dots, j_l$ a j. Z uvedených dôvodov je vektor $\boldsymbol{c}'' = (c_{j_1}, \dots, c_{j_l}, 1)^\mathsf{T}$ riešením sústavy $\boldsymbol{X}'' \cdot \boldsymbol{c}'' = \boldsymbol{0}$. To znamená, že

$$\boldsymbol{x}_j = -(c_{j_1}\boldsymbol{x}_{j_1} + \dots + c_{j_l}\boldsymbol{x}_{j_l}) \in [\boldsymbol{x}_{j_1}, \dots, \boldsymbol{x}_{j_l}].$$

(c) je bezprostredným dôsledkom (b) a tvrdenia 4.4.3

Práve dokázané tvrdenie nám dáva priamy návod na riešenie otázky (3). Stačí pomocou ERO upraviť maticu $\boldsymbol{X} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ na maticu $\boldsymbol{Y}$ v stupňovitom tvare a zistiť v nej indexy $j_1 < \dots < j_k$ všetkých stĺpcov, v ktorých ležia vedúce prvky jej riadkov. Potom $\boldsymbol{x}_{j_1}, \dots, \boldsymbol{x}_{j_k}$ sú hľadané lineárne nezávislé vektory, ktoré generujú lineárny podpriestor $[\boldsymbol{x}_1, \dots, \boldsymbol{x}_n]$.

4.5.4. Príklad. Zo stĺpcov reálnej matice

$$\boldsymbol{X} = \begin{pmatrix} 1 & 1 & 3 & -1 & 1 \\ 2 & 0 & 2 & 1 & 3 \\ 1 & 1 & 3 & 2 & 4 \\ 2 & 0 & 2 & 0 & 2 \end{pmatrix}$$

treba vybrať lineárne nezávislé stĺpce, ktoré generujú lineárny obal všetkých stĺpcov matice $\boldsymbol{X}$. Matica $\boldsymbol{X}$ je riadkovo ekvivalentná s maticou

$$\boldsymbol{Y} = \begin{pmatrix} 1 & 1 & 3 & -1 & 1 \\ 0 & 1 & 2 & 2 & -3 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

v stupňovitom tvare. Vedúce prvky riadkov matice $\boldsymbol{Y}$ sa nachádzajú v stĺpcoch 1, 2 a 4. Hľadané vektory sú teda stĺpce 1, 2 a 4 matice $\boldsymbol{X}$. Zapísané vedľa seba tvoria maticu

$$\begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \\ 1 & 1 & 2 \\ 2 & 0 & 0 \end{pmatrix}.$$

I keď sme celý postup riešenia otázok (1), (2) a (3) vyložili len pre priestory stĺpcových vektorov K^m a týchto priestorov sa týkali aj všetky príklady, čitateľovi by už nemalo robiť ťažkosti modifikovať popísanú metódu aj na priestory riadkových vektorov K^m – či už transponovaním, príslušných matíc riadkových vektorov alebo nahradením elementárnych riadkových operácií stĺpcovými.

4.6 Lineárne nezávislé postupnosti a množiny

V tomto paragrafe stručne doplníme pojmy lineárnej závislosti a nezávislosti spôsobom, ktorý umožňuje ich použitie i v prípade nekonečných postupností a ľubovoľných (t. j. konečných aj nekonečných) množín vektorov. Nakoľko však tieto otázky zostávajú na okraji nášho záujmu, popri príslušných definíciách sa obmedzíme len na niekoľko jednoduchých zovšeobecnení výsledkov o lineárnej (ne)závislosti usporiadaných n-tíc.

Nekonečnú postupnosť $(\boldsymbol{u}_k)_{k=0}^{\infty} = (\boldsymbol{u}_0, \boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k, \dots)$ vektorov z priestoru V nazývame *lineárne nezávislou*, ak každá jej konečná podpostupnosť $(\boldsymbol{u}_{k_1}, \dots, \boldsymbol{u}_{k_n})$, kde $0 \leq k_1 < \ldots < k_n$, je lineárne nezávislá.

Dôkaz nasledujúceho jednoduchého tvrdenia prenechávame čitateľovi.

4.6.1. Tvrdenie. *Nekonečná postupnosť* $(\boldsymbol{u}_k)_{k=0}^{\infty}$ *vektorov z* V *je lineárne nezávislá práve vtedy, keď pre každé* $n \in \mathbb{N}$ *jej počiatočný úsek* $(\boldsymbol{u}_0, \boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ *je lineárne nezávislý.*

Napríklad postupnosť $(1, x, x^2, \dots, x^k, \dots)$ všetkých mocnín x je lineárne nezávislá postupnosť vo vektorovom priestore $K[x]$ všetkých polynómov v premennej x nad poľom K. Polynóm $f(x) = a_0 + a_1x + \ldots + a_nx^n$ je totiž (definitoricky) nulový práve vtedy, keď $a_0 = a_1 = \ldots = a_n = 0$. *Množina* $X \subseteq V$ sa nazýva *lineárne nezávislá*, ak pre ľubovoľné $n \in \mathbb{N}$ každá usporiadaná n-tica *navzájom rôznych* vektorov $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ z množiny X je lineárne nezávislá.

Ešte raz podčiarkujeme ono „navzájom rôznych“ – keby totiž $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ neboli navzájom rôzne vektory, nemohli by byť lineárne nezávislé.

Lineárna závislosť či nezávislosť usporiadanej n-tice vektorov nezávisí od ich poradia – zrejme usporiadaná n-tica $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ je lineárne nezávislá práve vtedy, keď je lineárne nezávislá usporiadaná n-tica $(\boldsymbol{u}_{\sigma(1)}, \dots, \boldsymbol{u}_{\sigma(n)})$, kde σ je ľubovoľná permutácia množiny $\{1, \dots, n\}$. Inak povedané, lineárna (ne)závislosť usporiadanej n-tice $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ navzájom rôznych vektorov je vlastnosťou množiny $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_n\}$. Čitateľ už iste ľahko nahliadne platnosť nasledujúceho očividného tvrdenia.

4.6.2. Tvrdenie. *Usporiadaná* n*-tica* $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ *navzájom rôznych vektorov z* V *je lineárne nezávislá práve vtedy, keď množina* $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_n\} \subseteq V$ *je lineárne nezávislá.*

Naše záverečné tvrdenie, ktoré dáva do súvisu lineárnu (ne)závislosť množiny s jej lineárnym obalom, je obdobou tvrdenia 4.4.3 Taktiež jeho dôkaz možno získať malou obmenou dôkazu spomínaného tvrdenia.

4.6.3. Tvrdenie. *Nech* $X \subseteq V$ *je lineárne nezávislá množina a* $\boldsymbol{v} \in V$*. Potom nasledujúce podmienky sú ekvivalentné:*

(i) $\boldsymbol{v} \in [X]$;
(ii) *množina* $X \cup \{\boldsymbol{v}\}$ *je lineárne závislá;*
(iii) $[X \cup \{\boldsymbol{v}\}] = [X]$.

Cvičenia

4.1. Nech S, T sú lineárne podpriestory vektorového priestoru V. Potom $S \cup T$ je lineárny podpriestor V práve vtedy, keď $S \subseteq T$ alebo $T \subseteq S$. Dokážte.

4.2. V každom z nasledujúcich prípadov rozhodnite, či daná podmnožina S vektorového priestoru V nad poľom K je jeho lineárnym podpriestorom. Svoje rozhodnutie zdôvodnite. Ak S nie je lineárny podpriestor, popíšte jeho lineárny obal $[S]$.

(a) $K = \mathbb{R}$, $V = \mathbb{R}$, $S = \langle -1, 1\rangle$;
(b) $K = \mathbb{C}$, $V = \mathbb{C}$, $S = \{z \in \mathbb{C};\ |z| = 1\}$;
(c) $K = \mathbb{R}$, $V = \mathbb{R}^2$, $S = \{(x, y) \in \mathbb{R}^2;\ x - 2y = 0\}$;
(d) $K = \mathbb{C}$, $V = \mathbb{C}^2$, $S = \{(x, y) \in \mathbb{C}^2;\ x + \mathrm{i}y = 1\}$;
(e) $K = \mathbb{R}$, $V = \mathbb{C}$, $S = \{x \in \mathbb{C};\ \operatorname{Re} x = \operatorname{Im} x\}$;
(f) $K = \mathbb{Q}$, $V = \mathbb{R}$, $S = \mathbb{Q}[\sqrt{3}] = \{a + b\sqrt{3};\ a, b \in \mathbb{Q}\}$;
(g) $K = \mathbb{Z}_2$, $V = \mathbb{Z}_2^3$, $S = \{(0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1)\}$;
(h) $K = \mathbb{Z}_3$, $V = \mathbb{Z}_3^2$, $S = \{(0, 0), (1, 2), (2, 1)\}$;
(i) K ľubovoľné, $V = K[x]$, $S = \{f(x) \in K[x];\ f(1) = 0\}$;
(j) K ľubovoľné, $V = K[x]$, $S = \{f(x) \in K[x];\ f(0) = 1\}$;
(k) K ľubovoľné, $V = K[x]$, $S = \{f(x) \in K[x]; f(0) = f(1)\}$;
(l) K ľubovoľné, $V = K[x]$, $S = \{a + bx + (a + b)x^2;\ a, b \in K\}$;
(m) K ľubovoľné, $V = K[x]$, $S = \{a + bx + (a + b + 1)x^2;\ a, b \in K\}$.

4.3. V každom z nasledujúcich prípadov rozhodnite, či uvedené vektory z vektorového priestoru V nad poľom K sú lineárne nezávislé. Svoje rozhodnutie odôvodnite.

(a) $K = \mathbb{Q}$, $V = \mathbb{Q}^2$, $\boldsymbol{u} = (0, 0)$, $\boldsymbol{v} = (1, 1)$;
(b) $K = \mathbb{Q}$, $V = \mathbb{Q}^3$, $\boldsymbol{w} = (1, 2, 3)^\mathsf{T}$;
(c) $K = \mathbb{R}$, $V = \mathbb{R}^4$, $\boldsymbol{x} = (0, 1, 0, 1)^\mathsf{T}$, $\boldsymbol{y} = (1, 0, 1, 0)^\mathsf{T}$, $\boldsymbol{z} = (1, 0, 0, 1)^\mathsf{T}$;
(d) $K = \mathbb{C}$, $V = \mathbb{C}^3$, $\boldsymbol{u} = (1, \mathrm{i}, -\mathrm{i})$, $\boldsymbol{v} = (2 + \mathrm{i}, 3 - \mathrm{i}, 1 + 2\mathrm{i})$, $\boldsymbol{w} = (2 + 2\mathrm{i}, 2 - \mathrm{i}, 2 + 2\mathrm{i})$;
(e) $K = \mathbb{Z}_5$, $V = \mathbb{Z}_5^4$, $\boldsymbol{x} = (1, 3, 2, 4)$, $\boldsymbol{y} = (2, 2, 1, 1)$, $\boldsymbol{z} = (4, 2, 4, 2)$;
(f) $K = \mathbb{Z}_7$, $V = \mathbb{Z}_7^4$, $\boldsymbol{x} = (1, 3, 2, 4)$, $\boldsymbol{y} = (2, 2, 1, 1)$, $\boldsymbol{z} = (4, 2, 4, 2)$;
(g) $K = \mathbb{R}$, $V = \mathbb{R}^{(3)}[x]$, $f_0(x) = 1$, $f_1(x) = x$, $f_2(x) = x(x-1)$, $f_3(x) = x(x-1)(x-2)$;
(h) $K = \mathbb{Z}_3$, $V = \mathbb{Z}_{13}[x]$, $f(x) = 5 + 12x$, $g(x) = 12 + 8x$;
(i) $K = \mathbb{R}$, $V = \mathbb{R}[x]$, $f(x) = 5 + 12x$, $g(x) = 12 + 8x$.

4.4. Nech V je vektorový priestor nad poľom K, $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n \in V$ sú lineárne nezávislé vektory a $\boldsymbol{A} \in K^{m \times n}$. Pre $i = 1, \dots, m$ označme $\boldsymbol{v}_i = a_{i1}\boldsymbol{u}_1 + \ldots + a_{in}\boldsymbol{u}_n = \boldsymbol{r}_i(\boldsymbol{A}) \cdot (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)^\mathsf{T}$. Potom vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_m \in V$ sú lineárne nezávislé práve vtedy, keď riadky matice $\boldsymbol{A}$ sú lineárne nezávislé vektory v K^n. Dokážte. Čo sa zmení, ak $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ sú lineárne závislé?

4.5. Nech V je vektorový priestor nad poľom K. V každom z nasledujúcich prípadov rozhodnite, či vektor $\boldsymbol{u} \in V$ patrí do lineárneho obalu vektorov $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$.

(a) $K = \mathbb{R}$, $V = \mathbb{R}^3$, $\boldsymbol{x} = (1,1,2)^\mathsf{T}$, $\boldsymbol{y} = (-2,1,-1)^\mathsf{T}$, $\boldsymbol{z} = (0,1,1)^\mathsf{T}$, $\boldsymbol{u} = (1,2,-1)^\mathsf{T}$;

(b) $K = \mathbb{C}$, $V = \mathbb{C}^2$, $\boldsymbol{x} = (\mathrm{i}, 1+\mathrm{i})$, $\boldsymbol{y} = (1, 1-\mathrm{i})$, $\boldsymbol{z} = (\mathrm{i}, -\mathrm{i})$, $\boldsymbol{u} = (1+\mathrm{i}, 1-\mathrm{i})$;

(c) $K = \mathbb{Z}_3$, $V = \mathbb{Z}_3^3$, $\boldsymbol{x} = (0,1,2)$, $\boldsymbol{y} = (1,2,0)$, $\boldsymbol{z} = (2,0,1)$, $\boldsymbol{u} = (1,0,1)$;

(d) $K = \mathbb{Z}_5$, $V = \mathbb{Z}_5^3$, $\boldsymbol{x} = (0,1,2)$, $\boldsymbol{y} = (1,2,0)$, $\boldsymbol{z} = (2,0,1)$, $\boldsymbol{u} = (1,0,1)$.

4.6. V každej z úloh (a)–(i) cvičenia 4.3 vyberte z daných vektorov lineárne nezávislé vektory, ktoré generujú ten istý lineárny podpriestor ako pôvodné vektory. Riešte rovnaký problém pre vektory $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$, $\boldsymbol{u}$ v každej z úloh (a)–(d) cvičenia 4.5. Využite pri tom výsledky cvičení 4.3 a 4.5.

4.7. Doplňte chýbajúce dôkazy častí (a)–(e) tvrdenia 4.2.2

4.8. (a) Zovšeobecnite definíciu *priameho súčtu* na ľubovoľný konečný počet lineárnych podpriestorov daného vektorového priestoru.

(b) Nech $S_1, \ldots, S_n$ ($n \geq 2$) sú lineárne podpriestory vektorového priestoru V. Pre $i = 1, \ldots, n$ označme $T_i = S_1 + \ldots + S_{i-1} + S_{i+1} + \ldots + S_n$, t. j. $T_1 = S_2 + \ldots + S_n$, $T_2 = S_1 + S_3 + \ldots + S_n$, $\ldots$, $T_n = S_1 + \ldots + S_{n-1}$. Potom $S_1 + \ldots + S_n = S_1 \oplus \ldots \oplus S_n$, t. j. súčet podpriestorov $S_1, \ldots, S_n$ je priamy, práve vtedy, keď pre každé $i \leq n$ platí $S_i \cap T_i = \{\mathbf{0}\}$. Dokážte.

4.9. Nech V je vektorový priestor nad poľom K a $X \subseteq V$ je ľubovoľná podmnožina. Dokážte tzv. *podmienku zámeny*: $(\forall \boldsymbol{u}, \boldsymbol{v} \in V)(\boldsymbol{u} \in [X \cup \{\boldsymbol{v}\}] \smallsetminus [X] \;\Rightarrow\; \boldsymbol{v} \in [X \cup \{\boldsymbol{u}\}])$. Rozhodnite, či platí dokonca ekvivalencia $(\forall \boldsymbol{u}, \boldsymbol{v} \in V)(\boldsymbol{u} \in [X \cup \{\boldsymbol{v}\}] \smallsetminus [X] \;\Leftrightarrow\; \boldsymbol{v} \in [X \cup \{\boldsymbol{u}\}] \smallsetminus [X])$?

4.10. Dokážte tvrdenia 4.6.1, 4.6.2 a 4.6.3

4.11. Nech K je pole a $(p_k(x))_{k=0}^{\infty}$ je postupnosť polynómov z $K[x]$ taká, že pre $k \neq l$ majú polynómy $p_k(x)$, $p_l(x)$ rôzny stupeň. Dokážte, že potom ide o lineárne nezávislú postupnosť.

5. Báza a dimenzia

V tejto kapitole sa oboznámime s pojmom *bázy* vektorového priestoru, čo nám v niektorých vektorových priestoroch umožní zaviesť *súradnice*. Ďalej budeme definovať *dimenziu* vektorového priestoru a odvodíme jej základné vlastnosti. V nasledujúcej kapitole si potom okrem iného dokážeme, že dimenzia je základný štruktúrny invariant tzv. *konečnorozmerných* vektorových priestorov.

I v tejto kapitole V označuje nejaký vektorový priestor nad pevným poľom K.

5.1 Steinitzova veta a konečnorozmerné priestory

Začneme jedným technickým výsledkom kľúčového významu.

5.1.1. Tvrdenie. (Steinitzova veta) *Nech $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m \in V$. Ak vektory $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ sú lineárne nezávislé a všetky patria do lineárneho obalu $[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m]$, tak $n \leq m$.*

Dôkaz. Keďže $\boldsymbol{u}_j \in [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m]$ pre každé $j \leq n$, existujú $\boldsymbol{c}_j = (c_{1j}, \ldots, c_{mj})^\mathsf{T} \in K^m$ také, že

$$\boldsymbol{u}_j = c_{1j}\boldsymbol{v}_1 + \ldots + c_{mj}\boldsymbol{v}_m = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) \cdot \boldsymbol{c}_j.$$

Inak povedané

$$(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) \cdot \boldsymbol{C},$$

kde $\boldsymbol{C} = (c_{ij})_{m\times n}$ je matica so stĺpcami $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$.

Predpokladajme, že $m < n$. Potom podľa tvrdenia 3.3.6 má homogénna sústava $\boldsymbol{C} \cdot \boldsymbol{x} = \boldsymbol{0}$ aspoň jedno riešenie $\boldsymbol{x} = (x_1, \ldots, x_n)^\mathsf{T} \neq \boldsymbol{0}$. Jednoduchým výpočtom dostávame

$$\begin{aligned} x_1\boldsymbol{u}_1 + \ldots + x_n\boldsymbol{u}_n &= (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) \cdot \boldsymbol{x} \\ &= (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) \cdot \boldsymbol{C} \cdot \boldsymbol{x} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) \cdot \boldsymbol{0} = \boldsymbol{0}, \end{aligned}$$

čo je v spore s lineárnou nezávislosťou vektorov $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$.

5.1.2. Tvrdenie. *Pre ľubovoľný vektorový priestor V nasledujúce podmienky sú ekvivalentné:*

(i) existuje konečná množina $X \subseteq V$ taká, že $[X] = V$;

(ii) každá lineárne nezávislá množina $Y \subseteq V$ je konečná.

Dôkaz. (i) $\Rightarrow$ (ii): Nech $X \subseteq V$ je konečná množina, ktorá generuje V. Podľa Steinitzovej vety pre ľubovoľné lineárne nezávislé vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ platí $n \leq \# X$, teda každá lineárne nezávislá množina $Y \subseteq V$ je konečná.

(ii) $\Rightarrow$ (i): Budeme dokazovať logicky ekvivalentnú implikáciu $\neg$ (i) $\Rightarrow \neg$ (ii).

Predpokladajme, že žiadna konečná podmnožina priestoru V negeneruje V. Potom vo V môžeme zostrojiť postupnosť vektorov $(\boldsymbol{y}_n)_{n=0}^{\infty}$ takú, že $\boldsymbol{y}_0 \neq \boldsymbol{0}$ a pre každé $n > 0$ platí $\boldsymbol{y}_n \notin [\boldsymbol{y}_0, \dots, \boldsymbol{y}_{n-1}]$. Podľa tvrdenia 4.4.1 je každý počiatočný úsek $(\boldsymbol{y}_0, \dots, \boldsymbol{y}_n)$ tejto postupnosti lineárne nezávislý, takže celá postupnosť je lineárne nezávislá podľa tvrdenia 4.6.1 Teda vo V existuje nekonečná lineárne nezávislá množina, napr. $Y = \{\boldsymbol{y}_n;\ n \in \mathbb{N}\}$.

Hovoríme, že vektorový priestor V je *konečnorozmerný* , ak spĺňa niektorú (teda nevyhnutne obe) z ekvivalentných podmienok (i), (ii) práve dokázaného tvrdenia. V opačnom prípade hovoríme, že V je *nekonečnorozmerný* vektorový priestor.

5.2 Báza a dimenzia konečnorozmerného priestoru

Nech V je konečnorozmerný vektorový priestor. *Bázou* priestoru V nazývame každú lineárne nezávislú usporiadnú n-ticu $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ vektorov z V, ktorá generuje celý priestor V. Stručne tiež hovoríme, že vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ tvoria bázu priestoru V.

Nasledujúce tvrdenie je priamym dôsledkom vety 4.4.4

5.2.1. Tvrdenie. *Nech V je konečnorozmerný vektorový priestor. Potom*

(a) ľubovoľnú lineárne nezávislú usporiadanú k-ticu $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ vektorov z V možno doplniť do nejakej bázy $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \dots, \boldsymbol{u}_n)$ priestoru V;

(b) z ľubovoľnej generujúcej usporiadanej m-tice $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ vektorov z V možno vybrať nejakú bázu $(\boldsymbol{v}_{i_1}, \dots, \boldsymbol{v}_{i_n})$ priestoru V.

5.2.2. Veta. *Nech V je konečnorozmerný vektorový priestor. Potom*

(a) V má aspoň jednu bázu;

(b) ľubovoľné dve bázy priestoru V majú rovnaký počet prvkov.

Dôkaz. (a) je bezprostredným dôsledkom predchádzajúceho tvrdenia, ktoré nám dokonca dáva dva varianty dôkazu: jeden doplnením prázdnej množiny (ktorá je lineárne nezávislá) na bázu vo V, druhý výberom bázy z nejakej konečnej generujúcej množiny vo V.

(b) je bezprostredným dôsledkom Steinitzovej vety. Ak sú totiž $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$, $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ dve bázy vo V, tak, keďže $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ je lineárne nezávislá a

$(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ generuje celý priestor V, musí platiť $n \leq m$. Nakoľko však i $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ je lineárne nezávislá a $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ generuje celé V, platí tiež $m \leq n$. Teda $m = n$.

Práve dokázaná veta nám umožňuje korektne definovať *dimenziu* alebo tiež *rozmer* konečnorozmerného vektorového priestoru V ako počet prvkov jeho ľubovoľnej bázy. Dimenziu vektorového priestoru V značíme $\dim V$. Ak $\dim V = n$, hovoríme, že V je *n-rozmerný* vektorový priestor. Ak V je nekonečnorozmerný priestor, kladieme $\dim V = \infty$. V prípade, že bude potrebné zdôrazniť úlohu poľa K, budeme používať podrobnejšie označenie $\dim_K V$.

Teda V je konečnorozmerný práve vtedy, keď $\dim V < \infty$.

Dôkaz nasledujúceho tvrdenia prenechávame ako cvičenie čitateľovi.

5.2.3. Tvrdenie. *Nech $\dim V = n$, $\boldsymbol{v}_1, \dots, \boldsymbol{v}_m \in V$. Potom ľubovoľné dve z nasledujúcich podmienok implikujú tretiu:*
(i) vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_m$ sú lineárne nezávislé;
(ii) $[\boldsymbol{v}_1, \dots, \boldsymbol{v}_m] = V$;
(iii) $m = n$.

To okrem iného znamená, že na overenie, či n vektorov $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ tvorí bázu n-rozmerného vektorového priestoru V, stačí overiť len jednu (a to ľubovoľnú) z podmienok (i), (ii).

5.3 Súradnice vektora vzhľadom na danú bázu

Nasledujúca veta je špeciálnym prípadom vety 4.4.2

5.3.1. Veta. *Vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ tvoria bázu vektorového priestoru V práve vtedy, keď každý vektor $\boldsymbol{x} \in V$ možno jednoznačne vyjadriť v tvare lineárnej kombinácie*

$$\boldsymbol{x} = c_1\boldsymbol{u}_1 + \ldots + c_n\boldsymbol{u}_n.$$

Uvedomme si, že existencia aspoň jedného vyjadrenia $\boldsymbol{x} = c_1\boldsymbol{u}_1 + \ldots + c_n\boldsymbol{u}_n$ je ekvivalentná s podmienkou, že vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ generujú V. Jednoznačnosť tohto vyjadrenia je zasa ekvivalentná s lineárnou nezávislosťou vektorov $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$.

Teda $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ je bázou V vtedy a len vtedy, keď pre každé $\boldsymbol{x} \in V$ existuje práve jedno $\boldsymbol{c} = (c_1, \dots, c_n)^{\mathsf{T}} \in K^n$ také, že

$$\boldsymbol{x} = c_1\boldsymbol{u}_1 + \ldots + c_n\boldsymbol{u}_n = \boldsymbol{\alpha} \cdot \boldsymbol{c}.$$

Tento jednoznačne určený stĺpcový vektor $\boldsymbol{c} \in K^n$ budeme nazývať *súradnice vektora $\boldsymbol{x}$ vzhľadom na bázu $\boldsymbol{\alpha}$* a označovať

$$\boldsymbol{c} = (\boldsymbol{x})_{\boldsymbol{\alpha}}.$$

Teda každá báza $\boldsymbol{\alpha}$ v n-rozmernom vektorovom priestore V definuje *súradnicové zobrazenie* $\boldsymbol{x} \mapsto (\boldsymbol{x})_{\boldsymbol{\alpha}}$ z V do stĺpcového vektorového priestoru K^n.

Jednoduchý dôkaz nasledujúceho tvrdenia prenechávame ako cvičenie čitateľovi.

5.3.2. Tvrdenie. *Nech $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ je báza konečnorozmerného vektorového priestoru V. Potom príslušné súradnicové zobrazenie $V \to K^n$ je bijektívne a zachováva lineárne kombinácie, t. j. pre ľubovoľné $a, b \in K$, $\boldsymbol{x}, \boldsymbol{y} \in V$ platí*

$$(a\boldsymbol{x} + b\boldsymbol{y})_{\boldsymbol{\alpha}} = a(\boldsymbol{x})_{\boldsymbol{\alpha}} + b(\boldsymbol{y})_{\boldsymbol{\alpha}}.$$

K nemu inverzné zobrazenie $K^n \to V$ je dané predpisom $\boldsymbol{c} \mapsto \boldsymbol{\alpha} \cdot \boldsymbol{c}$.

V označení posledného tvrdenia teda pre ľubovoľné $\boldsymbol{x} \in V$, $\boldsymbol{c} \in K^n$ platí

$$\boldsymbol{x} = \boldsymbol{\alpha} \cdot (\boldsymbol{x})_{\boldsymbol{\alpha}}, \qquad (\boldsymbol{\alpha} \cdot \boldsymbol{c})_{\boldsymbol{\alpha}} = \boldsymbol{c}.$$

Prvá rovnosť ukazuje, ako možno vektor $\boldsymbol{x}$ zrekonštruovať z danej bázy $\boldsymbol{\alpha}$ a jeho súradníc $(\boldsymbol{x})_{\boldsymbol{\alpha}}$ v tejto báze; druhá zachytáva zrejmý fakt, že súradnice lineárnej kombinácie $\sum_{i=1}^n c_i \boldsymbol{u}_i$ v báze $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ tvorí vektor $(c_1, \dots, c_n)^\mathsf{T}$.

Práve zavedené súradnice by sme mohli podrobnejšie nazvať *stĺpcovými súradnicami* vzhľadom na danú bázu. Podobným spôsobom možno zaviesť i *riadkové súradnice* a dokázať pre ne analogické tvrdenia ako pre stĺpcové. V takom prípade je samozrejme vhodnejšie zapisovať príslušnú bázu ako stĺpcový vektor $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)^\mathsf{T}$ a v prípade riadkového priestoru $V = K^n$ ju stotožniť s maticou s riadkami $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$. Podrobnosti prenechávame na doplnenie čitateľovi.

5.3.3. Príklad. Označme $\boldsymbol{e}_i^{(n)} = \boldsymbol{s}_i(\boldsymbol{I}_n) \in K^n$ stĺpcový vektor pozostávajúci zo samých núl, okrem i-tej zložky, ktorá je 1. Potom $\boldsymbol{\varepsilon}^{(n)} = \big(\boldsymbol{e}_1^{(n)}, \dots, \boldsymbol{e}_n^{(n)}\big)$ je báza stĺpcového vektorového priestoru K^n. Nazývame ju *kanonickou bázou* tohto priestoru. Túto bázu možno zrejmým spôsobom stotožniť s jednotkovou maticou $\boldsymbol{I}_n$. Pokiaľ nebude hroziť nedorozumenie, budeme horný index (n) vynechávať a príslušnú bázu označovať stručne $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$. Pre ľubovoľný vektor $\boldsymbol{x} = (x_1, \dots, x_n)^\mathsf{T} \in K^n$ platí

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n,$$

preto $(\boldsymbol{x})_{\boldsymbol{\varepsilon}} = \boldsymbol{x}$, t. j. každý vektor $\boldsymbol{x} \in K^n$ splýva so svojimi vlastnými súradnicami v kanonickej báze.

Kanonická báza riadkového vektorového priestoru K^n je tvorená riadkami jednotkovej matice $\boldsymbol{I}_n$ a značíme ju rovnako ako v predchádzajúcom prípade $\boldsymbol{\varepsilon}^{(n)} = \left(\boldsymbol{e}_1^{(n)}, \ldots, \boldsymbol{e}_n^{(n)}\right)^{\mathsf{T}}$ alebo stručne $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)^{\mathsf{T}}$, len s tým rozdielom, že $\boldsymbol{\varepsilon}^{(n)} = \boldsymbol{\varepsilon}$ je teraz stĺpec vektorov a každé $\boldsymbol{e}_i$ je riadok pozostávajúci zo samých núl, okrem i-teho miesta, ktoré je 1.

V predošlom príklade je, okrem iného, zahrnutý aj dôkaz nasledujúceho očakávaného výsledku.

5.3.4. Veta. *Pre ľubovoľné $n \in \mathbb{N}$ platí* $\dim K^n = n$.

5.3.5. Príklad. Stĺpce matice

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

tvoria bázu $\boldsymbol{\alpha}$ (stĺpcového) vektorového priestoru K^4 (presvedčte sa o tom s využitím tvrdenia 5.2.3 a vety 5.3.4 Súradnice vektora $\boldsymbol{x} = (x_1, x_2, x_3, x_4)^{\mathsf{T}} \in K^n$ v báze $\boldsymbol{\alpha}$ sú dané vzťahom

$$(\boldsymbol{x})_{\boldsymbol{\alpha}} = (x_4, x_3 - x_4, x_2 - x_3, x_1 - x_2)^{\mathsf{T}}.$$

Platí totiž

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_4 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + (x_3 - x_4) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + (x_2 - x_3) \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + (x_1 - x_2) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Overte.

5.3.6. Príklad. Označme $\boldsymbol{\xi}^{(n)} = (1, x, \ldots, x^n)$ usporiadanú $(n+1)$-ticu prvých $n+1$ mocnín premennej x. Ľahko nahliadneme, že $\boldsymbol{\xi}^{(n)}$ je báza vektorového priestoru $K^{(n)}[x]$ všetkých polynómov stupňa $\leq n$ v premennej x nad poľom K. Súradnice polynómu $f(x) = \sum_{i=0}^{n} a_i x^i$ v tejto báze tvorí vektor

$$(f)_{\boldsymbol{\xi}^{(n)}} = (a_0, a_1, \ldots, a_n)^{\mathsf{T}} \in K^{n+1}.$$

Teda $\dim K^{(n)}[x] = n+1$. Na druhej strane vektorový priestor $K[x]$ všetkých polynómov v premennej x nad poľom K zrejme nie je konečnorozmerný, teda $\dim K[x] = \infty$.

5.3.7. Príklad. Nech $m, n \in \mathbb{N}$. Pre ľubovoľné $1 \leq k \leq m$, $1 \leq l \leq n$ označme $\boldsymbol{E}_{kl}^{(m,n)} = \boldsymbol{E}_{kl} = (\delta_{ik}\delta_{jl})_{m\times n}$ maticu typu $m \times n$ nad poľom K, pozostávajúcu zo samých núl, okrem miesta (k, l), na ktorom je 1. Zrejme každú maticu $\boldsymbol{A} = (a_{kl}) \in K^{m\times n}$ možno jednoznačne vyjadriť v tvare

$$\boldsymbol{A} = \sum_{k=1}^{m}\sum_{l=1}^{n} a_{kl}\boldsymbol{E}_{kl},$$

z čoho vyplýva, že matice $\boldsymbol{E}_{kl}^{(m,n)}$, $1 \leq k \leq m$, $1 \leq l \leq n$, tvoria bázu vektorového priestoru $K^{m\times n}$ všetkých matíc typu $m \times n$ nad poľom K. Jej špeciálnym prípadom je kanonická báza $\boldsymbol{\varepsilon}^{(n)}$ v priestore K^n. Dostávame tak ďalší očakávaný vzťah: $\dim K^{m\times n} = mn$.

5.3.8. Príklad. Pole $\mathbb{C}$ všetkých komplexných čísel je rozšírením poľa $\mathbb{R}$ všetkých reálnych čísel. Teda $\mathbb{C}$ možno považovať za vektorový priestor nad poľom $\mathbb{R}$ (príklad 1.6.1). Každé komplexné číslo z možno jednoznačne vyjadriť v tvare

$$z = a + b\mathrm{i} = a1 + b\mathrm{i},$$

kde $a = \operatorname{Re} z$, $b = \operatorname{Im} z$ sú reálne čísla, nazývané *reálna* resp. *imaginárna časť* komplexného čísla z, a i je *imaginárna jednotka*. To znamená, že komplexné čísla (t. j. vektory) 1, i tvoria bázu vektorového priestoru $\mathbb{C}$ nad poľom $\mathbb{R}$. Súradnicové zobrazenie vzhľadom na túto bázu je dané vzťahom

$$(z)_{(1,\mathrm{i})} = \begin{pmatrix} \operatorname{Re} z \\ \operatorname{Im} z \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(z + \bar{z}) \\ \frac{1}{2\mathrm{i}}(z - \bar{z}) \end{pmatrix},$$

kde $\bar{z} = a - b\mathrm{i}$ je číslo *komplexne združené* k číslu $z = a + b\mathrm{i}$. Teda $\dim_{\mathbb{R}} \mathbb{C} = 2$. Na druhej strane každé pole K, uvažované ako vektorový priestor nad sebou samým má dimenziu 1, t. j. $\dim_K K = 1$. Špeciálne $\dim_{\mathbb{R}} \mathbb{R} = 1$ aj $\dim_{\mathbb{C}} \mathbb{C} = 1$.

5.4 Dimenzia prieniku, súčtu a súčinu vektorových priestorov

V tomto paragrafe preskúmame niektoré základné vlastnosti dimenzie, uvažovanej ako zobrazenie definované na všetkých vektorových priestoroch nad pevným poľom K.

Na začiatok si uvedomme, že ľubovoľný lineárny podpriestor S vektorového priestoru V je i sám vektorovým priestorom nad tým istým poľom, teda pojmy ako báza podpriestoru S a dimenzia podpriestoru S majú dobre definovaný význam. Zrejme každý podpriestor konečnorozmerného vektorového priestoru je i sám konečnorozmerný.

5.4.1. Veta. *Nech S, $T \subseteq V$ sú konečnorozmerné lineárne podpriestory vektorového priestoru V. Potom*

$$\dim(S+T) = \dim S + \dim T - \dim(S \cap T).$$

Dôkaz. Označme $\dim S = m$, $\dim T = n$, $\dim(S \cap T) = k$. Nech $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ je báza podpriestoru $S \cap T$. Doplňme túto bázu do bázy $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{m-k}$ podpriestoru S, a taktiež do bázy $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{w}_1, \dots, \boldsymbol{w}_{n-k}$ podpriestoru T. Dokážeme, že vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{m-k}, \boldsymbol{w}_1, \dots, \boldsymbol{w}_{n-k}$ tvoria bázu podpriestoru $S+T$. Tým budeme hotoví, lebo potom naozaj platí

$$\begin{aligned}\dim(S+T) &= k + (m-k) + (n-k) = m+n-k \\ &= \dim S + \dim T - \dim(S \cap T).\end{aligned}$$

Keďže vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{m-k}, \boldsymbol{w}_1, \dots, \boldsymbol{w}_{n-k}$ zrejme generujú podpriestor $S+T$ (premyslite si detaily), zostáva dokázať, že sú tiež lineárne nezávislé. Nech $a_1, \dots, a_k$, $b_1, \dots, b_{m-k}$, $c_1, \dots, c_{n-k}$ sú skaláry také, že

$$a_1\boldsymbol{u}_1 + \ldots + a_k\boldsymbol{u}_k + b_1\boldsymbol{v}_1 + \ldots + b_{m-k}\boldsymbol{v}_{m-k} + c_1\boldsymbol{w}_1 + \ldots + c_{n-k}\boldsymbol{w}_{m-k} = \mathbf{0}.$$

Potom

$$a_1\boldsymbol{u}_1 + \ldots + a_k\boldsymbol{u}_k + b_1\boldsymbol{v}_1 + \ldots + b_{m-k}\boldsymbol{v}_{m-k} = -(c_1\boldsymbol{w}_1 + \ldots + c_{n-k}\boldsymbol{w}_{m-k}),$$

pričom vektor na ľavej strane patrí do S a vektor na pravej do T. Túto spoločnú hodnotu $\boldsymbol{z} \in S \cap T$ možno vyjadriť ako lineárnu kombináciu len vektorov $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$. Z jednoznačnosti vyjadrenia $\boldsymbol{z}$ v báze $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{m-k}$ podpriestoru S tak dostávame $b_1 = \ldots = b_{m-k} = 0$. Preto

$$a_1\boldsymbol{u}_1 + \ldots + a_k\boldsymbol{u}_k + c_1\boldsymbol{w}_1 + \ldots + c_{m-k}\boldsymbol{w}_{n-k} = \mathbf{0}.$$

Z lineárnej nezávislosti bázy $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$, $\boldsymbol{w}_1, \dots, \boldsymbol{w}_{n-k}$ podpriestoru T potom vyplýva $a_1 = \ldots = a_k = 0$, $c_1 = \ldots = c_{n-k} = 0$. Teda $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{m-k}$, $\boldsymbol{w}_1, \dots, \boldsymbol{w}_{n-k}$ sú lineárne nezávislé vektory.

5.4.2. Dôsledok. *Nech S, T sú lineárne podpriestory vektorového priestoru V. Potom $S \cap T = \{\mathbf{0}\}$, t. j. súčet $S+T$ je direktný, práve vtedy, keď*

$$\dim(S+T) = \dim S + \dim T.$$

Práve dokázané vzťahy pre dimenzie konečnorozmerných podpriestorov nejakého vektorového priestoru nápadne pripomínajú vzťah

$$\#(X \cup Y) = \#X + \#Y - \#(X \cap Y)$$

pre počty prvkov konečných množín z paragrafu 0.2, ktorý sa v prípade disjunktných množín redukuje na rovnosť

$$\#\,(X \cup Y) = \#\,X + \#\,Y.$$

To znamená, že konečnorozmerné vektorové priestory (hoci v typickom prípade priestorov nenulovej dimenzie nad nekonečným poľom ide o nekonečné množiny) sa správajú do značnej miery podobne ako konečné množiny. Dimenzia $\dim V$ konečnorozmerného priestoru V je tak akousi mierou jeho „veľkosti“, podobne ako počet prvkov $\#\,X$ je mierou veľkosti konečnej množiny X. Direktný (priamy) súčet lineárnych podpriestorov je tak analógiou zjednotenia disjunktných množín.

Na rozdiel od multiplikatívneho charakteru počtu prvkov karteziánskeho súčinu konečných množín, ktorý je daný formulou

$$\#\,(X \times Y) = \#\,X \cdot \#\,Y,$$

sa však dimenzia priameho súčinu konečnorozmerných vektorových priestorov (pozri príklad 1.6.4) správa aditívne, t. j. do značnej miery podobne ako logaritmus.

5.4.3. Tvrdenie. *Nech V, W sú konečnorozmerné vektorové priestory nad poľom K. Potom pre dimenziu ich priameho súčinu platí*

$$\dim(V \times W) = \dim V + \dim W.$$

Dôkaz. Nech $\boldsymbol{v}_1, \dots, \boldsymbol{v}_m$ je báza priestoru V a $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$ je báza priestoru W. Stačí overiť, že vektory $(\boldsymbol{v}_1, \mathbf{0}), \dots, (\boldsymbol{v}_m, \mathbf{0}), (\mathbf{0}, \boldsymbol{w}_1), \dots, (\mathbf{0}, \boldsymbol{w}_n)$ tvoria bázu priameho súčinu $V \times W$. Podrobnosti prenechávame čitateľovi.

V dôsledku toho pre konečnorozmerné priestory $V_1, \dots, V_k$ nad poľom K platí

$$\dim(V_1 \times \ldots \times V_k) = \dim V_1 + \ldots + \dim V_k,$$

a pre k-tu priamu mocninu V^k priestoru V máme

$$\dim V^k = k \dim V.$$

Poznámka. Ak obvyklým spôsobom rozšírime aritmetiku prirodzených čísel aj na symbol ∞, t. j. položíme $n + \infty = \infty + n = \infty + \infty = \infty$ pre $n \in \mathbb{N}$, $0 \cdot \infty = \infty \cdot 0 = 0$ a $n \cdot \infty = \infty \cdot n = \infty \cdot \infty = \infty$ pre $n > 0$, ľahko nahliadneme, že vzťahy dokázané v tomto paragrafe zostávajú v platnosti aj pre nekonečnorozmerné priestory.

5.5 Usporiadané a neusporiadané bázy

Ak $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ je báza vektorového priestoru V, tak $(\boldsymbol{u}_{\sigma(1)}, \dots, \boldsymbol{u}_{\sigma(n)})$ je tiež báza V pre ľubovoľnú premutáciou σ množiny $\{1, \dots, n\}$. Inak povedané, vlastnosť „byť bázou vektorového priestoru“ nezávisí od poradia vektorov v báze – nie je to ani tak vlastnosť príslušnej usporiadanej n-tice $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ ako skôr množiny $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_n\}$. Na druhej strane je rozumné považovať bázy $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ a $(\boldsymbol{u}_{\sigma(1)}, \dots, \boldsymbol{u}_{\sigma(n)})$, kde σ je neidentická permutácia, za rôzne. Prislúchajú im totiž rôzne súradnicové zobrazenia. Napr. $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$, $\boldsymbol{\eta} = (\boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_1)$ sú bázy stĺpcového priestoru K^3, líšiace sa len poradím svojich vektorov. Pre súradnice ľubovoľného vektora $\boldsymbol{x} = (x_1, x_2, x_3)^{\mathsf{T}} \in K^3$ v týchto bázach však platí:

$$(\boldsymbol{x})_{\boldsymbol{\varepsilon}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \qquad (\boldsymbol{x})_{\boldsymbol{\eta}} = \begin{pmatrix} x_2 \\ x_3 \\ x_1 \end{pmatrix}.$$

Teda $(\boldsymbol{x})_{\boldsymbol{\varepsilon}} \neq (\boldsymbol{x})_{\boldsymbol{\eta}}$, okrem prípadu, keď $x_1 = x_2 = x_3$.

Doteraz študované bázy by sme vlastne mali presnejšie nazývať *konečnými usporiadanými bázami.* To naznačuje možnosti uvažovať jednak o nekonečných, jednak o „neusporiadaných“ bázach. Keďže v centre nášho záujmu naďalej zostávajú iba konečnorozmerné priestory, oboch týchto otázok sa len letmo dotkneme.

Hovoríme, že nekonečná postupnosť $(\boldsymbol{u}_k)_{k=0}^{\infty} = (\boldsymbol{u}_0, \boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k, \dots)$ je *báza*, presnejšie *usporiadaná báza* vektorového priestoru V, ak je lineárne nezávislá a generuje celý priestor V.

Treba zdôrazniť, že podmienka generovania priestoru V hovorí, že každý vektor $\boldsymbol{x} \in V$ možno vyjadriť ako *konečnú* lineárnu kombináciu $\boldsymbol{x} = \sum_{k=0}^{n} c_k \boldsymbol{u}_k$, kde $n \in \mathbb{N}$ a $c_0, \dots, c_n \in K$, prvkov príslušnej bázy. „Nekonečné lineárne kombinácie“ tvaru $\sum_{k=0}^{\infty} c_k \boldsymbol{u}_k$ sme zatiaľ nedefinovali a len samotná algebraická štruktúra vektorého priestoru nám to vo všeobecnosti ani neumožňuje.

Nasledujúce tvrdenie, ktorého dôkaz neuvádzame, je obdobou vety 5.3.1

5.5.1. Tvrdenie. *Postupnosť vektorov $(\boldsymbol{u}_k)_{k=0}^{\infty}$ je bázou vektorového priestoru V práve vtedy, keď každý vektor $\boldsymbol{x} \in V$ možno jednoznačne až na nulové členy vyjadriť v tvare lineárnej kombinácie*

$$\boldsymbol{x} = c_0 \boldsymbol{u}_0 + c_1 \boldsymbol{u}_1 + \ldots + c_n \boldsymbol{u}_n,$$

kde $n \in \mathbb{N}$ a $c_0, c_1, \dots, c_n \in K$.

Uvedomme si podstatnosť vsuvky „až na nulové členy“. Vzhľadom na premennú hodnotu n dĺžky príslušnej lineárnej kombinácie možno napr. vektor $\boldsymbol{u}_0 - \boldsymbol{u}_1 \in V$ písať aj v tvare $\boldsymbol{u}_0 - \boldsymbol{u}_1 + 0\boldsymbol{u}_2 + 0\boldsymbol{u}_3$ a pod.

Na druhej strane, pri danej báze $\boldsymbol{\alpha} = (\boldsymbol{u}_k)_{k=0}^{\infty}$ priestoru V každý vektor $\boldsymbol{x} \in V$ jednoznačne určuje postupnosť skalárov $(c_k)_{k=0}^{\infty} \in K^{\mathbb{N}}$ takú, že $c_k = 0$ pre všetky k až na konečný počet, t. j. $(c_k)_{k=0}^{\infty} \in K^{(\mathbb{N})}$ (pozri príklad 4.1.3 (a)), a platí

$$\boldsymbol{x} = \sum_{k=0}^{\infty} c_k \boldsymbol{u}_k.$$

Všimnite si, že takéto lineárne kombinácie obsahujú len konečne mnoho nenulových sčítancov, takže s ich definíciou nie je žiaden problém. Uvedenú postupnosť $(c_k)_{k=0}^{\infty}$ potom nazývame *súradnicami vektora* $\boldsymbol{x}$ *vzhľadom na bázu* $\boldsymbol{\alpha}$ a označujeme ju $(\boldsymbol{x})_{\boldsymbol{\alpha}}$. (Vzhľadom na to, že nemienime ďalej rozvíjať príslušnú teóriu pre nekonečnorozmerné priestory, nemá zmysel bližšie špecifikovať, či tým mienime „riadkovú“ alebo „stĺpcovú“ postupnosť $(c_k)_{k=0}^{\infty}$.)

5.5.2. Príklad. Postupnosť $\boldsymbol{\xi} = (x^n)_{n=0}^{\infty} = (1, x, x^2, \dots, x^n, \dots)$ všetkých mocnín premennej x je bázou priestoru $K[x]$ všetkých polynómov v premennej x nad poľom K. (Presvedčte sa o tom.) Súradnicami polynómu

$$f(x) = a_0 + a_1 x + \dots + a_n x^n = \sum_{i=0}^{n} a_i x^i \in K[x]$$

v tejto báze je postupnosť

$$(f)_{\boldsymbol{\xi}} = (a_0, a_1, \dots, a_n, 0, 0, \dots) \in K^{(\mathbb{N})}.$$

Podmnožinu X vektorového priestoru V nazývame *bázou*, presnejšie *neusporiadanou bázou* priestoru V, ak X je lineárne nezávislá a $[X] = V$. Používa sa tiež názov *Hamelova báza*.

Aj v prípade Hamelových báz platí obdoba tvrdení 5.3.1 a 5.5.1, čo umožňuje zaviesť na priestore V s takouto bázou súradnicové zobrazenie $V \to K^{(X)}$ (pripomíname, že $K^{(X)}$ označuje vektorový priestor všetkých zobrazení $f\colon X \to K$ takých, že $f(\boldsymbol{x}) = 0$ pre všetky $\boldsymbol{x} \in X$ až na konečný počet – pozri príklad 4.1.3). *Súradnicami vektora* $\boldsymbol{v} \in V$ *vzhľadom na bázu* X nazývame jednoznačne určené zobrazenie $f \in K^{(X)}$, pre ktoré platí

$$\boldsymbol{v} = \sum_{\boldsymbol{x} \in X} f(\boldsymbol{x})\boldsymbol{x}.$$

(Vzhľadom na konečný počet nenulových sčítancov je uvedená lineárna kombinácia dobre definovaná.) I tieto súradnice označujeme obvyklým spôsobom $(\boldsymbol{v})_X = f$.

Zostáva otázka, či aj každý nekonečnorozmerný vektorový priestor má bázu, podobne ako konečnorozmerné priestory resp. nekonečnorozmerné priestory polynómov $K[x]$. Inak povedané, radi by sme vedieť, či vôbec každý vektorový priestor má bázu. Na základe základných axióm teórie množín nemožno na túto otázku odpovedať. Až prijatie tzv. *axiómy výberu*, postulujúcej platnosť istého princípu platného pre konečné množiny aj pre nekonečné množiny, nám umožňuje dať na uvedenú otázku kladnú odpoveď. Teda za predpokladu axiómy výberu má každý vektorový priestor nad ľubovoľným poľom Hamelovu bázu. Na druhej strane pre väčšinu nekonečnorozmerných priestorov nám toto tvrdenie zaručuje skutočne len existenciu takejto bázy a nič viac. Nedáva nám nijakú konkrétnu bázu ani návod ako ju zostrojiť.

K príkladom vektorových priestorov, v ktorých nevieme nijako rozumne popísať Hamelovu bázu, hoci jej existenciu máme zaručenú, patria priestory K^X, kde X je nekonečná množina, priestor $\mathcal{C}\langle a, b\rangle$ všetkých spojitých funkcií z netriviálneho uzavretého intervalu $\langle a, b\rangle$ do množiny $\mathbb{R}$, no taktiež polia $\mathbb{R}$ či $\mathbb{C}$ uvažované ako vektorové priestory nad poľom $\mathbb{Q}$. Nie je to však až taká chyba, lebo v mnohých nekonečnorozmerných priestoroch študovaných vo funkcionálnej analýze sú užitočnejšie iné typy „báz“, umožňujúce vyjadrovať vektory z priestoru napr. v tvare istých „nekonečných lineárnych kombinácií“ $\sum_{n=0}^{\infty} c_n \boldsymbol{x}_n$ prvkov „bázy“.

5.6 Fyzika v n-rozmernom priestore*

Na záver kapitoly si dovolíme jedno odbočenie od hlavnej témy. Keď sa už toľko bavíme o dimenzii, môžeme spolu trochu porozmýšľať, ako sa trojrozmernosť „nášho“ priestoru prejavuje v matematickej podobe niektorých fyzikálnych zákonov. Na základe toho sa pokúsime o extrapoláciu týchto zákonov za hranice trojrozmerného priestoru. Inak povedané, podnikneme spolu metafyzikálny (nie metafyzický) myšlienkový experiment, v ktorom sa pokúsime trochu pošpekulovať nad otázkou, ako by asi mohla vyzerať „fyzika v n-rozmernom priestore“. Samozrejme, nie je jasné, či by pre $n \neq 3$ v n-rozmernom priestore mohli existovať vôbec nejakí „fyzici“, t. j. či by tú „fyziku“ mal kto pestovať. Touto otázkou sa však zaoberať nebudeme, hoci naše úvahy nám aj na ňu naznačia istú odpoveď. Ale nebudeme predbiehať.

Ak sa len trochu hlbšie zamyslíme nad charakterom priestoru, do ktorého sme nevdojak vrhnutí, uvedomíme si, že je plný záhad. Je konečný (ohraničený) alebo nekonečný (neohraničený)? Je diskrétny (pozostávajúci z akýchsi najmenších, ďalej už nedeliteľných častí) alebo spojitý (súvislý a donekonečna deliteľný)? Keďže skúsenosť nám na tieto otázky nedáva jednoznačnú odpoveď, filozofi sa oddávna pokúšali zodpovedať ich na základe špekulatívnych úvah. Aktuálne nekonečno, či už smerom do diaľky (t. j. smerom k čoraz väčším rozmerom) alebo smerom do hĺbky (t. j. smerom k čoraz menším rozmerom) sa však vymyká našim predstavám. Rovnako problematická je však predstava ohraničeného priestoru ako i predstava akejsi najmenšej, ďalej už nedeliteľnej priestorovej oblasti. Priestor si totiž nepredstavujeme ako súcno, t. j. ako „niečo“, ale ako prázdnu formu, naplnenú súcnami. Za hranicou, ohraničujúcou „celý priestor“, by už nemohlo byť absolútne nič, čo si však nedokážeme predstaviť inak, ako prázdny priestor. Podobne, akákoľvek malá priestorová oblasť, je aspoň myšlienkovo (hoc nie nutne fyzikálne) ďalej deliteľná na menšie časti.

Moderná fyzika sa s podobnými otázkami nevysporadúva nijakou definitívnou odpoveďou. Namiesto toho konštruuje rôzne matematické modely a na ich základe získava predpovede, ktoré možno porovnať s výsledkami experimentov. Tým sa tieto modely čiastočne potvrdzujú alebo falzifikujú. Navyše hypotéza zakriveného priestoru oddeľuje otázky (ne)konečnosti a (ne)ohraničenosti. Zakrivený priestor môže byť (sám v sebe) neohraničený a pritom mať konečný objem. Ale tak, ako zakrivená guľová plocha poukazuje na existenciu trojrozmerného (nezakriveného) priestoru, zakrivený konečne veľký trojrozmerný priestor vyvoláva otázku existencie nejakého viacrozmerného, neohraničeného a nezakriveného priestoru.

My sa však na tomto mieste nemienime zaoberať otázkou konečnosti či nekonečnosti priestoru, či už smerom k čoraz väčším alebo čoraz menším vzdialenostiam. Svoju pozornosť upriamime na omnoho tvrdšiu hranicu priestoru, ktorú predstavuje jeho trojrozmernosť. Na túto hranicu narazíme, keď sa pokúsime uskutočniť štyri rôzne, navzájom kolmé úsečky, vychádzajúce z jedného bodu. Priestor nám také niečo nedovolí. Pritom existencia takýchto úsečiek nevedie nevyhnutne k sporu, ich uskutočneniu nebránia nijaké logické zákony, ale len a len priestor. Aby sme si uvedomili rozdiel medzi priestorovou nepredstaviteľnosťou a logickou nemožnosťou, pokúsime sa vmyslieť do postavenia akýchsi plochých bytostí, obývajúcich dvojrozmerný priestor, t. j. rovinu. V rovine možno uskutočniť len dve rôzne navzájom kolmé úsečky vychádzajúce z daného bodu. Naši „dvojrozmerní ľudkovia“ by si zrejme nevedeli predstaviť tri takéto úsečky, čo pre nás nepredstavuje nijaký problém. Podobne nejaké bytosti, obývajúce n-rozmerný priestor, kde $n \geq 3$, by si asi vedeli predstaviť n navzájom kolmých úsečiek. Teda ich existencia je *logicky možná*.

Vráťme sa však k pôvodnej otázke: ako sa prejavuje trojrozmernosť nášho priestoru v matematickej podobe niektorých fyzikálnych zákonov a akú „fyziku“ by asi objavili „fyzici“ v n-rozmernom priestore.

Samozrejme, nebudeme sa zaoberať uvedenými otázkami v celej ich šírke, len sa pokúsime ilustrovať naznačenú problematiku na príklade Newtonovho gravitačného zákona. Úplne analogicky by sme mohli postupovať i v prípade Coulombovho zákona pre elektrostatickú silu.

Podľa Newtonovho gravitačného zákona centrálne symetrické teleso o hmotnosti M vytvára okolo seba centrálne symetrické gravitačné pole, ktoré na hmotný bod o hmotnosti m vo vzdialenosti r od stredu telesa pôsobí silou

$$F = \varkappa \frac{mM}{r^2},$$

kde $\varkappa \approx 6,672 \cdot 10^{-11}\,\mathrm{kg}^{-1}\,\mathrm{m}^3\,\mathrm{s}^{-2}$ je gravitačná konštanta, ktorej hodnotu možno stanoviť experimentálne. Gravitačná hmotnosť je priamo definovaná ako miera gravitačného účinku telesa, čo vyjadruje priama úmernosť uvedenej sily hmotnostiam oboch telies. Z centrálnej symetrie gravitačného poľa, ktorá je dôsledkom izotropie (homogenity) priestoru, vyplýva, že uvedená sila závisí len od vzájomnej vzdialenosti oboch telies a nie od ďalších parametrov ich vzájomnej polohy, napr. od smeru. Navyše je rozumné predpokladať, že gravitačná sila bude slabnúť so vzdialenosťou r. Na prvý pohľad však nie je jasné, prečo by mala slabnúť akurát nepriamo úmerne jej druhej mocnine. Ukážeme si, že práve to je dôsledkom trojrozmernosti priestoru. Rovnako oprávnene však možno tvrdiť, že trojrozmernosť priestoru je dôsledkom príslušnej podoby v ňom platného gravitačného zákona.

Gravitačné pole si znázorňujeme geometricky pomocou kriviek nazývaných *siločiary*. Tie majú v prípade centrálne symetrického poľa v izotropnom priestore tvar polpriamok vychádzajúcich zo stredu zdroja. Veľkosť príťažlivej sily pôsobiacej na hmotný bod je (okrem jeho hmotnosti) priamo úmerná hustote týchto siločiar v danom mieste. Keďže na povrchu guľovej plochy s polomerom r, opísanej okolo stredu príťažlivosti je hustota siločiar všade rovnaká, táto hustota klesá so vzdialenosťou r nepriamo úmerne plošnému obsahu povrchu danej guľovej plochy. Tento obsah má hodnotu $4\pi r^2$. To znamená, že veľkosť príťažlivej sily F je nepriamo úmerná druhej mocnine vzdialenosti r.

Pod $(n-1)$-rozmernou sférou rozumieme povrch n-rozmernej gule v n-rozmernom priestore. Ak si jej stred zvolíme za počiatok súradnej sústavy, tak $(n-1)$-rozmernú sféru s polomerom r možno stotožniť s množinou

$$S^{(n-1)}(r) = \{(x_1, \ldots, x_n) \in \mathbb{R}^n;\ x_1^2 + \ldots + x_n^2 = r^2\}.$$

Veľkosť $(n-1)$-rozmerného povrchu sféry $S^{(n-1)}(r)$ je priamo úmerná mocnine r^{n-1}. Rovnakou úvahou ako v predchádzajúcom odstavci tak možno odvodiť nasledujúci tvar Newtonovho gravitačného zákona v n-rozmernom priestore:

$$F = \varkappa_n \frac{mM}{r^{n-1}},$$

kde $\varkappa_n$ je gravitačná konštanta, M je hmotnosť centrálne symetrického telesa vytvárajúceho príslušné gravitačné pole, m je hmotnosť hmotného bodu a r jeho vzdialenosť od stredu príťažlivosti. Špeciálne si uvedomme, že v jednorozmernom priestore, t. j. na priamke, gravitačná sila nezávisí na vzdialenosti (siločiary sa nemajú kam rozptýliť, ich hustota sa so vzdialenosťou nemení).

Práca, ktorú je potrebné vynaložiť na premiestnenie hmotného bodu s hmotnosťou m zo vzdialenosti $r_1 > 0$ do vzdialenosti $r_2 > r_1$ od stredu príťažlivosti, je daná integrálom

$$A = \int_{r_1}^{r_2} F \, \mathrm{d}r = \varkappa_n mM \int_{r_1}^{r_2} r^{1-n} \, \mathrm{d}r.$$

Pre jednotlivé hodnoty n dostávame

$$\begin{aligned} A &= \varkappa_1 mM(r_2 - r_1), && \text{ak } n = 1, \\ A &= \varkappa_2 mM \ln \frac{r_2}{r_1}, && \text{ak } n = 2, \\ A &= \varkappa_n \frac{mM}{n-2}\left(\frac{1}{r_1^{n-2}} - \frac{1}{r_2^{n-2}}\right), && \text{ak } n \geq 3. \end{aligned}$$

Ďalšou analýzou uvedených vzťahov (čo už nebudeme robiť) možno zistiť, že pohyb v jedno- a dvojrozmernom priestore, t. j. na priamke a v rovine, by bol nesmierne energeticky náročný. Vymaniť sa z gravitačného poľa daného telesa ($r_2 \to \infty$) by si vyžiadalo nekonečne veľkú energiu. Navyše v rovine je v takomto poli možný len pohyb po kruhových uzavretých orbitách, alebo po neuzavretých orbitách (tvaru ružice), pričom oba typy sú stabilné. Uzavretá dráha prechádzajúca v rôznych vzdialenostiach od stredu príťažlivosti nie je možná. (Prípadom $n = 1$ sa ani nemusíme zaoberať, lebo na priamke jednoducho „nie je dosť miesta“ na pohyb bodu po uzavretej orbite „okolo“ iného bodu – zrážka by bola nevyhnutná.) Gravitačné pôsobenie v n-rozmernom priestore je pre $n \geq 4$ zasa ďaleko od zdroja také slabé a blízko zdroja také silné, že iné uzavreté orbity ako kruhové nie sú možné, a i tie sú nestabilné. Aj tá najmenšia odchýlka od kruhovej dráhy (zapríčinená napr. pôsobením ďalších planét) by spôsobila zmenu kruhovej dráhy na špirálovú (napr. únik Zeme od Slnka alebo pád naň). Nestabilita pre $n = 4$ má špeciálny charakter: za veľmi idealizovaných predpokladov si možno predstaviť prechod z jednej kruhovej orbity na inú v dôsledku dvoch po sebe nasledujúcich presne zladených „drgnutí“ v opačných smeroch (pri ktorých sa zachová energia a moment hybnosti). Každopádne však stabilné kruhové orbity sú možné len v dimenziách 2 a 3 a stabilné eliptické orbity len v dimenzii 3.

K podobným efektom by dochádzalo aj pôsobením elektrostatickej sily, pod vplyvom ktorej sa elektróny pohybujú okolo jadra atómu. Dvojrozmerný atóm by bol natoľko stabilný, že by sa mohol len veľmi ťažko ionizovať, teda v dvojrozmernom priestore by asi nemohlo dochádzať ku vzniku chemických zlúčenín. Vo viac než trojrozmernom priestore by zas nemohli existovať stabilné atómy – pri najmenšej odchýlke by elektrón po špirálovej dráhe z atómu unikol alebo spadol na jadro. Jemnejšia analýza kvantovomechanických javov ukazuje, že pre $n \geq 5$ – aj bez pôsobenia vyvolávajúceho malú odchýlku – by elektróny v obale atómu samovoľne prechádzali na čoraz vzdialenejšie orbity, teda atóm vo viac než štvorrozmernom priestore by sa spontánne ionizoval. Prípad $n = 4$ je opäť singulárny: moment hybnosti obiehajúceho elektrónu by mohol nadobúdať len jedinú pevne stanovenú hodnotu.

Pokiaľ teda uznáme oprávnenosť vykonanej extrapolácie fyzikálnych zákonov trojrozmerného sveta aj na svety iných rozmerov (čo je zrejme najproblematickejšie miesto našich úvah), dochádzame k záveru, že tak stabilné systémy planét obiehajúcich okolo centrálnych hviezd ako aj stabilné a jednako zlučovania schopné atómy pozostávajúce z jadra a elektrónového obalu sú pravdepodobne možné len v trojrozmernom priestore.

Cvičenia

5.1. Dokážte tvrdenie 5.2.3 (*Návod*: Použite Stenitzovu vetu 5.1.1 a tvrdenie 5.2.1)

5.2. Vyberte z daných vektorov bázu vektorového priestoru V nad poľom K (ak je to možné; ak to nie je možné, zdôvodnite prečo):

(a) $K = \mathbb{R}$, $V = \mathbb{R}^3$, $\boldsymbol{x} = (2,2,3)^\mathsf{T}$, $\boldsymbol{y} = (-1,0,2)^\mathsf{T}$, $\boldsymbol{z} = (0,2,7)^\mathsf{T}$, $\boldsymbol{u} = (1,2,7)^\mathsf{T}$, $\boldsymbol{v} = (3,2,1)^\mathsf{T}$;

(b) $K = \mathbb{C}$, $V = \mathbb{C}^4$, $\boldsymbol{x} = (1,\mathrm{i},1,\mathrm{i})$, $\boldsymbol{y} = (\mathrm{i},1,\mathrm{i},1)$, $\boldsymbol{z} = (1,1,1,1)$, $\boldsymbol{u} = (1,0,0,1)$, $\boldsymbol{v} = (0,\mathrm{i},\mathrm{i},0)$;

(c) $K = \mathbb{Z}_5$, $V = \mathbb{Z}_5^3$, $\boldsymbol{x} = (0,1,2,3)^\mathsf{T}$, $\boldsymbol{y} = (1,2,3,4)^\mathsf{T}$, $\boldsymbol{z} = (2,1,0,4)^\mathsf{T}$, $\boldsymbol{u} = (1,3,0,2)^\mathsf{T}$, $\boldsymbol{v} = (3,4,0,1)^\mathsf{T}$;

(d) $K = \mathbb{Q}$, $V = \mathbb{Q}^{(3)}[x]$, $f_0(x) = 1$, $f_1(x) = 1+x$, $f_2(x) = (1+x)^2$, $f_3(x) = (1+x)^3$.

5.3. Doplňte uvedené vektory do bázy vektorového priestoru V nad poľom K (ak je to možné; ak to nie je možné, zdôvodnite prečo):

(a) $K = \mathbb{R}$, $V = \mathbb{R}^{(2)}[x]$, $g(x) = 1 + 2x + 7x^2$, $h(x) = 1 + x$;

(b) $K = \mathbb{C}$, $V = \mathbb{C}^4$, $\boldsymbol{u} = (1+\mathrm{i}, 1-\mathrm{i}, 2, 2\mathrm{i})^\mathsf{T}$, $\boldsymbol{v} = (1+3\mathrm{i}, 3-\mathrm{i}, 4+2\mathrm{i}, -2+4\mathrm{i})^\mathsf{T}$;

(c) $K = \mathbb{Z}_7$, $V = \mathbb{Z}_7^{(3)}[x]$, $f_0(x) = 5 + 6x + 5x^2 + 6x^3$, $f_1(x) = 6 + 5x + 6x^2 + 5x^3$;

(d) $K = \mathbb{Z}_{11}$, $V = \mathbb{Z}_{11}^{(3)}[x]$, $f_0(x) = 5 + 6x + 5x^2 + 6x^3$, $f_1(x) = 6 + 5x + 6x^2 + 5x^3$.

5.4. V každej z úloh cvičení 5.2 a 5.3 určte dimenziu lineárneho podpriestoru generovaného všetkými danými vektormi.

5.5. Dokážte tvrdenie 5.3.2

5.6. Podrobne dokážte, že $\boldsymbol{\varepsilon}^{(n)} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ je báza vektorového priestoru K^n.

5.7. Doplňte vynechané podrobnosti v príklade 5.3.5

5.8. Dokážte, že uvedená konečná postupnosť vektorov $\boldsymbol{\beta}$ tvorí bázu vektorového priestoru V nad poľom K a nájdite súradnice vektorov $\boldsymbol{x}$, $\boldsymbol{y}$ v tejto báze.

(a) $K = \mathbb{R}$, $V = \mathbb{R}^3$, $\boldsymbol{\beta} = \big((1,2,3)^\mathsf{T}, (1,-1,1)^\mathsf{T}, (2,1,0)^\mathsf{T}\big)$, $\boldsymbol{x} = (1,1,1)^\mathsf{T}$, $\boldsymbol{y} = (0,1,-2)^\mathsf{T}$;

(b) $K = \mathbb{C}$, $V = \mathbb{C}^{(2)}[z]$, $\boldsymbol{\beta} = (1+\mathrm{i}, 1+\mathrm{i}z, \mathrm{i}-z^2)$, $\boldsymbol{x} = f(z) = z$, $\boldsymbol{y} = g(z) = 1+z^2$;

(c) $K = \mathbb{Z}_2$, $V = \mathbb{Z}_2^4$, $\boldsymbol{\beta} = \big((1,1,0,0)^{\mathsf{T}}, (0,0,1,1)^{\mathsf{T}}, (1,0,0,1)^{\mathsf{T}}, (0,1,1,1)^{\mathsf{T}}\big)$, $\boldsymbol{x} = (1,0,0,0)^{\mathsf{T}}$, $\boldsymbol{y} = (1,1,1,0)^{\mathsf{T}}$;

(d) $K = \mathbb{R}$, $V = \mathbb{C}$, $\boldsymbol{\beta} = (1+\mathrm{i}, 1-\mathrm{i})$, $\boldsymbol{x} = 1$, $\boldsymbol{y} = \mathrm{i}$.

5.9. Dokážte, že vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{m-k}, \boldsymbol{w}_1, \dots, \boldsymbol{w}_{n-k}$ z dôkazu vety 5.4.1 naozaj generujú lineárny podpriestor $S + T$.

5.10. Zovšeobecnite dôsledok 5.4.2 na súčet ľubovoľného konečného počtu lineárnych podpriestorov a dokážte toto zovšeobecnenie.

5.11. Doplňte vynechané podrobnosti v dôkaze tvrdenia 5.4.3

5.12. Nech $\boldsymbol{A} \in K^{m \times n}$ je matica v stupňovitom tvare. Dokážte, že

(a) jej nenulové riadky tvoria bázu lineárneho podpriestoru $[\boldsymbol{r}_1(\boldsymbol{A}), \dots, \boldsymbol{r}_m(\boldsymbol{A})] \subseteq K^{1 \times n}$;

(b) jej stĺpce, v ktorých ležia vedúce prvky jej riadkov, tvoria bázu lineárneho podpriestoru

$$[\boldsymbol{s}_1(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A})] \subseteq K^{m \times 1}.$$

(c) Odvoďte z (a) a (b), že pre matice v stupňovitom tvare platí

$$\dim[\boldsymbol{r}_1(\boldsymbol{A}), \dots, \boldsymbol{r}_m(\boldsymbol{A})] = \dim[\boldsymbol{s}_1(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A})].$$

5.13. Nech $\boldsymbol{A}, \boldsymbol{B} \in K^{m \times n}$. Dokážte postupne nasledujúce tvrdenia:

(a) $\boldsymbol{A} \sim \boldsymbol{B} \;\Rightarrow\; [\boldsymbol{r}_1(\boldsymbol{A}), \dots, \boldsymbol{r}_m(\boldsymbol{A})] = [\boldsymbol{r}_1(\boldsymbol{B}), \dots, \boldsymbol{r}_m(\boldsymbol{B})]$.

(b) Ak $\boldsymbol{A} \sim \boldsymbol{B}$ a $\boldsymbol{B}$ je navyše v stupňovitom tvare, tak jej nenulové riadky tvoria bázu lineárneho podpriestoru $[\boldsymbol{r}_1(\boldsymbol{A}), \dots, \boldsymbol{r}_m(\boldsymbol{A})] \subseteq K^{1 \times n}$.

(c) Na základe (b) sformulujte postup, ako možno úpravou vhodnej matice pomocou ERO nájsť k daným (riadkovým) vektorom $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m \in K^n$ nejakú bázu ich lineárneho obalu $[\boldsymbol{x}_1, \dots, \boldsymbol{x}_m]$. Táto báza nie je spravidla (až na veľmi špeciálne prípady) vybraná z vektorov $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$, na druhej strane však možno úpravou príslušnej matice na redukovaný stupňovitý tvar dosiahnuť veľmi jednoduchý a prehľadný tvar tejto bázy.

(d) Riešte analogickú úlohu ako v (c) pre stĺpcové vektory.

5.14. S využitím cvičenia 5.13(b) nanovo dokážte jednoznačnosť redukovaného stupňovitého tvaru danej matice, t. j. pre ľubovoľné matice $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{K}^{m \times n}$ v redukovanom stupňovitom tvare platí $\boldsymbol{A} \sim \boldsymbol{B} \Rightarrow \boldsymbol{A} = \boldsymbol{B}$ (porovnaj s cvičením 3.13) (*Návod*: Uvedomte si, že sa stačí obmedziť na matice s nenulovými riadkami, a ďalej postupujte indukciou podľa počtu riadkov m.)

5.15. S použitím cvičenia 5.14 dokážte zosilnenie tvrdenia z cvičenia 5.13(a) do podoby ekvivalencie, t. j. pre ľubovoľné $\boldsymbol{A}, \boldsymbol{B} \in K^{m \times n}$ platí $\boldsymbol{A} \sim \boldsymbol{B} \;\Leftrightarrow\; [\boldsymbol{r}_1(\boldsymbol{A}), \dots, \boldsymbol{r}_m(\boldsymbol{A})] = [\boldsymbol{r}_1(\boldsymbol{B}), \dots, \boldsymbol{r}_m(\boldsymbol{B})]$.

5.16. S využitím výsledkov cvičenia 5.13 nájdite pre uvedené vektory z vektorového priestoru V nad poľom K „čo najjednoduchšiu“ bázu ich lineárneho obalu a doplňte ju (ak treba) do „čo najjednoduchšej“ bázy celého priestoru V:

(a) $K = \mathbb{R}$, $V = \mathbb{R}^4$, $\boldsymbol{x} = (2,0,0,3)$, $\boldsymbol{y} = (4,-1,4,0)$, $\boldsymbol{z} = (2,-1,4,3)$, $\boldsymbol{u} =$

$(-2, 2, -8, -9)$;
(b) $K = \mathbb{R}$, $V = \mathbb{R}^4$, $\boldsymbol{x} = (0, 0, 2, -1)^\mathsf{T}$, $\boldsymbol{y} = (3, -1, 2, 0)^\mathsf{T}$, $\boldsymbol{z} = (-3, 1, 2, -2)^\mathsf{T}$, $\boldsymbol{u} = (2, -1, 1, -2)^\mathsf{T}$;
(c) $K = \mathbb{C}$, $V = \mathbb{C}^3$, $\boldsymbol{x} = (\mathrm{i}, 1, 1 + \mathrm{i})$, $\boldsymbol{y} = (1 + \mathrm{i}, 1 - \mathrm{i}, 2)$, $\boldsymbol{z} = (2, -\mathrm{i}, 3 - 2\mathrm{i})^\mathsf{T}$;
(d) $K = \mathbb{R}$, $V = \mathbb{R}^{(3)}[x]$, $f(x) = 2 + x + x^2 - x^3$, $g(x) = 2x^2 - x^3$, $h(x) = 1 - x + 2x^2$;
(e) $K = \mathbb{Z}_5$, $V = \mathbb{Z}_5^3$, $\boldsymbol{x} = (2, 4, 3)$, $\boldsymbol{y} = (0, 1, 2)$, $\boldsymbol{z} = (4, 0, 0)$;
(f) $K = \mathbb{Z}_7$, $V = \mathbb{Z}_7^{(2)}[x]$, $f(x) = 2 + 4x + 3x^2$, $g(x) = x + 2x^2$, $h(x) = 4$.

5.17. Nech $q \neq \pm 1$ je ľubovoľné reálne číslo. Pre $k, n \in \mathbb{N}$, $k \leq n$, definujme *q-binomický koeficient* ako výraz

$$\binom{n}{k}_q = \frac{(q^n - 1)(q^{n-1} - 1) \dots (q^{n-k+1} - 1)}{(q^k - 1)(q^{k-1} - 1) \dots (q - 1)}.$$

Špeciálne pre $k = 0$ sa tým myslí $\binom{n}{0}_q = 1$. Pre $k > n$ navyše kladieme $\binom{n}{k}_q = 0$. Potom pre ľubovoľné $k \leq n$ platí:

(a) $\binom{n}{k}_q = \binom{n}{n-k}_q$, (b) $\binom{n}{k} = \lim\limits_{q \to 1} \binom{n}{k}_q$, (c) $\binom{n}{0}_q = \binom{n}{n}_q = 1$,

(d) $\binom{n+1}{k}_q = \binom{n}{k-1}_q + q^k \binom{n}{k}_q = q^{n-k+1} \binom{n}{k-1}_q + \binom{n}{k}_q$, ak $1 \leq k \leq n$.

Dokážte. Rovnosti (c) a (d) sa nazývajú pravidlami *q-Pascalovho trojuholníka* pre q-binomické koeficienty (porovnaj s cvičením 0.18).

5.18. Nech pole K je konečné a má práve q prvkov. Pre $k, n \in \mathbb{N}$ označme $C_q(n, k)$ počet všetkých k-rozmerných lineárnych podpriestorov vektorového priestoru K^n. Dokážte postupne nasledujúce tvrdenia:
(a) Pre ľubovoľné n platí $C_q(n, 0) = C_q(n, n) = 1$ a $C_q(n, k) = 0$ pre $k > n$.

(b) $C_q(n, k)$ sa rovná počtu všetkých matíc $\boldsymbol{A} \in K^{k \times n}$ v redukovanom stupňovitom tvare, ktoré majú všetky riadky nenulové. (*Návod*: Uvažujte K^n ako priestor riadkových vektorov a na základe cvičení 5.13 a 5.14 reprezentujte každý jeho k-rozmerný lineárny podpriestor jednoznačne určenou bázou, ktorej vektory, zapísané ako riadky pod sebou, tvoria maticu v redukovanom stupňovitom tvare.)

(c) Pre $1 \leq k \leq n$ platí $C_q(n + 1, k) = C_q(n, k - 1) + q^k C_q(n, k)$. To spolu s (a) zabezpečuje, že čísla $C_q(n, k)$ vyhovujú rovnakým podmienkam q-Pascalovho trojuholníka ako q-binomické koeficienty $\binom{n}{k}_q$. (*Návod*: Využite (b); matice $\boldsymbol{A} \in K^{k \times (n+1)}$ v redukovanom stupňovitom tvare s nenulovými riadkami rozdeľte do dvoch skupín podľa toho, či vedúci prvok posledného riadku leží alebo neleží v poslednom stĺpci – ukážte, že prvých je $C_q(n, k - 1)$ a druhých $q^k C_q(n, k)$.)

(d) Odvoďte z (a) a (c) rovnosť $C_q(n, k) = \binom{n}{k}_q$ pre všetky $k, n \in \mathbb{N}$.

(e) Koľko k-rozmerných lineárnych podpriestorov majú vektorové priestory $\mathbb{Z}_p^n$ nad poľom $\mathbb{Z}_p$ pre $p = 2, 3, 5, 7$, $n = 0, 1, 2, 3, 4, 5$, $0 \leq k \leq n$? Zostrojte počiatočné úseky príslušných p-Pascalových trojuholníkov.

5.19. Nech K je pole a $(p_k(x))_{k=0}^{\infty}$ je postupnosť polynómov z $K[x]$ taká, že stupeň polynómu $p_k(x)$ je práve k. Dokážte, že postupnosť $(p_k(x))$ je bázou vektorového priestoru $K[x]$ (využite cvičenie 4.11).

6. Lineárne zobrazenia

Zatiaľ sme sa pri štúdiu lineárnej algebry sústredili zakaždým na štruktúru jedného, izolovaného vektorového priestoru. Doteraz sme si nevybudovali pojmy, ktoré by nám umožnili štúdium vzťahov medzi viacerými vektorovými priestormi. V tejto kapitole hodláme zaplniť túto medzeru. Zavedieme a bližšie preskúmame pojem *lineárneho zobrazenia*, ktorý nám umožní porovnávať štruktúry rôznych vektorových priestorov nad tým istým pevne zvoleným poľom. Voľne povedané, pôjde o zobrazenia medzi vektorovými priestormi, ktoré zachovávajú ich lineárnu štruktúru.

6.1 Lineárne zobrazenia

Nech U, V sú vektorové priestory nad tým istým poľom K. Hovoríme, že $\varphi\colon V \to U$ je *lineárne zobrazenie*, ak φ zachováva operácie vektorového súčtu a skalárneho násobku, t. j. ak pre ľubovoľné $\boldsymbol{x}, \boldsymbol{y} \in V$, $c \in K$ platí

$$\begin{aligned}\varphi(\boldsymbol{x}+\boldsymbol{y}) &= \varphi(\boldsymbol{x}) + \varphi(\boldsymbol{y}),\\ \varphi(c\boldsymbol{x}) &= c\varphi(\boldsymbol{x}).\end{aligned}$$

Prvú z uvedených vlastností niekedy nazývame *aditivitou* a druhú *homogenitou.* Ako cvičenie si dokážte, že lineárne zobrazenia zachovávajú nulu a opačné vektory, t. j. pre lineárne zobrazenie $\varphi\colon V \to U$ a $\boldsymbol{x} \in V$ platí

$$\varphi(\mathbf{0}) = \mathbf{0}, \qquad \varphi(-\boldsymbol{x}) = -\varphi(\boldsymbol{x}).$$

Zrejme pre každý vektorový priestor V identita $\mathrm{id}_V\colon V \to V$ je lineárne zobrazenie. Taktiež pre ľubovoľné vektorové priestory U, V nad poľom K zobrazenie $\mathbf{0}\colon V \to U$, ktoré každému vektoru $\boldsymbol{x} \in V$ priradí nulový vektor $\mathbf{0} \in U$, je lineárne. Komutatívnosť operácie súčinu v poli a jeho distributívnosť vzhľadom na sčítanie znamená, že pre ľubovoľný pevný skalár $a \in K$ je priradením $x \mapsto ax$ definované lineárne zobrazenie $K \to K$. Čoskoro sa zoznámime aj s menej triviálnymi príkladmi lineárnych zobrazení.

No zatiaľ si ešte všimnime jednu odlišnosť v použití názvu „lineárne zobrazenie“ v lineárnej algebre oproti matematickej analýze, kde sa pod lineárnou funkciou $\mathbb{R} \to \mathbb{R}$ rozumie ľubovoľná funkcia tvaru $f(x) = ax + b$, kde $a, b \in \mathbb{R}$. Ľahko sa možno presvedčiť, že takéto $f\colon \mathbb{R} \to \mathbb{R}$ je lineárne zobrazenie v zmysle našej definície práve vtedy, keď $b = 0$. Neskôr zavedieme širšiu triedu zobrazení medzi vektorovými priestormi, ktorá zahŕňa aj takéto „v zmysle matematickej analýzy lineárne“ zobrazenia.

Lineárne zobrazenia možno charakterizovať ako zobrazenia medzi vektorovými priestormi (nad tým istým poľom), ktoré zachovávajú lineárne kombinácie. Jednoduchý dôkaz tohto pozorovania prenechávame čitateľovi. (*Návod*: Pozrite sa na dôkaz tvrdenia 4.1.2)

6.1.1. Tvrdenie. *Nech U, V sú vektorové priestory nad poľom K a $\varphi\colon V \to U$ je ľubovoľné zobrazenie. Nasledujúce podmienky sú ekvivalentné:*
(i) *φ je lineárne zobrazenie;*
(ii) *pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$, $a, b \in K$ platí $\varphi(a\boldsymbol{x} + b\boldsymbol{y}) = a\varphi(\boldsymbol{x}) + b\varphi(\boldsymbol{y})$;*
(iii) *pre ľubovoľné $n \in \mathbb{N}$ a všetky $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n \in V$, $c_1, \dots, c_n \in K$ platí $\varphi(c_1\boldsymbol{x}_1 + \ldots + c_n\boldsymbol{x}_n) = c_1\varphi(\boldsymbol{x}_1) + \ldots + c_n\varphi(\boldsymbol{x}_n)$.*

Nasledujúce dve tvrdenia zachytávajú významné vlastnosti lineárnych zobrazení: kompozícia lineárnych zobrazení je opäť lineárne zobrazenie a obrazy i vzory lineárnych podpriestorov v lineárnych zobrazeniach sú tiež lineárnymi podpriestormi.

6.1.2. Tvrdenie. *Nech U, V, W sú vektorové priestory nad poľom K a $\psi\colon W \to V$, $\varphi\colon V \to U$ sú lineárne zobrazenia. Potom aj ich kompozícia $\varphi \circ \psi\colon W \to U$ je lineárne zobrazenie.*

Dôkaz. Overíme, že i zložené zobrazenie $\varphi \circ \psi$ zachováva lineárne kombinácie. Nech $\boldsymbol{x}, \boldsymbol{y} \in W$, $a, b \in K$. S využitím linearity zobrazení ψ a φ postupne dostávame

$$\begin{aligned}(\varphi \circ \psi)(a\boldsymbol{x} + b\boldsymbol{y}) &= \varphi(\psi(a\boldsymbol{x} + b\boldsymbol{y})) = \varphi(a\psi\boldsymbol{x} + b\psi\boldsymbol{y}) \\ &= a\varphi(\psi\boldsymbol{x}) + b\varphi(\psi\boldsymbol{y}) = a(\varphi \circ \psi)(\boldsymbol{x}) + b(\varphi \circ \psi)(\boldsymbol{y}).\end{aligned}$$

Podľa tvrdenia 6.1.1 to znamená, že zobrazenie $\varphi \circ \psi$ je lineárne.

6.1.3. Tvrdenie. *Nech U, V sú vektorové priestory nad poľom K a $\varphi\colon V \to U$ je lineárne zobrazenie.*
(a) *Ak S je lineárny podpriestor priestoru V, tak $\varphi(S)$ je lineárny podpriestor priestoru U.*
(b) *Ak T je lineárny podpriestor priestoru U, tak $\varphi^{-1}(T)$ je lineárny podpriestor priestoru V.*

Dôkaz. (a) Keďže $\mathbf{0} \in S$, $\mathbf{0} = \varphi(\mathbf{0}) \in \varphi(S)$. Overíme, že obraz $\varphi(S)$ je uzavretý vzhľadom na lineárne kombinácie. Nech $a, b \in K$, $\boldsymbol{u}, \boldsymbol{v} \in \varphi(S)$. Potom existujú $\boldsymbol{x}, \boldsymbol{y} \in S$ také, že $\boldsymbol{u} = \varphi(\boldsymbol{x})$, $\boldsymbol{v} = \varphi(\boldsymbol{y})$. S využitím linearity φ dostávame:

$$a\boldsymbol{u} + b\boldsymbol{v} = a\varphi(\boldsymbol{x}) + b\varphi(\boldsymbol{y}) = \varphi(a\boldsymbol{x} + b\boldsymbol{y}) \in \varphi(S),$$

lebo $a\boldsymbol{x} + b\boldsymbol{y} \in S$, nakoľko $S \subseteq V$ je lineárny podpriestor.

(b) Keďže $\varphi(\mathbf{0}) = \mathbf{0} \in T$, $\mathbf{0} \in \varphi^{-1}(T)$. Ukážeme, že aj vzor $\varphi^{-1}(T)$ je uzavretý vzhľadom na lineárne kombinácie. Zvoľme $a, b \in K$, $\boldsymbol{x}, \boldsymbol{y} \in \varphi^{-1}(T)$. To znamená, že $\varphi(\boldsymbol{x}), \varphi(\boldsymbol{y}) \in T$. Z linearity φ vyplýva

$$\varphi(a\boldsymbol{x} + b\boldsymbol{y}) = a\varphi(\boldsymbol{x}) + b\varphi(\boldsymbol{y}) \in T,$$

lebo $T \subseteq U$ je lineárny podpriestor. Preto $a\boldsymbol{x} + b\boldsymbol{y} \in \varphi^{-1}(T)$.

6.1.4. Príklad. Nech K je pole. Distributívnosť súčinu matíc vzhľadom na ich súčet a jeho zameniteľnosť s operáciou skalárneho násobku (pozri odstavec 2.2.2) vlastne hovorí, že pre pevné $m, n, p \in \mathbb{N}$ a ľubovoľnú maticu $\boldsymbol{A} \in K^{m\times n}$ je priradením $\boldsymbol{X} \mapsto \boldsymbol{A} \cdot \boldsymbol{X}$ definované lineárne zobrazenie medzi vektorovými priestormi matíc $K^{n\times p} \to K^{m\times p}$. Podobne je priradením $\boldsymbol{Y} \mapsto \boldsymbol{Y} \cdot \boldsymbol{A}$ definované lineárne zobrazenie $K^{p\times m} \to K^{p\times n}$. Špeciálne pre $p = 1$ je takto definované lineárne zobrazenie $\boldsymbol{x} \mapsto \boldsymbol{A} \cdot \boldsymbol{x}$ medzi stĺpcovými vektorovými priestormi $K^n \to K^m$, resp. lineárne zobrazenie $\boldsymbol{y} \mapsto \boldsymbol{y} \cdot \boldsymbol{A}$ medzi riadkovými vektorovými priestormi $K^m \to K^n$. Neskôr uvidíme, že každé lineárne zobrazenie medzi *konečnorozmernými* vektorovými priestormi nad K má „v podstate“ takúto podobu.

6.1.5. Príklad. Nech K je pole. Pre $m, n \in \mathbb{N}$ a pevné $1 \leq i \leq m$, $1 \leq j \leq n$ sú predpismi $\boldsymbol{A} \mapsto \boldsymbol{r}_i(\boldsymbol{A})$, $\boldsymbol{A} \mapsto \boldsymbol{s}_j(\boldsymbol{A})$ definované lineárne zobrazenia $K^{m\times n} \to K^{1\times n}$ resp. $K^{m\times n} \to K^{m\times 1}$. Takisto $\boldsymbol{A} \mapsto \boldsymbol{A}^\mathsf{T}$ je lineárne zobrazenie $K^{m\times n} \to K^{n\times m}$.

6.1.6. Príklad. Nech V je vektorový priestor nad poľom K, X je množina a $x \in X$ je pevne zvolený prvok. Pripomeňme, že V^X je vektorový priestor všetkých funkcií $f\colon X \to V$ (pozri 1.6.5) Dosadenie prvku x do funkcie f, t. j. priradenie $f \mapsto f(x)$, je lineárne zobrazenie $V^X \to V$. Podobne, pre ľubovoľnú podmnožinu $Y \subseteq X$ je zúženie $f \mapsto f \restriction Y$ lineárne zobrazenie $V^X \to V^Y$.

6.1.7. Príklad. Označme V množinu všetkých konvergentných postupností reálnych čísel. Zrejme V je lineárny podpriestor vektorového priestoru $\mathbb{R}^\mathbb{N}$ všetkých postupností reálnych čísel. Potom zobrazenie $V \to \mathbb{R}$, ktoré postupnosti $\boldsymbol{a} = (a_n)_{n=0}^\infty \in V$ priradí jej limitu $\lim_{n\to\infty} a_n$, je lineárne.

6.1.8. Príklad. (a) Nech $X \subseteq \mathbb{R}$ a V označuje množinu všetkých zobrazení $X \to \mathbb{R}$, ktoré majú v pevne zvolenom *vnútornom* bode a množiny X konečnú deriváciu. Zrejme V je lineárny podpriestor vektorového priestoru $\mathbb{R}^X$. Potom zobrazenie $V \to \mathbb{R}$, ktoré funkcii $f \in V$ priradí jej deriváciu $f'(a)$ v bode a, je lineárne.

(b) Nech $X \subseteq \mathbb{R}$ je ľubovoľný netriviálny interval. Pripomeňme, že $\mathcal{D}(X)$ označuje lineárny podpriestor vektorového priestoru $\mathbb{R}^X$, tvorený všetkými funkciami $f\colon X \to \mathbb{R}$, ktoré majú v každom bode $x \in X$ konečnú deriváciu (pozri príklad 4.1.3 (c). Potom derivácia, t. j. priradenie $f \mapsto f'$, je lineárne zobrazenie $\mathcal{D}(X) \to \mathbb{R}^X$.

6.1.9. Príklad. Pre reálne čísla $a < b$ označuje $\mathcal{C}\langle a, b\rangle$ lineárny podpriestor vektorového priestoru $\mathbb{R}^{\langle a,b\rangle}$, tvorený všetkými spojitými funkciami $f\colon \langle a, b\rangle \to \mathbb{R}$ (pozri príklad 4.1.3 (b)).

(a) Určitý integrál $f \mapsto \int_a^b f(x)\,\mathrm{d}x$ je lineárne zobrazenie $\mathcal{C}\langle a, b\rangle \to \mathbb{R}$.

(b) Podobne, na určitý integrál ako funkciu hornej medze, ktorý funkcii $f \in \mathcal{C}\langle a, b\rangle$ priradí jej primitívnu funkciu $F \in \mathcal{C}\langle a, b\rangle$ danú predpisom

$$F(x) = \int_a^x f(t)\,\mathrm{d}t, \qquad (a \le x \le b),$$

sa možno dívať ako na lineárne zobrazenie $\mathcal{C}\langle a, b\rangle \to \mathcal{C}\langle a, b\rangle$.

6.2 Jadro a obraz lineárneho zobrazenia

Nech $\varphi\colon V \to U$ je lineárne zobrazenie medzi vektorovými priestormi nad poľom K. Jeho *jadrom* nazývame množinu

$$\operatorname{Ker}\varphi = \varphi^{-1}(\mathbf{0}) = \{\boldsymbol{x} \in V;\ \varphi(\boldsymbol{x}) = \mathbf{0}\}.$$

Obrazom lineárneho zobrazenia φ nazývame, v zhode s paragrafom 0.3, množinu

$$\operatorname{Im}\varphi = \varphi(V) = \{\varphi(\boldsymbol{x});\ \boldsymbol{x} \in V\}.$$

(Označenie pochádza z anglických slov *kernel* a *image*.)

Keďže $\{\mathbf{0}\}$ je lineárny podpriestor priestoru U a V je lineárny podpriestor priestoru V, ako špeciálny prípad tvrdenia 6.1.3 dostávame nasledujúci výsledok.

6.2.1. Tvrdenie. *Nech $\varphi\colon V \to U$ je lineárne zobrazenie medzi vektorovými priestormi nad poľom K. Potom $\operatorname{Ker}\varphi$ je lineárny podpriestor priestoru V a $\operatorname{Im}\varphi$ je lineárny podpriestor priestoru U.*

Pomocou pojmov jadra a obrazu možno charakterizovať injektívne resp. surjektívne lineárne zobrazenia.

6.2.2. Veta. *Nech $\varphi\colon V \to U$ je lineárne zobrazenie. Potom*
(a) φ je injektívne práve vtedy, keď $\operatorname{Ker}\varphi = \{\mathbf{0}\}$;
(b) φ je surjektívne práve vtedy, keď $\operatorname{Im}\varphi = U$.

Dôkaz. (a) Ak φ je injektívne, tak $\mathbf{0} \in V$ je jediný prvok priestoru V, ktorý sa zobrazí na $\mathbf{0} \in U$. Teda $\operatorname{Ker}\varphi = \{\mathbf{0}\}$. Naopak, ak φ nie je injektívne, tak existujú $\boldsymbol{x}, \boldsymbol{y} \in V$ také, že $\boldsymbol{x} \neq \boldsymbol{y}$ a $\varphi(\boldsymbol{x}) = \varphi(\boldsymbol{y})$. Potom $\boldsymbol{x} - \boldsymbol{y} \neq \mathbf{0}$ a $\varphi(\boldsymbol{x}-\boldsymbol{y}) = \varphi(\boldsymbol{x})-\varphi(\boldsymbol{y}) = \mathbf{0}$, teda $\boldsymbol{x}-\boldsymbol{y} \in \operatorname{Ker}\varphi$. Inak povedané, $\operatorname{Ker}\varphi \neq \{\mathbf{0}\}$.

(b) je priamo definícia surjektívnosti.

Treba poznamenať, že zakiaľ časť (b) uvedeného tvrdenia je triviálna a platí aj bez predpokladu linearity zobrazenia φ, o časti (a) to už povedať nemožno. Pre všeobecné zobrazenie $\varphi\colon V \to U$ by sa totiž mohlo stať, že $\mathbf{0} \in V$ je jediný prvok, ktorý sa zobrazí na $\mathbf{0} \in U$, no φ aj tak nie je injektívne. Stále by totiž mohol existovať nejaký vektor $\mathbf{0} \neq \boldsymbol{u} \in U$ a dva rôzne vektory $\boldsymbol{x}, \boldsymbol{y} \in V$ také, že $\varphi(\boldsymbol{x}) = \boldsymbol{u} = \varphi(\boldsymbol{y})$. Spomínané tvrdenie teda hovorí, že lineárne zobrazenia majú značne homogénnu štruktúru, takže ich injektivitu nemusíme zisťovať „všade“ – dá sa rozpoznať už podľa množiny vzorov jediného prvku $\mathbf{0} \in U$.

6.2.3. Veta. *Nech $\varphi\colon V \to U$ je lineárne zobrazenie, pričom vektorový priestor V je konečnorozmerný. Potom aj* $\operatorname{Ker}\varphi$ *a* $\operatorname{Im}\varphi$ *sú konečnorozmerné priestory a platí*

$$\dim V = \dim \operatorname{Ker}\varphi + \dim \operatorname{Im}\varphi.$$

Dôkaz. Stačí dokázať uvedenú rovnosť pre dimenzie, konečný rozmer podpriestorov $\operatorname{Ker}\varphi$ a $\operatorname{Im}\varphi$ je už jej dôsledkom.

Označme $k = \dim \operatorname{Ker}\varphi$, $l = \dim V - k$. Zrejme $l \geq 0$. Nech $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ je nejaká báza priestoru $\operatorname{Ker}\varphi$. Doplňme ju do bázy $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_l$ priestoru V. Potom $\varphi(\boldsymbol{u}_1) = \ldots = \varphi(\boldsymbol{u}_k) = \mathbf{0}$. Dokážeme, že vektory $\varphi(\boldsymbol{v}_1), \dots, \varphi(\boldsymbol{v}_l)$ tvoria bázu priestoru $\operatorname{Im}\varphi$, z čoho už vyplýva požadovaná rovnosť.

Najprv dokážeme, že vektory $\varphi(\boldsymbol{v}_1), \dots, \varphi(\boldsymbol{v}_l)$ generujú podpriestor $\operatorname{Im}\varphi \subseteq U$. Každý vektor $\boldsymbol{w} \in \operatorname{Im}\varphi$ možno vyjadriť v tvare $\boldsymbol{w} = \varphi(\boldsymbol{x})$ pre nejaké $\boldsymbol{x} \in V$, ktoré je vhodnou lineárnou kombináciou $\boldsymbol{x} = a_1\boldsymbol{u}_1 + \ldots + a_k\boldsymbol{u}_k + b_1\boldsymbol{v}_1 + \ldots + b_l\boldsymbol{v}_l$ vektorov $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_l$ bázy priestoru V. Potom

$$\begin{aligned}\boldsymbol{w} = \varphi(\boldsymbol{x}) &= \varphi(a_1\boldsymbol{u}_1 + \ldots + a_k\boldsymbol{u}_k + b_1\boldsymbol{v}_1 + \ldots + b_l\boldsymbol{v}_l)\\ &= a_1\varphi(\boldsymbol{u}_1) + \ldots + a_k\varphi(\boldsymbol{u}_k) + b_1\varphi(\boldsymbol{v}_1) + \ldots + b_l\varphi(\boldsymbol{v}_l)\\ &= b_1\varphi(\boldsymbol{v}_1) + \ldots + b_l\varphi(\boldsymbol{v}_l),\end{aligned}$$

teda $\boldsymbol{w} \in [\varphi(\boldsymbol{v}_1), \dots, \varphi(\boldsymbol{v}_l)]$.

Zostáva dokázať lineárnu nezávislosť vektorov $\varphi(\boldsymbol{v}_1), \dots, \varphi(\boldsymbol{v}_l)$. Nech $c_1, \dots, c_l$ sú skaláry také, že $c_1\varphi(\boldsymbol{v}_1) + \ldots + c_l\varphi(\boldsymbol{v}_l) = \mathbf{0}$. Potom $\varphi(c_1\boldsymbol{v}_1 + \ldots + c_l\boldsymbol{v}_l) = \mathbf{0}$, teda $c_1\boldsymbol{v}_1 + \ldots + c_l\boldsymbol{v}_l \in \operatorname{Ker}\varphi$. Preto sa tento vektor musí dať vyjadriť ako lineárna kombinácia $c_1\boldsymbol{v}_1 + \ldots + c_l\boldsymbol{v}_l = d_1\boldsymbol{u}_1 + \ldots + d_k\boldsymbol{u}_k$ vektorov bázy $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ podpriestoru $\operatorname{Ker}\varphi$. Z lineárnej nezávislosti bázy $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_l$ priestoru V vyplýva $c_1 = \ldots = c_l = 0 = d_1 = \ldots = d_k$, teda aj nezávislosť vektorov $\varphi(\boldsymbol{v}_1), \dots, \varphi(\boldsymbol{v}_l)$.

Dimenziu obrazu $\operatorname{Im}\varphi$ nazývame *hodnosťou* lineárneho zobrazenia φ a značíme ju

$$h(\varphi) = \dim \operatorname{Im}\varphi.$$

Lineárne zobrazenie $\varphi\colon V \to V$ vektorového priestoru V do seba nazývame *lineárnym operátorom* alebo *lineárnou transformáciou.*

Ako sme spomínali v paragrafe 0.5, transformácia $f\colon X \to X$ konečnej množiny X je injektívna práve vtedy, keď je surjektívna. Ako dôsledok práve dokázanej vety dostávame analogický výsledok aj pre lineárne transformácie konečnorozmerných vektorových priestorov.

6.2.4. Dôsledok. *Nech $\varphi\colon V \to V$ je lineárna transformácia konečnorozmerného vektorového priestoru V. Potom φ je injektívna práve vtedy, keď je surjektívna.*

Dôkaz. Nech $\dim V = n$. Potom φ je injektívne práve vtedy, keď $\dim \operatorname{Ker}\varphi = 0$, a surjektívne práve vtedy, keď $\dim \operatorname{Im}\varphi = n$. Keďže $\dim \operatorname{Ker}\varphi + \dim \operatorname{Im}\varphi = n$, obe tieto podmienky sú ekvivalenté.

6.3 Lineárne izomorfizmy

Bijektívne lineárne zobrazenie $\varphi\colon V \to U$ medzi vektorovými priestormi V, U nad tým istým poľom K nazývame *lineárny izomorfizmus.* Hovoríme, že vektorové priestory V, U sú *lineárne izomorfné* alebo len krátko *izomorfné*, označenie $V \cong U$, ak existuje nejaký lineárny izomorfizmus $\varphi\colon V \to U$.

6.3.1. Tvrdenie. *Nech U, V, W sú vektorové priestory nad poľom K.*
(a) $\mathrm{id}_V\colon V \to V$ *je lineárny izomorfizmus.*
(b) Ak $\varphi\colon V \to U$ je lineárny izomorfizmus, tak aj $\varphi^{-1}\colon U \to V$ je lineárny izomorfizmus.
(c) Ak $\psi\colon W \to V$, $\varphi\colon V \to U$ sú lineárne izomorfizmy, tak aj $\varphi\circ\psi\colon W \to U$ je lineárny izomorfizmus.

Dôkaz. (a) je triviálne, (c) vyplýva z toho, že kompozícia bijekcií je bijekcia a kompozícia lineárnych zobrazení je lineárne zobrazenie. Zostáva dokázať (b). Treba overiť, že inverzné zobrazenie $\varphi^{-1}\colon U \to V$ k lineárnej bijekcii $\varphi\colon V \to U$ je tiež lineárne (jeho bijektívnosť je totiž zrejmá).

Zvoľme $\boldsymbol{u}, \boldsymbol{v} \in U$, $a, b \in K$. Máme dokázať rovnosť

$$\varphi^{-1}(a\boldsymbol{u} + b\boldsymbol{v}) = a\varphi^{-1}(\boldsymbol{u}) + b\varphi^{-1}(\boldsymbol{v}),$$

ktorá je ekvivalentná s podmienkou

$$a\boldsymbol{u} + b\boldsymbol{v} = \varphi\big(a\varphi^{-1}(\boldsymbol{u}) + b\varphi^{-1}(\boldsymbol{v})\big).$$

Vďaka linearite φ a vzťahu $\varphi \circ \varphi^{-1} = \mathrm{id}_U$ naozaj dostávame

$$\varphi\big(a\varphi^{-1}(\boldsymbol{u}) + b\varphi^{-1}(\boldsymbol{v})\big) = a\varphi\varphi^{-1}(\boldsymbol{u}) + b\varphi\varphi^{-1}(\boldsymbol{v}) = a\boldsymbol{u} + b\boldsymbol{v}.$$

Z práve dokázaného tvrdenia okamžite vyplýva nasledujúci dôsledok.

6.3.2. Dôsledok. *Pre vektorové priestory U, V, W nad tým istým poľom platí:*
(a) $V \cong V$;
(b) $V \cong U \;\Rightarrow\; U \cong V$;
(c) $W \cong V \;\&\; V \cong U \;\Rightarrow\; W \cong U$.

Hovoríme, že vzťah izomorfnosti $\cong$ je *reflexívny*, *symetrický* a *tranzitívny*, t. j. je vzťahom *ekvivalencie*. Z formálneho hľadiska s ním teda môžeme narábať podobne ako so vzťahom rovnosti $=$.

Izomorfné vektorové priestory majú rovnakú štruktúru, líšia sa nanajvýš označením svojich prvkov, nie však vzťahmi medzi nimi. Preto ich možno v prípade potreby stotožniť, či nahradiť jeden vektorový priestor jeho izomorfnou kópiou. Z toho dôvodu je dôležité mať k dispozícii vhodnú triedu vektorových priestorov nad daným poľom, ktorá by pre každý vektorový priestor obsahovala nejaký priestor s ním izomorfný.

6.3.3. Príklad. (a) Nech V je konečnorozmerný vektorový priestor nad poľom K, $\dim V = n$ a $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ je nejaká jeho báza. Potom tvrdenie 5.3.2 vlastne hovorí, že súradnicové zobrazenie $\boldsymbol{x} \mapsto (\boldsymbol{x})_{\boldsymbol{\beta}}$ je lineárny izomorfizmus $V \to K^n$.

(b) Podobne možno nahliadnuť, že (i v nekonečnorozmernom prípade) určuje Hamelova báza X vektorového priestoru V súradnicové zobrazenie $\boldsymbol{v} \mapsto (\boldsymbol{v})_X$, ktoré je lineárnym izomorfizmom $V \to K^{(X)}$ (pozri záver paragrafu 5.5).

Na záver tohto paragrafu ešte ukážeme, že typ izomorfizmu daného konečnorozmerného priestoru je jednoznačne určený jeho dimenziou.

6.3.4. Veta. *Nech U, V sú konečnorozmerné vektorové priestory nad poľom K. Potom*

$$V \cong U \;\Leftrightarrow\; \dim V = \dim U.$$

Dôkaz. Nech $V \cong U$ a $\varphi\colon V \to U$ je lineárny izomorfizmus. Potom $\operatorname{Ker}\varphi = \{\mathbf{0}\}$ a $\operatorname{Im}\varphi = U$ Podľa vety 6.2.3 o dimenzii jadra a obrazu

$$\dim V = \dim \operatorname{Ker}\varphi + \dim \operatorname{Im}\varphi = 0 + \dim U = \dim U.$$

Naopak, nech $\dim V = \dim U = n$. Podľa príkladu 6.3.3 (a) platí $V \cong K^n \cong U$.

Teda konečnorozmerný vektorový priestor V nad poľom K je izomorfný so stĺpcovým (no rovnako aj s riadkovým) vektorovým priestorom K^n práve vtedy, keď $n = \dim V$. Pritom každá báza $\boldsymbol{\beta}$ priestoru V určuje jeden takýto izomorfizmus $V \to K^n$ – je ním súradnicové zobrazenie $\boldsymbol{x} \mapsto (\boldsymbol{x})_{\boldsymbol{\beta}}$.

6.4 Matica lineárneho zobrazenia

Uvažujme nejaké lineárne zobrazenie $\varphi\colon K^n \to K^m$. V priestore K^n máme kanonickú bázu $\boldsymbol{\varepsilon}^{(n)} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$. Keďže obrazy $\varphi(\boldsymbol{e}_j)$ vektorov tejto bázy sú stĺpcové vektory z priestoru K^m, môžeme vytvoriť maticu

$$\boldsymbol{A} = (\varphi(\boldsymbol{e}_1), \dots, \varphi(\boldsymbol{e}_n)) \in K^{m\times n},$$

ktorej stĺpcami sú práve tieto vektory, t. j. platí $\boldsymbol{s}_j(\boldsymbol{A}) = \varphi(\boldsymbol{e}_j)$ pre $1 \le j \le n$. Ukážeme, ako možno obraz $\varphi(\boldsymbol{x})$ ľubovoľného vektora $\boldsymbol{x} = (x_1, \dots, x_n)^{\mathsf{T}} \in K^n$ vypočítať len zo znalosti tejto matice. Uvedomme si, že $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n$, a počítajme

$$\begin{aligned}\varphi(\boldsymbol{x}) &= \varphi(x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n) = x_1\varphi(\boldsymbol{e}_1) + \ldots + x_n\varphi(\boldsymbol{e}_n)\\ &= (\boldsymbol{s}_1(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A})) \cdot \begin{pmatrix} x_1\\ \vdots \\ x_n\end{pmatrix} = \boldsymbol{A}\cdot\boldsymbol{x}\end{aligned}$$

Teda *každé* lineárne zobrazenie $\varphi\colon K^n \to K^m$ má tvar $\varphi(\boldsymbol{x}) = \boldsymbol{A}\cdot\boldsymbol{x}$ pre vhodnú maticu $\boldsymbol{A} \in K^{m\times n}$. Keďže každý konečnorozmerný vektorový priestor nad poľom K je izomorfný s priestorom K^n pre $n = \dim V$, pri voľbe pevných báz v konečnorozmerných priestoroch U, V bude možné ľubovoľné lineárne zobrazenie $\varphi\colon V \to U$ zakódovať pomocou vhodnej matice $\boldsymbol{A}$.

Nech U, V sú konečnorozmerné vektorové priestory nad poľom K, $\dim U = m$, $\dim V = n$ a $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$, $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ sú bázy v U, resp. vo V. *Maticou lineárneho zobrazenia $\varphi\colon V \to U$ vzhľadom na bázy $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$* nazývame maticu

$$\boldsymbol{A} = \big((\varphi\boldsymbol{v}_1)_{\boldsymbol{\alpha}}, \dots, (\varphi\boldsymbol{v}_n)_{\boldsymbol{\alpha}}\big) \in K^{m\times n},$$

ktorej stĺpce sú tvorené súradnicami obrazov $\varphi(\boldsymbol{v}_j)$ vektorov bázy $\boldsymbol{\beta}$ vzhľadom na bázu $\boldsymbol{\alpha}$, t. j. platí $\boldsymbol{s}_j(\boldsymbol{A}) = (\varphi\boldsymbol{v}_j)_{\boldsymbol{\alpha}}$ pre $1 \le j \le n$. Túto maticu značíme tiež

$$\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}.$$

(Všimnite si obrátené poradie znakov báz voči poradiu vektorových priestorov v označení zobrazenia $\varphi\colon V \to U$.)

Maticu $\boldsymbol{A}$ zo začiatku tohto paragrafu by sme teda mohli nazvať *maticou lineárneho zobrazenia $\varphi\colon K^n \to K^m$ vzhľadom na kanonické bázy $\boldsymbol{\varepsilon}^{(n)}$*,

$\boldsymbol{\varepsilon}^{(m)}$. Pokiaľ nepovieme inak, budeme pod maticou lineárneho zobrazenia $\varphi\colon K^n \to K^m$ medzi stĺpcovými vektorovými priestormi vždy rozumieť maticu $(\varphi)_{\boldsymbol{\varepsilon}^{(m)},\boldsymbol{\varepsilon}^{(n)}}$ zobrazenia φ vzhľadom na kanonické bázy.

Pri štúdiu lineárnych transformácií $\varphi\colon V \to V$ konečnorozmerného vektorového priestoru V budeme spravidla vzory i obrazy vektorov z V vyjadrovať v tej istej báze. *Maticou lineárnej transformácie* $\varphi\colon V \to V$ vzhľadom na bázu $\boldsymbol{\alpha}$ priestoru V teda rozumieme maticu

$$(\varphi)_{\boldsymbol{\alpha}} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\alpha}}.$$

Pri dôkaze nasledujúcej vety bude potrebné si uvedomiť, že $(\boldsymbol{v}_j)_{\boldsymbol{\beta}} = \boldsymbol{e}_j^{(n)}$ pre ľubovoľný vektor $\boldsymbol{v}_j$ bázy $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ priestoru V. Z toho je zrejmé, že pre každú bázu $\boldsymbol{\beta}$ n-rozmerného vektorového priestoru V platí

$$(\mathrm{id}_V)_{\boldsymbol{\beta},\boldsymbol{\beta}} = \boldsymbol{I}_n.$$

6.4.1. Veta. *Nech $\varphi\colon V \to U$ je lineárne zobrazenie medzi konečnorozmernými vektorovými priestormi nad poľom K, $\dim V = n$, $\dim U = m$ a $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ sú bázy priestorov U resp. V. Potom pre všetky $\boldsymbol{x} \in V$ platí*

$$(\varphi\boldsymbol{x})_{\boldsymbol{\alpha}} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot (\boldsymbol{x})_{\boldsymbol{\beta}}$$

a $\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ je jediná matica s touto vlastnosťou.

Dôkaz. Nech $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$. Zvoľme ľubovoľný vektor $\boldsymbol{x} \in V$ a označme $(\boldsymbol{x})_{\boldsymbol{\beta}} = (x_1, \ldots, x_n)^{\mathsf{T}}$, teda $\boldsymbol{x} = x_1 e \boldsymbol{v}_1 + \cdots + x_n \boldsymbol{v}_n$. Podobne ako v špeciálnom prípade zo začiatku tohto paragrafu, i teraz dostávame

$$\begin{aligned}
\varphi(\boldsymbol{x}) &= \varphi(x_1\boldsymbol{v}_1 + \ldots + x_n\boldsymbol{v}_n) = x_1\varphi(\boldsymbol{v}_1) + \ldots + x_n\varphi(\boldsymbol{v}_n),\\
(\varphi\boldsymbol{x})_{\boldsymbol{\alpha}} &= \big(x_1\varphi(\boldsymbol{v}_1) + \ldots + x_n\varphi(\boldsymbol{v}_n)\big)_{\boldsymbol{\alpha}} = x_1(\varphi\boldsymbol{v}_1)_{\boldsymbol{\alpha}} + \ldots + x_n(\varphi\boldsymbol{v}_n)_{\boldsymbol{\alpha}}\\
&= \big((\varphi\boldsymbol{v}_1)_{\boldsymbol{\alpha}}, \ldots, (\varphi\boldsymbol{v}_n)_{\boldsymbol{\alpha}}\big) \cdot \begin{pmatrix} x_1\\ \vdots \\ x_n \end{pmatrix} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot (\boldsymbol{x})_{\boldsymbol{\beta}}.
\end{aligned}$$

Zostáva ukázať, že pre ľubovoľnú maticu $\boldsymbol{A} \in K^{m\times n}$ platí

$$\big(\forall \boldsymbol{x} \in V\big)\big((\varphi\boldsymbol{x})_{\boldsymbol{\alpha}} = \boldsymbol{A} \cdot (\boldsymbol{x})_{\boldsymbol{\beta}}\big) \;\Rightarrow\; \boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}.$$

Za uvedeného predpokladu voľbou $\boldsymbol{x} = \boldsymbol{v}_j$ dostávame

$$(\varphi\boldsymbol{v}_j)_{\boldsymbol{\alpha}} = \boldsymbol{A} \cdot (\boldsymbol{v}_j)_{\boldsymbol{\beta}} = \boldsymbol{A} \cdot \boldsymbol{e}_j = \boldsymbol{s}_j(\boldsymbol{A})$$

pre každé $1 \le j \le n$. Teda matice $\boldsymbol{A}$ a $(\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ majú rovnaké stĺpce, preto sa rovnajú.

Matica $(\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ lineárneho zobrazenia $\varphi\colon V \to U$ z n-rozmerného vektorového priestoru V do m-rozmerného vektorového priestoru U nad poľom K vzhľadom na bázy $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$ je teda medzi všetkými maticami $\boldsymbol{A} \in K^{m\times n}$ jednoznačne určená podmienkou

$$(\varphi\boldsymbol{x})_{\boldsymbol{\alpha}} = \boldsymbol{A}\cdot(\boldsymbol{x})_{\boldsymbol{\beta}}$$

pre každé $\boldsymbol{x} \in V$.

Ďalej si ukážeme, že skladanie lineárnych zobrazení zodpovedá násobeniu matíc, čo umožňuje vypočítať maticu kompozície lineárnych zobrazení $\varphi\circ\psi$ len zo znalosti matíc jednotlivých zobrazení φ a ψ. Tento výsledok definitívne „ospravedlňuje“ spôsob, akým sme súčin matíc definovali v odstavci 2.2.2

6.4.2. Veta. *Nech U, V, W sú konečnorozmerné vektorové priestory nad poľom K, $\boldsymbol{\alpha}$ je báza U, $\boldsymbol{\beta}$ je báza V a $\boldsymbol{\gamma}$ je báza W. Potom pre ľubovoľné lineárne zobrazenia $\psi\colon W\to V$, $\varphi\colon V\to U$ platí*

$$(\varphi\circ\psi)_{\boldsymbol{\alpha},\boldsymbol{\gamma}} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}\cdot(\psi)_{\boldsymbol{\beta},\boldsymbol{\gamma}}.$$

Dôkaz. Označme $\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ maticu lineárneho zobrazenia φ vzhľadom na bázy $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$ a $\boldsymbol{B} = (\psi)_{\boldsymbol{\beta},\boldsymbol{\gamma}}$ maticu lineárneho zobrazenia ψ vzhľadom na bázy $\boldsymbol{\gamma}$, $\boldsymbol{\beta}$. Dokážeme rovnosť

$$(\varphi\circ\psi)_{\boldsymbol{\alpha},\boldsymbol{\gamma}} = \boldsymbol{A}\cdot\boldsymbol{B}.$$

Na základe definície matíc $\boldsymbol{A}$, $\boldsymbol{B}$ pre ľubovoľné $\boldsymbol{x}\in W$ platí

$$\begin{aligned}\big((\varphi\circ\psi)\boldsymbol{x}\big)_{\boldsymbol{\alpha}} &= \big(\varphi(\psi\boldsymbol{x})\big)_{\boldsymbol{\alpha}} = \boldsymbol{A}\cdot(\psi\boldsymbol{x})_{\boldsymbol{\beta}}\\ &= \boldsymbol{A}\cdot\big(\boldsymbol{B}\cdot(\boldsymbol{x})_{\boldsymbol{\gamma}}\big) = (\boldsymbol{A}\cdot\boldsymbol{B})\cdot(\boldsymbol{x})_{\boldsymbol{\gamma}}.\end{aligned}$$

Keďže matica kompozície $\varphi\circ\psi$ vzhľadom na bázy $\boldsymbol{\gamma}$, $\boldsymbol{\alpha}$ je podľa vety 6.4.1 touto podmienkou určená jednoznačne, dôkaz je hotový.

V nasledujúcich príkladoch sa zoznámime s niekoľkými dôležitými lineárnymi transformáciami roviny $\mathbb{R}^2$ a ich maticami vzhľadom na kanonickú bázu $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2)$.

6.4.3. Príklad. *Otočenie* roviny okolo počiatku o uhol $\alpha\in\mathbb{R}$ je lineárne zobrazenie $\boldsymbol{R}_\alpha\colon\mathbb{R}^2\to\mathbb{R}^2$. Homogenita je zrejmá na prvý pohľad. O aditivite sa presvedčíme nasledujúcou úvahou. Ak otočíme rovnobežník vektorov $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^2$ o uhol α, dostaneme tak rovnobežník prislúchajúci vektorom $\boldsymbol{R}_\alpha(\boldsymbol{x}),\boldsymbol{R}_\alpha(\boldsymbol{y})$. Pritom uhlopriečka prvého rovnobežníka prejde na uhlopriečku druhého. Teda $\boldsymbol{R}_\alpha(\boldsymbol{x}+\boldsymbol{y}) = \boldsymbol{R}_\alpha(\boldsymbol{x})+\boldsymbol{R}_\alpha(\boldsymbol{y})$ (nakreslite si obrázok). Maticu tohto lineárneho zobrazenia vzhľadom na kanonickú bázu $\boldsymbol{\varepsilon}$ budeme značiť rovnako $\boldsymbol{R}_\alpha$, teda pre $\boldsymbol{x}\in\mathbb{R}^2$ budeme písať $\boldsymbol{R}_\alpha(\boldsymbol{x}) = \boldsymbol{R}_\alpha\cdot\boldsymbol{x}$. Jej stĺpce získame otočením vektorov $\boldsymbol{e}_1 = (1,0)^\mathsf{T}$, $\boldsymbol{e}_2 = (0,1)^\mathsf{T}$ o uhol α. Z definície goniometrických funkcií sínus a kosínus pomocou jednotkovej kružnice priamo dostávame

$$\boldsymbol{R}_\alpha \cdot \boldsymbol{e}_1 = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}, \qquad \boldsymbol{R}_\alpha \cdot \boldsymbol{e}_2 = \begin{pmatrix} \cos\left(\frac{\pi}{2}+\alpha\right) \\ \sin\left(\frac{\pi}{2}+\alpha\right) \end{pmatrix} = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}.$$

To znamená, že

$$\boldsymbol{R}_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

a obrazom ľubovoľného vektora $(x,y)^\mathsf{T} \in \mathbb{R}^2$ v otočení $\boldsymbol{R}_\alpha$ je vektor

$$\boldsymbol{R}_\alpha \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\alpha - y\sin\alpha \\ x\sin\alpha + y\cos\alpha \end{pmatrix}.$$

Na základe znalostí matíc $\boldsymbol{R}_\alpha$ je už jednoduché napísať matice otočení v $\mathbb{R}^3$ okolo súradných osí vzhľadom na kanonickú bázu $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$. Otočenia okolo osí $x = [\boldsymbol{e}_1]$, $y = [\boldsymbol{e}_2]$ resp $z = [\boldsymbol{e}_3]$ o uhol α majú postupne matice

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}, \qquad \begin{pmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix}, \qquad \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Na orientáciu otočení používame *pravidlo pravej ruky*, t. j. prsty pravej ruky ukazujú smer otáčania a palec určuje orientáciu osi: v prvom prípade prsty od vektora $\boldsymbol{e}_2$ k $\boldsymbol{e}_3$ a palec v smere vektora $\boldsymbol{e}_1$, v druhom prsty od $\boldsymbol{e}_3$ k $\boldsymbol{e}_1$ a palec v smere $\boldsymbol{e}_2$, v treťom prsty od $\boldsymbol{e}_1$ k $\boldsymbol{e}_2$ a palec v smere $\boldsymbol{e}_3$.

6.4.4. Príklad. *Osová súmernosť* roviny podľa ľubovoľnej priamky prechádzajúcej počiatkom definuje zobrazenie $\boldsymbol{S}_\alpha \colon \mathbb{R}^2 \to \mathbb{R}^2$, kde $\alpha \in \mathbb{R}$ je uhol, ktorý zviera os súmernosti s osou x. Obdobnou úvahou ako v prípade otočení možno nahliadnuť, že i $\boldsymbol{S}_\alpha$ je lineárne zobrazenie. Jeho maticu vzhľadom na kanonickú bázu $\boldsymbol{\varepsilon}$ budeme značiť rovnako $\boldsymbol{S}_\alpha$. Zrejme matica súmernosti podľa osi x je

$$\boldsymbol{S}_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

a osovú súmernosť $\boldsymbol{S}_\alpha$ možno dostať ako kompozíciu otočenia $\boldsymbol{R}_{-\alpha}$, osovej súmernosti $\boldsymbol{S}_0$ a otočenia $\boldsymbol{R}_\alpha$, t. j.

$$\boldsymbol{S}_\alpha = \boldsymbol{R}_\alpha \cdot \boldsymbol{S}_0 \cdot \boldsymbol{R}_{-\alpha}.$$

Po vynásobení príslušných matíc z toho s využitím pár trigonometrických vzorcov dostávame

$$\boldsymbol{S}_\alpha = \begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix}.$$

Teda osová súmernosť $\boldsymbol{S}_\alpha$ zobrazí vektor $(x,y)^\mathsf{T} \in \mathbb{R}^2$ do vektora

$$\boldsymbol{S}_\alpha \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos 2\alpha + y\sin 2\alpha \\ x\sin 2\alpha - y\cos 2\alpha \end{pmatrix}.$$

6.4.5. Príklad. *Rovnoľahlosť* alebo tiež *homotetia* so stredom v počiatku a s koeficientom podobnosti $0 \neq c \in \mathbb{R}$ je opäť lineárne zobrazenie $\mathbb{R}^2 \to \mathbb{R}^2$ s maticou $c\boldsymbol{I}_2 = \operatorname{diag}(c,c)$. Tento príklad možno zrejmým spôsobom zovšeobecniť na ľubovoľnú dimenziu n.

6.4.6. Príklad. *Skosenie v smere osi x s parametrom a.* Pre ľubovoľné $a \in \mathbb{R}$ je priradením

$$\varphi \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + ay \\ y \end{pmatrix} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

definovaná lineárna transformácia roviny, ktorá posúva každú jej „vodorovnú vrstvu“ $\{(x,y);\ y = s\}$, $s \in \mathbb{R}$, o vektor $as\boldsymbol{e}_1$ (nakreslite si obrázok). Analogické lineárne transformácie fungujú aj vo viacrozmerných priestoroch $\mathbb{R}^n$ ako aj v konečnorozmerných vektorových priestoroch nad ľubovoľným poľom.

6.4.7. Príklad. *Galileova transformácia* „roviny“, alebo skôr „časopriamky“ $\mathbb{R}^2$ je z formálneho hľadiska totožná so skosením v smere priestorovej osi x s parametrom $a = -v$:

$$\begin{aligned} t' &= t \\ x' &= -vt + x, \end{aligned} \qquad \text{t. j.} \qquad \begin{pmatrix} t' \\ x' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -v & 1 \end{pmatrix} \cdot \begin{pmatrix} t \\ x \end{pmatrix}.$$

Ak k týmto rovniciam ešte doplníme $y' = y$, $z' = z$, dostaneme Galileovu transformáciu „časopriestoru“ $\mathbb{R}^4$. Vektor $(t,x,y,z)^\mathsf{T}$ interpretujeme ako súradnice času t a polohy (x,y,z), ktoré nejakej okamžitej bodovej udalosti v trojrozmernom fyzikálnom priestore priradí pozorovateľ P, a vektor $(t',x',y',z')^\mathsf{T}$ ako súradnice, ktoré jej priradí pozorovateľ P', ktorý sa vzhľadom na P pohybuje rovnomerne priamočiaro rýchlosťou v v smere osi x (pričom počiatky ich súradných sústav sú zhodne stanovené okamihom a miestom ich stretnutia, kedy tiež splývajú ich príslušné súradné osi). Galileova transformácia potom udáva vzťah medzi týmito súradnicami, ku ktorému dospejeme na základe princípov klasickej mechaniky. Je v nej zachytený newtonovský princíp absolútneho času a priestoru, rovnakého pre všetkých pozorovateľov nezávisle od ich pohybu, a galileovský princíp relatívnosti pohybu, ktorý sa prejavuje v rovnocennosti súradných sústav ľubovoľných navzájom rovnomerne priamočiaro sa pohybujúcich pozorovateľov. Je známe, že Galileova transformácia sa veľmi dobre zhoduje so skutočnosťou pre rýchlosti v z „bežného života“, ktoré sú malé v porovnaní s rýchlosťou svetla. Pre rýchlosti blízke rýchlosti svetla však stráca svoju platnosť a treba ju nahradiť tzv. Lorentzovou transformáciou, s ktorou sa zoznámime neskôr.

6.5 Priestory lineárnych zobrazení

Nech U, V sú vektorové priestory nad poľom K. Ak zabudneme na štruktúru vektorového priestoru vo V, t. j. V budeme považovať len za množinu, môžeme vytvoriť vektorový priestor U^V *všetkých* zobrazení $f\colon V \to U$ s operáciami súčtu a skalárneho násobku definovanými po zložkách (pozri odstavec 1.6.5). Potom pre množinu $\mathcal{L}(V,U)$ všetkých *lineárnych* zobrazení $\varphi\colon V \to U$, samozrejme, platí $\mathcal{L}(V,U) \subseteq U^V$.

6.5.1. Tvrdenie. *Nech U, V sú vektorové priestory nad poľom K. Potom $\mathcal{L}(V,U)$ je lineárny podpriestor vektorového priestoru U^V. Teda $\mathcal{L}(V,U)$ je vektorový priestor nad poľom K.*

Dôkaz. Treba overiť, že pre ľubovoľné $a, b \in K$ lineárna kombinácia $\vartheta = a\varphi + b\psi$ lineárnych zobrazení $\varphi, \psi\colon V \to U$, definovaná pre $\boldsymbol{x} \in V$ predpisom

$$\vartheta(\boldsymbol{x}) = a\varphi(\boldsymbol{x}) + b\psi(\boldsymbol{x}),$$

je tiež lineárne zobrazenie $\vartheta\colon V \to U$. Zvoľme $\boldsymbol{x}, \boldsymbol{y} \in V$, $c, d \in K$. Priamym výpočtom dostávame

$$\begin{aligned}\vartheta(c\boldsymbol{x} + d\boldsymbol{y}) &= a\varphi(c\boldsymbol{x} + d\boldsymbol{y}) + b\psi(c\boldsymbol{x} + d\boldsymbol{y}) = a(c\varphi(\boldsymbol{x}) + d\varphi(\boldsymbol{y})) + b(c\psi(\boldsymbol{x}) + d\psi(\boldsymbol{y}))\\ &= c(a\varphi(\boldsymbol{x}) + b\psi(\boldsymbol{x})) + d(a\varphi(\boldsymbol{y}) + b\psi(\boldsymbol{y})) = c\vartheta(\boldsymbol{x}) + d\vartheta(\boldsymbol{y}).\end{aligned}$$

Teda $\vartheta \in \mathcal{L}(V,U)$.

6.5.2. Tvrdenie. *Nech U, V sú konečnorozmerné vektorové priestory nad poľom K a $\dim U = m$, $\dim V = n$. Potom*

$$\mathcal{L}(V,U) \cong K^{m\times n},$$

teda $\dim \mathcal{L}(V,U) = mn$.

Dôkaz. Zvoľme bázy $\boldsymbol{\alpha}$ v priestore U a $\boldsymbol{\beta}$ v priestore V. Z výsledkov predchádzajúceho paragrafu vyplýva, že priradenie $\varphi \mapsto (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ je bijekcia $\mathcal{L}(V,U) \to K^{m\times n}$. Stačí teda dokázať, že je to aj lineárne zobrazenie, to znamená, že pre $a, b \in K$, $\varphi, \psi \in \mathcal{L}(V,U)$ platí

$$(a\varphi + b\psi)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = a(\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}} + b(\psi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}.$$

To je však zrejmé. Čitateľovi, ktorý má nejaké pochybnosti, odporúčame, aby si jednoduchým výpočtom porovnal stĺpce matíc na ľavej a pravej strane.

Na maticu $(\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ sa teda možno dívať ako na súradnice vektora $\varphi \in \mathcal{L}(V,U)$ v priestore $K^{m\times n}$, vzhľadom na dvojicu báz $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$.

Lineárne zobrazenie $\varphi\colon V \to K$ z vektorového priestoru V do poľa K sa nazýva *lineárny funkcionál* alebo *lineárna forma* na V. Vektorový priestor $\mathcal{L}(V,K)$ všetkých lineárnych foriem na V sa nazýva *duálny priestor* alebo len krátko *duál* vektorového priestoru V. Budeme používať označenie $\mathcal{L}(V,K) = V^*$

V poli K budeme vždy uvažovať len kanonickú bázu pozostávajúcu z jediného vektora $1 \in K$. Ľubovoľná báza $\boldsymbol{\beta}$ v konečnorozmernom priestore V dimenzie n tak určuje lineárny izomorfizmus $V^* \to K^{(1\times n)}$ daný predpisom $\varphi \mapsto (\varphi)_{1,\boldsymbol{\beta}}$. Nakoľko aj $\boldsymbol{x} \mapsto (\boldsymbol{x})_{\boldsymbol{\beta}}$ je lineárny izomorfizmus $V \to K^{(n\times 1)}$ a $K^{(1\times n)} \cong K^{(n\times 1)}$, dokázali sme tým nasledujúce tvrdenie.

6.5.3. Tvrdenie. *Pre ľubovoľný konečnorozmerný vektorový priestor V nad poľom K platí $V^* \cong V$.*

Uvedomme si, že matica $(\varphi)_{1,\boldsymbol{\beta}}$ lineárneho funkcionálu $\varphi\colon V \to K$ je riadkový vektor z priestoru $K^{1\times n}$. To nám odkrýva nový pohľad na vektorový priestor riadkov $K^{1\times n}$. Pri voľbe kanonickej bázy $\boldsymbol{\varepsilon}$ v stĺpcovom priestore $K^{n\times 1}$ možno riadkový priestor $K^{1\times n}$ stotožniť s duálom $\left(K^{n\times 1}\right)^*$ stĺpcového priestoru $K^{n\times 1}$.

Ešte raz podčiarknime, že izomorfizmus konečnorozmerného priestoru V a jeho duálu V^* závisí od výberu bázy vo V. Na druhej strane, pre ľubovoľný vektorový priestor V možno definovať kanonické, t. j. od výberu bázy nezávislé zobrazenie z priestoru V do jeho *druhého duálu* V^{**} dané predpisom $\boldsymbol{x} \mapsto \widehat{\boldsymbol{x}}$, kde

$$\widehat{\boldsymbol{x}}(\varphi) = \varphi(\boldsymbol{x})$$

pre $\boldsymbol{x} \in V$, $\varphi \in V^*$.

6.5.4. Veta. *Nech V je vektorový priestor nad poľom K. Potom*
(a) $\boldsymbol{x} \mapsto \widehat{\boldsymbol{x}}$ je injektívne lineárne zobrazenie $V \to V^{**}$;
(b) ak V je konečnorozmerný, tak $\boldsymbol{x} \mapsto \widehat{\boldsymbol{x}}$ je lineárny izomorfizmus $V \to V^{**}$.

Dôkaz. Najprv ukážeme, že pre každé $\boldsymbol{x} \in V$ vôbec platí $\widehat{\boldsymbol{x}} \in V^{**}$, t. j. $\widehat{\boldsymbol{x}}$ je lineárny funkcionál na priestore V^*. Pre $\varphi, \psi \in V^*$, $a, b \in K$ jednoduchý výpočet dáva

$$\widehat{\boldsymbol{x}}(a\varphi + b\psi) = (a\varphi + b\psi)(\boldsymbol{x}) = a\varphi(\boldsymbol{x}) + b\psi(\boldsymbol{x}) = a\widehat{\boldsymbol{x}}(\varphi) + b\widehat{\boldsymbol{x}}(\psi).$$

(a) Pre $\boldsymbol{x} \in V$ označme $\widehat{\boldsymbol{x}} = \eta(\boldsymbol{x})$. Dokážeme, že $\eta\colon V \to V^{**}$ je lineárne zobrazenie. Zvoľme $\boldsymbol{x}, \boldsymbol{y} \in V$ $c, d \in K$. Treba overiť, že zobrazenia $\eta(c\boldsymbol{x} + d\boldsymbol{y})$ a $c\eta(\boldsymbol{x}) + d\eta(\boldsymbol{y})$ sa rovnajú, t. j. že každému $\varphi \in V^*$ priradia tú istú hodnotu. Počítajme

$$\begin{aligned}\eta(c\boldsymbol{x} + d\boldsymbol{y})(\varphi) &= \varphi(c\boldsymbol{x} + d\boldsymbol{y}) = c\varphi(\boldsymbol{x}) + d\varphi(\boldsymbol{y}) \\ &= c\widehat{\boldsymbol{x}}(\varphi) + d\widehat{\boldsymbol{y}}(\varphi) = (c\eta(\boldsymbol{x}) + d\eta(\boldsymbol{y}))(\varphi).\end{aligned}$$

Injektívnosť zobrazenia η spoznáme podľa jeho jadra. Stačí si uvedomiť, že pre ľubovoľné $\mathbf{0} \neq \boldsymbol{x} \in V$ existuje $\varphi \in V^*$ také, že $\varphi(\boldsymbol{x}) \neq 0$. [1]

Potom

$$\eta(\boldsymbol{x})(\varphi) = \widehat{\boldsymbol{x}}(\varphi) = \varphi(\boldsymbol{x}) \neq 0,$$

teda $\eta(\boldsymbol{x})$ sa nerovná identicky nulovému zobrazeniu $\mathbf{0}\colon V^* \to K$, ktoré je nulou vo V^{**}. Preto $\boldsymbol{x} \notin \operatorname{Ker} \eta$ pre $\boldsymbol{x} \neq \mathbf{0}$, čiže $\operatorname{Ker} \eta = \{\mathbf{0}\}$.

(b) Ak V je konečnorozmerný, tak s vyžitím vety 6.2.3, už dokázanej časti (a) a tvrdenia 6.5.3 dostávame

$$\dim \operatorname{Im} \eta = \dim V - \dim \operatorname{Ker} \eta = \dim V = \dim V^* = \dim V^{**}.$$

Preto $\operatorname{Im} \eta = V^{**}$, čiže η je i surjektívne.

Každý vektor $\boldsymbol{x} \in V$ tak definuje lineárny funkcionál $\widehat{\boldsymbol{x}}$ na duálnom priestore V^*. Pritom konečnorozmerný vektorový priestor V možno prostredníctvom priradenia $\boldsymbol{x} \mapsto \widehat{\boldsymbol{x}}$ *prirodzene* stotožniť s duálom priestoru V^*. Napr. stĺpcový priestor $K^{n\times 1}$ možno stotožniť s duálom riadkového priestoru $K^{1\times n}$. Vo všeobecnom prípade možno V stotožniť s lineárnym podpriestorom $\operatorname{Im} \eta = \{\widehat{\boldsymbol{x}};\, \boldsymbol{x} \in V\}$ jeho druhého duálu V^{**}.

Poznámka. Prostriedkami, ktoré presahujú rámec tohto kurzu, možno dokázať, že $V \ncong V^*$ a $V \ncong V^{**}$ pre každý nekonečnorozmerný vektorový priestor V. Zjednodušene povedané, duál V^* je vždy „podstatne väčší“ než pôvodný priestor V a druhý duál V^{**} je „ešte väčší“. Teda ani $\boldsymbol{x} \mapsto \widehat{\boldsymbol{x}}$ nemôže byť surjektívne zobrazenie $V \to V^{**}$.

Cvičenia

6.1. Nech U, V sú vektorové priestory nad poľom K a $\varphi\colon V \to U$ je lineárne zobrazenie. Potom $\varphi(\mathbf{0}) = \mathbf{0}$ a pre každé $\boldsymbol{x} \in V$ platí $\varphi(-\boldsymbol{x}) = -\varphi(\boldsymbol{x})$. Dokážte.

6.2. Dokážte tvrdenie 6.1.1

6.3. Nech $\boldsymbol{A} \in K^{m\times n}$. Uvažujme lineárne zobrazenie $\varphi\colon K^m \to K^n$ dané predpisom $\varphi(\boldsymbol{x}) = \boldsymbol{A} \cdot \boldsymbol{x}$. Ako súvisí jadro $\operatorname{Ker} \varphi$ s homogénnou sústavou lineárnych rovníc $\boldsymbol{A} \cdot \boldsymbol{x} = \mathbf{0}$?

6.4. (a) V príkladoch 6.1.5–9 podrobne overte, že v nich definované zobrazenia sú naozaj lineárne.

(b) Pre každé z uvedených lineárnych zobrazení určte jeho jadro a obraz.

6.5. Zdôvodnite, prečo žiadne z nasledujúcich zobrazení $\mathbb{R} \to \mathbb{R}$ nie je lineárne:

(a) $\varphi(x) = 1 - 3x$, (b) $\psi(x) = |x|$,
(c) $f(x) = x^2$, (d) $g(x) = x^3$,

[1] V konečnorozmernom prípade ide o očividné tvrdenie. V nekonečnorozmernom prípade však jeho dôkaz využíva existenciu Hamelovej bázy vo V, prípadne si vyžaduje nejaký iný vhodný dôsledok axiómy výberu (pozri záverečnú časť paragrafu 5.5 a cvičenie 6.15).

(e) $h(x) = \sqrt{|x|}$, (f) $k(x) = \sqrt[3]{x}$,
(g) $l(x) = \ln(1 + x^2)$ (h) $s(x) = \sin x$,
(i) $\eta(x) = \mathrm{e}^x$, (j) $\sigma(x) = \sinh x = \frac{1}{2}(\mathrm{e}^x - \mathrm{e}^{-x})$.

6.6. (a) Uvažujte $\mathbb{C}$ ako vektorový priestor nad poľom $\mathbb{R}$. Ukážte, že komplexná konjugácia $z \mapsto \overline{z}$ je lineárne zobrazenie $\mathbb{C} \to \mathbb{C}$.

(b) Uvažujte $\mathbb{C}$ ako vektorový priestor nad poľom $\mathbb{C}$. Ukážte, že komplexná konjugácia $z \mapsto \overline{z}$ *nie je* lineárne zobrazenie $\mathbb{C} \to \mathbb{C}$.

6.7. Lineárne zobrazenie $\varphi\colon \mathbb{R}^3 \to \mathbb{R}^2$ je pre $\boldsymbol{x} = (x_1, x_2, x_3)^\mathsf{T} \in \mathbb{R}^3$ dané predpisom $\varphi(\boldsymbol{x}) = (x_1 + 2x_2, 3x_1 + 7x_2 - x_3)^\mathsf{T}$.

(a) Napíšte maticu lineárneho zobrazenia φ vzhľadom na kanonické bázy $\boldsymbol{\varepsilon}^{(3)}$ v $\mathbb{R}^3$, $\boldsymbol{\varepsilon}^{(2)}$ v $\mathbb{R}^2$.

(b) Napíšte maticu φ vzhľadom na bázy $\boldsymbol{\beta} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{v})$ v $\mathbb{R}^3$, $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \boldsymbol{u}_2)$ v $\mathbb{R}^2$, kde $\boldsymbol{v} = (-2, 1, 1)^\mathsf{T}$, $\boldsymbol{u}_1 = (1, 3)^\mathsf{T}$, $\boldsymbol{u}_2 = (2, 7)^\mathsf{T}$.

(c) Napíšte priamo nejaké bázy lineárnych podpriestorov $\operatorname{Ker}\varphi \subseteq \mathbb{R}^3$ a $\operatorname{Im}\varphi \subseteq \mathbb{R}^2$.

6.8. Báza $\boldsymbol{\beta}$ vektorového priestoru $\mathbb{R}^3$ je tvorená stĺpcami matice $\begin{pmatrix} 1 & 1 & 2 \\ 2 & 0 & 2 \\ 3 & -1 & 1 \end{pmatrix}$ a báza $\boldsymbol{\alpha}$ vektorového priestoru $\mathbb{R}^4$ je tvorená stĺpcami matice $\begin{pmatrix} 1 & 2 & 3 & 1 \\ 1 & 0 & -1 & 0 \\ 1 & -1 & 1 & -1 \\ 0 & 1 & 2 & 1 \end{pmatrix}$. Lineárne zobrazenie $\psi\colon \mathbb{R}^3 \to \mathbb{R}^4$ má vzhľadom na tieto bázy maticu $\boldsymbol{A} = (\psi)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

(a) Napíšte priamo nejaké bázy lineárnych podpriestorov $\operatorname{Ker}\psi \subseteq \mathbb{R}^3$ a $\operatorname{Im}\psi \subseteq \mathbb{R}^4$.

(b) Napíšte maticu ψ vzhľadom na kanonické bázy $\boldsymbol{\varepsilon}^{(3)}$ v $\mathbb{R}^3$, $\boldsymbol{\varepsilon}^{(4)}$ v $\mathbb{R}^{(3)}$.

6.9. Nech $n \in \mathbb{N}$. Uvažujme deriváciu $f \mapsto f'$ ako lineárnu transformáciu $D\colon \mathbb{R}^{(n)}[x] \to \mathbb{R}^{(n)}[x]$.

(a) Určte lineárne podpriestory $\operatorname{Ker} D$ a $\operatorname{Im} D$ v $\mathbb{R}^{(n)}[x]$ a nájdite nejaké ich bázy. Zistite dimenzie oboch podpriestorov a overte vzťah $\dim \operatorname{Ker} D + \dim \operatorname{Im} D = n$.

(b) Napíšte maticu lineárnej transformácie D vzhľadom na bázu $\boldsymbol{\xi}^{(n)} = (1, x, x^2, \ldots, x^n)$.

(c) Napíšte maticu D vzhľadom na bázu $\boldsymbol{\zeta}^{(n)} = \left(1, \frac{1}{1!}x, \frac{1}{2!}x^2, \ldots, \frac{1}{n!}x^n\right)$.

6.10. *Operátor diferencie* $\Delta\colon \mathbb{R}^{(n)}[x] \to \mathbb{R}^{(n)}[x]$ je daný predpisom $\Delta(f)(x) = f(x+1) - f(x)$.

(a) Dokážte, že Δ je lineárny operátor.

(b) Určte lineárne podpriestory $\operatorname{Ker}\Delta$ a $\operatorname{Im}\Delta$ v $\mathbb{R}^{(n)}[x]$ a nájdite nejaké ich bázy. Zistite dimenzie oboch podpriestorov.

(c) Napíšte maticu lineárnej transformácie Δ vzhľadom na bázu $\boldsymbol{\xi}^{(n)} = (1, x, x^2, \ldots, x^n)$.

(d) Dokážte, že polynómy $1, x, [x]_2, \ldots, [x]_n$, kde $[x]_k = x(x-1)\ldots(x-k+1)$, tvoria bázu vektorového priestoru $\mathbb{R}^{(n)}[x]$.

(e) Napíšte maticu Δ vzhľadom na bázu $\boldsymbol{\eta}^{(n)} = (1, x, [x]_2, \ldots, [x]_n)$.

(f) Nájdite nejakú bázu vektorového priestoru $\mathbb{R}^{(n)}[x]$, vzhľadom na ktorú má Δ maticu tvaru $\begin{pmatrix} \mathbf{0}_{n,1} & \boldsymbol{I}_n \\ 0 & \mathbf{0}_{1,n} \end{pmatrix}$.

6.11. Pre ľubovoľné $\alpha, \beta \in \mathbb{R}$ vynásobením príslušných matíc dokážte uvedené rovnosti a interpretujte ich v reči lineárnych transformácií roviny $\mathbb{R}^2$:

$$\boldsymbol{R}_\alpha \cdot \boldsymbol{R}_\beta = \boldsymbol{R}_{\alpha+\beta}, \qquad \boldsymbol{R}_\alpha \cdot \boldsymbol{S}_\beta = \boldsymbol{S}_{\beta+\alpha/2},$$
$$\boldsymbol{S}_\alpha \cdot \boldsymbol{S}_\beta = \boldsymbol{R}_{2(\beta-\alpha)}, \qquad \boldsymbol{S}_\beta \cdot \boldsymbol{R}_\alpha = \boldsymbol{S}_{\beta-\alpha/2}.$$

6.12. Predpismi $\varphi(x_1,x_2,x_3)^\mathsf{T} = (x_1-2x_3, 2x_1+x_2+3x_4, -x_1-x_4, x_3+2x_4)^\mathsf{T}$, $\psi(x_1,x_2,x_3,x_4)^\mathsf{T} = (2x_1-x_3, 3x_3+x_4, x_2+x_4)^\mathsf{T}$ sú dané lineárne zobrazenia $\varphi\colon \mathbb{R}^3 \to \mathbb{R}^4$, $\psi\colon \mathbb{R}^4 \to \mathbb{R}^3$. Napíšte matice lineárnych zobrazení $\psi \circ \varphi\colon \mathbb{R}^3 \to \mathbb{R}^3$, $\varphi \circ \psi\colon \mathbb{R}^4 \to \mathbb{R}^4$ vzhľadom na kanonické bázy vektorových priestorov $\mathbb{R}^3$ resp. $\mathbb{R}^4$.

6.13. Nech S, T sú lineárne podpriestory vektorového priestoru V.

(a) Priradením $(\boldsymbol{x},\boldsymbol{y}) \mapsto \boldsymbol{x}+\boldsymbol{y}$ je definované surjektívne lineárne zobrazenie $\mu\colon S\times T \to S+T$. Dokážte.

(b) Nájdite jadro $\operatorname{Ker}\mu$ a dokážte, že μ je injektívne práve vtedy, keď $S\cap T = \{\mathbf{0}\}$.

(c) Odvoďte z toho, že vektorové priestory $S\times T$ a $S+T$ sú izomorfné práve vtedy, keď $S+T = S\oplus T$.

(d) Zovšeobecnite na ľubovoľný konečný počet lineárnych podpriestorov (pozri cvičenie 4.8).

6.14. *Galileova transformácia* „časopriestoru“ $\mathbb{R}^4$ je daná priradením $(t,x,y,z)^\mathsf{T} \mapsto (t',x',y',z')^\mathsf{T}$, kde $t'=t$, $x'=x-v_xt$, $y'=y-v_yt$, $z'=z-v_zt$, pričom vektor $\boldsymbol{v}=(v_x,v_y,v_z)^\mathsf{T} \in \mathbb{R}^3$ interpretujeme ako rýchlosť (pozri príklad 6.4.7).

(a) Napíšte maticu $\boldsymbol{G}_{\boldsymbol{v}}$ Galileovej transformácie.

(b) Uvažujme dvoch pozorovateľov P a P'. Vysvetlite, za akých podmienok je uvedenou transformáciou sprostredkovaný vzťah medzi časmi a priestorovými súradnicami udalostí z hľadiska pozorovateľov P a P'. (Riešte otázku v rámci klasickej fyziky, bez ohľadu na relativistické efekty.)

(c) Vynásobením matíc dokážte vzťah $\boldsymbol{G}_{\boldsymbol{u}} \cdot \boldsymbol{G}_{\boldsymbol{v}} = \boldsymbol{G}_{\boldsymbol{u}+\boldsymbol{v}}$. Vysvetlite, prečo sa tento vzťah nazýva *klasickým pravidlom skladania rýchlostí.*

6.15. (a) Dokážte nasledujúcu konečnorozmernú verziu *Hahnovej-Banachovej vety*: Nech V je konečnorozmerný vektorový priestor nad poľom K, $S\subset V$ je vlastný lineárny podpriestor a $\psi\colon S\to K$ je lineárny funkcionál. Potom pre každé $\boldsymbol{x}\in V\smallsetminus S$ existuje lineárny funkcionál $\varphi\in V^*$ taký, že $\varphi(\boldsymbol{x})\neq 0$ a $\varphi\restriction S = \psi$.

(b) Zovšeobecnite tvrdenie (a) aj na nekonečnorozmerné vektorové priestory. (*Návod*: Predpokladajte, že V má Hamelovu bázu X takú, že $\boldsymbol{x}\in X$ a $X\cap S$ je Hamelova báza podpriestoru S – pozri záverečnú časť paragrafu 5.5).

(c) Pre každý vektor $\mathbf{0}\neq\boldsymbol{x}\in V$ existuje lineárny funkcionál $\varphi\in V^*$ taký, že $\varphi(\boldsymbol{x})\neq 0$. Dokážte jednak priamo podľa vzoru (a) resp. (b), jednak z nich odvoďte ako špeciálny prípad.

6.16. Nech K je konečné pole, ktoré má práve q prvkov. Potom q je mocninou nejakého prvočísla. Dokážte. (*Návod*: Ukážte, že K nesie prirodzenú štruktúru vektorového priestoru nad poľom $\mathbb{Z}_p$, kde $p=\operatorname{char}K$.)

6.17. Nech K je konečné pole, $q=\#K$ a U, V sú vektorové priestory nad K s konečnými rozmermi $\dim U = m$, $\dim V = n$. Akú dimenziu a koľko prvkov má vektorový priestor $\mathcal{L}(V,U)$ všetkých lineárnych zobrazení $V\to U$? Porovnajte s dimenziou a počtom prvkov vektorového priestoru U^V všetkých zobrazení $V\to U$. Aká je pravdepodobnosť, že náhodne vybrané zobrazenie $f\in U^V$ bude lineárne?

6.18. Nech $(V_i)_{i\in I}$ je súbor vektorových priestorov nad poľom K indexovaný množinou I. Označme $\prod_{i\in I} V_i$ množinu všetkých funkcií f s definičným oborom Itakých, že $f(i) \in V_i$ pre každé $i \in I$. Na množine $W = \prod_{i\in I} V_i$ definujme operácie súčtu a skalárneho násobku po zložkách, t. j. $(f+g)(i) = f(i) + g(i)$ a $(cf)(i) = cf(i)$ pre $f, g \in W$, $c \in K$, $i \in I$.

(a) Dokážte, že $W = \prod_{i\in I} V_i$ s práve definovanými operáciami tvorí vektorový priestor nad poľom K. Ako vyzerá $\mathbf{0} \in W$ a opačný prvok k prvku $f \in W$? Vektorový priestor $\prod_{i\in I} V_i$ nazývame *priamym súčinom* systému vektorových priestorov $(V_i)_{i\in I}$.

(b) Predpokladajme, že $I = \{1,\dots,n\}$ je konečná množina. Dokážte, že oba priame súčiny $V_1 \times \dots, \times V_n$ a $\prod_{i\in I} V_i$ sú lineárne izomorfné. Snažte sa nájsť čo najprirodzenejší izomorfizmus.

(c) Predpokladajme, že $V_i = V$ pre každé $i \in I$. Dokážte (presnejšie, uvedomte si, lebo dokazovať nie je veľmi čo), že priamy súčin $\prod_{i\in I} V_i$ splýva s priestorom funkcií V^I (pozri príklad 1.6.5). Z toho dôvodu nazývame priestor V^I aj *priamou mocninou vektorového priestoru* V.

(d) Pre každé $j \in I$ označme $\pi_j\colon \prod_{i\in I} V_i \to V_j$ zobrazenie dané predpisom $\pi_j(f) = f(j)$ pre $f \in \prod_{i\in I} V_i$. Dokážte, že tzv. *kanonická projekcia* π_j je surjektívne lineárne zobrazenie. Ako vyzerá jadro $\operatorname{Ker}\pi_j$?

(e) Pre každé $j \in I$ označme $\sigma_j\colon V_j \to \prod_{i\in I} V_i$ zobrazenie dané predpisom $\sigma_j(\boldsymbol{x})(j) = \boldsymbol{x}$ a $\sigma_j(\boldsymbol{x})(i) = \mathbf{0}$, ak $i \neq j$, pre $\boldsymbol{x} \in V_j$. Dokážte, že tzv. *kanonické vnorenie* σ_j je injektívne lineárne zobrazenie, teda je to izomorfizmus priestoru V_j na lineárny podpriestor $\operatorname{Im}\sigma_j \subseteq \prod_{i\in I} V_i$. Ako vyzerá obraz $\operatorname{Im}\sigma_j$? Dokážte, že pre každé $j \in I$ platí $W = \operatorname{Im}\sigma_j \oplus \operatorname{Ker}\pi_j$.

6.19. V označení cvičenia 6.18 stotožníme každý vektor $\boldsymbol{x} \in V_j$ s funkciou $\sigma_j(\boldsymbol{x}) \in W$ a následne aj každý z priestorov V_j s jeho izomorfnou kópiou $\operatorname{Im}\sigma_j \subseteq W$.

(a) Predpokladajme, že $I = \{1,\dots,n\}$ je konečná množina. Dokážte, že $W = \prod_{i\in I} V_i$ je priamym súčtom svojich lineárnych podpriestorov $V_1,\dots,V_n$.

(b) Predpokladajme, že I je nekonečná množina. Dokážte, že zjednotenie podpriestorov V_i, $i \in I$ *negeneruje* vektorový priestor $W = \prod_{i\in I} V_i$.

(c) Označme $S = \bigoplus_{i\in I} V_i$ lineárny podpriestor priameho súčinu $\prod_{i\in I} V_i$ generovaný zjednotením jeho lineárnych podpriestorov $V_i = \operatorname{Im}\sigma_i$, $i \in I$. Dokážte, že S pozostáva zo všetkých funkcií $f \in W$, pre ktoré je množina $\{i \in I;\ f(i) \neq \mathbf{0}\}$ konečná. Vektorový priestor $S = \bigoplus_{i\in I} V_i$ nazývame *priamym súčtom* systému vektorových priestorov $(V_i)_{i\in I}$.

(d) Dokážte, že každý prvok $f \in S$ možno jednoznačne až na poradie a nulové sčítance vyjadriť v tvare tzv. *formálneho súčtu* $f = \boldsymbol{x}_1 + \dots + \boldsymbol{x}_m$ konečného počtu sčítancov $\boldsymbol{x}_k \in V_{i_k}$, kde indexy $i_1,\dots,i_k \in I$ sú navzájom rôzne.

(e) Sformulujte pravidlá pre rovnosť a sčítanie formálnych súčtov $\boldsymbol{x}_1 + \dots + \boldsymbol{x}_m$, $\boldsymbol{y}_1 + \dots + \boldsymbol{y}_n$ tak, aby boli v zhode s rovnosťou vektorov a sčítaním v priestore $\bigoplus_{i\in I} V_i$.

(f) Predpokladajme, že $V_i = V$ pre každé $i \in I$. Dokážte, že priamy súčet $\bigoplus_{i\in I} V_i$ splýva s podpriestorom $V^{(I)}$ priamej mocniny V^I z príkladu 4.1.3.

6.20. Nech X je Hamelova báza vektorového priestoru V. Dokážte, že priestor V je izomorfný s priamym súčtom $\bigoplus_{\boldsymbol{x}\in X}[\boldsymbol{x}]$ svojich jednorozmerných lineárnych podpriestorov $[\boldsymbol{x}]$, $\boldsymbol{x} \in X$, a tým pádom tiež s vektorovým priestorom $K^{(X)}$.

7. Inverzné matice a zmena bázy

V tejto kapitole zavedieme pojem *inverznej matice* k danej štvorcovej matici a dáme ho do súvisu s pojmom inverzného lineárneho zobrazenia. Ďalej sa naučíme počítať inverzné matice a matice prechodu z jednej súradnej bázy do druhej. Nakoniec preskúmame vplyv zmeny bázy na maticu lineárneho zobrazenia. Začneme však s pojmom *hodnosti matice*, ktorý nám umožní rozhodnúť o existencii inverznej matice a – ako uvidíme neskôr – bude nám ešte veľakrát užitočný.

V celej kapitole K označuje pevné pole, m, n, p sú kladné celé čísla.

7.1 Hodnosť matice

V tomto paragrafe je potrebné rozlišovať medzi vektorovými priestormi riadkových resp. stĺpcových vektorov. Nebudeme teda používať nešpecifikované označenie K^n, ale priestor riadkových vektorov budeme značiť $K^{1\times n}$ a priestor stĺpcových vektorov $K^{n\times 1}$.

Pripomeňme, že $\boldsymbol{r}_i(\boldsymbol{A}) \in K^{1\times n}$ označuje i-tý riadok a $\boldsymbol{s}_j(\boldsymbol{A}) \in K^{m\times 1}$ zase j-tý stĺpec matice $\boldsymbol{A} = (a_{ij})_{m\times n}$. Túto maticu teda môžeme zapísať v blokových tvaroch

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{r}_1(\boldsymbol{A}) \\ \boldsymbol{r}_2(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{r}_m(\boldsymbol{A}) \end{pmatrix} = (\boldsymbol{s}_1(\boldsymbol{A}), \boldsymbol{s}_2(\boldsymbol{A}), \ldots, \boldsymbol{s}_n(\boldsymbol{A})).$$

Riadkovou hodnosťou $h_r(\boldsymbol{A})$ matice $\boldsymbol{A}$ nazývame dimenziu lineárneho podpriestoru vektorového priestoru $K^{1\times n}$ generovaného riadkami matice $\boldsymbol{A}$. Podobne, *stĺpcovou hodnosťou* $h_s(\boldsymbol{A})$ matice $\boldsymbol{A}$ nazývame dimenziu lineárneho podpriestoru vektorového priestoru $K^{m\times 1}$ generovaného stĺpcami matice $\boldsymbol{A}$. Teda

$$h_r(\boldsymbol{A}) = \dim[\boldsymbol{r}_1(\boldsymbol{A}), \boldsymbol{r}_2(\boldsymbol{A}), \ldots, \boldsymbol{r}_m(\boldsymbol{A})],$$
$$h_s(\boldsymbol{A}) = \dim[\boldsymbol{s}_1(\boldsymbol{A}), \boldsymbol{s}_2(\boldsymbol{A}), \ldots, \boldsymbol{s}_n(\boldsymbol{A})]$$

Označme $\varphi\colon K^{n\times 1} \to K^{m\times 1}$ lineárne zobrazenie dané predpisom $\varphi(\boldsymbol{x}) = \boldsymbol{A}\cdot\boldsymbol{x}$ pre $\boldsymbol{x} \in K^{n\times 1}$. Pripomeňme, že hodnosťou lineárneho zobrazenia φ nazývame dimenziu jeho obrazu, t. j. $h(\varphi) = \dim \operatorname{Im}\varphi$. V našom prípade zrejme platí $h(\varphi) = h_s(\boldsymbol{A})$, keďže lineárny podpriestor $\operatorname{Im}\varphi \subseteq K^{m\times 1}$ je generovaný stĺpcami matice $\boldsymbol{A}$.

7.1.1. Lema. *Nech* $\boldsymbol{A} \in K^{m\times n}$.

(a) *Nech matica* $\boldsymbol{B}$ *vznikne z matice* $\boldsymbol{A}$ *vykonaním jednej (inak ľubovoľnej) ERO. Potom* $[\boldsymbol{r}_1(\boldsymbol{A}), \boldsymbol{r}_2(\boldsymbol{A}), \dots, \boldsymbol{r}_m(\boldsymbol{A})] = [\boldsymbol{r}_1(\boldsymbol{B}), \boldsymbol{r}_2(\boldsymbol{B}), \dots, \boldsymbol{r}_m(\boldsymbol{B})]$.

(b) *Nech matica* $\boldsymbol{C}$ *vznikne z matice* $\boldsymbol{A}$ *vykonaním jednej (inak ľubovoľnej) ESO. Potom* $[\boldsymbol{s}_1(\boldsymbol{A}), \boldsymbol{s}_2(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A})] = [\boldsymbol{s}_1(\boldsymbol{C}), \boldsymbol{s}_2(\boldsymbol{C}), \dots, \boldsymbol{s}_n(\boldsymbol{C})]$.

Dôkaz. Zrejme pre ľubovoľné vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ v každom vektorovom priestore V a ľubovoľný skalár $c \in K$ platí:

$$\begin{aligned}
[\boldsymbol{u}_1, \dots, \boldsymbol{u}_i, \dots, \boldsymbol{u}_j, \dots, \boldsymbol{u}_k] &= [\boldsymbol{u}_1, \dots, \boldsymbol{u}_j, \dots, \boldsymbol{u}_i, \dots, \boldsymbol{u}_k],\\
[\boldsymbol{u}_1, \dots, \boldsymbol{u}_i, \dots, \boldsymbol{u}_k] &= [\boldsymbol{u}_1, \dots, c\boldsymbol{u}_i, \dots, \boldsymbol{u}_k] \qquad (\text{ak } c \neq 0),\\
[\boldsymbol{u}_1, \dots, \boldsymbol{u}_i, \dots, \boldsymbol{u}_j, \dots, \boldsymbol{u}_k] &= [\boldsymbol{u}_1, \dots, \boldsymbol{u}_i, \dots, c\boldsymbol{u}_i + \boldsymbol{u}_j, \dots, \boldsymbol{u}_k].
\end{aligned}$$

7.1.2. Tvrdenie. *Pre každú maticu* $\boldsymbol{A} \in K^{m\times n}$ *platí* $h_r(\boldsymbol{A}) = h_s(\boldsymbol{A})$.

Dôkaz. Upravme $\boldsymbol{A}$ pomocou ERO na redukovaný stupňovitý tvar $\boldsymbol{B} \in K^{m\times n}$ a označme k počet nenulových riadkov v matici $\boldsymbol{B}$. Podľa práve dokázanej lemy platí $[\boldsymbol{r}_1(\boldsymbol{A}), \dots, \boldsymbol{r}_m(\boldsymbol{A})] = [\boldsymbol{r}_1(\boldsymbol{B}), \dots, \boldsymbol{r}_m(\boldsymbol{B})]$. Preto tiež $h_r(\boldsymbol{A}) = h_r(\boldsymbol{B})$. Keďže nenulové riadky matice $\boldsymbol{B}$ sú zrejme lineárne nezávislé (rozmyslite si prečo), $h_r(\boldsymbol{B}) = k$, čo je vlastne počet stĺpcov matice $\boldsymbol{B}$, v ktorých sa nachádza vedúci prvok nejakého jej riadku. Označme $1 \leq j_1 < \ldots < j_k \leq n$ indexy týchto stĺpcov. Podľa tvrdenia 4.5.3 vektory $\boldsymbol{s}_{j_1}(\boldsymbol{A}), \dots, \boldsymbol{s}_{j_k}(\boldsymbol{A})$ sú lineárne nezávislé a platí $[\boldsymbol{s}_1(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A})] = [\boldsymbol{s}_{j_1}(\boldsymbol{A}), \dots, \boldsymbol{s}_{j_k}(\boldsymbol{A})]$. Preto tiež $h_s(\boldsymbol{A}) = k = h_r(\boldsymbol{A})$.

Keďže riadková a stĺpcová hodnosť ľubovoľnej matice $\boldsymbol{A}$ splývajú, túto ich spoločnú hodnotu budeme odteraz značiť jednoducho $h(\boldsymbol{A})$ a nazývať *hodnosťou matice* $\boldsymbol{A}$. Zrejme pre $\boldsymbol{A} \in K^{m\times n}$ je $h(\boldsymbol{A}) \leq \min(m, n)$.

Práve vykonané úvahy majú dva bezprostredné dôsledky.

7.1.3. Tvrdenie. *Nech* $\boldsymbol{A} \in K^{m\times n}$. *Potom* $h(\boldsymbol{A}) = h(\boldsymbol{A}^\mathsf{T})$.

7.1.4. Tvrdenie. *Nech* $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n \in K^{m\times 1}$ *sú ľubovoľné vektory a* $\boldsymbol{A} \in K^{m\times n}$ *je matica taká, že* $\boldsymbol{s}_j(\boldsymbol{A}) = \boldsymbol{u}_j$ *pre* $1 \leq j \leq n$. *Potom*

(a) $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ *sú lineárne nezávislé práve vtedy, keď* $h(\boldsymbol{A}) = n$;

(b) $[\boldsymbol{u}_1, \dots, \boldsymbol{u}_n] = K^{m\times 1}$ *práve vtedy, keď* $h(\boldsymbol{A}) = m$.

Všimnite si, že prípad (a) môže nastať iba vtedy, keď $n \leq m$; naopak, (b) môže nastať jedine za predpokladu $m \leq n$.

Sami si sformulujte a premyslite analogické tvrdenia pre riadkové vektory.

Ešte si dokážeme jeden odhad hodnosti súčinu matíc pomocou hodností jednotlivých činiteľov.

7.1.5. Tvrdenie. *Nech* $\boldsymbol{A} \in K^{m\times n}$, $\boldsymbol{B} \in K^{n\times p}$. *Potom*

$$h(\boldsymbol{A}\cdot\boldsymbol{B}) \leq \min\bigl(h(\boldsymbol{A}), h(\boldsymbol{B})\bigr).$$

Dôkaz. Označme $\varphi\colon K^n \to K^m$, $\psi\colon K^p \to K^n$ lineárne zobrazenia dané predpismi $\varphi(\boldsymbol{x}) = \boldsymbol{A}\cdot\boldsymbol{x}$ pre $\boldsymbol{x} \in K^n$ resp. $\psi(\boldsymbol{y}) = \boldsymbol{B}\cdot\boldsymbol{y}$ pre $\boldsymbol{y} \in K^p$. Zrejme $\mathrm{Im}(\varphi\circ\psi) \subseteq \mathrm{Im}\,\varphi$, preto

$$h(\boldsymbol{A}\cdot\boldsymbol{B}) = h(\varphi\circ\psi) \leq h(\varphi) = h(\boldsymbol{A}).$$

S využitím toho druhý potrebný odhad už dostaneme priamym výpočtom

$$h(\boldsymbol{A}\cdot\boldsymbol{B}) = h\bigl((\boldsymbol{A}\cdot\boldsymbol{B})^\mathsf{T}\bigr) = h\bigl(\boldsymbol{B}^\mathsf{T}\cdot\boldsymbol{A}^\mathsf{T}\bigr) \leq h\bigl(\boldsymbol{B}^\mathsf{T}\bigr) = h(\boldsymbol{B}).$$

7.2 Inverzné matice a inverzné lineárne zobrazenia

Nech $\boldsymbol{A} \in K^{n\times n}$, t. j. $\boldsymbol{A}$ je *štvorcová* matica typu $n\times n$. *Inverznou maticou* k matici $\boldsymbol{A}$ rozumieme maticu $\boldsymbol{B} \in K^{n\times n}$ takú, že

$$\boldsymbol{A}\cdot\boldsymbol{B} = \boldsymbol{I}_n = \boldsymbol{B}\cdot\boldsymbol{A}.$$

Zrejme k danej štvorcovej matici $\boldsymbol{A}$ existuje najviac jedna inverzná matica (rozmyslite si prečo). Túto jednoznačne určenú maticu (ak existuje) budeme značiť $\boldsymbol{A}^{-1}$.

Nasledujúca veta je bezprostredným dôsledkom súvisu medzi lineárnymi zobrazeniami a ich maticami.

7.2.1. Veta. *Nech U, V sú vektorové priestory nad poľom K a* $\dim U = \dim V = n$. *Nech ďalej $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ sú nejaké bázy v U, resp. vo V a $\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ je matica lineárneho zobrazenia $\varphi\colon V \to U$ vzhľadom na bázy $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$. Potom k matici $\boldsymbol{A}$ existuje inverzná matica $\boldsymbol{A}^{-1}$ práve vtedy, keď k zobrazeniu φ existuje inverzné zobrazenie φ^{-1}. V tom prípade $\boldsymbol{A}^{-1}$ je maticou lineárneho zobrazenia $\varphi^{-1}\colon U \to V$ vzhľadom na bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, t. j.*

$$\boldsymbol{A}^{-1} = \bigl((\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}\bigr)^{-1} = \bigl(\varphi^{-1}\bigr)_{\boldsymbol{\beta},\boldsymbol{\alpha}}.$$

Hovoríme, že štvorcová matica $\boldsymbol{A} \in K^{n\times n}$ je *regulárna*, ak k nej existuje inverzná matica $\boldsymbol{A}^{-1}$; v opačnom prípade hovoríme, že $\boldsymbol{A}$ je *singulárna*.

7.2.2. Veta. *Matica $\boldsymbol{A} \in K^{n\times n}$ je regulárna práve vtedy, keď $h(\boldsymbol{A}) = n$.*

Dôkaz. Označme $\varphi\colon K^n \to K^n$ lineárnu transformáciu danú predpisom $\varphi(\boldsymbol{x}) = \boldsymbol{A}\cdot\boldsymbol{x}$ pre $\boldsymbol{x} \in K^n$. K matici $\boldsymbol{A}$ existuje inverzná matica $\boldsymbol{A}^{-1}$ práve vtedy, keď k zobrazeniu φ existuje inverzné zobrazenie φ^{-1}, t. j. práve vtedy, keď φ je bijekcia. Podľa dôsledku 6.2.4 to nastane práve vtedy, keď φ je surjekcia, čiže $\operatorname{Im}\varphi = K^n$, čo je ekvivalentné s rovnosťou $\dim \operatorname{Im}\varphi = n$. Na dokončenie dôkazu si stačí spomenúť, že $h(\boldsymbol{A}) = h(\varphi) = \dim \operatorname{Im}\varphi$.

Z praktických dôvodov bude užitočné si uvedomiť, že na to, aby sme sa presvedčili, že matica $\boldsymbol{B} \in K^{n\times n}$ je inverzná k matici $\boldsymbol{A} \in K^{n\times n}$, stačí overiť len jednu (a to hocktorú) z rovností $\boldsymbol{A}\cdot\boldsymbol{B} = \boldsymbol{I}_n$, $\boldsymbol{B}\cdot\boldsymbol{A} = \boldsymbol{I}_n$.

7.2.3. Tvrdenie. *Pre ľubovoľné $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$ platí $\boldsymbol{A}\cdot\boldsymbol{B} = \boldsymbol{I}_n$ práve vtedy, keď $\boldsymbol{B}\cdot\boldsymbol{A} = \boldsymbol{I}_n$.*

Dôkaz. Označme $\varphi, \psi \colon K^n \to K^n$ lineárne transformácie dané pre $\boldsymbol{x} \in K^n$ predpismi $\varphi(\boldsymbol{x}) = \boldsymbol{A}\cdot\boldsymbol{x}$, resp. $\psi(\boldsymbol{x}) = \boldsymbol{B}\cdot\boldsymbol{x}$. Nech $\boldsymbol{A}\cdot\boldsymbol{B} = \boldsymbol{I}_n$. To nastane práve vtedy, keď $\varphi\circ\psi = \mathrm{id}_{K^n}$. Z toho vyplýva, že φ je surjekcia a ψ je injekcia (pozri paragraf 0.3). Keďže φ, ψ sú lineárne transformácie konečnorozmerného vektorového priestoru, podľa dôsledku 6.2.4 to znamená, že φ aj ψ sú bijekcie, teda lineárne izomorfizmy, a $\psi = \varphi^{-1}$. Potom však $\boldsymbol{B} = \boldsymbol{A}^{-1}$, preto tiež $\boldsymbol{B}\cdot\boldsymbol{A} = \boldsymbol{I}_n$. Obrátená implikácia vyplýva zo symetrie tvrdenia.

S využitím posledného tvrdenia si ako cvičenie overte nasledujúce vzorce, z ktorých už vyplýva zvyšok tvrdenia.

7.2.4. Tvrdenie. *Nech $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$ sú regulárne matice. Potom aj matice $\boldsymbol{A}^{-1}$, $\boldsymbol{A}\cdot\boldsymbol{B}$ a $\boldsymbol{A}^{\mathsf{T}}$ sú regulárne a platí:*

$$\left(\boldsymbol{A}^{-1}\right)^{-1} = \boldsymbol{A}, \qquad (\boldsymbol{A}\cdot\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\cdot\boldsymbol{A}^{-1}, \qquad \left(\boldsymbol{A}^{\mathsf{T}}\right)^{-1} = \left(\boldsymbol{A}^{-1}\right)^{\mathsf{T}}.$$

7.3 Výpočet inverznej matice

Prakticky všetky úlohy lineárnej algebry, s ktorými sme sa doteraz stretli, sme riešili tak, že sme danú situáciu reprezentovali nejakou vhodnou maticou, tú sme ďalej pomocou ERO upravili na redukovaný stupňovitý tvar a tento výsledný tvar sme potom interpretovali v závislosti na charaktere pôvodnej úlohy. Prezraďme už vopred, že zatiaľ sme všetky úlohy, ktoré sa riešia úpravou matíc pomocou ERO prípadne ESO, zďaleka nevyčerpali. Naopak, táto metóda nás bude v lineárnej algebre neustále sprevádzať.

Skôr než pristúpime k ďalšiemu využitiu tejto metódy, tentoraz pri výpočte inverznej matice, však bude potrebné si uvedomiť, že ERO aj ESO možno realizovať pomocou násobenia matíc.

7.3.1. Tvrdenie. *Nech $\boldsymbol{A} \in K^{m\times n}$.*

(a) *Nech $\boldsymbol{B} \in K^{m\times n}$ vznikne z $\boldsymbol{A}$ vykonaním jednej (inak ľubovoľnej) ERO. Označme $\boldsymbol{E}$ maticu, ktorá vznikne z matice $\boldsymbol{I}_m$ vykonaním tej istej ERO. Potom $\boldsymbol{B} = \boldsymbol{E}\cdot\boldsymbol{A}$.*

(b) *Nech $\boldsymbol{C} \in K^{m\times n}$ vznikne z $\boldsymbol{A}$ vykonaním jednej (inak ľubovoľnej) ESO. Označme $\boldsymbol{F}$ maticu, ktorá vznikne z matice $\boldsymbol{I}_n$ vykonaním tej istej ESO. Potom $\boldsymbol{C} = \boldsymbol{A} \cdot \boldsymbol{F}$.*

Dôkaz. Možno overiť priamym výpočtom pre každý jednotlivý druh ERO resp. ESO. Ako cvičenie si to skúste napr. pre maticu $\boldsymbol{A}$ typu 3×2, resp. 3×3.

Štvorcové matice $\boldsymbol{E} \in K^{n\times n}$, ktoré vzniknú z jednotkovej matice $\boldsymbol{I}_n$ vykonaním jedinej ERO alebo ESO, nazývame *elementárne matice*. Posledná veta teda hovorí, že ľubovoľnú ERO (ESO) na matici $\boldsymbol{A}$ možno realizovať vynásobením matice $\boldsymbol{A}$ vhodnou elementárnou maticou $\boldsymbol{E}$ zľava (sprava).

Návod na výpočet inverznej matice k danej štvorcovej matici $\boldsymbol{A} \in K^{n\times n}$ si možno najľahšie zapamätať v tvare nasledujúcej schémy:

$$(\boldsymbol{A}\,|\,\boldsymbol{I}_n) \xrightarrow{\text{ERO}} (\boldsymbol{I}_n\,|\,\boldsymbol{A}^{-1}).$$

Tento postup má navyše tú výhodu, že sa nemusíme vopred starať, či inverzná matica k matici $\boldsymbol{A}$ existuje alebo nie. Ak $\boldsymbol{A}^{-1}$ existuje, tak ju nakoniec vypočítame, ak neexistuje, tak to odhalíme počas nášho výpočtu a ďalej v ňom nebudeme pokračovať. Celý postup si teraz vysvetlíme trochu podrobnejšie.

Bloková matica $(\boldsymbol{A}\,|\,\boldsymbol{I}_n)$ vznikne tak, že matice $\boldsymbol{A}$ a $\boldsymbol{I}_n$ jednoducho napíšeme vedľa seba. Túto maticu teraz budeme upravovať pomocou ERO tak, aby sme v ľavej časti z matice $\boldsymbol{A}$ dostali jednotkovú maticu $\boldsymbol{I}_n$. Akonáhle sa nám to podarí, matica v pravej časti výslednej blokovej matice je už hľadaná matica $\boldsymbol{A}^{-1}$. Ak sa nám to nepodarí, t. j. matica $\boldsymbol{A}$ nie je riadkovo ekvivalentná s jednotkovou maticou (čo nastane práve vtedy, keď $h(\boldsymbol{A}) < n$, a spoznáme to podľa toho, že sa nám v ľavej časti objaví nejaký nulový riadok), tak inverzná matica k matici $\boldsymbol{A}$ neexistuje.

Korektnosť uvedeného postupu vyplýva z nasledujúceho očividného tvrdenia a skutočnosti, že ERO možno reprezentovať násobením elementárnymi maticami zľava. Taktiež tu hrá úlohu fakt, že pre $\boldsymbol{B} \in K^{n\times n}$ platí $\boldsymbol{B} = \boldsymbol{A}^{-1}$ práve vtedy, keď $\boldsymbol{B} \cdot \boldsymbol{A} = \boldsymbol{I}_n$, uvedený v tvrdení 7.2.3

7.3.2. Tvrdenie. *Nech $\boldsymbol{A} \in K^{n\times n}$ a $\boldsymbol{E}_1, \boldsymbol{E}_2, \dots, \boldsymbol{E}_k \in K^{n\times n}$ sú elementárne matice také, že $\boldsymbol{E}_k \cdot \ldots \cdot \boldsymbol{E}_2 \cdot \boldsymbol{E}_1 \cdot \boldsymbol{A} = \boldsymbol{I}_n$. Potom $\boldsymbol{A}^{-1} = \boldsymbol{E}_k \cdot \ldots \cdot \boldsymbol{E}_2 \cdot \boldsymbol{E}_1$.*

Poznamenajme, že k rovnakému cieľu vedie tiež postup reprezentovaný schémou:

$$\left(\frac{\boldsymbol{A}}{\boldsymbol{I}_n}\right) \xrightarrow{\text{ESO}} \left(\frac{\boldsymbol{I}_n}{\boldsymbol{A}^{-1}}\right).$$

Rozmyslite si prečo a sformulujte príslušné tvrdenie.

Z práve vykonaných úvah vyplývajú nasledujúce tri dôsledky. Posledný z nich je čiastočným obrátením odhadu hodnosti súčinu matíc za predpokladu regularity aspoň jedného z činiteľov.

7.3.3. Tvrdenie. *Matica* $\boldsymbol{A} \in K^{n\times n}$ *je regulárna práve vtedy, keď ju možno rozložiť na súčin* $\boldsymbol{A} = \boldsymbol{E}_1 \cdot \ldots \cdot \boldsymbol{E}_k$ *konečného počtu elementárnych matíc* $\boldsymbol{E}_1, \ldots, \boldsymbol{E}_k \in K^{n\times n}$.

7.3.4. Tvrdenie. *Pre ľubovoľné* $\boldsymbol{A}, \boldsymbol{B} \in K^{m\times n}$ *platí:*
(a) $\boldsymbol{A}$ *je riadkovo ekvivalentná s* $\boldsymbol{B}$ *práve vtedy, keď existuje regulárna matica* $\boldsymbol{P} \in K^{m\times m}$ *taká, že* $\boldsymbol{A} = \boldsymbol{P} \cdot \boldsymbol{B}$;
(b) $\boldsymbol{A}$ *je stĺpcovo ekvivalentná s* $\boldsymbol{B}$ *práve vtedy, keď existuje regulárna matica* $\boldsymbol{Q} \in K^{n\times n}$ *taká, že* $\boldsymbol{A} = \boldsymbol{B} \cdot \boldsymbol{Q}$.

7.3.5. Tvrdenie. *Nech* $\boldsymbol{A} \in K^{m\times n}$, $\boldsymbol{P} \in K^{m\times m}$, $\boldsymbol{Q} \in K^{n\times n}$, *pričom* $\boldsymbol{P}, \boldsymbol{Q}$ *sú regulárne matice. Potom*

$$h(\boldsymbol{A}) = h(\boldsymbol{P} \cdot \boldsymbol{A}) = h(\boldsymbol{A} \cdot \boldsymbol{Q}) = h(\boldsymbol{P} \cdot \boldsymbol{A} \cdot \boldsymbol{Q}).$$

Trochu všeobecnejšie možno uvedené úvahy použiť na násobenie ľubovoľnej matice vhodného rozmeru maticou $\boldsymbol{A}^{-1}$ (ak existuje) zľava resp. sprava. Tieto operácie možno uskutočniť pre regulárnu $\boldsymbol{A} \in K^{n\times n}$ a ľubovoľné $\boldsymbol{B} \in K^{n\times m}$, $\boldsymbol{C} \in K^{m\times n}$ podľa nasledujúcich schém:

$$(\boldsymbol{A} \,|\, \boldsymbol{B}) \xrightarrow{\text{ERO}} (\boldsymbol{I}_n \,|\, \boldsymbol{A}^{-1} \cdot \boldsymbol{B}),$$

$$\left(\frac{\boldsymbol{A}}{\boldsymbol{C}}\right) \xrightarrow{\text{ESO}} \left(\frac{\boldsymbol{I}_n}{\boldsymbol{C} \cdot \boldsymbol{A}^{-1}}\right).$$

Ešte si všimnime, že v špeciálnom prípade sme niečo podobné vlastne robili už dávno, pri riešení sústav lineárnych rovníc úpravou na redukovaný stupňovitý tvar pomocou ERO. Aj tento postup totiž možno vyjadriť pomocou schémy

$$(\boldsymbol{A} \,|\, \boldsymbol{b}) \xrightarrow{\text{ERO}} (\boldsymbol{B} \,|\, \boldsymbol{c}),$$

ktorá má pre regulárnu $\boldsymbol{A} \in K^{n\times n}$ tvar

$$(\boldsymbol{A} \,|\, \boldsymbol{b}) \xrightarrow{\text{ERO}} (\boldsymbol{I}_n \,|\, \boldsymbol{A}^{-1} \cdot \boldsymbol{b}).$$

Ako vedľajší produkt našich úvah tak dostávame nasledujúci výsledok o riešení sústav n lineárnych rovníc o n neznámych.

7.3.6. Veta. *Nech* $\boldsymbol{A} \in K^{n\times n}$, $\boldsymbol{b} \in K^n$. *Ak* $\boldsymbol{A}$ *je regulárna, tak sústava* $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ *má jediné riešenie* $\boldsymbol{x} = \boldsymbol{A}^{-1} \cdot \boldsymbol{b}$.

7.4 Matica prechodu

Nech V je vektorový priestor nad poľom K a $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$, $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ sú jeho dve bázy. *Maticou prechodu* z bázy $\boldsymbol{\beta}$ do bázy $\boldsymbol{\alpha}$ nazývame maticu identického zobrazenia $\mathrm{id}_V \colon V \to V$ vzhľadom na bázy $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, ktorú značíme $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$. Teda

$$\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = (\mathrm{id}_V)_{\boldsymbol{\alpha},\boldsymbol{\beta}}.$$

Podľa definície matice lineárneho zobrazenia vzhľadom na dané bázy (pozri paragraf 6.4), stĺpce matice prechodu $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ sú tvorené súradnicami vektorov bázy $\boldsymbol{\beta}$ vzhľadom na bázu $\boldsymbol{\alpha}$, t. j. $\boldsymbol{s}_j(\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}) = (\boldsymbol{v}_j)_{\boldsymbol{\alpha}}$ pre $1 \leq j \leq n$. Teda

$$\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \big((\boldsymbol{v}_1)_{\boldsymbol{\alpha}}, (\boldsymbol{v}_2)_{\boldsymbol{\alpha}}, \dots, (\boldsymbol{v}_n)_{\boldsymbol{\alpha}}\big),$$

a podľa vety 6.4.1 je táto matica jednoznačne určená podmienkou transformácie súradníc

$$(\boldsymbol{x})_{\boldsymbol{\alpha}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot (\boldsymbol{x})_{\boldsymbol{\beta}}$$

pre ľubovoľné $\boldsymbol{x} \in V$.

Ak do zrejmej rovnosti $\boldsymbol{x} = \boldsymbol{\alpha} \cdot (\boldsymbol{x})_{\boldsymbol{\alpha}}$ (pozri paragraf 5.3) budeme za $\boldsymbol{x}$ postupne dosadzovať vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ bázy $\boldsymbol{\beta}$, s využitím vzťahu pre stĺpce súčinu matíc z paragrafu 2.3 dostaneme

$$\boldsymbol{v}_j = \boldsymbol{\alpha} \cdot (\boldsymbol{v}_j)_{\boldsymbol{\alpha}} = \boldsymbol{\alpha} \cdot \boldsymbol{s}_j(\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}) = \boldsymbol{s}_j(\boldsymbol{\alpha} \cdot \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}})$$

pre každé $1 \leq j \leq n$. Tým sme dostali ďalší dôležitý vzťah, ktorý jednoznačne charakterizuje maticu prechodu $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$:

$$\boldsymbol{\alpha} \cdot \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \boldsymbol{\beta}.$$

(Podotýkame, že súčin $\boldsymbol{\alpha} \cdot \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ treba chápať v zmysle paragrafu 2.3)

Priradenie $\boldsymbol{u}_i \mapsto \boldsymbol{v}_i$ pre $i \leq n$ možno jednoznačne rozšíriť do bijektívnej lineárnej transformácie $\vartheta \colon V \to V$. Skrátene píšeme $\vartheta(\boldsymbol{\alpha}) = \boldsymbol{\beta}$. Matica transformácie ϑ vzhľadom na bázu $\boldsymbol{\alpha}$ je opäť

$$(\vartheta)_{\boldsymbol{\alpha}} = (\vartheta)_{\boldsymbol{\alpha},\boldsymbol{\alpha}} = \big((\vartheta\boldsymbol{u}_1)_{\boldsymbol{\alpha}}, \dots, (\vartheta\boldsymbol{u}_n)_{\boldsymbol{\alpha}}\big) = \big((\boldsymbol{v}_1)_{\boldsymbol{\alpha}}, \dots, (\boldsymbol{v}_n)_{\boldsymbol{\alpha}}\big) = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}.$$

Zhrnutím vykonaných úvah dostávame štyri ekvivalentné charakterizácie matice prechodu.

7.4.1. Tvrdenie. *Nech $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$, $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ sú dve bázy n-rozmerného vektorového priestoru V nad poľom K. Potom pre ľubovoľnú maticu $\boldsymbol{P} \in K^{n\times n}$ nasledujúce podmienky sú ekvivalentné:*

(i) $\boldsymbol{P} = (\mathrm{id}_V)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$, *t. j. $\boldsymbol{P}$ je matica prechodu z bázy $\boldsymbol{\beta}$ do bázy $\boldsymbol{\alpha}$;*

(ii) $(\boldsymbol{x})_{\boldsymbol{\alpha}} = \boldsymbol{P} \cdot (\boldsymbol{x})_{\boldsymbol{\beta}}$ *pre každé $\boldsymbol{x} \in V$;*

(iii) $\boldsymbol{\alpha} \cdot \boldsymbol{P} = \boldsymbol{\beta}$;
(iv) $\boldsymbol{P} = (\vartheta)_{\boldsymbol{\alpha}}$, *kde* $\vartheta\colon V \to V$ *je lineárna transformácia taká, že* $\vartheta(\boldsymbol{\alpha}) = \boldsymbol{\beta}$.

Z definície matice prechodu a vety 6.4.2 okamžite vyplývajú nasledujúce rovnosti.

7.4.2. Tvrdenie. *Nech* $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ *sú bázy konečnorozmerného vektorového priestoru* V *nad poľom* K. *Potom*

$$\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\alpha}} = \boldsymbol{I}_n, \qquad \boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\alpha}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}^{-1}, \qquad \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot \boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\gamma}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\gamma}}.$$

Z druhej z uvedených podmienok vidno, že matica prechodu $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ je vždy *regulárna*. Taktiež naopak, každá regulárna matica $\boldsymbol{P} \in K^{n\times n}$ je maticou prechodu medzi vhodnou dvojicou báz.

7.4.3. Tvrdenie. *Nech* V *je* n*-rozmerný vektorový priestor nad poľom* K, $\boldsymbol{P} = (p_{ij}) \in K^{n\times n}$ *je ľubovoľná regulárna matica a* $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ *je nejaká báza vo* V. *Položme*

$$\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) = \boldsymbol{\alpha} \cdot \boldsymbol{P}, \qquad \boldsymbol{\gamma} = (\boldsymbol{w}_1, \dots, \boldsymbol{w}_n) = \boldsymbol{\alpha} \cdot \boldsymbol{P}^{-1},$$

t. j. pre $1 \le j \le n$ *platí* $\boldsymbol{v}_j = p_{1j}\boldsymbol{u}_1 + \ldots + p_{nj}\boldsymbol{u}_n$ *a* $\boldsymbol{w}_j = q_{1j}\boldsymbol{u}_1 + \ldots + q_{nj}\boldsymbol{u}_n$, *kde* $\boldsymbol{P}^{-1} = (q_{ij})_{n\times n}$. *Potom* $\boldsymbol{P}$ *je maticou prechodu z bázy* $\boldsymbol{\beta}$ *do bázy* $\boldsymbol{\alpha}$ *a taktiež z bázy* $\boldsymbol{\alpha}$ *do bázy* $\boldsymbol{\gamma}$, *čiže*

$$\boldsymbol{P} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \boldsymbol{P}_{\boldsymbol{\gamma},\boldsymbol{\alpha}}.$$

Špeciálne, $\boldsymbol{P}$ *je maticou prechodu z bázy* $(\boldsymbol{s}_1(\boldsymbol{P}), \dots, \boldsymbol{s}_n(\boldsymbol{P}))$ *do bázy* $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ *v* K^n *a taktiež z bázy* $\boldsymbol{\varepsilon}$ *do bázy* $(\boldsymbol{s}_1(\boldsymbol{P}^{-1}), \dots, \boldsymbol{s}_n(\boldsymbol{P}^{-1}))$.

V prípade, keď $V = K^n$ je priestor stĺpcových vektorov, možno každú jeho bázu $\boldsymbol{\alpha}$ stotožniť s príslušnou regulárnou maticou, ktorej stĺpcami sú vektory danej bázy. Pri takomto stotožnení je návod na výpočet matice prechodu obsiahnutý v nasledujúcom tvrdení.

7.4.4. Tvrdenie. *Nech* $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$, $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ *sú dve bázy stĺpcového vektorového priestoru* K^n. *Potom* $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \boldsymbol{\alpha}^{-1} \cdot \boldsymbol{\beta}$.

Dôkaz. Z podmienky $\boldsymbol{\alpha} \cdot \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \boldsymbol{\beta}$ okamžite vyplýva požadovaná rovnosť.

To nám dáva návod na výpočet matice prechodu pre bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ vektorového priestoru K^n podľa už známej schémy

$$(\boldsymbol{\alpha} \,|\, \boldsymbol{\beta}) \xrightarrow{\text{ERO}} (\boldsymbol{I}_n \,|\, \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}) = (\boldsymbol{\varepsilon} \,|\, \boldsymbol{\alpha}^{-1} \cdot \boldsymbol{\beta}).$$

7.5 Matice lineárneho zobrazenia vzhľadom na rôzne bázy

V tomto paragrafe sa budeme zaoberať vplyvom zmeny báz na maticu lineárneho zobrazenia, presnejšie, vzťahom medzi maticami daného lineárneho zobrazenia vzhľadom na rôzne dvojice báz.

7.5.1. Veta. *Nech V_1, V_2 sú konečnorozmerné vektorové priestory nad poľom K, $\varphi\colon V_1 \to V_2$ je lineárne zobrazenie, $\boldsymbol{\alpha}_1$, $\boldsymbol{\beta}_1$ sú dve bázy priestoru V_1 a $\boldsymbol{\alpha}_2$, $\boldsymbol{\beta}_2$ sú dve bázy priestoru V_2. Potom*

$$(\varphi)_{\boldsymbol{\beta}_2,\boldsymbol{\beta}_1} = \boldsymbol{P}_{\boldsymbol{\beta}_2,\boldsymbol{\alpha}_2} \cdot (\varphi)_{\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1} \cdot \boldsymbol{P}_{\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1}.$$

Dôkaz. Označme $\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1}$, $\boldsymbol{B} = (\varphi)_{\boldsymbol{\beta}_2,\boldsymbol{\beta}_1}$ matice lineárneho zobrazenia φ vzhľadom na bázy $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, resp. bázy $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$. Pre ľubovoľné $\boldsymbol{x} \in V_1$ platí:

$$\begin{aligned}\boldsymbol{B} \cdot (\boldsymbol{x})_{\boldsymbol{\beta}_1} &= (\varphi\boldsymbol{x})_{\boldsymbol{\beta}_2} = \boldsymbol{P}_{\boldsymbol{\beta}_2,\boldsymbol{\alpha}_2} \cdot (\varphi\boldsymbol{x})_{\boldsymbol{\alpha}_2} \\ &= \boldsymbol{P}_{\boldsymbol{\beta}_2,\boldsymbol{\alpha}_2} \cdot \boldsymbol{A} \cdot (\boldsymbol{x})_{\boldsymbol{\alpha}_1} = \boldsymbol{P}_{\boldsymbol{\beta}_2,\boldsymbol{\alpha}_2} \cdot \boldsymbol{A} \cdot \boldsymbol{P}_{\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1} \cdot (\boldsymbol{x})_{\boldsymbol{\beta}_1}.\end{aligned}$$

Na základe vety 6.4.1 z toho okamžite vyplýva dokazovaná rovnosť

$$\boldsymbol{B} = \boldsymbol{P}_{\boldsymbol{\beta}_2,\boldsymbol{\alpha}_2} \cdot \boldsymbol{A} \cdot \boldsymbol{P}_{\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1}.$$

Poslednú transformačnú formulku si možno najľahšie zapamätať pomocou nasledujúceho diagramu:

$$\begin{tikzcd}
(V_1,\boldsymbol{\alpha}_1) \arrow[r, "\boldsymbol{A}"] & (V_2,\boldsymbol{\alpha}_2) \arrow[d, "\boldsymbol{P}_{\boldsymbol{\beta}_2,\boldsymbol{\alpha}_2}"] \\
(V_1,\boldsymbol{\beta}_1) \arrow[u, "\boldsymbol{P}_{\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1}"] \arrow[r, "\boldsymbol{B}"'] & (V_2,\boldsymbol{\beta}_2)
\end{tikzcd}$$

Nezabudnite, že zobrazenia skladáme „v obrátenom poradí“, a tomu musí zodpovedať aj „obrátené poradie“ násobenia matíc!

7.5.2. Príklad. Nech $\varphi\colon K^n \to K^m$ je lineárne zobrazenie a $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ sú nejaké bázy priestorov K^m resp. K^n. Označme $\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$, $\boldsymbol{M} = (\varphi)_{\boldsymbol{\varepsilon}^{(m)},\boldsymbol{\varepsilon}^{(n)}}$ matice zobrazenia φ vzhľadom na bázy $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$ resp. vzhľadom na kanonické bázy $\boldsymbol{\varepsilon}^{(n)}$, $\boldsymbol{\varepsilon}^{(m)}$. Podľa poslednej vety platí:

$$\begin{aligned}\boldsymbol{A} &= \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\varepsilon}^{(m)}} \cdot \boldsymbol{M} \cdot \boldsymbol{P}_{\boldsymbol{\varepsilon}^{(n)},\boldsymbol{\beta}}, \\ \boldsymbol{M} &= \boldsymbol{P}_{\boldsymbol{\varepsilon}^{(m)},\boldsymbol{\alpha}} \cdot \boldsymbol{A} \cdot \boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\varepsilon}^{(n)}}.\end{aligned}$$

Ak stotožníme každú bázu s regulárnou maticou, ktorej stĺpce sú vektory tejto bázy, tak uvedené rovnosti nadobudnú tvar

$$\begin{aligned}\boldsymbol{A} &= \boldsymbol{\alpha}^{-1} \cdot \boldsymbol{I}_m \cdot \boldsymbol{M} \cdot \boldsymbol{I}_n^{-1} \cdot \boldsymbol{\beta} = \boldsymbol{\alpha}^{-1} \cdot \boldsymbol{M} \cdot \boldsymbol{\beta}, \\ \boldsymbol{M} &= \boldsymbol{I}_m^{-1} \cdot \boldsymbol{\alpha} \cdot \boldsymbol{A} \cdot \boldsymbol{\beta}^{-1} \cdot \boldsymbol{I}_n = \boldsymbol{\alpha} \cdot \boldsymbol{A} \cdot \boldsymbol{\beta}^{-1},\end{aligned}$$

umožňujúci priamy výpočet jednej z matíc $\boldsymbol{A}$, $\boldsymbol{M}$ na základe znalosti báz $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ a druhej z nich.

Položme si teraz obrátenú otázku. Za akých podmienok sú matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{m\times n}$ maticami toho istého lineárneho zobrazenia $\varphi\colon V \to U$ vzhľadom na nejaké dve (možno no nie nutne rôzne) dvojice báz konečnorozmerných vektorových priestorov U, V? Odpoveď na ňu dáva nasledujúca veta.

7.5.3. Veta. *Nech U je m-rozmerný a V je n-rozmerný vektorový priestor nad poľom K. Potom pre ľubovoľné matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{m\times n}$ nasledujúce podmienky sú ekvivalentné:*

(i) $\boldsymbol{A}, \boldsymbol{B}$ sú maticami toho istého lineárneho zobrazenia $\varphi\colon V \to U$ vzhľadom na nejaké dve (možno no nie nutne rôzne) dvojice báz priestorov U, V;

(ii) existujú regulárne matice $\boldsymbol{P} \in K^{m\times m}$, $\boldsymbol{Q} \in K^{n\times n}$ také, že $\boldsymbol{B} = \boldsymbol{P}\cdot\boldsymbol{A}\cdot\boldsymbol{Q}$;

(iii) $h(\boldsymbol{A}) = h(\boldsymbol{B})$.

Dôkaz. Ekvivalencia (i) $\Leftrightarrow$ (ii) je priamym dôsledkom vety 7.5.1 a tvrdenia 7.4.3 Implikácia (ii) $\Rightarrow$ (iii) vyplýva z tvrdenia 7.3.5

Zostáva dokázať (iii) $\Rightarrow$ (ii). Označme $h = h(\boldsymbol{A}) = h(\boldsymbol{B})$. Pomocou ERO upravíme $\boldsymbol{A}$ aj $\boldsymbol{B}$ na redukovaný trojuholníkový tvar $\boldsymbol{A}' = \boldsymbol{P}_1\cdot\boldsymbol{A}$, resp. $\boldsymbol{B}' = \boldsymbol{P}_2\cdot\boldsymbol{B}$, kde $\boldsymbol{P}_1$, $\boldsymbol{P}_2$ sú regulárne matice. Zrejme $\boldsymbol{A}'$, $\boldsymbol{B}'$ majú rovnaký počet nenulových riadkov rovný h. $\boldsymbol{A}'$ aj $\boldsymbol{B}'$ možno ďalej pomocou ESO upraviť na blokový tvar

$$\boldsymbol{A}'' = \boldsymbol{A}'\cdot\boldsymbol{Q}_1 = \begin{pmatrix} \boldsymbol{I}_h & \boldsymbol{0}_{h,n-h} \\ \boldsymbol{0}_{m-h,h} & \boldsymbol{0}_{m-h,n-h} \end{pmatrix} = \boldsymbol{B}'\cdot\boldsymbol{Q}_2 = \boldsymbol{B}'',$$

kde $\boldsymbol{Q}_1$, $\boldsymbol{Q}_2$ sú regulárne matice. Stačí pomocou vedúcich prvkov jednotlivých riadkov vynulovať prípadné ďalšie nenulové prvky týchto riadkov a, ak treba, vymeniť poradie niektorých stĺpcov. Potom $\boldsymbol{P}_1\cdot\boldsymbol{A}\cdot\boldsymbol{Q}_1 = \boldsymbol{P}_2\cdot\boldsymbol{B}\cdot\boldsymbol{Q}_2$, teda $\boldsymbol{B} = \boldsymbol{P}_2^{-1}\cdot\boldsymbol{P}_1\cdot\boldsymbol{A}\cdot\boldsymbol{Q}_1\cdot\boldsymbol{Q}_2^{-1}$ a matice $\boldsymbol{P} = \boldsymbol{P}_2^{-1}\cdot\boldsymbol{P}_1$, $\boldsymbol{Q} = \boldsymbol{Q}_1\cdot\boldsymbol{Q}_2^{-1}$ sú zrejme regulárne.

Na základe dôkazu tejto vety okamžite dostávame záverečný výsledok.

7.5.4. Veta. *Pre každé lineárne zobrazenie $\varphi\colon V \to U$ medzi konečnorozmernými vektorovými priestormi nad poľom K možno zvoliť bázu $\boldsymbol{\beta}$ priestoru V a bázu $\boldsymbol{\alpha}$ priestoru U tak, že φ má vzhľadom na bázy $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$ maticu v blokovom tvare*

$$(\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \begin{pmatrix} \boldsymbol{I}_h & \boldsymbol{0}_{h,n-h} \\ \boldsymbol{0}_{m-h,h} & \boldsymbol{0}_{m-h,n-h} \end{pmatrix},$$

kde $n = \dim V$, $m = \dim U$ a $h = h(\varphi)$.

Skúste si túto vetu dokázať priamo a bližšie špecifikovať bázy $\boldsymbol{\beta}$ a $\boldsymbol{\alpha}$. (*Návod*: Spomeňte si na dôkaz vety 6.2.3 o dimenzii jadra a obrazu.)

7.6 Pohyblivé bázy

Lineárnu transformáciu konečnorozmerného vektorového priestoru V je často výhodné popisovať zadaním obrazov vektorov nejakej bázy. Pritom pre rôzne transformácie môžu byť výhodné rôzne bázy. Napr. pri sledovaní nejakého pohybujúceho sa objektu je prirodzené udávať jeho polohu vzhľadom na bázu spojenú s „nehybným“ pozorovateľom. Zmenu orientácie takého objektu spôsobenú jeho rotáciou je však výhodnejšie udávať vzhľadom na bázu spojenú s týmto objektom.

Uvažujme bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ vektorového priestoru V. Nech $\vartheta\colon V \to V$ je lineárna transformácia taká, že $\vartheta(\boldsymbol{\alpha}) = \boldsymbol{\beta}$. Ak prechod od vektora $\boldsymbol{x} \in V$ k jeho obrazu $\vartheta(\boldsymbol{x}) \in V$ chápeme ako výsledok nejakého pohybu (napr. otočenia) v priestore V, tak bázu $\boldsymbol{\beta}$ možno považovať za novú polohu premiestnenej bázy $\boldsymbol{\alpha}$. Ak $\boldsymbol{\gamma}$ je ďalšia báza vo V a $\eta\colon V \to V$ je lineárna transformácia taká, že $\eta(\boldsymbol{\beta}) = \boldsymbol{\gamma}$, tak bázu $\boldsymbol{\gamma}$ môžeme chápať jednak ako bázu $\boldsymbol{\beta}$ premiestnenú transformáciou η, jednak ako pôvodnú bázu $\boldsymbol{\alpha}$ premiestnenú transformáciou $\eta \circ \vartheta$.

Pre maticu zloženej transformácie $\eta \circ \vartheta$ vzhľadom na bázu $\boldsymbol{\alpha}$ dostávame popri obvyklom vyjadrení pomocou „pevnej“ bázy $\boldsymbol{\alpha}$

$$(\eta \circ \vartheta)_{\boldsymbol{\alpha}} = (\eta)_{\boldsymbol{\alpha}} \cdot (\vartheta)_{\boldsymbol{\alpha}}$$

taktiež vyjadrenie pomocou „pohyblivej“ bázy $\boldsymbol{\alpha}$, t. j. pomocou báz $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$,

$$(\eta \circ \vartheta)_{\boldsymbol{\alpha}} = (\vartheta)_{\boldsymbol{\alpha}} \cdot (\eta)_{\boldsymbol{\beta}}.$$

ktoré je zaujímavé obráteným poradím činiteľov. Stačí si uvedomiť, že zo vzťahov $\vartheta(\boldsymbol{\alpha}) = \boldsymbol{\beta}$, $\eta(\boldsymbol{\beta}) = \boldsymbol{\gamma}$ vyplýva $(\eta \circ \vartheta)(\boldsymbol{\alpha}) = \boldsymbol{\gamma}$, a ďalej podľa časti (iv) tvrdenia 7.4.1 a tvrdenia 7.4.2 tiež

$$(\eta \circ \vartheta)_{\boldsymbol{\alpha}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\gamma}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot \boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\gamma}} = (\vartheta)_{\boldsymbol{\alpha}} \cdot (\eta)_{\boldsymbol{\beta}}.$$

Toto pozorovanie možno matematickou indukciou rozšíriť na ľubovoľný konečný počet báz a lineárnych transformácií.

7.6.1. Veta. *Nech $\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k$ sú bázy konečnorozmerného vektorového priestoru V a $\vartheta_1, \dots, \vartheta_k\colon V \to V$ sú lineárne transformácie také, že $\vartheta_i(\boldsymbol{\alpha}_{i-1}) = \boldsymbol{\alpha}_i$ pre $1 \le i \le k$. Potom maticu zloženej transformácie $\vartheta = \vartheta_k \circ \ldots \circ \vartheta_1$ vzhľadom na pôvodnú bázu $\boldsymbol{\alpha}_0$ možno vyjadriť ako súčin*

$$(\vartheta)_{\boldsymbol{\alpha}_0} = (\vartheta_1)_{\boldsymbol{\alpha}_0} \cdot \ldots \cdot (\vartheta_k)_{\boldsymbol{\alpha}_{k-1}}$$

matíc $(\vartheta_i)_{\boldsymbol{\alpha}_{i-1}}$ dielčích transformácií ϑ_i vzhľadom na jednotlivé polohy $\boldsymbol{\alpha}_{i-1}$ pohyblivej bázy $\boldsymbol{\alpha}_0$.

7.6.2. Príklad. Nájdeme maticu lineárnej transformácie $\Omega\colon \mathbb{R}^3 \to \mathbb{R}^3$ ktorá vznikne zložením otočenia Ω_1 okolo osi $z = [\boldsymbol{e}_3]$ o uhol $\pi/4$ s otočením Ω_2 okolo osi so smerovým vektorom $\Omega_1(\boldsymbol{e}_1)$ o uhol $\pi/3$.

Lineárna transformácia Ω_1 má v báze $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ maticu

$$(\Omega_1)_{\boldsymbol{\varepsilon}} = \begin{pmatrix} \cos(\pi/4) & -\sin(\pi/4) & 0 \\ \sin(\pi/4) & \cos(\pi/4) & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 & 0 \\ \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Lineárna transformácia Ω_2 má v báze $\boldsymbol{\beta} = \Omega_1(\boldsymbol{\varepsilon}) = (\Omega_1(\boldsymbol{e}_1), \Omega_1(\boldsymbol{e}_2), \boldsymbol{e}_3)$ (t. j. v „novej“ báze $\boldsymbol{\varepsilon}$) maticu

$$(\Omega_2)_{\boldsymbol{\beta}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\pi/3) & -\sin(\pi/3) \\ 0 & \sin(\pi/6) & \cos(\pi/6) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -\sqrt{3}/2 \\ 0 & \sqrt{3}/2 & 1/2 \end{pmatrix}.$$

Preto

$$\begin{aligned} (\Omega)_{\boldsymbol{\varepsilon}} &= (\Omega_2 \circ \Omega_1)_{\boldsymbol{\varepsilon}} = (\Omega_1)_{\boldsymbol{\varepsilon}} \cdot (\Omega_2)_{\boldsymbol{\beta}} \\ &= \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 & 0 \\ \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -\sqrt{3}/2 \\ 0 & \sqrt{3}/2 & 1/2 \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/4 & \sqrt{6}/4 \\ \sqrt{2}/2 & \sqrt{2}/4 & -\sqrt{6}/4 \\ 0 & \sqrt{3}/2 & 1/2 \end{pmatrix}. \end{aligned}$$

Na základe rovnosti $(\Omega)_{\boldsymbol{\varepsilon}} = (\Omega_2)_{\boldsymbol{\varepsilon}} \cdot (\Omega_1)_{\boldsymbol{\varepsilon}}$ je už teraz hračkou vypočítať aj samotnú maticu otočenia Ω_2 vzhľadom na kanonickú bázu $\boldsymbol{\varepsilon}$:

$$\begin{aligned} (\Omega_2)_{\boldsymbol{\varepsilon}} = (\Omega)_{\boldsymbol{\varepsilon}} \cdot (\Omega_1)_{\boldsymbol{\varepsilon}}^{-1} &= \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/4 & \sqrt{6}/4 \\ \sqrt{2}/2 & \sqrt{2}/4 & -\sqrt{6}/4 \\ 0 & \sqrt{3}/2 & 1/2 \end{pmatrix} \cdot \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ -\sqrt{2}/2 & \sqrt{2}/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3/4 & 1/4 & \sqrt{6}/4 \\ 1/4 & 3/4 & -\sqrt{6}/4 \\ -\sqrt{6}/4 & \sqrt{6}/4 & 1/2 \end{pmatrix}. \end{aligned}$$

Cvičenia

7.1. Nech U, V sú vektorové priestory nad poľom K a $\varphi\colon V \to U$ je lineárne zobrazenie. Potom pre ľubovoľné vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n \in V$ platí $\varphi[\boldsymbol{v}_1, \dots, \boldsymbol{v}_n] = [\varphi(\boldsymbol{v}_1), \dots, \varphi(\boldsymbol{v}_n)]$. Dokážte. Odvoďte z toho, že ak $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ generujú V, tak $\varphi(\boldsymbol{v}_1), \dots, \varphi(\boldsymbol{v}_n)$ generujú

$\operatorname{Im}\varphi$. Špeciálne, stĺpce matice $\boldsymbol{A} \in K^{m\times n}$ generujú lineárny podpriestor $\operatorname{Im}\varphi \subseteq K^m$, kde $\varphi\colon K^n \to K^m$ je dané predpisom $\varphi(\boldsymbol{x}) = \boldsymbol{A}\cdot\boldsymbol{x}$.

7.2. Určte hodnosť matice $\boldsymbol{A}$ nad poľom K:

(a) $K = \mathbb{R}$, $\boldsymbol{A} = \begin{pmatrix} 2 & 2 & -1 \\ 3 & 0 & 1 \end{pmatrix}$;

(b) $K = \mathbb{C}$, $\boldsymbol{A} = \begin{pmatrix} 1+\mathrm{i} & 1+3\mathrm{i} & 2 \\ 3+\mathrm{i} & 5+5\mathrm{i} & 4-2\mathrm{i} \end{pmatrix}$;

(c) $K = \mathbb{Z}_7$, $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 5 & 4 \\ 1 & 3 & 0 \end{pmatrix}$;

(d) $K = \mathbb{Z}_{17}$, $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 5 & 4 \\ 1 & 3 & 0 \end{pmatrix}$.

7.3. (a) Dokážte vzorce z tvrdenia 7.2.4

(b) Dokážte, že pre ľubovoľnú maticu $\boldsymbol{A} \in K^{n\times n}$ a $k, l \in \mathbb{N}$ platí $\boldsymbol{A}^k \cdot \boldsymbol{A}^l = \boldsymbol{A}^{k+l}$ a $(\boldsymbol{A}^k)^l = \boldsymbol{A}^{kl}$.

(c) Pre regulárnu maticu $\boldsymbol{A} \in K^{n\times n}$ rozšírte definíciu jej mocniny $\boldsymbol{A}^k$ na ľubovovoľný celočíselný exponent k a dokážte rovnosti $\boldsymbol{A}^k \cdot \boldsymbol{A}^l = \boldsymbol{A}^{k+l}$, $(\boldsymbol{A}^k)^l = \boldsymbol{A}^{kl}$ pre všetky $k, l \in \mathbb{Z}$ (porovnaj s cvičením 2.13).

(d) Za akých okolností platí pre matice $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$ rovnosť $(\boldsymbol{A}\cdot\boldsymbol{B})^k = \boldsymbol{A}^k \cdot \boldsymbol{B}^k$ pre ľubovoľné $k \in \mathbb{N}$?

(e) Za akých okolností platí pre regulárne matice $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$ rovnosť $(\boldsymbol{A}\cdot\boldsymbol{B})^k = \boldsymbol{B}^k \cdot \boldsymbol{A}^k$ pre ľubovoľné $k \in \mathbb{Z}$?

(f) Nájdite príklad regulárnych matíc $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{2\times 2}$ takých, že všetky štyri matice $(\boldsymbol{A}\cdot\boldsymbol{B})^{-2}$, $(\boldsymbol{B}\cdot\boldsymbol{A})^{-2}$, $\boldsymbol{A}^{-2}\cdot\boldsymbol{B}^{-2}$, $\boldsymbol{B}^{-2}\cdot\boldsymbol{A}^{-2}$ sú rôzne. Dá sa táto úloha riešiť aj bez počítania inverzných matíc?

7.4. Dokážte tvrdenie 7.3.5.

7.5. Zistite, či uvedená matica $\boldsymbol{A}$ nad poľom K je regulárna; v tom prípade vypočítajte k nej inverznú maticu $\boldsymbol{A}^{-1}$:

(a) $K = \mathbb{Q}$, $\boldsymbol{A} = \begin{pmatrix} 3 & 2 & 1 \\ 4 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}$;

(b) $K = \mathbb{R}$, $\boldsymbol{A} = \begin{pmatrix} 1 & \sqrt{2} & \sqrt{6} \\ 0 & 1 & \sqrt{3} \\ 0 & 0 & 1 \end{pmatrix}$;

(c) $K = \mathbb{C}$, $\boldsymbol{A} = \begin{pmatrix} 1+2\mathrm{i} & 1-\mathrm{i} \\ \mathrm{i}-2 & 1 \end{pmatrix}$;

(d) $K = \mathbb{C}$, $\boldsymbol{A} = \begin{pmatrix} 1+2\mathrm{i} & 1-\mathrm{i} \\ \mathrm{i}-2 & 1+\mathrm{i} \end{pmatrix}$;

(e) $K = \mathbb{Z}_2$, $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$;

(f) $K = \mathbb{Z}_3$, $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.

7.6. Pre matice $\boldsymbol{A}, \boldsymbol{B}$ nad poľom $\mathbb{R}$ vypočítajte maticu $\boldsymbol{C} = \boldsymbol{A}^{-1}\cdot\boldsymbol{B}$ (ak $\boldsymbol{A}$ je regulárna):

(a) $\boldsymbol{A} = \begin{pmatrix} -2 & 1 \\ 3 & -2 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 2 & 1 & -1 \\ 3 & 2 & -5 \end{pmatrix}$;

(b) $\boldsymbol{A} = \begin{pmatrix} 2 & 1 & 3 \\ 1 & -2 & 0 \\ 3 & 0 & 3 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 2 & 1 \\ 2 & 2 \\ -2 & 0 \end{pmatrix}$

Ako skúšku správnosti vypočítajte maticu $\boldsymbol{A}\cdot\boldsymbol{C}$ – mali by ste dostať $\boldsymbol{B}$.

7.7. Pre matice $\boldsymbol{A}, \boldsymbol{B}$ nad poľom $\mathbb{R}$ vypočítajte maticu $\boldsymbol{C} = \boldsymbol{A}\cdot\boldsymbol{B}^{-1}$ (ak $\boldsymbol{B}$ je regulárna)

(a) $\boldsymbol{A} \begin{pmatrix} -2 & 1 & 1 \\ -3 & 0 & 5 \\ 0 & 2 & 3 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 2 & 5 \\ 1 & 0 & 0 \end{pmatrix}$;

(b) $\boldsymbol{A} = \begin{pmatrix} 2 & 5 \\ 1 & 0 \\ 7 & 4 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix}$.

Ako skúšku správnosti vypočítajte maticu $\boldsymbol{C}\cdot\boldsymbol{B}$ – mali by ste dostať $\boldsymbol{A}$.

7.8. Ná základe skúseností nadobudnutých v cvičeniach 7.5, 7.6 a 7.7 dokážte tvrdenie 7.3.1

7.9. Nájdite inverzné matice k maticiam $\boldsymbol{R}_\alpha$, $\boldsymbol{S}_\alpha$ z príkladov 6.4.3, 6.4.4 Vysvetlite geometrický význam získaných výsledkov. Riešte rovnakú úlohu pre matice otočení okolo súradných osí v $\mathbb{R}^3$. Všimnite si, že pre všetky uvažované matice $\boldsymbol{A}$ platí $\boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathsf{T}}$.

7.10. Bázy $\boldsymbol{\alpha}, \boldsymbol{\beta}$ vektorového priestoru $\mathbb{R}^3$ sú tvorené stĺpcami matíc $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ resp. $\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$.

(a) Nájdite matice prechodu $\boldsymbol{P}_{\boldsymbol{\varepsilon},\boldsymbol{\alpha}}$, $\boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\varepsilon}}$, $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ a $\boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\alpha}}$.

(b) Vektor $\boldsymbol{x} \in \mathbb{R}^3$ má vzhľadom na bázu $\boldsymbol{\beta}$ súradnice $(2,7,1)^{\mathsf{T}}$. Nájdite vektor $\boldsymbol{x}$ ako aj jeho súradnice vzhľadom na bázu $\boldsymbol{\alpha}$.

7.11. Báza $\boldsymbol{\alpha}$ vektorového priestoru $\mathbb{R}^3$ je tvorená stĺpcami matice $\left(\begin{smallmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ 0 & 2 & 1 \end{smallmatrix}\right)$. Nájdite bázy $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ priestoru $\mathbb{R}^3$, ak poznáte matice prechodu $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \left(\begin{smallmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 0 & 1 \end{smallmatrix}\right)$ a $\boldsymbol{P}_{\boldsymbol{\gamma},\boldsymbol{\alpha}} = \left(\begin{smallmatrix} 0 & 1 & 0 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \end{smallmatrix}\right)$. Vypočítajte matice prechodu $\boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\gamma}}$ a $\boldsymbol{P}_{\boldsymbol{\gamma},\boldsymbol{\beta}}$.

7.12. Lineárne zobrazenie $\varphi\colon \mathbb{R}^3 \to \mathbb{R}^4$ je dané predpisom $\varphi(x,y,z)^{\mathsf{T}} = (x+y, x-y, 3x+y+z, z)^{\mathsf{T}}$.

(a) Nájdite maticu zobrazenia φ vzhľadom na bázy $\boldsymbol{\beta}$ priestoru $\mathbb{R}^3$ a $\boldsymbol{\alpha}$ priestoru $\mathbb{R}^4$ tvorené stĺpcami matíc $\left(\begin{smallmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{smallmatrix}\right)$ resp. $\left(\begin{smallmatrix} 0 & 1 & 0 & 1 \\ 1 & -1 & 1 & -1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{smallmatrix}\right)$

(b) Nech $\boldsymbol{x} = (1,-4,3)^{\mathsf{T}} \in \mathbb{R}^3$. Nájdite súradnice vektora $\varphi(\boldsymbol{x}) \in \mathbb{R}^4$ vzhľadom na bázu $\boldsymbol{\alpha}$.

(c) Vektor $\boldsymbol{y} \in \mathbb{R}^3$ má vzhľadom na bázu $\boldsymbol{\beta}$ súradnice $(1,1,1)^{\mathsf{T}}$. Nájdite vektor $\varphi(\boldsymbol{y}) \in \mathbb{R}^4$.

(d) Každú z úloh (b), (c) možno riešiť dvoma spôsobmi – vysvetlite ako.

(e) Určte hodnosť h zobrazenia ψ a nájdite nejaké bázy priestorov $\mathbb{R}^3$, $\mathbb{R}^4$, vzhľadom na ktoré má matica φ blokový tvar $\left(\begin{smallmatrix} \boldsymbol{I}_h & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{smallmatrix}\right)$. Napíšte explicitne túto maticu.

7.13. Lineárne zobrazenie $\psi\colon \mathbb{R}^4 \to \mathbb{R}^3$ má vzhľadom na bázy $\boldsymbol{\beta}$ priestoru $\mathbb{R}^4$ a $\boldsymbol{\alpha}$ priestoru $\mathbb{R}^3$ tvorené stĺpcami matíc $\left(\begin{smallmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & -1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{smallmatrix}\right)$ resp. $\left(\begin{smallmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{smallmatrix}\right)$ maticu $\boldsymbol{A} = \left(\begin{smallmatrix} 0 & 5 & 0 & 7 \\ 1 & 1 & 2 & 2 \\ 5 & 2 & 9 & 0 \end{smallmatrix}\right)$.

(a) Nájdite maticu zobrazenia ψ vzhľadom na kanonické bázy $\boldsymbol{\varepsilon}^{(4)}$, $\boldsymbol{\varepsilon}^{(3)}$.

(b) Nech $\boldsymbol{u} = (2,1,0,-1)^{\mathsf{T}} \in \mathbb{R}^4$. Nájdite súradnice vektora $\psi(\boldsymbol{u}) \in \mathbb{R}^3$ vzhľadom na bázu $\boldsymbol{\alpha}$.

(c) Vektor $\boldsymbol{v} \in \mathbb{R}^4$ má vzhľadom na bázu $\boldsymbol{\beta}$ súradnice $(1,-1,1,-1)^{\mathsf{T}}$. Nájdite vektor $\psi(\boldsymbol{v}) \in \mathbb{R}^3$.

(d) Každú z úloh (b), (c) možno riešiť dvoma spôsobmi – vysvetlite ako.

(e) Určte hodnosť h zobrazenia ψ a nájdite nejaké bázy priestorov $\mathbb{R}^4$, $\mathbb{R}^3$, vzhľadom na ktoré má matica ψ blokový tvar $\left(\begin{smallmatrix} \boldsymbol{I}_h & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{smallmatrix}\right)$. Napíšte explicitne túto maticu.

7.14. Dokážte, že $\boldsymbol{\zeta} = \boldsymbol{\zeta}^{(n)} = (1, 1+x, (1+x)^2, \dots, (1+x)^n)$ je bázou vektorového priestoru $\mathbb{R}^{(n)}[x]$ a nájdite matice prechodu $\boldsymbol{P}_{\boldsymbol{\xi},\boldsymbol{\zeta}}$, $\boldsymbol{P}_{\boldsymbol{\zeta},\boldsymbol{\xi}}$, kde $\boldsymbol{\xi} = \boldsymbol{\xi}^{(n)} = (1, x, x^2, \dots, x^n)$ je kanonická báza priestoru $\mathbb{R}^{(n)}[x]$.

7.15. Podľa cvičenia 6.10 je $\boldsymbol{\eta}^{(n)} = \boldsymbol{\eta} = \big(1, x, [x]_2, \dots, [x]_n\big)$ bázou vektorového priestoru $\mathbb{R}^{(n)}[x]$. Pre $0 \le k, m \le n$ označme $s(m,k)$ prvok na mieste (k,m) matice prechodu $\boldsymbol{P}_{\boldsymbol{\xi},\boldsymbol{\eta}}$ a $S(m,k)$ prvok na mieste (k,m) matice prechodu $\boldsymbol{P}_{\boldsymbol{\eta},\boldsymbol{\xi}}$. To znamená, že pre $0 \le m \le n$ platí $[x]_m = \sum_{k=0}^{n} s(m,k) x^k$ a $x^m = \sum_{k=0}^{n} S(m,k)[x]_k$. Koeficienty $s(m,k)$, $S(m,k)$ sa nazývajú *Stirlingove čísla prvého* resp. *druhého druhu*.

(a) Dokážte, že Stirlingove čísla prvého druhu vyhovujú podmienkam $s(0,0) = 1$, $s(n,0) = 0$ pre $n \ge 1$, $s(n,n) = 1$ a $s(n,k) = 0$ pre $k > n$.

(b) Dokážte, že pre $1 \le k \le n$ platí $s(n+1,k) = s(n,k-1) - ns(n,k)$. (*Návod*: Vyjadrite polynóm $[x]_{n+1} = [x]_n(x-n)$ dvojakým spôsobom a porovnajte koeficienty pri x^k.)

(c) Označme $c(n,k)$ počet všetkých permutácií n-prvkovej množiny, pozostávajúcich z práve k disjunktných cyklov (vrátane 1-cyklov) – pozri cvičenie 0.14. Dokážte, že čísla $c(n,k)$ vyhovujú podmienkam $c(0,0) = 1$, $c(n,0) = 0$ pre $n \geq 1$, $c(n,n) = 1$ a $c(n,k) = 0$ pre $k > n$.

(d) Kombinatorickou úvahou dokážte rovnosť $c(n+1,k) = c(n,k-1) + nc(n,k)$ pre $1 \leq k \leq n$. (*Návod*: Všetky permutácie $(n+1)$-prvkovej množiny $\{0,1,\ldots,n\}$ pozostávajúce z k disjunktných cyklov rozdeľte do dvoch skupín podľa toho, či obsahujú alebo neobsahujú 1-cyklus (0). Ukážte, že prvých je $c(n,k-1)$ a druhých $nc(n,k)$.)

(e) Na základe (a), (b), (c) a (d) dokážte, že pre všetky $k,n \in \mathbb{N}$ platí $c(n,k) = (-1)^{n-k}s(n,k) = |s(n,k)|$. Čísla $c(n,k)$ sa nazývajú *znamienkovo prosté Stirlingove čísla prvého druhu.*

(f) Uvedomte si, že (c), (d) sú vlastne akýmisi pravidlami modifikovaného Pascalovho trojuholníka pre čísla $c(n,k)$. Vypočítajte hodnoty týchto čísel pre $n \leq 5$, $0 \leq k \leq n$.

(g) Dokážte, že Stirlingove čísla druhého druhu vyhovujú podmienkam $S(0,0) = 1$, $S(n,0) = 0$ pre $n \geq 1$, $S(n,n) = 1$ a $S(n,k) = 0$ pre $k > n$.

(h) Dokážte, že pre $1 \leq k \leq n$ platí $S(n+1,k) = S(n,k-1) + kS(n,k)$. (*Návod*: S využitím rovnosti $[x]_{k+1} = [x]_k(x-k)$ vyjadrite polynóm $x^{n+1} = x^n x$ dvojakým spôsobom a porovnajte koeficienty pri $[x]_k$.)

(i) Označme $\overline{S}(n,k)$ počet všetkých rozkladov n-prvkovej množiny, pozostávajúcich z práve k disjunktných množín – pozri paragraf 0.6. Dokážte, že čísla $\overline{S}(n,k)$ vyhovujú podmienkam $\overline{S}(0,0) = 1$, $\overline{S}(n,0) = 0$ pre $n \geq 1$, $\overline{S}(n,n) = 1$ a $\overline{S}(n,k) = 0$ pre $k > n$.

(j) Kombinatorickou úvahou dokážte rovnosť $\overline{S}(n+1,k) = \overline{S}(n,k-1) + k\overline{S}(n,k)$ pre $1 \leq k \leq n$. (*Návod*: Všetky rozklady $(n+1)$-prvkovej množiny $\{0,1,\ldots,n\}$ na k disjunktných podmnožín rozdeľte do dvoch skupín podľa toho, či obsahujú alebo neobsahujú jednoprvkovú množinu $\{0\}$. Ukážte, že prvých je $\overline{S}(n,k-1)$ a druhých $k\overline{S}(n,k)$.)

(k) Na základe (g), (h), (i) a (j) dokážte, že pre všetky $k,n \in \mathbb{N}$ platí $\overline{S}(n,k) = S(n,k)$.

(l) Uvedomte si, že (g), (h) sú vlastne akýmisi pravidlami modifikovaného Pascalovho trojuholníka pre čísla $S(n,k)$. Vypočítajte hodnoty týchto čísel pre $n \leq 5$, $0 \leq k \leq n$.

7.16. (a) Nech $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ sú dve bázy konečnorozmerného vektorového priestoru V a $\vartheta\colon V \to V$ je lineárna transformácia taká, že $\vartheta(\boldsymbol{\alpha}) = \boldsymbol{\beta}$. Potom $(\vartheta)_{\boldsymbol{\beta}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$. Dokážte.

(b) Nech $\boldsymbol{\gamma}$ je tretia báza vo V a $\eta\colon V \to V$ je lineárna transformácia taká, že $\eta(\boldsymbol{\beta}) = \boldsymbol{\gamma}$. Odvoďte nasledujúce vyjadrenie pre maticu zloženej transformácie $\eta \circ \vartheta$ vo výslednej báze $\boldsymbol{\gamma}$ pomocou jej *predchádzajúcich* polôh $\boldsymbol{\alpha}$ a $\boldsymbol{\beta}$: $(\eta \circ \vartheta)_{\boldsymbol{\gamma}} = (\vartheta)_{\boldsymbol{\beta}} \circ (\eta)_{\boldsymbol{\gamma}}$.

(c) Zovšeobecnite (b) na ľubovoľný počet k lineárnych transformácií a $k+1$ báz podľa vzoru vety 7.6.1

7.17. (a) Pre ľubovoľnú dvojicu vektorov $\boldsymbol{e}_i \neq \boldsymbol{e}_j$ kanonickej bázy $\boldsymbol{\varepsilon}$ v $\mathbb{R}^3$ nájdite maticu $(\Omega)_{\boldsymbol{\varepsilon}}$ transformácie Ω, ktorá vznikne zložením otočenia Ω_1 okolo osi $[\boldsymbol{e}_i]$ o uhol φ s otočením Ω_2 okolo osi $[\boldsymbol{e}_j]$ o uhol ψ.

(b) Na základe (a) nájdite maticu otočenia o uhol ψ okolo priamky, ktorá leží v rovine $[\boldsymbol{e}_i,\boldsymbol{e}_j]$ a zviera s osou $[\boldsymbol{e}_i]$ uhol φ, vzhľadom na bázu $\boldsymbol{\varepsilon}$. (Všetky otočenia orientujte podľa pravidla pravej ruky.)

(c) Riešte úlohy (a) a (b) pre niekoľko konkrétnych volieb bázických vektorov $\boldsymbol{e}_i$, $\boldsymbol{e}_j$ a uhlov φ, ψ.

7.18. Nech $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ je nejaká báza vektorového priestoru V nad poľom K a $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ je ľubovoľná usporiadaná m-tica vektorov z V. Uvedomte si, že maticu $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \big((\boldsymbol{v}_1)_{\boldsymbol{\alpha}}, \dots, (\boldsymbol{v}_m)_{\boldsymbol{\alpha}}\big) \in K^{n\times m}$ možno definovať aj za takýchto všeobecnejších pomienok a dokážte nasledujúce tvrdenia:

(a) Pre hodnosť matice $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ platí $h(\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}) = \dim[\boldsymbol{v}_1, \dots, \boldsymbol{v}_m]$.

(b) Ak $m = n$, tak matica $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ je regulárna práve vtedy, keď aj $\boldsymbol{\beta}$ je báza priestoru V.

(c) Ak $m = n$, tak pre jednoznačne určené (nie nutne bijektívne) lineárne zobrazenie $\vartheta\colon V \to V$ také, že $\vartheta(\boldsymbol{u}_i) = \boldsymbol{v}_i$ pre $i \le n$ stále platí $(\vartheta)_{\boldsymbol{\alpha}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$.

8. Afinné podpriestory a afinné zobrazenia

Keď sme v paragrafe 4.1, odvolávajúc sa na geometrický názor, ilustrovali pojem lineárneho podpriestoru, ako príklad sme uviedli, že netriválne vlastné lineárne podpriestory „nášho“ trojrozmerného vektorového priestoru $\mathbb{R}^3$ sú práve priamky a roviny *prechádzajúce počiatkom* $\mathbf{0}$. Kritickejší čitateľ mohol vtedy oprávnene zapochybovať o adekvátnosti a prirodzenosti tohto pojmu, či aspoň pocítiť potrebu zaviesť taký pojem podpriestoru, ktorý by napr. v $\mathbb{R}^3$ zahŕňal *všetky* priamky a roviny, nielen tie prechádzajúce počiatkom. Podobne sme v paragrafe 6.1 hneď po definícii pojmu lineárneho zobrazenia boli nútení učiniť poznámku o jeho odlišnosti od pojmu lineárnej funkcie používaného v matematickej analýze. Vzápätí sme prijali záväzok, že sa s týmto nedostatkom v príhodný čas vyrovnáme.

Ten čas práve nastal. Spomínané medzery zaplníme definíciami pojmu *afinného podpriestoru* alebo tiež *lineárnej variety* a pojmu *afinného zobrazenia*. *Afinita* znamená *príbuznosť, spriaznenosť*. Čitateľ sám uvidí, že objekty označené prívlastkom „afinný“ sú úzko spriaznené so zodpovedajúcimi objektmi nesúcimi prívlastok „lineárny“. Ťažiskom kapitoly bude klasifikácia vzájomnej polohy afinných podpriestorov vo vektorovom priestore.

8.1 Body a vektory

Na vektory, čiže na prvky vektorových priestorov – aspoň pokiaľ ide o konečnorozmerné vektorové priestory nad $\mathbb{R}$, – sa dívame ako na orientované úsečky s počiatkom v bode $\mathbf{0}$. Už táto veta prezrádza, že *pôvodne* sa na prvky takéhoto priestoru dívame ako na *body* a celý priestor chápeme ako *homogénny*, t. j. všetky body považujeme za rovnocenné a nevyčleňujeme v ňom nijaký privilegovaný bod za počiatok. Až na základe tohto pôvodného porozumenia dokážeme po vyčlenení nejakého počiatku O (ktorým sa môže stať ľubovoľný bod homogénneho priestoru) nahradiť bod A príslušného priestoru orientovanou úsečkou $\overrightarrow{OA}$ a následne abstrahovať od jej polohy, to znamená uvidieť za ňou *vektor* $\boldsymbol{u} = \overrightarrow{OA}$, daný len jej veľkosťou, smerom a orientáciou, ktorý možno umiestniť do ľubovoľného bodu priestoru – nielen do počiatku.

Afinným priestorom nad poľom K rozumieme vektorový priestor V nad týmto poľom, pri pohľade na ktorý sme sa vrátili k onomu pôvodnému porozumeniu jeho štruktúre a prvkom. Tie sa z vektorov stali opäť bodmi a počiatok (t. j. nulový vektor) stratil svoje výsadné postavenie – stal sa z neho bod ako každý iný.

Formálnu definíciu afinného priestoru nad poľom K tu uvádzať nebudeme. Sme totiž toho názoru, že matematická formalizácia rozdielu medzi oboma spomínanými pohľadmi na prvky vektorového priestoru by v tejto chvíli vniesla do veci viac zmätku než svetla. Celkom postačí, keď úlohu prepínača medzi oboma pohľadmi zveríme dvojiciam slov „bod“ – „vektor“ a „afinný“,– „lineárny“, prípadne „afinný“ – „vektorový“. Na druhej strane však pred nami vyvstáva potreba formálnej definície podmnožín vektorového priestoru, ktoré sú „vernými kópiami“ lineárnych podpriestorov – nemusia však prechádzať počiatkom, ale možu byť umiestnené „kdekoľvek“.

8.2 Afinné podpriestory

V celom tomto a nasledujúcich dvoch paragrafoch V označuje nejaký pevný, no inak ľubovoľný, vektorový priestor nad poľom K a m, n sú prirodzené čísla.

Kvôli pohodliu čitateľa budeme písmenami $\boldsymbol{p}$, $\boldsymbol{q}$, $\boldsymbol{r}$ (možno s indexmi) značiť výlučne body, $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ označujú zasa výlučne vektory, kým $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ môžu podľa potreby označovať body i vektory. Taktiež sa dohodneme, že rozdiel dvoch bodov budeme chápať ako vektor, kým súčet bodu a vektora ako bod.

Nech $\boldsymbol{p}, \boldsymbol{q} \in V$, $\boldsymbol{p} \neq \boldsymbol{q}$. *Priamkou* prechádzajúcou alebo tiež určenou bodmi $\boldsymbol{p}, \boldsymbol{q}$ rozumieme množinu $\ell(\boldsymbol{p}, \boldsymbol{q})$, ktorú dostaneme tak, že do bodu $\boldsymbol{p}$ umiestnime všetky možné skalárne násobky vektora $\boldsymbol{q} - \boldsymbol{p}$. Typický bod priamky $\ell(\boldsymbol{p}, \boldsymbol{q})$ má teda tvar

$$\boldsymbol{x} = \boldsymbol{p} + t(\boldsymbol{q} - \boldsymbol{p}) = (1 - t)\boldsymbol{p} + t\boldsymbol{q},$$

kde $t \in K$, čiže

$$\ell(\boldsymbol{p}, \boldsymbol{q}) = \{s\boldsymbol{p} + t\boldsymbol{q};\ s, t \in K \ \&\ s + t = 1\} \subseteq V.$$

Tento výraz má, samozrejme, zmysel aj pre $\boldsymbol{p} = \boldsymbol{q}$, vtedy však nejde o priamku ale o jednobodovú množinu $\ell(\boldsymbol{p}, \boldsymbol{p}) = \{\boldsymbol{p}\}$. Z uvedeného tvaru ihneď vidíme, že

$$\ell(\boldsymbol{p}, \boldsymbol{q}) = \ell(\boldsymbol{q}, \boldsymbol{p})$$

pre ľubovoľné $\boldsymbol{p}, \boldsymbol{q} \in V$.

Lineárnu kombináciu, t. j. výraz tvaru

$$t_0\boldsymbol{p}_0 + t_1\boldsymbol{p}_1 + \ldots + t_n\boldsymbol{p}_n = \sum_{i=0}^{n} t_i\boldsymbol{p}_i,$$

kde $n \in \mathbb{N}$, $\boldsymbol{p}_0, \ldots, \boldsymbol{p}_n \in V$, $t_0, t_1, \ldots, t_n \in K$, nazývame *afinnou* alebo tiež *barycentrickou*[1] *kombináciou* bodov $\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$, ak platí $t_0 + t_1 + \ldots +$

[1] *Barycentrum* znamená *ťažisko*.

$t_n = 1$. Výsledok afinnej kombinácie bodov budeme chápať ako bod; iné lineárne kombinácie bodov ako afinné sa v našich úvahách nevyskytnú. (Ešte si všimnite, že každá afinná kombinácia je neprázdna, t. j. obsahuje aspoň jeden člen.)

Neprázdnu podmnožinu M vektorového priestoru V nazývame jeho *afinným podpriestorom*, prípadne *lineárnou varietou* vo V, ak pre všetky body $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r} \in M$ a každý skalár $s \in K$ platí

$$s\boldsymbol{p} + (1-s)\boldsymbol{q} \in M \qquad \text{a} \qquad \boldsymbol{p} - \boldsymbol{q} + \boldsymbol{r} \in M.$$

Inak povedané, $\emptyset \neq M \subseteq V$ je afinný podpriestor, ak M je uzavretá vzhľadom na afinné kombinácie uvedených dvoch typov. Prvá podmienka znamená, že pre všetky $\boldsymbol{p}, \boldsymbol{q} \in M$ platí $\ell(\boldsymbol{p}, \boldsymbol{q}) \subseteq M$, t. j. M s každou dvojicou bodov obsahuje celú priamku nimi určenú. Druhú podmienku dodávame len kvôli poliam charakteristiky 2; ak $\operatorname{char} K \neq 2$, tak už vyplýva z prvej, takže je vlastne zbytočná. Na druhej strane, napr. vo vektorovom priestore V nad poľom $\mathbb{Z}_2$ pre všetky body $\boldsymbol{p}, \boldsymbol{q} \in V$ platí $\ell(\boldsymbol{p}, \boldsymbol{q}) = \{\boldsymbol{p}, \boldsymbol{q}\}$, teda len prvej podmienke by vyhovovala *každá podmnožina* $M \subseteq V$. Podrobnejšie o tom pojednáva nasledujúce tvrdenie, ktoré je očividne analógiou tvrdenia 4.1.2

8.2.1. Tvrdenie. *Pre ľubovoľnú neprázdnu množinu $M \subseteq V$ nasledujúce podmienky sú ekvivalentné:*

(i) M je afinný podpriestor vo V, t. j. pre ľubovoľné $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r} \in M$, $s \in K$ platí $s\boldsymbol{p} + (1-s)\boldsymbol{q} \in M$ a $\boldsymbol{p} - \boldsymbol{q} + \boldsymbol{r} \in M$;

(ii) M je uzavretá vzhľadom na ľubovoľné afinné kombinácie trojíc bodov, t. j. pre všetky $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r} \in M$, $s, t \in K$ platí $s\boldsymbol{p} + t\boldsymbol{q} + (1-s-t)\boldsymbol{r} \in M$;

(iii) M je uzavretá vzhľadom na akékoľvek afinné kombinácie, t. j. pre všetky $n \in \mathbb{N}$, body $\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_n \in M$ a skaláry $t_0, t_1, \ldots, t_n \in K$ také, že $t_0 + t_1 + \ldots + t_n = 1$, platí $t_0\boldsymbol{p}_0 + t_1\boldsymbol{p}_1 + \ldots + t_n\boldsymbol{p}_n \in M$;

Ak $\operatorname{char} K \neq 2$, tak uvedené podmienky sú navyše ekvivalentné s podmienkou

(i^-) pre ľubovoľné $\boldsymbol{p}, \boldsymbol{q} \in M$, $s \in K$ platí $s\boldsymbol{p} + (1-s)\boldsymbol{q} \in M$.

Dôkaz. Implikácie (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) sú zrejmé aj bez predpokladu $\operatorname{char} K \neq 2$. Dokážeme implikáciu (i) $\Rightarrow$ (iii); pri dôkaze vyjde navyše najavo, že pre $\operatorname{char} K \neq 2$ stačí na odvodenie záveru (iii) slabšia podmienka (i^-) miesto (i).

Predpokladajme (i) (teda tým skôr (i^-)) a pripusťme, že podmienka (iii) neplatí. Označme n najmenšie prirodzené číslo, pre ktoré to nastane. Potom $n \geq 2$ a pre všetky $k < n$ podmienka (iii) platí, čiže M je uzavretá na afinné kombinácie $\leq n$ bodov. Nech $\boldsymbol{p}_0, \ldots, \boldsymbol{p}_n \in M$, $t_0, \ldots, t_n \in K$ sú také, že $t_0 + \ldots + t_n = 1$ a $t_0\boldsymbol{p}_0 + \ldots + t_n\boldsymbol{p}_n \notin M$. Treba zvážiť dve možnosti.

(a) Ak $t_i \neq 1$ pre aspoň jedno $i \leq n$, tak bez ujmy na všeobecnosti môžeme predpokladať, že $t_0 \neq 1$. Označme

$$\boldsymbol{q} = \frac{t_1}{1-t_0}\boldsymbol{p}_1 + \ldots + \frac{t_n}{1-t_0}\boldsymbol{p}_n.$$

Keďže

$$\frac{t_1}{1-t_0} + \ldots + \frac{t_n}{1-t_0} = \frac{t_1 + \ldots + t_n}{1-t_0} = 1,$$

$\boldsymbol{q} \in M$, lebo $\boldsymbol{q}$ je afinnou kombináciou n bodov z M. Potom

$$t_0\boldsymbol{p}_0 + t_1\boldsymbol{p}_1 + \ldots + t_n\boldsymbol{p}_n = t_0\boldsymbol{p}_0 + (1-t_0)\boldsymbol{q} \in M$$

vyplýva už z podmienky (i$^-$). To je však spor.

(b) Ak $t_i = 1$ pre všetky $i \leq n$, tak ide o afinnú kombináciu $\boldsymbol{p}_0 + \boldsymbol{p}_1 + \ldots + \boldsymbol{p}_{n-1} + \boldsymbol{p}_n$ a $t_1 + \ldots + t_{n-1} = -1$. Potom $\boldsymbol{q} = -\boldsymbol{p}_1 - \ldots - \boldsymbol{p}_{n-1}$ je afinnou kombináciou $n-1$ bodov z M, teda $\boldsymbol{q} \in M$. Podľa druhej z podmienok v (i) máme

$$\boldsymbol{p}_0 + \boldsymbol{p}_1 + \ldots + \boldsymbol{p}_{n-1} + \boldsymbol{p}_n = \boldsymbol{p}_0 - \boldsymbol{q} + \boldsymbol{p}_n \in M,$$

čo je opäť spor.

Ak $\operatorname{char} K \neq 2$, možno sa zaobísť bez tejto podmienky. Keďže $\frac{1}{2} + \frac{1}{2} = 1$, už z (i$^-$) vyplýva $\frac{1}{2}\boldsymbol{p}_0 + \frac{1}{2}\boldsymbol{p}_n \in M$. Nakoľko $2 + (-1) = 1$, opäť len z (i$^-$) dostávame

$$\boldsymbol{p}_0 + \boldsymbol{p}_1 + \ldots + \boldsymbol{p}_{n-1} + \boldsymbol{p}_n = 2\left(\frac{1}{2}\boldsymbol{p}_0 + \frac{1}{2}\boldsymbol{p}_n\right) - \boldsymbol{q} \in M.$$

Poznámka. Ak $\operatorname{char} K = \infty$, tak možnosť (b) zrejme nemôže nastať, teda v uvedenom dôkaze stačí uvažovať len možnosť (a). Zároveň vidno, že v druhej časti bodu (b) je podstatný predpoklad $\operatorname{char} K \neq 2$. Bez neho by sme totiž nevedeli zaručiť existenciu prvku $1/2 = 2^{-1} \in K$ inverzného k prvku $2 = 1 + 1 \in K$.

Nasledujúca veta ukazuje, že afinné podpriestory skutočne nie sú ničím iným, než lineárnymi podpriestormi posunutými do ľubovoľného bodu príslušného vektorového priestoru.

8.2.2. Veta. *Nech $M \subseteq V$. Potom M je afinný podpriestor vo V práve vtedy, keď existuje bod $\boldsymbol{p} \in V$ a lineárny podpriestor $S \subseteq V$ taký, že*

$$M = \boldsymbol{p} + S = \{\boldsymbol{p} + \boldsymbol{u};\ \boldsymbol{u} \in S\}.$$

V tom prípade pre všetky $\boldsymbol{q}, \boldsymbol{r} \in M$, $\boldsymbol{u} \in S$ platí

$$\boldsymbol{q} - \boldsymbol{r} \in S, \qquad \boldsymbol{q} + \boldsymbol{u} \in M, \qquad M = \boldsymbol{q} + S,$$
$$S = \{\boldsymbol{x} - \boldsymbol{q};\ \boldsymbol{x} \in M\} = \{\boldsymbol{x} - \boldsymbol{y};\ \boldsymbol{x}, \boldsymbol{y} \in M\}.$$

Dôkaz. Nech $M \subseteq V$ je afinný podpriestor a $\boldsymbol{p} \in M$ je jeho ľubovoľný bod. Položme
$$S = \{\boldsymbol{x} - \boldsymbol{p};\ \boldsymbol{x} \in M\}.$$
Potom zrejme $M = \boldsymbol{p} + S$. Stačí teda dokázať, že $S \subseteq V$ je lineárny podpriestor. Keďže $\boldsymbol{p} \in M$, platí $\mathbf{0} = \boldsymbol{p} - \boldsymbol{p} \in S$. Ukážeme uzavretosť S na lineárne kombinácie. Nech $\boldsymbol{u}, \boldsymbol{v} \in S$, $a, b \in K$. Potom $\boldsymbol{u} = \boldsymbol{x} - \boldsymbol{p}$, $\boldsymbol{v} = \boldsymbol{y} - \boldsymbol{p}$ pre nejaké $\boldsymbol{x}, \boldsymbol{y} \in M$. Jednoduchý výpočet dáva

$$a\boldsymbol{u} + b\boldsymbol{v} = a(\boldsymbol{x} - \boldsymbol{p}) + b(\boldsymbol{y} - \boldsymbol{p}) = a\boldsymbol{x} + b\boldsymbol{y} + (1 - a - b)\boldsymbol{p} - \boldsymbol{p}.$$

Prvé tri sčítance tvoria afinnú kombináciu bodov z M, teda $a\boldsymbol{x} + b\boldsymbol{y} + (1 - a - b)\boldsymbol{p} \in M$; preto tiež $a\boldsymbol{u} + b\boldsymbol{v} \in S$

Nech naopak $M = \boldsymbol{p} + S$ pre nejaký bod $\boldsymbol{p} \in V$ a lineárny podpriestor $S \subseteq V$. Podľa tvrdenia 8.2.1 stačí ukázať uzavretosť M na afinné kombinácie trojíc. Nech $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in M$, $s, t \in K$. Potom $\boldsymbol{x} = \boldsymbol{p} + \boldsymbol{u}$, $\boldsymbol{y} = \boldsymbol{p} + \boldsymbol{v}$, $\boldsymbol{z} = \boldsymbol{p} + \boldsymbol{w}$ pre nejaké $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in S$. Počítajme

$$\begin{aligned} s\boldsymbol{x} + t\boldsymbol{y} + (1 - s - t)\boldsymbol{z} &= s(\boldsymbol{p} + \boldsymbol{u}) + t(\boldsymbol{p} + \boldsymbol{v}) + (1 - s - t)(\boldsymbol{p} + \boldsymbol{w}) \\ &= \boldsymbol{p} + s\boldsymbol{u} + t\boldsymbol{v} + (1 - s - t)\boldsymbol{w}. \end{aligned}$$

Keďže $t\boldsymbol{u} + s\boldsymbol{v} + (1 - s - t)\boldsymbol{w} \in S$, dostávame $s\boldsymbol{x} + t\boldsymbol{y} + (1 - s - t)\boldsymbol{z} \in M$.

Ďalšie tri podmienky možno teraz overiť priamymi výpočtami, ktoré prenechávame čitateľovi; štvrtá z nich okamžite vyplýva.

8.2.3. Dôsledok. *Každý lineárny podpriestor S vektorového priestoru V je jeho afinným podpriestorom. Afinný podpriestor M vektorového priestoru V je jeho lineárnym podpriestorom práve vtedy, keď $\mathbf{0} \in M$.*

Zameraním alebo tiež *smerovým podpriestorom* afinného podpriestoru $M \subseteq V$ nazývame lineárny podpriestor

$$\operatorname{Dir} M = \{\boldsymbol{x} - \boldsymbol{y};\ \boldsymbol{x}, \boldsymbol{y} \in M\} \subseteq V.$$

(Označenie pochádza z anglického slova *direction*). Podľa vety 8.2.2 je $\operatorname{Dir} M$ jediný lineárny podpriestor vo V taký, že $M = \boldsymbol{p} + \operatorname{Dir} M$ pre nejaké (pre každé) $\boldsymbol{p} \in M$. Taktiež pre každé $\boldsymbol{p} \in M$ platí

$$\operatorname{Dir} M = \{\boldsymbol{x} - \boldsymbol{p};\ \boldsymbol{x} \in M\}.$$

Pre každú usporiadanú $(n+1)$-ticu bodov $(\boldsymbol{p}_0, \ldots, \boldsymbol{p}_n)$, vektorového priestoru V, prípadne pre jeho konečnú neprázdnu podmnožinu $\{\boldsymbol{p}_0, \ldots, \boldsymbol{p}_n\}$, označme

$$\ell(\boldsymbol{p}_0, \ldots, \boldsymbol{p}_n) = \{t_0\boldsymbol{p}_0 + \ldots + t_n\boldsymbol{p}_n;\ t_0, \ldots, t_n \in K \ \&\ t_0 + \ldots + t_n = 1\}$$

množinu všetkých afinných kombinácií bodov $\boldsymbol{p}_0, \dots, \boldsymbol{p}_n$. Z práve dokázaného tvrdenia vyplýva, že $\ell(\boldsymbol{p}_0, \dots, \boldsymbol{p}_n)$ je najmenší afinný podpriestor vo V, ktorý obsahuje všetky body $\boldsymbol{p}_0, \dots, \boldsymbol{p}_n$; nazývame ho *afinný obal* bodov $\boldsymbol{p}_0, \dots, \boldsymbol{p}_n$ alebo tiež *afinný podpriestor generovaný* bodmi $\boldsymbol{p}_0, \dots, \boldsymbol{p}_n$.

Vo všeobecnosti možno pre ľubovoľnú (i nekonečnú) neprázdnu množinu $X \subseteq V$ definovať jej *afinný obal* $\ell(X)$, nazývaný tiež *afinný podpriestor generovaný* množinou X, ako množinu všetkých (konečných) afinných kombinácií bodov z X. Opäť platí, že $\ell(X)$ je najmenší afinný podpriestor vo V, pre ktorý $X \subseteq \ell(X)$.

8.2.4. Tvrdenie. *Nech* $\boldsymbol{p}_0, \boldsymbol{p}_1, \dots, \boldsymbol{p}_n \in V$. *Potom*

$$\begin{aligned} \ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \dots, \boldsymbol{p}_n) &= \boldsymbol{p}_0 + [\boldsymbol{p}_1 - \boldsymbol{p}_0, \dots, \boldsymbol{p}_n - \boldsymbol{p}_0], \\ \operatorname{Dir} \ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \dots, \boldsymbol{p}_n) &= [\boldsymbol{p}_1 - \boldsymbol{p}_0, \dots, \boldsymbol{p}_n - \boldsymbol{p}_0]. \end{aligned}$$

Dôkaz. prenechávame ako cvičenie čitateľovi.

Dimenziou alebo tiež *rozmerom* afinného podpriestoru $M \subseteq V$, označenie $\dim M$, nazývame dimenziu jeho zamerania, teda

$$\dim M = \dim \operatorname{Dir} M.$$

Body $\boldsymbol{p}_0, \boldsymbol{p}_1, \dots, \boldsymbol{p}_n$ vektorového priestoru V nazývame *afinne nezávislé*, ak vektory $\boldsymbol{p}_1 - \boldsymbol{p}_0, \dots, \boldsymbol{p}_n - \boldsymbol{p}_0$ sú lineárne nezávislé. Z nasledujúceho očividného tvrdenia okrem iného vyplýva, že body $\boldsymbol{p}_0, \boldsymbol{p}_1, \dots, \boldsymbol{p}_n \in V$ sú afinne nezávislé práve vtedy, keď pre nejaké (pre každé) $0 \le k \le n$ vektory $\boldsymbol{p}_j - \boldsymbol{p}_k$, kde $0 \le j \le n$ a $j \ne k$, sú lineárne nezávislé. Inak povedané, platí

8.2.5. Tvrdenie. *Body* $\boldsymbol{p}_0, \boldsymbol{p}_1, \dots, \boldsymbol{p}_n \in V$ *sú afinne nezávislé práve vtedy, keď*

$$\dim \ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \dots, \boldsymbol{p}_n) = n$$

Zrejme 0-rozmerné afinné podpriestrory vo V sú práve všetky body $\boldsymbol{p} \in V$ (presnejšie, všetky jednobodové podmnožiny vo V). Tieto afinné podpriestory nazývame tiež *triviálne*. Jednorozmené afinné podpriestory vo V nazývame *priamkami*. Každá priamka má naozaj tvar $\ell(\boldsymbol{p}, \boldsymbol{q})$ pre nejaké afinne nezávislé (t. j. rôzne) body $\boldsymbol{p}, \boldsymbol{q} \in V$. Dvojrozmerné afinné podpriestory vo V nazývame *rovinami*. Taktiež samotný priestor V je svojim *nevlastným* afinným podpriestorom. Ak $\dim V = n$, tak $(n-1)$-rozmerné afinné podpriestory vo V nazývame *nadrovinami*.

Kým pojmy „bod“, „priamka“ a „rovina“ sú absolútne v tom zmysle, že závisia len na dimenzii príslušného afinného podpriestoru, pojem nadroviny je relatívny, lebo závisí na vzťahu dimenzií afinného podpriestoru a celého priestoru. Napríklad ak $\dim V = 1$ (t. j. ak samotné V je priamka), tak každý bod vo V je zároveň nadrovinou. Nadrovinami v dvojrozmernom priestore

(t. j. v rovine) sú zasa všetky priamky. V trojrozmernom priestore V pojmy roviny a nadroviny splývajú. V štvorrozmernom priestore sú zasa nadrovinami trojrozmerné podpriestory; atď. Ešte poznamenajme, že v 0-rozmernom (t. j. jednobodovom) priestore V niet priamok, rovín ani nadrovín.

8.3 Prienik a spojenie afinných podpriestorov

V tomto paragrafe mierne zovšeobecníme niektoré výsledky paragrafov 4.3 a 5.4 o prieniku a súčte lineárnych podpriestorov do podoby použiteľnej pre afinné podpriestory.

8.3.1. Tvrdenie. *Nech M, $N \subseteq V$ sú afinné podpriestory. Potom $M \cap N$ je afinný podpriestor vo V práve vtedy, keď $M \cap N \neq \emptyset$. V tom prípade*

$$\operatorname{Dir}(M \cap N) = \operatorname{Dir} M \cap \operatorname{Dir} N.$$

Dôkaz. Ak $M \cap N = \emptyset$, tak to samozrejme nie je afinný podpriestor. Nech $M \cap N \neq \emptyset$. Označme $S = \operatorname{Dir} M$, $T = \operatorname{Dir} N$ príslušné smerové podpriestory. Zvoľme ľubovoľný bod $\boldsymbol{p} \in M \cap N$. Stačí dokázať rovnosť

$$M \cap N = \boldsymbol{p} + (S \cap T).$$

Zvoľme $\boldsymbol{q} \in M \cap N$. K nemu existujú $\boldsymbol{u} \in S$, $\boldsymbol{v} \in T$ také, že $\boldsymbol{q} = \boldsymbol{p} + \boldsymbol{u} = \boldsymbol{p} + \boldsymbol{v}$. Potom $\boldsymbol{u} = \boldsymbol{v} \in S \cap T$ a $\boldsymbol{q} \in \boldsymbol{p} + (S \cap T)$. Teda $M \cap N \subseteq \boldsymbol{p} + (S \cap T)$. Obrátená inklúzia je triviálna.

Neprázdnosť prieniku $M \cap N$ možno zaručiť za predpokladu, že lineárny priestor $\operatorname{Dir} M + \operatorname{Dir} N$ je „dosť veľký“.

8.3.2. Tvrdenie. *Nech M, $N \subseteq V$ sú afinné podpriestory. Potom*

$$\operatorname{Dir} M + \operatorname{Dir} N = V \;\Rightarrow\; M \cap N \neq \emptyset.$$

Dôkaz. Označme $S = \operatorname{Dir} M$, $T = \operatorname{Dir} N$. Zvoľme ľubovoľné $\boldsymbol{p} \in M$, $\boldsymbol{q} \in N$. Keďže $S + T = V$, existujú vektory $\boldsymbol{u} \in S$, $\boldsymbol{v} \in T$ také, že $\boldsymbol{q} - \boldsymbol{p} = \boldsymbol{u} + \boldsymbol{v}$. Potom

$$\boldsymbol{q} = \boldsymbol{p} + (\boldsymbol{q} - \boldsymbol{p}) = \boldsymbol{p} + \boldsymbol{u} + \boldsymbol{v}$$

V dôsledku toho $\boldsymbol{p} + \boldsymbol{u} = \boldsymbol{q} - \boldsymbol{v} \in M \cap N$, lebo $\boldsymbol{p} + \boldsymbol{u} \in M$ a $\boldsymbol{q} - \boldsymbol{v} \in N$.

Spojením afinných podpriestorov M, $N \subseteq V$, označenie $M \sqcup N$, nazývame afinný obal ich zjednotenia. Teda

$$M \sqcup N = \ell(M \cup N).$$

Zrejme $M \sqcup N$ je najmenší afinný podpriestor vo V, ktorý obsahuje M aj N, a pre lineárne podpriestory S, $T \subseteq V$ platí $S \sqcup T = S + T$.

8.3.3. Tvrdenie. *Nech M, $N \subseteq V$ sú afinné podpriestory.*
(a) Ak $M \cap N \neq \emptyset$, tak

$$\operatorname{Dir}(M \sqcup N) = \operatorname{Dir} M + \operatorname{Dir} N,$$

$$M \sqcup N = M + \operatorname{Dir} N = N + \operatorname{Dir} M.$$

(b) Ak $M \cap N = \emptyset$, tak pre ľubovoľné $\boldsymbol{p} \in M$, $\boldsymbol{q} \in N$ platí

$$\operatorname{Dir}(M \sqcup N) = [\boldsymbol{q} - \boldsymbol{p}] + \operatorname{Dir} M + \operatorname{Dir} N,$$

$$M \sqcup N = M + ([\boldsymbol{q} - \boldsymbol{p}] + \operatorname{Dir} N) = N + ([\boldsymbol{q} - \boldsymbol{p}] + \operatorname{Dir} M).$$

Poznámka. Stojí za zmienku, že obe rovnosti z (b) sú splnené aj za predpokladu $M \cap N \neq \emptyset$. V tom prípade však pre ľubovoľné $\boldsymbol{r} \in M \cap N$ platí

$$\boldsymbol{q} - \boldsymbol{p} = (\boldsymbol{r} - \boldsymbol{p}) + (\boldsymbol{q} - \boldsymbol{r}) \in \operatorname{Dir} M + \operatorname{Dir} N,$$

takže vektor $\boldsymbol{q} - \boldsymbol{p}$ možno v príslušných vzťahoch vynechať. Rovnako tomu bude i v príklade 8.3.5

Dôkaz. Stačí dokázať len (b), lebo (a) z neho vyplýva vo svetle našej poznámky. Označme $S = \operatorname{Dir} M$, $T = \operatorname{Dir} N$ a zvoľme $\boldsymbol{p} \in M$, $\boldsymbol{q} \in N$. Budeme dokazovať iba rovnosť

$$M \sqcup N = \boldsymbol{p} + [\boldsymbol{q} - \boldsymbol{p}] + S + T;$$

zvyšok je už jej bezprostredným dôsledkom.

Každý bod $\boldsymbol{r} \in M \sqcup N$ je afinnou kombináciou

$$\boldsymbol{r} = \sum_{i=0}^{m} s_i \boldsymbol{p}_i + \sum_{j=0}^{n} t_j \boldsymbol{q}_j$$

kde $\boldsymbol{p}_0, \ldots, \boldsymbol{p}_m \in M$, $\boldsymbol{q}_0, \ldots, \boldsymbol{q}_n \in N$, $s_0, \ldots, s_m, t_0, \ldots, t_n \in K$ a $\sum_i s_i + \sum_j t_j = 1$. Potom $\boldsymbol{p}_i - \boldsymbol{p} \in S$, $\boldsymbol{q}_j - \boldsymbol{q} \in T$ pre $i \leq m$, $j \leq n$. Označme $s = s_0 + \ldots + s_m$, $t = t_0 + \ldots + t_n$ a počítajme

$$\begin{aligned} \boldsymbol{r} &= (s\boldsymbol{p} + t\boldsymbol{q}) + \sum_{i=0}^{m} s_i(\boldsymbol{p}_i - \boldsymbol{p}) + \sum_{j=0}^{n} t_j(\boldsymbol{q}_j - \boldsymbol{q}) \\ &= \boldsymbol{p} + t(\boldsymbol{q} - \boldsymbol{p}) + \sum_{i=0}^{m} s_i(\boldsymbol{p}_i - \boldsymbol{p}) + \sum_{j=0}^{n} t_j(\boldsymbol{q}_j - \boldsymbol{q}) \in \boldsymbol{p} + [\boldsymbol{q} - \boldsymbol{p}] + S + T, \end{aligned}$$

keďže $s = 1 - t$. Teda $M \sqcup N \subseteq \boldsymbol{p} + [\boldsymbol{q} - \boldsymbol{p}] + S + T$. Obrátená inklúzia je triviálna.

8.3.4. Dôsledok. *Nech M, $N \subseteq V$ sú konečnorozmerné afinné podpriestory. Potom*

$$\dim(M \sqcup N) = \begin{cases} \dim M + \dim N - \dim(M \cap N), & \text{ak } M \cap N \neq \emptyset, \\ \dim M + \dim N - \dim(\operatorname{Dir} M \cap \operatorname{Dir} N) + 1, & \text{ak } M \cap N = \emptyset. \end{cases}$$

8.3.5. Príklad. Vo vektorovom priestore V uvažujme konečnorozmerné afinné podpriestory

$$M = \boldsymbol{p} + [\boldsymbol{u}_1, \dots, \boldsymbol{u}_m], \qquad N = \boldsymbol{q} + [\boldsymbol{v}_1, \dots, \boldsymbol{v}_n].$$

Potom

$$M \sqcup N = \begin{cases} \boldsymbol{p} + [\boldsymbol{u}_1, \dots, \boldsymbol{u}_m, \boldsymbol{v}_1, \dots, \boldsymbol{v}_n], & \text{ak } M \cap N \neq \emptyset, \\ \boldsymbol{p} + [\boldsymbol{q} - \boldsymbol{p}, \boldsymbol{u}_1, \dots, \boldsymbol{u}_m, \boldsymbol{v}_1, \dots, \boldsymbol{v}_n], & \text{ak } M \cap N = \emptyset, \end{cases}$$

$$\dim(M \sqcup N) = \begin{cases} \dim[\boldsymbol{u}_1, \dots, \boldsymbol{u}_m, \boldsymbol{v}_1, \dots, \boldsymbol{v}_n], & \text{ak } M \cap N \neq \emptyset, \\ \dim[\boldsymbol{q} - \boldsymbol{p}, \boldsymbol{u}_1, \dots, \boldsymbol{u}_m, \boldsymbol{v}_1, \dots, \boldsymbol{v}_n]), & \text{ak } M \cap N = \emptyset. \end{cases}$$

Ak navyše predpokladáme, že tak vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_m$ ako aj vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ sú lineárne nezávislé, tak

$$\dim(M \sqcup N) = \begin{cases} m + n - k, & \text{ak } M \cap N \neq \emptyset, \\ m + n - k + 1, & \text{ak } M \cap N = \emptyset, \end{cases}$$

kde $k = \dim([\boldsymbol{u}_1, \dots, \boldsymbol{u}_m] \cap [\boldsymbol{v}_1, \dots, \boldsymbol{v}_n])$.

8.3.6. Príklad. V stĺpcovom priestore $\mathbb{R}^4$ sú dané vektory $\boldsymbol{x} = (1, 2, 3, 4)^\mathsf{T}$, $\boldsymbol{y} = (0, -3, 1, -1)^\mathsf{T}$, $\boldsymbol{z} = (1, 1, 0, 0)^\mathsf{T}$, $\boldsymbol{u} = (0, -2, 4, 3)^\mathsf{T}$, $\boldsymbol{v} = (2, 6, 2, 5)^\mathsf{T}$, $\boldsymbol{w} = (0, 0, 1, 1)^\mathsf{T}$ a bližšie neurčené body $\boldsymbol{p}, \boldsymbol{q}$. Potom $S = [\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}]$, $T = [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}]$ sú lineárne podpriestory a $M = \boldsymbol{p} + S$, $N = \boldsymbol{q} + N$ sú afinné podpriestory v $\mathbb{R}^4$. Nájdeme dimenzie lineárnych podpriestorov $S + T$, $S \cap T$ a afinných podpriestorov $M \cap N$, $M \sqcup N$ v závislosti na $\boldsymbol{p}$, $\boldsymbol{q}$.

Lineárny podpriestor $S + T$ je generovaný stĺpcami blokovej matice

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 1 & 0 & 2 & 0 \\ 2 & -3 & 1 & -2 & 6 & 0 \\ 3 & 1 & 0 & 4 & 2 & 1 \\ 4 & -1 & 0 & 3 & 5 & 1 \end{array}\right),$$

pričom stĺpce ľavého bloku generujú lineárny podpriestor S a stĺpce pravého bloku lineárny podpriestor T. Táto matica je riadkovo ekvivalentná s blokovou maticou

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 2 & 0 \\ 0 & 1 & -3 & 4 & -4 & 0 \\ 0 & 0 & 3 & -3 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

v stupňovitom tvare, ktorej riadky majú vedúce prvky v stĺpcoch 1, 2, 3 a 6. Hneď vidíme, že vektory $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ tvoria bázu S a vektory $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{w}$ bázu $S+T$. Dopravovaním pravého bloku na riadkovo ekvivalentný stupňovitý tvar

$$\begin{pmatrix} 4 & -4 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

sa možno presvedčiť, že i vektory $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ sú lineárne nezávislé, teda tvoria bázu T. Zhrnutím dostávame $\dim S = \dim T = 3$, $\dim(S+T) = 4$. Odtiaľ podľa vety 5.4.1 vyplýva $\dim(S \cap T) = 3 + 3 - 4 = 2$. Takže $S + T = \mathbb{R}^4$, a bez toho, že by sme čokoľvek ďalej počítali, z tvrdenia 8.3.2 vieme, že nezávisle na bodoch $\boldsymbol{p}$, $\boldsymbol{q}$ platí $M \cap N \neq \emptyset$. Preto $\dim(M \cap N) = \dim(S \cap T) = 2$ podľa tvrdenia 8.3.1 S použitím tvrdenia 8.3.3 a dôsledku 8.3.4 dostávame $\dim(M \sqcup N) = \dim(S+T) = 4$.

8.4 Vzájomná poloha afinných podpriestorov

V tomto paragrafe podáme sľúbenú klasifikáciu vzájomnej polohy dvojíc netriviálnych vlastných afinných podpriestorov vo vektorovom priestore V. (Hoci to nie je z logického hľadiska nevyhnutné, aby sme sa vyhli triedeniu trivialít, body a celý priestor V z našich úvah vylučujeme.) Táto téma prirodzeným spôsobom rozširuje látku stredoškolskej geometrie, zahŕňajúcu klasifikáciu vzájomnej polohy priamok v rovine resp. priamok a rovín v (trojrozmernom) priestore.

Polohu netriviálnych vlastných lineárnych variet M, $N \subseteq V$ budeme klasifikovať na základe dvoch kritérií:

(A) Ak platí $\operatorname{Dir} M \subseteq \operatorname{Dir} N \vee \operatorname{Dir} N \subseteq \operatorname{Dir} M$, hovoríme, že M, N sú *rovnobežné* a píšeme $M \parallel N$.
V opačnom prípade, t. j. ak platí $\operatorname{Dir} M \not\subseteq \operatorname{Dir} N \;\&\; \operatorname{Dir} N \not\subseteq \operatorname{Dir} M$, hovoríme, že M, N *nie sú rovnobežné*, a píšeme $M \nparallel N$.

(B) Ak platí $M \cap N \neq \emptyset$, hovoríme, že M, N *sa pretínajú.*
V opačnom prípade, t. j. ak $M \cap N = \emptyset$, hovoríme, že M, N *sa nepretínajú*, alebo, že sú *disjunktné.*

Celkovo teda dostávame štyri možnosti:

(1) $M \parallel N$ & $M \cap N \neq \emptyset$, čiže M, N sú rovnobežné a pretínajú sa.
Ľahko možno nahliadnuť, že v takom prípade platí $\operatorname{Dir} M \subseteq \operatorname{Dir} N \Leftrightarrow M \subseteq N$ a $\operatorname{Dir} N \subseteq \operatorname{Dir} M \Leftrightarrow N \subseteq M$. Teda $M \subseteq N$ alebo $N \subseteq M$. Hovoríme, že jedna z lineárnych variet M, N je *podvarietou* druhej, alebo, že M, N sú vo vzťahu *inklúzie*.

(2) $M \parallel N$ & $M \cap N = \emptyset$, čiže M, N sú rovnobežné a nepretínajú sa.
Tento prípad nazývame vzťahom *pravej rovnobežnosti*.

(3) $M \nparallel N$ & $M \cap N \neq \emptyset$, čiže M, N nie sú rovnobežné a pretínajú sa.
Hovoríme, že M, N sú *rôznobežné*.

(4) $M \nparallel N$ & $M \cap N = \emptyset$, čiže M, N nie sú rovnobežné a nepretínajú sa.
V tomto prípade ešte rozlišujeme dve ďalšie možnosti:
(4a) Ak $\operatorname{Dir} M \cap \operatorname{Dir} N = \{\mathbf{0}\}$, hovoríme, že M, N sú *mimobežné*.
(4b) Ak $\operatorname{Dir} M \cap \operatorname{Dir} N \neq \{\mathbf{0}\}$, hovoríme, že M, N sú *čiastočne rovnobežné*.

Prípady (1), (2), (3) sú nám dobre známe zo stredoškolskej planimetrie, s prípadom (4) sa však v rovine stretnúť nemožno – dve priamky v rovine buď splývajú alebo sú to pravé rovnobežky alebo rôznobežky. Zo stredoškolskej stereometrie, okrem prípadov (1), (2), (3), ktoré sa realizujú vo vzájomných polohách dvojíc priamok, dvojíc rovín ako i priamky a roviny v trojrozmernom priestore, poznáme aj prípad (4a) – ide o prípad mimobežných priamok. S prípadom (4b), t. j. s prípadom čiastočnej rovnobežnosti sme sa však dosiaľ nestretli a nedokážeme ho spojiť so žiadnou názornou geometrickou predstavou. Nie je to náhoda. Platí totiž nasledujúce tvrdenie.

8.4.1. Tvrdenie. *Nech M, $N \subseteq V$ sú čiastočne rovnobežné lineárne variety. Potom $\dim M \geq 2$, $\dim N \geq 2$ a $\dim V \geq 4$.*

Dôkaz. Označme $S = \operatorname{Dir} M$, $T = \operatorname{Dir} N$. Potom $S \cap T$ je netriválny vlastný lineárny podpriestor každého zo zameraní S, T. Teda $\dim(S \cap T) \geq 1$, $\dim M = \dim S \geq 2$, $\dim N = \dim T \geq 2$ a taktiež

$$\dim(S \cap T) \leq \min(\dim S, \dim T) - 1.$$

S použitím vety 5.4.1 z toho vyplýva

$$\begin{aligned}\dim(S+T) &= \dim S + \dim T - \dim(S \cap T)\\ &\geq \dim S + \dim T - \min(\dim S, \dim T) + 1\\ &= \max(\dim S, \dim T) + 1 \geq 3.\end{aligned}$$

Keďže $M \cap N = \emptyset$, podľa tvrdenia 8.3.2 je $S + T$ vlastný lineárny podpriestor vo V. Preto

$$\dim V \geq \dim(S+T) + 1 \geq 4.$$

Na druhej strane v ľubovoľnom vektorovom priestore V dimenzie ≥ 4 nie je ťažké nájsť príklady čiastočne rovnobežných lineárnych variet. Presvedčte sa, že napr.

$$M = [\boldsymbol{e}_1, \boldsymbol{e}_2], \qquad N = \boldsymbol{e}_4 + [\boldsymbol{e}_2, \boldsymbol{e}_3]$$

sú čiastočne rovnobežné roviny v K^4. Skúste nájsť iné príklady.

8.5 Afinné zobrazenia

Nech U, V sú vektorové priestory nad tým istým poľom K. Hovoríme, že $f\colon V \to U$ je *afinné zobrazenie*, ak pre ľubovoľné body $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r} \in V$ a skalár $s \in K$ platí

$$f(s\boldsymbol{p} + (1-s)\boldsymbol{q}) = sf(\boldsymbol{p}) + (1-s)f(\boldsymbol{q}),$$
$$f(\boldsymbol{p} - \boldsymbol{q} + \boldsymbol{r}) = f(\boldsymbol{p}) - f(\boldsymbol{q}) + f(\boldsymbol{r}).$$

Podobným spôsobom ako tvrdenie 8.2.1 možno dokázať, že afinné sú práve tie zobrazenia $f\colon V \to U$, ktoré zachovávajú všetky afinné kombinácie trojíc bodov, či, takisto, vôbec všetky afinné kombinácie; v prípade poľa charakteristiky $\neq 2$ stačí žiadať zachovávanie afinných kombinácií dvojíc.

8.5.1. Tvrdenie. *Nech U, V sú vektorové priestory nad poľom K. Potom pre ľubovoľné zobrazenie $f\colon V \to U$ nasledujúce podmienky sú ekvivalentné:*

(i) f je afinné zobrazenie;

(ii) pre všetky $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r} \in V$, $s, t \in K$ platí

$$f(s\boldsymbol{p} + t\boldsymbol{q} + (1-s-t)\boldsymbol{r}) = sf(\boldsymbol{p}) + tf(\boldsymbol{q}) + (1-s-t)f(\boldsymbol{r});$$

(iii) pre každé $n \in \mathbb{N}$ a všetky body $\boldsymbol{p}_0, \boldsymbol{p}_1 \dots, \boldsymbol{p}_n \in V$ a skaláry $t_0, t_1, \dots, t_n \in K$ také, že $t_0 + t_1 + \ldots + t_n = 1$, platí

$$f(t_0\boldsymbol{p}_0 + t_1\boldsymbol{p}_1 + \ldots + t_n\boldsymbol{p}_n) = t_0 f(\boldsymbol{p}_0) + t_1 f(\boldsymbol{p}_1) + \ldots + t_n f(\boldsymbol{p}_n).$$

Ak char $K \neq 2$, *tak uvedené podmienky sú navyše ekvivalentné s podmienkou*

(ii$^-$) pre všetky $\boldsymbol{p}, \boldsymbol{q} \in V$, $s \in K$ platí

$$f(s\boldsymbol{p} + (1-s)\boldsymbol{q}) = sf(\boldsymbol{p}) + (1-s)f(\boldsymbol{q}).$$

Posunutím alebo *transláciou* vektorového priestoru V o vektor $\boldsymbol{u} \in V$ nazývame zobrazenie $V \to V$ dané predpisom $\boldsymbol{x} \mapsto \boldsymbol{x} + \boldsymbol{u}$.

Zrejme kompozíciou posunutia o vektor $\boldsymbol{u} \in V$ a posunutia o vektor $\boldsymbol{v} \in V$ je posunutie o vektor $\boldsymbol{u}+\boldsymbol{v}$. Každé posunutie je bijektívne zobrazenie; inverzné zobrazenie k posunutiu o vektor $\boldsymbol{u}$ je posunutie o opačný vektor $-\boldsymbol{u}$.

Z nasledujúcej vety okrem iného vyplýva, že každé afinné zobrazenie možno dostať kompozíciou lineárneho zobrazenia a posunutia.

8.5.2. Veta. *Nech U, V sú vektorové priestory nad poľom K. Potom zobrazenie $f\colon V \to U$ je afinné práve vtedy, keď existuje vektor $\boldsymbol{u} \in U$ a lineárne zobrazenie $\varphi\colon V \to U$ také, že pre každé $\boldsymbol{x} \in V$ platí*

$$f(\boldsymbol{x}) = \varphi(\boldsymbol{x}) + \boldsymbol{u}.$$

Dôkaz. Treba dokázať dve veci:

(1) Pre ľubovoľný vektor $\boldsymbol{u} \in U$ a lineárne zobrazenie $\varphi\colon V \to U$ je predpisom $f(\boldsymbol{x}) = \varphi(\boldsymbol{x}) + \boldsymbol{u}$ dané afinné zobrazenie $f\colon V \to U$.

(2) Ak $f\colon V \to U$ je afinné zobrazenie, tak priradenie $\varphi(\boldsymbol{x}) = f(\boldsymbol{x}) - f(\boldsymbol{0})$ definuje lineárne zobrazenie $\varphi\colon V \to U$.

V jednom i druhom prípade možno zachovávanie príslušných afinných resp. lineárnych kombinácií overiť priamymi výpočtami, ktoré prenechávame čitateľovi.

Zrejme vektor $\boldsymbol{u} \in U$ ako aj lineárne zobrazenie φ sú podmienkou vety určené jednoznačne. Zobrazenie $\varphi = f - f(\boldsymbol{0})$ nazývame *lineárnou časťou* a vektor $\boldsymbol{u} = f(\boldsymbol{0})$ *absolútnym členom* afinného zobrazenia f. Píšeme tiež $f = \varphi + \boldsymbol{u}$.

Afinné zobrazenia sú tak zovšeobecnením funkcií $f\colon K \to K$ tvaru $f(x) = ax+b$, kde $a, b \in K$, ktoré (najmä v prípade $K = \mathbb{R}$) v matematickej analýze nazývame lineárnymi, na viacrozmerné vektorové priestory.

8.5.3. Dôsledok. *Nech U, V sú vektorové priestory nad poľom K. Potom*

(a) ľubovoľná translácia priestoru V je afinné zobrazenie;

(b) ľubovoľné lineárne zobrazenie $\varphi\colon V \to U$ je afinné;

(c) afinné zobrazenie $f\colon V \to U$ je lineárne práve vtedy, keď $f(\boldsymbol{0}) = \boldsymbol{0}$.

8.5.4. Tvrdenie. *Nech U, V, W sú vektorové priestory nad poľom K a $g\colon W \to V$, $f\colon V \to U$ sú afinné zobrazenia. Potom aj ich kompozícia $f \circ g\colon W \to U$ je afinné zobrazenie.*

Dôkaz. Hoci priamym výpočtom možno overiť, že $f \circ g$ zachováva afinné kombinácie, podáme radšej dôkaz založený na vete 8.5.2, ktorý nám poskytne informáciu navyše.

Nech $f = \varphi + \boldsymbol{u}$, $g = \psi + \boldsymbol{v}$, kde $\varphi\colon V \to U$, $\psi\colon W \to V$ sú lineárne zobrazenia a $\boldsymbol{u} = f(\boldsymbol{0})$, $\boldsymbol{v} = g(\boldsymbol{0})$. Potom pre $\boldsymbol{z} \in W$ s využitím linearity φ dostávame

$$(f \circ g)(\boldsymbol{z}) = \varphi(\psi(\boldsymbol{z}) + \boldsymbol{v}) + \boldsymbol{u} = (\varphi \circ \psi)(\boldsymbol{z}) + \varphi(\boldsymbol{v}) + \boldsymbol{u}.$$

Teda $f \circ g$ je afinné zobrazenie zložené z lineárneho zobrazenia $\varphi \circ \psi$ a posunutia o vektor $\varphi(\boldsymbol{v}) + \boldsymbol{u}$.

Vzorec odvodený v našom dôkaze stojí za zaznamenanie. Pre lineárne zobrazenia $\psi\colon W \to V$, $\varphi\colon V \to U$ a vektory $\boldsymbol{v} \in V$, $\boldsymbol{u} \in U$ platí

$$(\varphi + \boldsymbol{u}) \circ (\psi + \boldsymbol{v}) = (\varphi \circ \psi) + (\varphi\boldsymbol{v} + \boldsymbol{u}).$$

8.5.5. Tvrdenie. *Nech U, V sú vektorové priestory nad poľom K, $f\colon V \to U$ je afinné zobrazenie a $M \subseteq V$, $N \subseteq U$ sú afinné podpriestory. Potom $f(M)$ je afinný podpriestor v U a $f^{-1}(N)$ je afinný podpriestor vo V alebo prázdna množina.*

Dôkaz. Nech $f = \varphi + \boldsymbol{u}$, kde φ je lineárna časť f a $\boldsymbol{u} = f(\boldsymbol{0})$. Nech ďalej $M = \boldsymbol{p} + S$, $N = \boldsymbol{q} + T$, kde $\boldsymbol{p} \in M$, $\boldsymbol{q} \in N$ a $S \subseteq V$, $T \subseteq U$ sú lineárne podpriestory. Potrebný záver vyplýva z tvrdení 6.1.3, a vety 8.2.2 a nasledujúcich rovností

$$f(M) = f(\boldsymbol{p}) + \varphi(S),$$
$$f^{-1}(N) = \begin{cases} \boldsymbol{z} + \varphi^{-1}(T), & \text{kde } \boldsymbol{z} \in V \text{ je ľubovoľné také, že } \varphi(\boldsymbol{z}) = \boldsymbol{q} - \boldsymbol{u}, \\ \emptyset, & \text{ak neexistuje } \boldsymbol{z} \in V \text{ také, že } \varphi(\boldsymbol{z}) = \boldsymbol{q} - \boldsymbol{u}, \end{cases}$$

ktorých dôkaz prenechávame čitateľovi.

Keďže každé posunutie je bijekcia, afinné zobrazenie $f = \varphi + \boldsymbol{u}\colon V \to U$ s lineárnou časťou φ je injektívne práve vtedy, keď φ je injektívne. Podobne, f je surjektívne práve vtedy, keď φ je surjektívne. Z toho už priamo vyplývajú ďalšie tri výsledky.

Prvý z nich zovšeobecňuje vetu 6.2.3 o dimenzii jadra a obrazu.

8.5.6. Veta. *Nech $f\colon V \to U$ je afinné zobrazenie, pričom V je konečnorozmerný vektorový priestor. Potom pre ľubovoľné $\boldsymbol{y} \in \operatorname{Im} f$ platí*

$$\dim V = \dim f^{-1}(\boldsymbol{y}) + \dim \operatorname{Im} f.$$

Afinnou transformáciou vektorového priestoru V nazývame ľubovoľné afinné zobrazenie $f\colon V \to V$. Aj pre afinné transformácie platí obdoba dôsledku 6.2.4

8.5.7. Dôsledok. *Nech $f\colon V \to V$ je afinná transformácia konečnorozmerného vektorového priestoru V. Potom f je injektívna práve vtedy, keď je surjektívna.*

8.5.8. Tvrdenie. *Nech $f\colon V \to U$ je afinné zobrazenie s lineárnou časťou φ a absolútnym členom $\boldsymbol{u} = f(\mathbf{0})$. Potom f je bijektívne práve vtedy, keď φ je bijektívne. V tom prípade aj inverzné zobrazenie $f^{-1}\colon U \to V$ je afinné a platí*

$$f^{-1} = \varphi^{-1} - \varphi^{-1}(\boldsymbol{u}).$$

Teda f^{-1} je kompozíciou lineárneho zobrazenia φ^{-1} a posunutia o vektor $-\varphi^{-1}(\boldsymbol{u})$.

Nech U, V sú konečnorozmerné vektorové priestory a $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ sú bázy v U resp. vo V. *Maticou* afinného zobrazenia $f\colon V \to U$ s lineárnou časťou φ a absolútnym členom $\boldsymbol{u}$ vzhľadom na bázy $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$ nazývame blokovú maticu

$$(f)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \big((\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}} \,|\, (\boldsymbol{u})_{\boldsymbol{\alpha}}\big).$$

Ak teda $\dim U = m$, $\dim V = n$, $\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ je matica lineárneho zobrazenia φ v bázach $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$, $\boldsymbol{\alpha}$ a $\boldsymbol{a} = (\boldsymbol{u})_{\boldsymbol{\alpha}}$ sú súradnice vektora $\boldsymbol{u}$ v báze $\boldsymbol{\alpha}$, tak maticou afinného zobrazenia f v bázach $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$ je bloková matica

$$(f)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \big((\varphi\boldsymbol{v}_1)_{\boldsymbol{\alpha}}, \dots, (\varphi\boldsymbol{v}_n)_{\boldsymbol{\alpha}} \,|\, (\boldsymbol{u})_{\boldsymbol{\alpha}}\big) = (\boldsymbol{A} \,|\, \boldsymbol{a}) \in K^{m\times(n+1)}.$$

Súradnice bodu $\boldsymbol{x} \in V$ v báze $\boldsymbol{\beta}$ a súradnice jeho obrazu $f(\boldsymbol{x}) \in U$ v báze $\boldsymbol{\alpha}$ sú tak spojené rovnosťou

$$(f\boldsymbol{x})_{\boldsymbol{\alpha}} = (\varphi)_{\alpha,\boldsymbol{\beta}} \cdot (\boldsymbol{x})_{\boldsymbol{\beta}} + (\boldsymbol{u})_{\boldsymbol{\alpha}} = \boldsymbol{A} \cdot (\boldsymbol{x})_{\boldsymbol{\alpha}} + \boldsymbol{a}.$$

Samozrejme, ak f je lineárne zobrazenie, t. j. ak $f = \varphi$ a $\boldsymbol{u} = \mathbf{0}$, nemá význam rozširovať maticu $(\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ o nulový stĺpec.

Z tvrdenia 8.5.4, presnejšie z formuly odvodenej počas jeho dôkazu, a z tvrdenia 8.5.8 s použitím výsledkov paragrafov 6.4 a 7.2 vyplýva náš záverečný výsledok.

8.5.9. Tvrdenie. *Nech U, V, W sú konečnorozmerné vektorové priestory nad poľom K a $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ sú nejaké bázy priestorov U, V, resp. W.*

(a) *Ak $g\colon W \to V$, $f\colon V \to U$ sú afinné zobrazenia, ktoré majú v príslušných bázach matice $(g)_{\boldsymbol{\beta},\boldsymbol{\gamma}} = (\boldsymbol{B} \,|\, \boldsymbol{b})$, $(f)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = (\boldsymbol{A} \,|\, \boldsymbol{a})$, tak ich kompozícia $f \circ g\colon W \to U$ má v bázach $\boldsymbol{\gamma}$, $\boldsymbol{\alpha}$ maticu*

$$(f \circ g)_{\boldsymbol{\alpha},\boldsymbol{\gamma}} = (\boldsymbol{A} \cdot \boldsymbol{B} \,|\, \boldsymbol{A} \cdot \boldsymbol{b} + \boldsymbol{a}).$$

(b) *Ak $f\colon V \to U$ je afinná bijekcia s maticou $(f)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = (\boldsymbol{A} \,|\, \boldsymbol{a})$ v bázach $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, tak k nej inverzné zobrazenie je afinná bijekcia $f^{-1}\colon U \to V$, ktorá má v bázach $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ maticu*

$$\big(f^{-1}\big)_{\boldsymbol{\beta},\boldsymbol{\alpha}} = \big(\boldsymbol{A}^{-1} \,|\, -\boldsymbol{A}^{-1} \cdot \boldsymbol{a}\big).$$

Cvičenia

8.1. Dokážte postupne záverečné štyri podmienky z vety 8.2.2

8.2. Nech V je vektorový priestor nad poľom K.

(a) Tri afinne nezávislé body $\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2 \in V$ sa nazývajú *nekolineárne*. Dokážte, že $\boldsymbol{p}_0$, $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ sú *kolineárne* práve vtedy, keď ležia na jednej priamke.

(b) Podobne, štyri afinne nezávislé body $\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3 \in V$ sa nazývajú *nekomplanárne*. Dokážte, že $\boldsymbol{p}_0$, $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ $\boldsymbol{p}_3$ sú *komplanárne* práve vtedy, keď ležia v jednej rovine.

8.3. Nech V je vektorový priestor nad poľom K a $\boldsymbol{p}_0, \boldsymbol{p}_1, \dots, \boldsymbol{p}_n, \boldsymbol{q} \in V$. Dokážte postupne nasledujúce tvrdenia:

(a) Body $\boldsymbol{p}_0, \boldsymbol{p}_1, \dots, \boldsymbol{p}_n$ sú afinne nezávislé práve vtedy, keď pre nejaké (ľubovoľné) $i \leq n$ sú lineárne nezávislé vektory $\boldsymbol{p}_j - \boldsymbol{p}_i$, kde $j \in \{0, 1, \dots, n\} \smallsetminus \{i\}$.

(b) $\boldsymbol{q} \in \ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \dots, \boldsymbol{p}_n)$ práve vtedy, keď $\boldsymbol{q} - \boldsymbol{p}_0 \in [\boldsymbol{p}_1 - \boldsymbol{p}_0, \dots, \boldsymbol{p}_n - \boldsymbol{p}_0]$. Odvoďte z toho obe rovnosti z tvrdenia 8.2.4

(c) Ak body $\boldsymbol{p}_0, \boldsymbol{p}_1, \dots, \boldsymbol{p}_n$ sú afinne nezávislé, tak $\boldsymbol{q} \in \ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \dots, \boldsymbol{p}_n)$ práve vtedy, keď vektory $\boldsymbol{p}_1 - \boldsymbol{p}_0, \dots, \boldsymbol{p}_n - \boldsymbol{p}_0, \boldsymbol{q} - \boldsymbol{p}_0$ sú lineárne závislé.

8.4. Vo vektorovom priestore $\mathbb{R}^4$ sú dané body $\boldsymbol{p}_0 = (1,1,2,2)^\mathsf{T}$, $\boldsymbol{p}_1 = (0,1,0,1)^\mathsf{T}$, $\boldsymbol{p}_2 = (1,2,0,3)^\mathsf{T}$ a $\boldsymbol{q} = (0,2,-4,2)^\mathsf{T}$, $\boldsymbol{r} = (-1,2,-4,1)^\mathsf{T}$.

(a) Zistite, či platí $\boldsymbol{q} \in \ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2)$, $\boldsymbol{r} \in \ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2)$.

(b) Vypočítajte dimenzie afinných podpriestorov $\ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2)$, $\ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{q})$, $\ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{r})$.

(c) Vypočítajte dimenzie afinných podpriestorov $\ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2) \cap \ell(\boldsymbol{q}, \boldsymbol{r})$, $\ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2) \sqcup \ell(\boldsymbol{q}, \boldsymbol{r})$ a určte vzájomnú polohu afinných podpriestorov $\ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2)$, $\ell(\boldsymbol{q}, \boldsymbol{r})$.

(d) Vypočítajte dimenzie afinných podpriestorov $\ell(\boldsymbol{p}_0, \boldsymbol{p}_1) \cap \ell(\boldsymbol{q}, \boldsymbol{r})$, $\ell(\boldsymbol{p}_0, \boldsymbol{p}_1) \sqcup \ell(\boldsymbol{q}, \boldsymbol{r})$ a určte vzájomnú polohu afinných podpriestorov $\ell(\boldsymbol{p}_0, \boldsymbol{p}_1)$, $\ell(\boldsymbol{q}, \boldsymbol{r})$.

8.5. (a) Nájdite príklad troch priamok v $\mathbb{R}^3$ tak, aby ľubovoľné dve z nich boli mimobežné.

(b) Dokážte, že priamka a rovina v trojrozmernom vektorovom priestore nemôžu byť mimobežné.

(c) Vo vektorovom priestore $\mathbb{R}^4$ nájdite príklad mimobežnej priamky a roviny.

(d) Dokážte, že dve roviny vo štvorrozmernom vektorovom priestore nemôžu byť mimobežné.

(e) Nájdite príklad dvoch mimobežných rovín v $\mathbb{R}^5$.

8.6. Nech $\boldsymbol{p}_0$, $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, $\boldsymbol{p}_3$, $\boldsymbol{p}_4$ sú ľubovoľné afinne nezávislé body vo vektorovom priestore V nad poľom K. Potom roviny $\ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}_2)$, $\boldsymbol{p}_4 + [\boldsymbol{p}_2 - \boldsymbol{p}_0, \boldsymbol{p}_3 - \boldsymbol{p}_0]$ sú čiastočne rovnobežné. Dokážte.

8.7. *Repér* vo vektorovom priestore sa zvykne definovať ako usporiadaná $(n+1)$-tica afinne nezávislých bodov $\boldsymbol{\rho} = (\boldsymbol{r}_0, \boldsymbol{r}_1, \dots, \boldsymbol{r}_n) \in V^{n+1}$, taká, že $\ell(\boldsymbol{r}_0, \boldsymbol{r}_1, \dots, \boldsymbol{r}_n) = V$, prípadne ako usporiadaná $(n+1)$-tica $(\boldsymbol{r}, \boldsymbol{\beta}) = (\boldsymbol{r}, \boldsymbol{v}_1, \dots, \boldsymbol{v}_n) \in V^{n+1}$, pozostávajúca z ľubovoľného bodu $\boldsymbol{r} \in V$ a bázy $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ vektorového priestoru V. Dokážte nasledujúce dve tvrdenia:

(a) Nech $\boldsymbol{\rho} = (\boldsymbol{r}_0, \boldsymbol{r}_1, \dots, \boldsymbol{r}_n)$ je usporiadaná $(n+1)$-tica bodov z V. Potom $\boldsymbol{\rho}$ je repér vo V v zmysle prvej definície práve vtedy, keď $\boldsymbol{\beta} = (\boldsymbol{r}_1 - \boldsymbol{r}_0, \dots, \boldsymbol{r}_n - \boldsymbol{r}_0)$ je báza vektorového priestoru V, t. j. práve vtedy, keď $(\boldsymbol{r}_0, \boldsymbol{\beta})$ je repér v zmysle druhej definície.

(b) Nech $\boldsymbol{r}$ je bod z V a $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ je usporiadaná n-tica vektorov z V. Potom $(\boldsymbol{r}, \boldsymbol{\beta})$ je repér v zmysle druhej definície práve vtedy, keď $\boldsymbol{\rho} = (\boldsymbol{r}, \boldsymbol{r} + \boldsymbol{v}_1, \ldots, \boldsymbol{r} + \boldsymbol{v}_n)$ je repér v zmysle prvej definície.

Z toho dôvodu nie je potrebné rozlišovať medzi repérmi v zmysle jednej či druhej definície.

8.8. Nech $\boldsymbol{\rho} = (\boldsymbol{r}, \boldsymbol{\beta})$ je repér vo vektorovom priestore V nad poľom K. *Afinnými* alebo tiež *barycentrickými súradnicami bodu* $\boldsymbol{x} \in V$ vzhľadom na repér $\boldsymbol{\rho}$ nazývame súradnice vektora $\boldsymbol{x} - \boldsymbol{r}$ vzhľadom na bázu $\boldsymbol{\beta}$, čiže $(\boldsymbol{x})_{\boldsymbol{\rho}} = (\boldsymbol{x} - \boldsymbol{r})_{\boldsymbol{\beta}}$. Ak je repér $\boldsymbol{\rho}$ známy z kontextu, hovoríme len o *afinných (barycentrických) súradniciach bodu* $\boldsymbol{x}$. Dokážte postupne nasledujúce tvrdenia:

(a) $(\boldsymbol{0}, \boldsymbol{\varepsilon})$, kde $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ je kanonická báza, je repér v K^n a pre každé $\boldsymbol{x} \in K^n$ platí $(\boldsymbol{x})_{(\boldsymbol{0}, \boldsymbol{\varepsilon})} = \boldsymbol{x}$.

(b) Body repéru $\boldsymbol{\rho} = (\boldsymbol{r}_0, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_n)$ majú vzhľadom na tento repér afinné súradnice $(\boldsymbol{r}_0)_{\boldsymbol{\rho}} = \boldsymbol{0}$, $(\boldsymbol{r}_1)_{\boldsymbol{\rho}} = \boldsymbol{e}_1$, ..., $(\boldsymbol{r}_n)_{\boldsymbol{\rho}} = \boldsymbol{e}_n$.

(c) Ak $\dim V = n$ a $\boldsymbol{\rho} = (\boldsymbol{r}, \boldsymbol{\beta})$ je repér vo V, tak predpisom $\boldsymbol{x} \mapsto (\boldsymbol{x})_{\boldsymbol{\rho}}$ je definované bijektívne afinné zobrazenie $V \to K^n$ a pre každé $\boldsymbol{x} \in V$, $\boldsymbol{c} = (c_1, \ldots, c_n)^{\mathsf{T}} \in K^n$ platí

$$\boldsymbol{x} = \boldsymbol{r} + \boldsymbol{\beta} \cdot (\boldsymbol{x})_{\boldsymbol{\rho}}, \qquad\qquad (\boldsymbol{r} + \boldsymbol{\beta} \cdot \boldsymbol{c})_{\boldsymbol{\rho}} = \boldsymbol{c}.$$

8.9. V $\mathbb{R}^3$ sú dané body $\boldsymbol{r}_0 = (5,2,1)^{\mathsf{T}}$, $\boldsymbol{r}_1 = (0,2,1)^5$, $\boldsymbol{r}_2 = (5,0,2)^{\mathsf{T}}$, $\boldsymbol{r}_3 = (5,2,0)^{\mathsf{T}}$.

(a) Dokážte, že $\boldsymbol{\rho} = (\boldsymbol{r}_0, \boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3)$ je repér v $\mathbb{R}^3$.

(b) Nájdite afinné súradnice bodov $\boldsymbol{x} = (4,4,-3)^{\mathsf{T}}$, $\boldsymbol{y} = (-5,-2,-1)^{\mathsf{T}}$, $\boldsymbol{z} = (0,0,0)^{\mathsf{T}}$ vzhľadom na repér $\boldsymbol{\rho}$.

(c) Nájdite body $\boldsymbol{p}$, $\boldsymbol{q}$, $\boldsymbol{r}$, ak poznáte ich afinné súradnice $(\boldsymbol{p})_{\boldsymbol{\rho}} = (0,2,1)^{\mathsf{T}}$, $(\boldsymbol{q})_{\boldsymbol{\rho}} = (-1,1,-1)^{\mathsf{T}}$, $(\boldsymbol{r})_{\boldsymbol{\rho}} = (0,0,0)^{\mathsf{T}}$.

8.10. Nech $\boldsymbol{\pi} = (\boldsymbol{p}, \boldsymbol{\alpha})$, $\boldsymbol{\rho} = (\boldsymbol{r}, \boldsymbol{\beta})$ sú dva repéry vo vektorovom priestore V nad poľom K. Potom afinné súradnice ľubovoľného bodu $\boldsymbol{x} \in V$ vzhľadom na tieto repéry sú zviazané vzťahom

$$(\boldsymbol{x})_{\boldsymbol{\pi}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot (\boldsymbol{x} - \boldsymbol{p})_{\boldsymbol{\beta}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot \big((\boldsymbol{x})_{\boldsymbol{\rho}} - (\boldsymbol{p})_{\boldsymbol{\rho}}\big),$$

kde $\boldsymbol{P}_{\alpha,\boldsymbol{\beta}}$ je matica prechodu z bázy $\boldsymbol{\beta}$ do bázy $\boldsymbol{\alpha}$. Dokážte.

8.11. Dokážte tvrdenie 8.5.1 (*Návod*: Modifikujte dôkaz tvrdenia 8.2.1)

8.12. Dokážte podmienky (1), (2) z dôkazu vety 8.5.2

8.13. Doplňte vynechané dôkazy oboch rovností z dôkazu tvrdenia 8.5.5

8.14. Na základe tvrdenia 8.5.5 doplňte dôkazy vety 8.5.6, dôsledku 8.5.7 a tvrdenia 8.5.8

8.15. Predpokladajme, že dvaja pozorovatelia P a P' popisujú udalosti v čase a v trojrozmernom priestore vzhľadom na po dvoch rovnobežné a rovnako orientované súradné osi x, y, z, resp. x', y', z', pričom počiatok súradnej sústavy pozorovateľa P' má z hľadiska pozorovateľa P v čase $t = t_0$, zodpovedajúcom času $t' = 0$ pozorovateľa P', súradnice $(x_0, y_0, z_0)^{\mathsf{T}}$. Nech navyše pozorovateľ P' sa vzhľadom na pozorovateľa P pohybuje rovnomerne priamočiaro rýchlosťou $\boldsymbol{v} = (v_x, v_y, v_z)^{\mathsf{T}}$ (pozri príklad 6.4.7 a cvičenie 6.14).

(a) Odvoďte tvar *Galileovej transformácie*, ktorou sú za týchto okolností v klasickej

(t. j. v nerelativistickej) fyzike zviazané časopriestorové súradnice bodových udalostí z hľadiska pozorovateľov P resp. P':

$$t' = t - t_0, \qquad x' = x - x_0 - v_x t, \qquad y' = y - y_0 - v_y t, \qquad z' = z - z_0 - v_z t.$$

Nahliadnite, že ide o afinnú transformáciu s maticou $(\boldsymbol{G_v} \mid -\boldsymbol{s}_0)$, kde $\boldsymbol{G_v}$ je matica Galileovej transformácie z cvičenia 6.14 a $\boldsymbol{s}_0 = (t_0, x_0, y_0, z_0)^{\mathsf{T}}$.

(b) Nech $f, g\colon \mathbb{R}^4 \to \mathbb{R}^4$ sú Galileove transformácie s maticami $(\boldsymbol{G_v} \mid -\boldsymbol{s}_0)$ resp. $(\boldsymbol{G_w} \mid -\boldsymbol{s}_1)$, kde $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$, $\boldsymbol{s}_0, \boldsymbol{s}_1 \in \mathbb{R}^4$. Nájdite rozšírenú maticu kompozície afinných zobrazení $f \circ g$ a maticu inverzného zobrazenia f^{-1}. Dokážte, že ide opäť o Galileove transformácie uvedeného typu a vysvetlite fyzikálny význam získaných výsledkov.

8.16. Nech U, V sú vektorové priestory nad poľom K. Označme $\mathcal{A}(V, U)$ množinu všetkých *afinných* zobrazení $f\colon V \to U$. Dokážte postupne nasledujúce tvrdenia:

(a) $\mathcal{A}(V, U)$ s operáciami súčtu a skalárneho násobku definovanými po zložkách tvorí *lineárny* podpriestor vektorového priestoru U^V. $\mathcal{A}(V, U)$ navyše obsahuje vektorový priestor $\mathcal{L}(V, U)$ všetkých lineárnych zobrazení $fi\colon V \to U$ ako svoj lineárny podpriestor. (Pozri príklad 1.6.5 a tvrdenie 6.5.1)

(b) Priradením $f \mapsto (f - f(\mathbf{0}), f(\mathbf{0}))$ je definovaný lineárny izomorfizmus vektorových priestorov $\mathcal{A}(V, U) \to \mathcal{L}(V, U) \times U$ (pozri príklad 1.6.4).

(c) Nech $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ sú nejaké bázy priestorov U resp. V. Potom priradením $f \mapsto (f)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ je daný lineárny izomorfizmus vektorových priestorov $\mathcal{A}(V, U) \to K^{m\times(n+1)}$.

(d) Predpokladajme, že U, V sú konečnorozmerné a $\dim U = m$, $\dim V = n$. Odvoďte, či už z (b) alebo z (c), že potom aj $\mathcal{A}(V, U)$ je konečnorozmerný a $\dim \mathcal{A}(V, U) = m(n + 1)$.

(e) Ak V je konečnorozmerný, tak jeho duál $V^* = \mathcal{L}(V, K)$ tvorí nadrovinu v $\mathcal{A}(V, K)$ (pozri text tesne pred tvrdením 6.5.3).

9. Afinné podpriestory a sústavy lineárnych rovníc

V tejto kapitole sa opäť pozrieme cez prizmu toho, čo sme sa dosiaľ naučili, na sústavy lineárnych rovníc. Uvidíme, že množina riešení každej takej sústavy tvorí afinný (v homogénnom prípade dokonca lineárny) podpriestor niektorého stĺpcového vektorového priestoru K^n. Taktiež naopak, ukážeme, že každý afinný podpriestor v K^n možno popísať ako podpriestor riešení vhodnej sústavy lineárnych rovníc. Na vyjadrenie množiny riešení tejto sústavy pomocou parametrov sa potom možno dívať ako na *parametrické rovnice* príslušného afinného podpriestoru. Získané znalosti nám umožnia v konkrétnych prípadoch určiť vzájomnú polohu afinných podpriestorov.

V celej kapitole K označuje pevné, inak ľubovoľné, pole; m, n sú ľubovoľné, pevne zvolené prirodzené čísla.

9.1 Podpriestor riešení homogénnej sústavy a jeho báza

Nech $\boldsymbol{A} \in K^{m\times n}$, $\boldsymbol{b} \in K^m$. Uvažujme homogénnu sústavu lineárnych rovníc s maticou $\boldsymbol{A}$

$$\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{0},$$

a nehomogénnu sústavu s maticou $\boldsymbol{A}$ a pravou stranou $\boldsymbol{b}$

$$\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}.$$

Množiny ich riešení označíme

$$\mathcal{R}(\boldsymbol{A}) = \{\boldsymbol{x} \in K^n;\ \boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{0}\},$$

resp.

$$\mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b}) = \{\boldsymbol{x} \in K^n;\ \boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}\}.$$

Predpisom $\varphi(\boldsymbol{x}) = \boldsymbol{A} \cdot \boldsymbol{x}$ je definované lineárne zobrazenie $\varphi\colon K^n \to K^m$, pričom $\mathcal{R}(\boldsymbol{A}) = \operatorname{Ker}\varphi$. Z toho okamžite vyplýva

9.1.1. Tvrdenie. *Pre ľubovoľnú maticu $\boldsymbol{A} \in K^{m\times n}$ množina $\mathcal{R}(\boldsymbol{A})$ riešení homogénnej sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{0}$ tvorí lineárny podpriestor vektorového priestoru K^n.*

Doteraz sme homogénnu sústavu riešili úpravou jej matice $\boldsymbol{A}$ na redukovaný stupňovitý tvar $\boldsymbol{B}$. Z tohto tvaru sme potom vyčítali, ktoré neznáme si zvolíme za parametre a ktoré neznáme si vyjadríme pomocou nich. Presnejšie, neznámu x_j sme si zvolili za parameter práve vtedy, keď sa v j-tom stĺpci matice $\boldsymbol{B}$ nenachádzal vedúci prvok žiadneho riadku matice $\boldsymbol{B}$; ak sa v j-tom stĺpci nachádzal vedúci prvok nejakého riadku, tak neznámu x_j sme si vyjadrili pomocou týchto parametrov.

Uvedená veta nám umožňuje alternatívny popis množiny riešení $\mathcal{R}(\boldsymbol{A})$ – kedže ide o lineárny podpriestor v K^n, môžeme ho najúspornejšie popísať zadaním (niektorej) jeho bázy. Každú bázu priestoru $\mathcal{R}(\boldsymbol{A})$ nazývame tiež *fundamentálnym systémom riešení* sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{0}$. Potom každé riešenie príslušnej homogénnej sústavy možno jednoznačne vyjadriť ako lineárnu kombináciu vektorov z fundamentálneho systému riešení, a tiež naopak, každá lineárna kombinácia vektorov fundamentálneho systému je riešením príslušnej sústavy. Fundamentálny systém riešení nájdeme nasledujúcim postupom.

Maticu $\boldsymbol{A}$ upravíme pomocou ERO na redukovaný stupňovitý tvar $\boldsymbol{B} \in K^{m\times n}$. Množinu $\{1,\dots,n\}$ rozdelíme na dve podmnožiny J, J', podľa toho, či sa v j-tom stĺpci matice $\boldsymbol{B}$ nachádza alebo nenachádza vedúci prvok nejakého jej riadku. Označme k počet prvkov množiny J' a zapíšme ju v tvare $J' = \{j_1 < j_2 < \dots < j_k\}$. Pre každý index $j_l \in J'$ zostrojíme vektor $\boldsymbol{v}_l = (v_{1l},\dots,v_{nl})^{\mathsf{T}} \in K^n$ takto: Zvolíme $v_{j_l l} = 1$ a $v_{j_i l} = 0$ pre $i \neq l$. Pre $j \in J$ vypočítame hodnoty v_{jl} k uvedeným hodnotám parametrov $v_{j_1 l},\dots,v_{j_k l}$ tak, aby celý vektor $\boldsymbol{v}_l$ vyhovoval podmienke $\boldsymbol{B} \cdot \boldsymbol{v}_l = \boldsymbol{0}$. Potom vektory $\boldsymbol{v}_1, \boldsymbol{v}_2,\dots,\boldsymbol{v}_k$ tvoria bázu podpriestoru riešení $\mathcal{R}(\boldsymbol{A})$. Pritom zrejme platí $k = n - h(\boldsymbol{A})$.

Namiesto dôkazu posledného tvrdenia si celý postup ozrejmíme na príklade.

9.1.2. Príklad. Predpokladajme, že sme maticu $\boldsymbol{A}$ pomocou ERO už upravili na redukovaný stupňovitý tvar

$$\boldsymbol{B} = \begin{pmatrix} 1 & 2 & 0 & 0 & -1/3 \\ 0 & 0 & 1 & 0 & 1/2 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Vedúce prvky riadkov sa nachádzajú v stĺpcoch 1, 3 a 4. Teda neznáme x_2 a x_5 si zvolíme za parametre a neznáme x_1, x_3 a x_4 si vyjadríme pomocou nich. Naša prvá voľba je $x_2 = 1$, $x_5 = 0$. Tomu zodpovedá vektor $\boldsymbol{v}_1 = (-2,1,0,0,0)^{\mathsf{T}}$. Druhá voľba parametrov je $x_2 = 0$, $x_5 = 1$. Tomu zodpovedá vektor $\boldsymbol{v}_2 = (1/3,0,-1/2,2,1)^{\mathsf{T}}$. Potom vektory $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ tvoria bázu podpriestoru (fundamentálny systém) riešení $\mathcal{R}(\boldsymbol{A}) = \mathcal{R}(\boldsymbol{B}) \subseteq \mathbb{R}^5$.

Ak sa „z estetických dôvodov“ chceme vyhnúť zlomkom vo výsledku, stačí miesto vektorov obsahujúcich ako súradnice zlomky vziať ich vhodné nenulové skalárne násobky. V našom prípade stačí nahradiť bázový vektor $\boldsymbol{v}_2$ „krajším“ bázovým vektorom $6\boldsymbol{v}_2 = (2, 0, -3, 12, 6)^{\mathsf{T}}$.

Pre „veľkosť“ podpriestoru riešení homogénnej sústavy z našich úvah vyplýva

9.1.3. Tvrdenie. *Pre ľubovoľnú maticu $\boldsymbol{A} \in K^{m\times n}$ platí*

$$\dim \mathcal{R}(\boldsymbol{A}) = n - h(\boldsymbol{A}).$$

9.2 Podpriestor riešení nehomogénnej sústavy

Prejdime teraz k otázke štruktúry množiny riešení $\mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b})$ nehomogénnej sústavy $\boldsymbol{A}\cdot\boldsymbol{x} = \boldsymbol{b}$.

9.2.1. Tvrdenie. *Nech $\boldsymbol{A} \in K^{m\times n}$, $\boldsymbol{b} \in K^m$.*
(a) Ak $\boldsymbol{y}, \boldsymbol{z} \in \mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b})$, tak $\boldsymbol{y} - \boldsymbol{z} \in \mathcal{R}(\boldsymbol{A})$.
(b) Ak $\boldsymbol{z} \in \mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b})$, $\boldsymbol{x} \in \mathcal{R}(\boldsymbol{A})$, tak $\boldsymbol{z} + \boldsymbol{x} \in \mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b})$.

Dôkaz. Možno overiť priamym výpočtom.

Z uvedeného tvrdenia vyplýva, že na popis množiny $\mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b})$ všetkých riešení nehomogénnej sústavy stačí poznať podpriestor $\mathcal{R}(\boldsymbol{A})$ všetkých riešení príslušnej homogénnej sústavy, t. j. nejaký fundamentálny systém $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$ jej riešení, a ľubovoľné jedno riešenie $\boldsymbol{z} \in \mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b})$ nehomogénnej sústavy. Tvrdenie možno potom schématicky zapísať v niektorom z tvarov

$$\begin{aligned}\mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b}) &= \boldsymbol{z} + \mathcal{R}(\boldsymbol{A}) = \{\boldsymbol{z} + \boldsymbol{x};\ \boldsymbol{x} \in \mathcal{R}(\boldsymbol{A})\}\\ &= \boldsymbol{z} + [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n] = \{\boldsymbol{z} + c_1\boldsymbol{v}_1 + \ldots + c_k\boldsymbol{v}_k;\ c_1, \ldots, c_k \in K\}.\end{aligned}$$

Každý si môže vybrať ten, ktorý sa mu najväčšmi pozdáva.

S využitím pojmov predchádzajúcej kapitoly možno naše úvahy zhrnúť do nasledujúcej podoby.

9.2.2. Tvrdenie. *Nech $\boldsymbol{A} \in K^{m\times n}$, $\boldsymbol{b} \in K^m$. Ak sústava $\boldsymbol{A}\cdot\boldsymbol{x} = \boldsymbol{b}$ má aspoň jedno riešenie, tak $\mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b})$ je afinný podpriestor v K^n so zameraním $\mathcal{R}(\boldsymbol{A})$. To znamená,*

$$\operatorname{Dir}\mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b}) = \mathcal{R}(\boldsymbol{A}) \quad \text{a} \quad \dim\mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b}) = \dim\mathcal{R}(\boldsymbol{A}) = n - h(\boldsymbol{A}).$$

9.3 Frobeniova veta a riešenie nehomogénnej sústavy

Odpoveď na otázku riešiteľnosti nehomogénnej sústavy možno dať porovnaním hodnosti jej základnej a rozšírenej matice.

Začneme pozorovaním, že sústava $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ má aspoň jedno riešenie $\boldsymbol{z} \in K^n$ práve vtedy, keď $\boldsymbol{b} \in \operatorname{Im} \varphi$ (kde $\varphi(\boldsymbol{x}) = \boldsymbol{A} \cdot \boldsymbol{x}$). Ak tento prípad nastane, tak, ako sme už ukázali, $\mathcal{R}(\boldsymbol{A} \,|\, \boldsymbol{b}) = \boldsymbol{z} + \mathcal{R}(\boldsymbol{A})$.

9.3.1. Veta. (Frobenius) *Nech $\boldsymbol{A} \in K^{m\times n}$, $\boldsymbol{b} \in K^m$. Potom nehomogénna sústava $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ má riešenie práve vtedy, keď $h(\boldsymbol{A} \,|\, \boldsymbol{b}) = h(\boldsymbol{A})$.*

Dôkaz. Z poznámky vyslovenej tesne pred vetou vyplýva, že sústava $\boldsymbol{A}\cdot\boldsymbol{x} = \boldsymbol{b}$ má riešenie práve vtedy, keď $\boldsymbol{b} \in [\boldsymbol{s}_1(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A})]$, t. j. práve vtedy keď

$$[\boldsymbol{s}_1(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A}), \boldsymbol{b}] = [\boldsymbol{s}_1(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A})].$$

Keďže

$$\begin{aligned} h(\boldsymbol{A} \,|\, \boldsymbol{b}) &= \dim[\boldsymbol{s}_1(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A}), \boldsymbol{b}], \\ h(\boldsymbol{A}) &= \dim[\boldsymbol{s}_1(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A})] \end{aligned}$$

a druhý z týchto podpriestorov je podpriestorom prvého, uvedená podmienka je zrejme ekvivalentná s rovnosťou $h(\boldsymbol{A} \,|\, \boldsymbol{b}) = h(\boldsymbol{A})$.

Frobeniova veta vlastne hovorí už známu vec: nehomogénna sústava $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ nemá riešenie práve vtedy, keď sa pri úprave jej rozšírenej matice $(\boldsymbol{A} \,|\, \boldsymbol{b})$ na redukovaný stupňovitý tvar objaví nejaký riadok tvaru $(0, \dots, 0 \,|\, d) \in K^{n+1}$, kde $0 \neq d \in K$. Takýto riadok totiž zodpovedá rovnici $0 = d$.

Ak upravíme pomocou ERO rozšírenú maticu $(\boldsymbol{A} \,|\, \boldsymbol{b})$ na redukovaný stupňovitý tvar $(\boldsymbol{B} \,|\, \boldsymbol{c})$, kde $\boldsymbol{B} \in K^{m\times n}$ a $\boldsymbol{c} \in K^m$, tak $\boldsymbol{B}$ je tiež v redukovanom stupňovitom tvare. Potom $\mathcal{R}(\boldsymbol{A} \,|\, \boldsymbol{b}) = \mathcal{R}(\boldsymbol{B} \,|\, \boldsymbol{c}) \neq \emptyset$ práve vtedy, keď sa žiaden vedúci prvok nejakého riadku matice $(\boldsymbol{B} \,|\, \boldsymbol{c})$ nenachádza v poslednom, t. j. $(n+1)$-om stĺpci. Bázu priestoru riešení $\mathcal{R}(\boldsymbol{A}) = \mathcal{R}(\boldsymbol{B})$ nájdeme postupom popísaným v paragrafe 9.1 Nech J, J' a k majú tam uvedený význam. Jedno riešenie $\boldsymbol{z} = (z_1, \dots, z_n)^{\mathsf{T}}$ nehomogénnej sústavy dostaneme voľbou parametrov $z_{j_1} = \ldots = z_{j_k} = 0$ pre $j_l \in J'$. Zvyšné hodnoty z_j potom vypočítame tak, aby $\boldsymbol{z}$ vyhovovalo podmienke $\boldsymbol{B} \cdot \boldsymbol{z} = \boldsymbol{c}$, t. j. $z_j = c_j$ pre $j \in J$.

9.3.2. Príklad. Predpokladajme, že sme maticu $(\boldsymbol{A} \,|\, \boldsymbol{b})$ pomocou ERO už upravili na redukovaný stupňovitý tvar

$$(\boldsymbol{B} \,|\, \boldsymbol{c}) = \left(\begin{array}{cccccc|c} 1 & 0 & 0 & 3 & 1/4 & 0 & 2 \\ 0 & 1 & 0 & 4 & 2 & -1 & -1 \\ 0 & 0 & 1 & 1 & -5 & 6 & -2/7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Vidíme, že $h(\boldsymbol{B} \,|\, \boldsymbol{c}) = h(\boldsymbol{B}) = 3$, teda $\mathcal{R}(\boldsymbol{A} \,|\, \boldsymbol{b}) = \mathcal{R}(\boldsymbol{B} \,|\, \boldsymbol{c}) \neq \emptyset$. Vedúce prvky riadkov sa nachádzajú v stĺpcoch 1, 2 a 3. Teda neznáme x_4, x_5 a x_6 si zvolíme za parametre a neznáme x_1, x_2 a x_3 si vyjadríme pomocou nich. Prvej voľbe $x_4 = 1$, $x_5 = x_6 = 0$ zodpovedá vektor $\boldsymbol{v}_1 = (-3, -4, -1, 1, 0, 0)^{\mathsf{T}}$. Druhá voľba $x_4 = 0$, $x_5 = 1$, $x_6 = 0$ nám dá vektor $\boldsymbol{v}_2 = (-1/4, -2, 5, 0, 1, 0)^{\mathsf{T}}$. Treťou voľbou $x_4 = x_5 = 0$, $x_6 = 1$ získame vektor $\boldsymbol{v}_3 = (0, 1, -6, 0, 0, 1)^{\mathsf{T}}$. Potom vektory $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ tvoria bázu podpriestoru riešení $\mathcal{R}(\boldsymbol{A}) = \mathcal{R}(\boldsymbol{B}) \subseteq \mathbb{R}^6$ príslušnej homogénnej sústavy. Konečne voľbou parametrov $x_4 = x_5 = x_6 = 0$ získame jedno riešenie $\boldsymbol{z} = (2, -1, -2/7, 0, 0, 0)^{\mathsf{T}}$ nehomogénnej sústavy.

Výsledok možno prehľadne zapísať do tabuľky (treba si však uvedomiť, že sme ju vypĺňali uvedeným postupom):

	$\boldsymbol{v}_1$	$\boldsymbol{v}_2$	$\boldsymbol{v}_3$	$\boldsymbol{z}$
x_1	-3	$-1/4$	0	2
x_2	-4	-2	1	-1
x_3	-1	5	-6	$-2/7$
x_4	1	0	0	0
x_5	0	1	0	0
x_6	0	0	1	0

Porovnajte túto tabuľku s maticou $(\boldsymbol{B}\,|\,\boldsymbol{c})$!

9.4 Parametrické a všeobecné rovnice afinných podpriestorov

Hoci sa v tomto i v nasledujúcom paragrafe obmedzíme len na afinné podpriestory stĺpcového vektorového priestoru K^n, naše úvahy majú širšiu platnosť. Zvolením pevnej bázy ich možno pomocou súradnicového zobrazenia zrejmým spôsobom preniesť aj na afinné podpriestory ľubovoľného konečnorozmerného vektorového priestoru V.

Každý afinný podpriestor $M \subseteq K^n$ má tvar

$$M = \boldsymbol{p} + [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k] = \boldsymbol{p} + [\boldsymbol{\alpha}]$$

pre nejaký bod $\boldsymbol{p} = (p_1, \ldots, p_n)^{\mathsf{T}} \in M$ a vhodnú usporiadanú k-ticu $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ vektorov z K^n, kde $\boldsymbol{u}_j = (u_{1j}, \ldots, u_{nj})^{\mathsf{T}}$. To znamená, že pre ľubovoľné $\boldsymbol{x} \in K^n$ platí $\boldsymbol{x} \in M$ práve vtedy, keď existuje $\boldsymbol{t} = (t_1, \ldots, t_k)^{\mathsf{T}} \in K^k$ také, že

$$\boldsymbol{x} = \boldsymbol{p} + \boldsymbol{\alpha} \cdot \boldsymbol{t},$$

kde usporiadanú k-ticu $\boldsymbol{\alpha}$ sme ako obyčajne stotožnili s maticou $(u_{ij}) \in K^{n\times k}$ so stĺpcami $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$. Rovnosť $\boldsymbol{x} = \boldsymbol{p} + \boldsymbol{\alpha} \cdot \boldsymbol{t}$ je maticovým zápisom *parametrických rovníc* afinného podpriestoru $M \subseteq K^n$. Vektor $\boldsymbol{t} \in K^n$ nazývame *vektorom parametrov* a jeho zložky $t_1, \ldots, t_k \in K$ *parametrami.* Po rozpísaní do zložiek

$$\begin{aligned} x_1 &= p_1 + u_{11}t_1 + u_{12}t_2 + \ldots + u_{1k}t_k \\ x_2 &= p_2 + u_{21}t_1 + u_{22}t_2 + \ldots + u_{2k}t_k \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= p_n + u_{n1}t_1 + u_{n2}t_n + \ldots + u_{nk}t_k \end{aligned}$$

dostaneme obvyklejší tvar, s akým sme sa v dimenzii $n = 2$ resp. $n = 3$ už stretli v stredoškolskej analytickej geometrii.

Ak navyše vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ sú lineárne nezávislé, čo možno vždy dosiahnuť vynechaním „nadbytočných“ vektorov, tak parametrické rovnice podpriestoru M nám priamo ukážu jeho dimenziu: $\dim M = k$.

Zápis afinného podpriestoru $M \subseteq K^n$ v tvare $M = \boldsymbol{p} + [\boldsymbol{\alpha}]$, kde $\boldsymbol{p} \in M$ a $\boldsymbol{\alpha}$ je nejaká usporiadaná k-tica, ktorá generuje jeho zameranie $\mathrm{Dir}\, M$ (môžeme si dovoliť predpokladať, že $\boldsymbol{\alpha}$ je dokonca báza v $\mathrm{Dir}\, M$), budeme nazývať jeho *parametrickým vyjadrením.* Parametrické vyjadrenie $M = \boldsymbol{p} + [\boldsymbol{\alpha}]$ afinného podpriestoru možno priamo prepísať do jeho parametrických rovníc $\boldsymbol{x} = \boldsymbol{p} + \boldsymbol{\alpha} \cdot \boldsymbol{t}$, ($\boldsymbol{t} \in K^k$). Taktiež naopak, z jeho parametrických rovníc možno okamžite získať jeho parametrické vyjadrenie. Súvis medzi týmito dvoma druhmi popisu je natoľko bezprostredný, že ich ani nemusíme príliš úzkostlivo rozlišovať. Na druhej strane, v predchádzajúcich paragrafoch sme videli, že každá sústava lineárnych rovníc $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ s rozšírenou maticou $(\boldsymbol{A} \,|\, \boldsymbol{b}) \in K^{m \times (n+1)}$ (pokiaľ má riešenie), popisuje afinný podpriestor $\mathcal{R}(\boldsymbol{A} \,|\, \boldsymbol{b}) \subseteq K^n$. Nájsť parametrické rovnice tohto podpriestoru už vieme, treba si to len uvedomiť. Ak totiž $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_k)$ je báza podpriestoru $\mathcal{R}(\boldsymbol{A})$ riešení príslušenej homogénnej sústavy a $\boldsymbol{z} \in \mathcal{R}(\boldsymbol{A} \,|\, \boldsymbol{b})$ je ľubovoľné jedno riešenie nehomogénnej sústavy – a jedno i druhé (pokiaľ existuje) naozaj vieme nájsť –, tak

$$\boldsymbol{x} = \boldsymbol{z} + \boldsymbol{\beta} \cdot \boldsymbol{t},$$

kde $\boldsymbol{t} \in K^k$ je vektor parametrov, sú parametrické rovnice afinného podpriestoru $\mathcal{R}(\boldsymbol{A} \,|\, \boldsymbol{b}) \subseteq K^n$.

Inak povedané, vyriešiť sústavu lineárnych rovníc $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ znamená vlastne nájsť nejaké (prípadne nie celkom hocaké ale v istom zmysle „pekné“) parametrické rovnice (alebo parametrické vyjadrenie) afinného podpriestoru $\mathcal{R}(\boldsymbol{A} \,|\, \boldsymbol{b}) \subseteq K^n$.

Často je však potrebné riešiť obrátenú úlohu: k parametricky zadanému afinnému podpriestoru $M \subseteq K^n$ nájsť jeho *všeobecné rovnice*, t. j. sústavu lineárnych rovníc $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ o n neznámych $x_1, \dots, x_n$ takú, že $M = \mathcal{R}(\boldsymbol{A} \,|\, \boldsymbol{b})$.

Nech teda $M = \boldsymbol{p} + [\boldsymbol{\alpha}]$ je afinný podpriestor v K^n, daný bodom $\boldsymbol{p} \in K^n$ a usporiadanou k-ticou $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ vektorov z K^n, ktorú stotožníme s maticou $\boldsymbol{\alpha} = (u_{ij}) \in K^{n \times k}$ so stĺpcami $\boldsymbol{u}_j$. Parametrické rovnice $\boldsymbol{x} = \boldsymbol{p} + \boldsymbol{\alpha} \cdot \boldsymbol{t}$ podpriestoru M, kde $\boldsymbol{x} = (x_1, \dots, \boldsymbol{x}_n)^\mathsf{T} \in K^n$ je vektor neznámych a $\boldsymbol{t} = (t_1, \dots, t_k)^\mathsf{T} \in K^k$ je vektor parametrov, možno prepísať do tvaru

$$\boldsymbol{I}_n \cdot \boldsymbol{x} = \boldsymbol{\alpha} \cdot \boldsymbol{t} \, + \, \boldsymbol{p},$$

ktorý možno reprezentovať blokovou maticou

$$(\boldsymbol{I}_n \,|\, \boldsymbol{\alpha} \,|\, \boldsymbol{p}).$$

Naša metóda bude založená na *eliminácii parametrov* $t_1, \dots, t_k$ úpravou tejto matice pomocou ERO. Maticu $(\boldsymbol{I}_n \,|\, \boldsymbol{\alpha} \,|\, \boldsymbol{p})$ budeme upravovať na riadkovo ekvivalentnú maticu tak, aby stredný blok vo výslednej matici bol v stupňovitom tvare. Môžu nastať dve možnosti

(1) $h(\boldsymbol{\alpha}) = n$, čo spoznáme podľa toho, že všetky riadky stredného bloku výslednej matice sú nenulové. V tom prípade $M = V$ a všeobecné rovnice tohto podpriestoru tvorí prázdna sústava (t. j. sústava ktorá neobsahuje žiadnu rovnicu). My sa jednoducho uspokojíme s konštatovaním $M = V$ a nijakými všeobecnými rovnicami sa ďalej nebudeme zaoberať.

(2) $h(\boldsymbol{\alpha}) < n$. Vtedy možno stredný blok výslednej matice rozdeliť do dvoch pod sebou umiestnených blokov $\left(\frac{\boldsymbol{D}}{\boldsymbol{0}}\right)$, kde horný blok $\boldsymbol{D}$ je stupňovitá matica typu $h(\boldsymbol{\alpha}) \times k$, ktorá má všetky riadky nenulové, teda dolný nulový blok má rozmer $(n - h(\boldsymbol{\alpha})) \times k$. Toto rozdelenie stredného bloku indukuje rozdelenie celej výslednej matice do blokov
$$\left(\begin{array}{c|c|c} \boldsymbol{A}' & \boldsymbol{D} & \boldsymbol{b}' \\ \hline \boldsymbol{A} & \boldsymbol{0} & \boldsymbol{b} \end{array}\right).$$
Potom $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ sú všeobecné rovnice afinného podpriestoru M, t. j. platí $M = \boldsymbol{p} + [\boldsymbol{\alpha}] = \mathcal{R}(\boldsymbol{A} \,|\, \boldsymbol{b})$.

Popísaný algoritmus možno stručne zhrnúť do schémy
$$(\boldsymbol{I}_n \,|\, \boldsymbol{\alpha} \,|\, \boldsymbol{p}) \xrightarrow{\text{ERO}} \left(\begin{array}{c|c|c} \boldsymbol{A}' & \boldsymbol{D} & \boldsymbol{b}' \\ \hline \boldsymbol{A} & \boldsymbol{0} & \boldsymbol{b} \end{array}\right),$$
kde $\boldsymbol{D}$ je matica v stupňovitom tvare s nenulovými riadkami (ktorých počet teda nutne je $h(\boldsymbol{D}) = h(\boldsymbol{\alpha})$). Ako vedľajší produkt takéhoto výpočtu, možno z k-tice $\boldsymbol{\alpha}$ vybrať bázu zamerania $\operatorname{Dir} M = [\boldsymbol{\alpha}]$: je tvorená vektormi $\boldsymbol{u}_{j_1}, \dots, \boldsymbol{u}_{j_l}$, kde $1 \le j_1 < \ldots < j_l \le k$ sú indexy tých stĺpcov matice $\boldsymbol{D}$, v ktorých sa nachádzajú vedúce prvky jej riadkov (pozri tvrdenie 4.5.3). Správnosť celého algoritmu vyplýva z nasledujúceho tvrdenia.

9.4.1. Tvrdenie. *Nech $\boldsymbol{B} \in K^{n\times m}$, $\boldsymbol{C} \in K^{n\times k}$ a $\boldsymbol{p} \in K^n$. Ak bloková matica $(\boldsymbol{B} \,|\, \boldsymbol{C} \,|\, \boldsymbol{p})$ je riadkovo ekvivalentná s blokovou maticou*
$$\left(\begin{array}{c|c|c} \boldsymbol{A}' & \boldsymbol{D} & \boldsymbol{b}' \\ \hline \boldsymbol{A} & \boldsymbol{0} & \boldsymbol{b} \end{array}\right),$$
kde $\boldsymbol{D}$ je matica v stupňovitom tvare s nenulovými riadkami, tak
$$\mathcal{R}(\boldsymbol{A} \,|\, \boldsymbol{b}) = \left\{\boldsymbol{x} \in K^m;\ (\exists \boldsymbol{t} \in K^k)(\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{C} \cdot \boldsymbol{t} + \boldsymbol{p})\right\}.$$

Dôkaz. Matica $(\boldsymbol{B}\,|\,\boldsymbol{C}\,|\,\boldsymbol{p})$ zodpovedá sústave $\boldsymbol{B}\cdot\boldsymbol{x}=\boldsymbol{C}\cdot\boldsymbol{t}+\boldsymbol{p}$ v neznámych $x_1,\dots,x_m$, $t_1,\dots,t_k$. Podobne matica

$$\left(\begin{array}{c|c|c} \boldsymbol{A}' & \boldsymbol{D} & \boldsymbol{b}' \\ \hline \boldsymbol{A} & \boldsymbol{0} & \boldsymbol{b} \end{array}\right)$$

zodpovedá sústave

$$\begin{aligned} \boldsymbol{A}'\cdot\boldsymbol{x} &= \boldsymbol{D}\cdot\boldsymbol{t}+\boldsymbol{b}' \\ \boldsymbol{A}\cdot\boldsymbol{x} &= \boldsymbol{b} \end{aligned}$$

v rovnakých neznámych. Vzhľadom na riadkovú ekvivalenciu príslušných matíc sú obe sústavy ekvivalentné.

Dokážeme, že pre ľubovoľné $\boldsymbol{x}\in K^m$ nasledujúce podmienky sú ekvivalentné:
(i) $\boldsymbol{A}\cdot\boldsymbol{x}=\boldsymbol{b}$;
(ii) $(\exists\boldsymbol{t}\in K^k)(\boldsymbol{A}'\cdot\boldsymbol{x}=\boldsymbol{D}\cdot\boldsymbol{t}+\boldsymbol{b}'\ \&\ \boldsymbol{A}\cdot\boldsymbol{x}=\boldsymbol{b})$;
(iii) $(\exists\boldsymbol{t}\in K^k)(\boldsymbol{B}\cdot\boldsymbol{x}=\boldsymbol{C}\cdot\boldsymbol{t}+\boldsymbol{p})$.
Keďže implikácie (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i) platia triviálne, vysvetlenie potrebuje iba implikácia
(i) $\Rightarrow$ (ii). Zrejme hodnosť matice $\boldsymbol{D}$ sa rovná počtu jej riadkov a ten je $\le n$. Preto tiež $h(\boldsymbol{D}\,|\,\boldsymbol{A}'\cdot\boldsymbol{x}-\boldsymbol{b}')=h(\boldsymbol{D})$ nezávisle na $\boldsymbol{x}$. Ale to podľa Frobeniovej vety 9.3.1 znamená, že sústava $\boldsymbol{D}\cdot\boldsymbol{t}=\boldsymbol{A}'\cdot\boldsymbol{x}-\boldsymbol{b}'$ (v neznámych $t_1,\dots,t_k$) má nejaké riešenie $\boldsymbol{t}\in K^k$ pre ľubovoľné $\boldsymbol{x}\in K^m$.

Poznámka. Z uvedeného dôkazu vyplýva, že tvrdenie 9.4.1 zostáva v platnosti, aj keď matica $\boldsymbol{D}$ nie je v stupňovitom tvare; stačí žiadať len lineárnu nezávislosť jej riadkov. Tú však možno najistejšie nahliadnuť práve úpravou príslušnej matice na stupňovitý tvar.

9.4.2. Príklad. Nájdeme všeobecné rovnice afinného podpriestorru $M=\boldsymbol{p}+[\boldsymbol{\alpha}]$ stĺpcového vektorového priestoru $\mathbb{Z}_{11}^5$ nad poľom $\mathbb{Z}_{11}$, kde $\boldsymbol{p}=(1,2,3,4,5)^{\mathsf{T}}$ a $\boldsymbol{\alpha}$ je tvorená stĺpcami matice

$$\begin{pmatrix} 2 & 1 & 0 & 1 \\ 1 & 5 & 2 & 6 \\ 0 & 2 & 7 & 0 \\ 1 & 6 & 0 & 9 \\ 0 & 3 & 5 & 8 \end{pmatrix}.$$

Budeme upravovať maticu

$$\left(\begin{array}{ccccc|cccc|c} 1 & 0 & 0 & 0 & 0 & 2 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 5 & 2 & 6 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 2 & 7 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & 1 & 6 & 0 & 9 & 4 \\ 0 & 0 & 0 & 0 & 1 & 0 & 3 & 5 & 8 & 5 \end{array}\right)$$

pomocou ERO tak, aby stredný blok nadobudol stupňovitý tvar. Po niekoľkých krokoch dostaneme

$$\left(\begin{array}{ccccc|cccc|c} 1 & 0 & 0 & 0 & 0 & 2 & 1 & 0 & 1 & 1 \\ 5 & 1 & 0 & 0 & 0 & 0 & 10 & 2 & 0 & 7 \\ 4 & 3 & 0 & 0 & 1 & 0 & 0 & 0 & 8 & 4 \\ 10 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 6 \\ 9 & 3 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 2 \end{array}\right).$$

Predovšetkým vidíme, že tretí vektor štvorice $\boldsymbol{\alpha}$ je lineárnou kombináciou predchádzajúcich dvoch, preto ho možno v príslušnom parametrickom vyjadrení podpriestoru M vynechať. Zvyšné tri vektory v $\boldsymbol{\alpha}$ sú lineárne nezávislé, čiže $\dim M = 3$. Konečne, všeobecné rovnice podpriestoru M vyzerajú takto

$$\begin{aligned} 10x_1 + 2x_2 + x_3 \phantom{{}+ x_4 + x_5} &= 6 \\ 9x_1 + 3x_2 \phantom{{}+ x_3} + x_4 + x_5 &= 2\,. \end{aligned}$$

Dosadením sa možno presvedčiť, že bod $\boldsymbol{p}$ skutočne vyhovuje tejto sústave a vektory štvorice $\boldsymbol{\alpha}$ príslušnej homogénnej sústave.

9.5 Rovnice prieniku a spojenia afinných podpriestorov

V tomto paragrafe sa pokúsime zostaviť všeobecné recepty, pomocou ktorých budeme vedieť napísať či už všeobecné alebo parametrické rovnice prieniku a spojenia dvoch afinných podpriestorov stĺpcového vektorového priestoru K^n. Pri tom vezmeme do úvahy tri možnosti zadania pôvodných podpriestorov:

(1) Oba podpriestory sú zadané všeobecnými rovnicami.
(2) Oba podpriestory sú zadané parametricky.
(3) Jeden podpriestor je zadaný pomocou všeobecných rovníc a druhý parametricky.

Každú situáciu budeme ilustrovať na jednom až dvoch konkrétnych príkladoch, v ktorých navyše vyšetríme i vzájomnú polohu oboch afinných podpriestorov.

(1) Nech afinné podpriestory M, $N \subseteq K^n$ majú všeobecné rovnice $\boldsymbol{A}\cdot\boldsymbol{x} = \boldsymbol{b}$ resp. $\boldsymbol{B}\cdot\boldsymbol{x} = \boldsymbol{c}$, kde $\boldsymbol{A} \in K^{m\times n}$, $\boldsymbol{b} \in K^m$, $\boldsymbol{B} \in K^{l\times n}$, $\boldsymbol{c} \in K^l$. Potom všeobecnými rovnicami prieniku $M \cap N$ je sústava

$$\begin{aligned} \boldsymbol{A}\cdot\boldsymbol{x} &= \boldsymbol{b} \\ \boldsymbol{B}\cdot\boldsymbol{x} &= \boldsymbol{c} \end{aligned}$$

s rozšírenou maticou

$$\left(\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{b} \\ \boldsymbol{B} & \boldsymbol{c} \end{array}\right).$$

Parametrické vyjadrenie prieniku $M \cap N$ možno získať vyriešením tejto sústavy.

Ak hľadáme parametrické vyjadrenie spojenia $M \sqcup N$, najprv vyriešením sústav s rozšírenými maticami $(\boldsymbol{A}\,|\,\boldsymbol{b})$ resp. $(\boldsymbol{B}\,|\,\boldsymbol{c})$ získame parametrické vyjadrenia jednotlivých podpriestorov M a N. Z nich na základe tvrdenia 8.3.3 (b) zostavíme parametrické vyjadrenie podpriestoru $M \sqcup N$ (pozri tiež príklad 8.3.5). Konečne, ak nás zaujímajú všeobecné rovnice podpriestoru $M \sqcup N$, môžeme ich odvodiť z jeho parametrických rovníc metódou opísanou v druhej časti predchádzajúceho paragrafu 9.4 (pozri tvrdenie 9.4.1 a príklad 9.4.2).

9.5.1. Príklad. Afinné podpriestory M, N vektorového priestoru $\mathbb{Q}^4$ sú dané sústavami

$$\begin{aligned} x_1 + x_2 - x_3 + x_4 &= 9 \\ x_1 - x_2 + x_3 - x_4 &= -3 \end{aligned}$$

resp.

$$\begin{aligned} x_1 + 3x_2 + 2x_3 - x_4 &= 0 \\ x_1 - 3x_2 - 2x_3 + x_4 &= 6\,. \end{aligned}$$

Ak dáme tieto sústavy dohromady, získame všeobecné rovnice prieniku. Ich riešenie však bude výhodné trochu odložiť a najprv upraviť rozšírené matice pôvodných sústav:

$$\left(\begin{array}{rrrr|r} 1 & 1 & -1 & 1 & 9 \\ 1 & -1 & 1 & -1 & -3 \end{array}\right) \sim \left(\begin{array}{rrrr|r} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & -1 & 1 & 6 \end{array}\right),$$

$$\left(\begin{array}{rrrr|r} 1 & 3 & 2 & -1 & 0 \\ 1 & -3 & -2 & 1 & 6 \end{array}\right) \sim \left(\begin{array}{rrrr|r} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 2/3 & -1/3 & -1 \end{array}\right).$$

Z upravených matíc okamžite dostávame parametrické vyjadrenie pôvodných podpriestorov (matica v hranatých zátvorkách označuje lineárny podpriestor generovaný jej stĺpcami)

$$M = \begin{pmatrix} 3 \\ 6 \\ 0 \\ 0 \end{pmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad N = \begin{pmatrix} 3 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \begin{bmatrix} 0 & 0 \\ -2 & 1 \\ 3 & 0 \\ 0 & 3 \end{bmatrix}.$$

Ak napíšeme obe upravené rozšírené matice všeobecných rovníc podpriestorov M a N do blokov pod seba, dostaneme rozšírenú maticu všeobecných rovníc podpriestoru $M \cap N$. Jej úpravou na redukovaný stupňovitý tvar vyjde

$$\left(\begin{array}{rrrr|r} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1/5 & 9/5 \\ 0 & 0 & 1 & -4/5 & -21/5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Odtiaľ už priamo vyplýva parametrické vyjadrenie

$$M \cap N = \begin{pmatrix} 3 \\ 9/5 \\ -21/5 \\ 0 \end{pmatrix} + \begin{bmatrix} 0 \\ -1 \\ 4 \\ 5 \end{bmatrix}.$$

Zistili sme, že dvojrozmerné afinné podpriestory M, N majú jednorozmerný prienik, teda sú *rôznobežné*. Preto tiež $\dim(M \sqcup N) = 2 + 2 - 1 = 3$.

Ak postavíme vedľa seba generátory smerových podpriestorov $\operatorname{Dir} M$ a $\operatorname{Dir} N$, úpravou príslušnej matice zistíme, že prvé tri sú lineárne nezávislé a posledný z nich je lineárnou kombináciou predchádzajúcich. Teda stĺpce matice

$$\boldsymbol{\beta} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & -2 \\ 1 & 0 & 3 \\ 0 & 1 & 0 \end{pmatrix}$$

tvoria bázu zamerania afinného podpriestoru $M \sqcup N$. Jeho parametrické vyjadrenie je

$$M \sqcup N = \boldsymbol{p} + [\boldsymbol{\beta}],$$

kde $\boldsymbol{p} = (3, 9/5, -21/5, 0)^{\mathsf{T}}$. Úpravou blokovej matice $(\boldsymbol{I}_4 \,|\, \boldsymbol{\beta} \,|\, \boldsymbol{p})$ podľa algoritmu z druhej časti paragrafu 9.4 (pozri poznámku za tvrdením 9.4.1) výmenou prvého a posledného riadku dostaneme všeobecné rovnice podpriestoru $M \sqcup N$:

$$x_1 = 3\,.$$

(2) Nech $M = \boldsymbol{p} + [\boldsymbol{\alpha}]$, $N = \boldsymbol{q} + [\boldsymbol{\beta}]$ sú parametrické vyjadrenia dvoch afinných podpriestorov v K^n. Potom, ako už vieme, $M \sqcup N = \boldsymbol{p} + [\boldsymbol{q} - \boldsymbol{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}]$ a podľa tvrdenia 4.5.3 vynechaním vhodných stĺpcov z blokovej matice $(\boldsymbol{q} - \boldsymbol{p}, \boldsymbol{\alpha}, \boldsymbol{\beta})$ možno dostať bázu zamerania $\operatorname{Dir}(M \sqcup N)$. Všeobecné rovnice podpriestoru $M \sqcup N$ dostaneme úpravou blokovej matice $(\boldsymbol{I}_n \,|\, \boldsymbol{q} - \boldsymbol{p}, \boldsymbol{\alpha}, \boldsymbol{\beta} \,|\, \boldsymbol{p})$, prípadne matice, v ktorej je prostredný blok nahradený bázou zamerania $\operatorname{Dir}(M \sqcup N)$, podľa algoritmu z paragrafu 9.4

Pokiaľ nás zaujímajú všeobecné rovnice prieniku $M \cap N$, najjednoduchšie ich získame tak, že parametrické rovnice každého z podpriestorov M, N prevedieme na všeobecné rovnice a tieto spojíme dohromady. Parametrické vyjadrenie prieniku $M \cap N$ dostaneme vyriešením jeho všeobecných rovníc.

Jestvuje aj iná cesta k parametrickým rovniciam prieniku $M \cap N$. Ako vedľajší produkt pri nej možno získať bázy zameraní $\operatorname{Dir} M$, $\operatorname{Dir} N$, $\operatorname{Dir}(M \sqcup N)$, teda aj parametrické rovnice spojenia $M \sqcup N$. Pri tejto metóde upravujeme blokovú maticu $(\boldsymbol{\alpha} \,|\, \boldsymbol{\beta} \,|\, \boldsymbol{q} - \boldsymbol{p})$ pomocou ERO na stupňovitý tvar

$$\left(\begin{array}{c|c|c} \boldsymbol{A}' & \boldsymbol{B}' & \boldsymbol{c}' \\ \hline \boldsymbol{0} & \boldsymbol{B} & \boldsymbol{c} \end{array}\right),$$

kde matica $\boldsymbol{A}'$ má všetky riadky nenulové (teda lineárne nezávislé a ich počet je $h(\boldsymbol{A}') = h(\boldsymbol{\alpha}) = \dim M$). Prienik $M \cap N$ je tvorený všetkými $\boldsymbol{x} = \boldsymbol{q} + \boldsymbol{\beta} \cdot \boldsymbol{t} \in N$, ktoré patria zároveň do M, t. j. existuje vektor parametrov $\boldsymbol{s}$ taký, že $\boldsymbol{x} = \boldsymbol{p} + \boldsymbol{\alpha} \cdot \boldsymbol{s}$. Hľadáme teda všetky vektory parametrov $\boldsymbol{t}$, ku ktorým existuje nejaký vektor parametrov $\boldsymbol{s}$ taký, že platí

$$\boldsymbol{\alpha} \cdot \boldsymbol{s} = \boldsymbol{\beta} \cdot \boldsymbol{t} + (\boldsymbol{q} - \boldsymbol{p}).$$

Podľa tvrdenia 9.4.1 (stačí v ňom zameniť poradie prvého a druhého zvislého bloku) k danému $\boldsymbol{t}$ existuje takéto $\boldsymbol{s}$ práve vtedy, keď $\boldsymbol{B} \cdot \boldsymbol{t} + \boldsymbol{c} = \boldsymbol{0}$. Vyriešením tejto sústavy získame parametrické vyjadrenie

$$\boldsymbol{t} = \boldsymbol{r} + \boldsymbol{\gamma} \cdot \boldsymbol{z},$$

kde $\boldsymbol{r} \in \mathcal{R}(\boldsymbol{B} \mid -\boldsymbol{c})$ a $\boldsymbol{\gamma}$ je báza lineárneho podpriestoru $\mathcal{R}(\boldsymbol{B})$, s vektorom parametrov $\boldsymbol{z}$, ktoré dosadíme do parametrických rovníc podpriestoru N. Dostaneme tak parametrické rovnice

$$\boldsymbol{x} = \boldsymbol{q} + \boldsymbol{\beta} \cdot (\boldsymbol{r} + \boldsymbol{\gamma} \cdot \boldsymbol{z}) = (\boldsymbol{q} + \boldsymbol{\beta} \cdot \boldsymbol{r}) + (\boldsymbol{\beta} \cdot \boldsymbol{\gamma}) \cdot \boldsymbol{z}$$

podpriestoru $M \cap N$.

Metóda zostavenia všeobecných rovníc prieniku $M \cap N$ ako i oboch typov rovníc spojenia $M \sqcup N$, popísaná v prvej časti bodu (2), je (aspoň dúfame) dostatočne jasná. V nasledujúcom príklade sa preto sústredíme len na nájdenie parametrických rovníc prieniku $M \cap N$ metódou z druhej časti a určenie vzájomnej polohy M a N.

9.5.2. Príklad. Nech

$$M = \begin{bmatrix} 1 & 1 \\ 1 & 4 \\ 1 & 1 \\ 2 & 5 \end{bmatrix}, \qquad N_1 = \begin{pmatrix} 0 \\ 2 \\ 2 \\ 3 \end{pmatrix} + \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 8 \\ 5 & 0 & 9 \\ 3 & 4 & 11 \end{bmatrix}, \qquad N_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix} + \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 8 \\ 5 & 0 & 9 \\ 3 & 4 & 11 \end{bmatrix}$$

sú afinné podpriestory v $\mathbb{R}^4$. Zrejme $\operatorname{Dir} N_1 = \operatorname{Dir} N_2$; označme tento lineárny podpriestor D. Obe úlohy o dvojiciach podpriestorov M, N_1 aj M, N_2 budeme riešiť súčasne. Platí

$$\left(\begin{array}{cc|ccc|cc} 1 & 1 & 1 & 2 & 3 & 0 & 1 \\ 1 & 4 & 2 & 2 & 8 & 2 & 1 \\ 1 & 1 & 5 & 0 & 9 & 2 & 2 \\ 2 & 5 & 3 & 4 & 11 & 3 & 2 \end{array}\right) \sim \left(\begin{array}{cc|ccc|cc} 1 & 1 & 1 & 2 & 3 & 0 & 1 \\ 0 & 3 & 1 & 0 & 5 & 2 & 0 \\ 0 & 0 & 4 & -2 & 6 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right).$$

Ak si z matice na pravej strane odmyslíme krajný pravý blok, po vynechaní rovnice $0 = 0$ z nej dostaneme sústavu

$$4t_1 - 2t_2 + 6t_3 = 0.$$

Lineárny podpriestor $\mathrm{Dir}\, M \cap D$ je tvorený práve všetkými lineárnyni kombináciami $\boldsymbol{\beta} \cdot \boldsymbol{t}$, kde $\boldsymbol{\beta}$ je matica generátorov D (a jeho báza, čo možno zistiť doupravením stredného bloku na stupňovitý tvar) a $\boldsymbol{t}$ vyhovuje uvedenej homogénnej rovnici. Teda $\dim(\mathrm{Dir}\, M \cap D) = \dim \mathrm{Dir}\, M = 2$. Preto $\mathrm{Dir}\, M \subseteq D$ a platí $M \parallel N_1$ aj $M \parallel N_2$.

Sústava

$$\begin{aligned} 4t_1 - 2t_2 + 6t_3 &= -2 \\ 0 &= -1, \end{aligned}$$

ktorej musí vyhovovať vektor parametrov $\boldsymbol{t} = (t_1, t_2, t_3)^{\mathsf{T}}$, aby ním určený bod z N_1 patril aj do M, nemá riešenie. Preto $M \cap N_1 = \emptyset$ a M, N_1 sú *pravé rovnobežky.*

Naopak, analogická sústava pre dvojicu M, N_2 vedie na jedinú, očividne riešiteľnú rovnicu

$$4t_1 - 2t_2 + 6t_3 = -1.$$

V dôsledku toho $M \subseteq N_2$.

(3) Nech afinný podpriestor $M \subseteq K^n$ je daný všeobecnými rovnicami $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ a afinný podpriestor $N = \boldsymbol{q} + [\boldsymbol{\beta}] \subseteq K^n$ je daný parametricky. Ak hľadáme všeobecné rovnice prieniku $M \cap N$, stačí nájsť všeobecné rovnice podpriestoru N a pridať ich k sústave $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$. Ich vyriešením potom možno dostať aj parametrické vyjadrenie $M \cap N$. Ak hľadáme popis spojenia $M \sqcup N$, najvýhodnejšie je vyriešiť všeobecné rovnice podpriestoru M a z parametrických vyjadrení oboch podpriestorov M, N zostaviť parametrické vyjadrenie $M \sqcup N$ podľa tvrdenia 8.3.3 a príkladu 8.3.5 Elimináciou parametrov odtiaľ dostaneme všeobecné rovnice podpriestoru $M \sqcup N$.

Iná metóda, ako nájsť parametrické vyjadrenie prieniku $M \cap N$ spočíva v dosadení parametrického vyjadrenia podpriestoru N do všeobecných rovníc podpriestoru M. Tým dostaneme sústavu

$$\boldsymbol{A} \cdot (\boldsymbol{q} + \boldsymbol{\beta} \cdot \boldsymbol{t}) = \boldsymbol{b},$$

alebo po úprave s ňou ekvivalentnú sústavu

$$(\boldsymbol{A} \cdot \boldsymbol{\beta}) \cdot \boldsymbol{t} = \boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{q},$$

ktorej musí vyhovovať vektor parametrov $\boldsymbol{t}$, aby ním určený bod $\boldsymbol{x} = \boldsymbol{q} + \boldsymbol{\beta} \cdot \boldsymbol{t} \in N$ patril aj do podpriestoru M, teda do prieniku $M \cap N$. Uvedenú sústavu vyriešime úpravou jej rozšírenej matice $(\boldsymbol{A} \cdot \boldsymbol{\beta} \,|\, \boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{q})$. Podobne ako v prípade (2) riešenie dostaneme v parametrickom tvare

$$\boldsymbol{t} = \boldsymbol{r} + \boldsymbol{\gamma} \cdot \boldsymbol{z},$$

kde $\boldsymbol{r} \in \mathcal{R}(\boldsymbol{A} \cdot \boldsymbol{\beta} \,|\, \boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{q})$ a $\boldsymbol{\gamma}$ je báza lineárneho podpriestoru $\mathcal{R}(\boldsymbol{A} \cdot \boldsymbol{\beta})$, s vektorom parametrov $\boldsymbol{z}$, ktoré dosadíme do parametrických rovníc podpriestoru N. Tak získame parametrické rovnice

$$\boldsymbol{x} = \boldsymbol{q} + \boldsymbol{\beta} \cdot (\boldsymbol{r} + \boldsymbol{\gamma} \cdot \boldsymbol{z}) = (\boldsymbol{q} + \boldsymbol{\beta} \cdot \boldsymbol{r}) + (\boldsymbol{\beta} \cdot \boldsymbol{\gamma}) \cdot \boldsymbol{z}$$

podpriestoru $M \cap N$.

I tentokrát sa v nasledujúcom príklade zameriame len na nájdenie parametrických rovníc prieniku $M \cap N$ druhou z opísaných metód a na určenie vzájomnej polohy M a N.

9.5.3. Príklad. Afinný podpriestor $M \subseteq \mathbb{R}^4$ má všeobecné rovnice

$$\begin{aligned} x_1 - x_2 + x_3 - x_4 &= 1 \\ x_1 + x_2 - x_3 + x_4 &= 3. \end{aligned}$$

Rozšírenú maticu tejto sústavy označíme $(\boldsymbol{A} \,|\, \boldsymbol{b})$. Afinný podpriestor $N \subseteq \mathbb{R}^4$ je určený ako afinný obal $N = \ell(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}, \boldsymbol{s})$ bodov $\boldsymbol{p} = (3, 0, 1, 1)^\mathsf{T}$, $\boldsymbol{q} = (4, -1, 2, 2)^\mathsf{T}$, $\boldsymbol{r} = (4, 1, 2, 0)^\mathsf{T}$ a $\boldsymbol{s} = (7, 3, 4, 5)^\mathsf{T}$. Jeho parametrické vyjadrenie potom je

$$N = \boldsymbol{p} + [\boldsymbol{q} - \boldsymbol{p}, \boldsymbol{r} - \boldsymbol{p}, \boldsymbol{s} - \boldsymbol{p}] = \begin{pmatrix} 3 \\ 0 \\ 1 \\ 1 \end{pmatrix} + \begin{bmatrix} 1 & 1 & 4 \\ -1 & 1 & 3 \\ 1 & 1 & 3 \\ 1 & -1 & 4 \end{bmatrix}.$$

Keďže

$$\boldsymbol{A} \cdot (\boldsymbol{q} - \boldsymbol{p}, \boldsymbol{r} - \boldsymbol{p}, \boldsymbol{s} - \boldsymbol{p}) = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 4 \\ -1 & 1 & 3 \\ 1 & 1 & 3 \\ 1 & -1 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 3 \\ 0 & 0 & 8 \end{pmatrix},$$

$$\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{p} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix},$$

bod tvaru $\boldsymbol{p} + t_1(\boldsymbol{q} - \boldsymbol{p}) + t_2(\boldsymbol{r} - \boldsymbol{p}) + t_3(\boldsymbol{s} - \boldsymbol{p}) \in N$ patrí do prieniku $M \cap N$ práve vtedy, keď príslušný vektor parametrov $\boldsymbol{t} = (t_1, t_2, t_3)^\mathsf{T}$ vyhovuje sústave s rozšírenou maticou

$$\left(\begin{array}{ccc|c} 2 & 2 & 3 & -2 \\ 0 & 0 & 8 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array}\right).$$

Podpriestor riešení tejto sústavy má parametrické vyjadrenie

$$\begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Dosadením do parametrického vyjadrenia N dostaneme

$$\begin{aligned} M \cap N &= \begin{pmatrix} 3 \\ 0 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 4 \\ -1 & 1 & 3 \\ 1 & 1 & 3 \\ 1 & -1 & 4 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + \left[\begin{pmatrix} 1 & 1 & 4 \\ -1 & 1 & 3 \\ 1 & 1 & 3 \\ 1 & -1 & 4 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right] \\ &= \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{bmatrix} 0 \\ 2 \\ 0 \\ -2 \end{bmatrix} \end{aligned}$$

a $\dim(M \cap N) = 1$. Ľahko nahliadneme, že hodnosť matice sústavy podpriestoru M je 2, preto tiež $\dim M = 4 - 2 = 2$, a $\dim N = 3$. Z toho dôvodu $M \cap N$ je vlastný podpriestor tak v M ako aj v N, čiže M, N sú *rôznobežné*.

Cvičenia

9.1. Nájdite nejaký fundamentálny systém riešení sústavy homogénnych lineárnych rovníc $\boldsymbol{A} \cdot \boldsymbol{x} = \mathbf{0}$ pre matice

(a) $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & -1 & -1 & 0 \\ 2 & 0 & -1 & 1 & -1 \\ 1 & 0 & 1 & 5 & -2 \end{pmatrix} \in \mathbb{Q}^{3\times 5}$;

(b) $\boldsymbol{A} = \begin{pmatrix} \sqrt{2} & 1 & \sqrt{3} \\ 1 & \sqrt{2} & 1 \\ 3 & 2 & 1+\sqrt{6} \end{pmatrix} \in \mathbb{R}^{3\times 3}$;

(c) $\boldsymbol{A} = \begin{pmatrix} 1+\mathrm{i} & 1-\mathrm{i} & 2+3\mathrm{i} \\ 5-\mathrm{i} & -1-5\mathrm{i} & 13 \end{pmatrix} \in \mathbb{C}^{2\times 3}$;

(d) $\boldsymbol{A} = \begin{pmatrix} 0 & 4 & 0 & 9 \\ 0 & 3 & 2 & 5 \\ 0 & 1 & 1 & 0 \end{pmatrix} \in \mathbb{Z}_{11}^{3\times 4}$.

Vyjadrite všeobecné riešenie každej sústavy ako lineárnu kombináciu fundamentálneho systému riešení.

9.2. Nájdite nejaké riešenie nehomogénnej sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ a fundamentálny systém riešení príslušnej homogénnej sústavy danej rozšírenou maticou

(a) $(\boldsymbol{A} \,|\, \boldsymbol{b}) = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 2 & 3 & 7 & 2 \\ 1 & 2 & 2 & 5 \end{array}\right) \in \mathbb{Q}^{3\times 4}$;

(b) $(\boldsymbol{A} \,|\, \boldsymbol{b}) = \left(\begin{array}{cc|c} 1 & \pi & \mathrm{e}-1 \\ \mathrm{e}^{-1} & -1 & \pi^2 \end{array}\right) \in \mathbb{R}^{2\times 3}$;

(c) $(\boldsymbol{A} \,|\, \boldsymbol{b}) = \left(\begin{array}{cccc|c} 1 & 2 & 1 & 0 & -1 \\ -1 & 1 & 2 & 1 & 2 \\ 2 & 7 & 5 & 1 & -1 \end{array}\right) \in \mathbb{R}^{3\times 5}$;

(d) $(\boldsymbol{A} \,|\, \boldsymbol{b}) = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 1 & 1 & 1 & 4 \\ 3 & 4 & 0 & 2 \end{array}\right) \in \mathbb{Z}_5^{3\times 4}$.

Vyjadrite všeobecné riešenie každej sústavy v tvare súčtu jedného jej riešenia a lineárnej kombinácie fundamentálneho systému riešení príslušnej homogénnej sústavy.

9.3. V každom z nasledujúcich prípadov napíšte parametrické rovnice afinného podpriestoru M vektorového priestoru $\mathbb{R}^3$ alebo $\mathbb{R}^4$ nad poľom $\mathbb{R}$ a nájdite jeho všeobecné rovnice:

(a) $M = \{(0,2,-1)\} \subseteq \mathbb{R}^3$;

(b) $M = (1,-1,2) + [(1,-5,4)] \subseteq \mathbb{R}^3$;

(c) $M = [(1,3,-1),(2,0,5)] \subseteq \mathbb{R}^3$;

(d) $M = \ell(\boldsymbol{e}_1 + \boldsymbol{e}_4, 2\boldsymbol{e}_2 - \boldsymbol{e}_3) \subseteq \mathbb{R}^4$

(e) $M = (0,2,-1,1) + [(3,1,10,-8),(3,5,8,-6)] \subseteq \mathbb{R}^4$;

(f) $M = \ell((1,1,1,0),(1,1,0,1),(1,0,1,1),(0,1,1,1)) \subseteq \mathbb{R}^4$;

(g) $M = \mathbb{R}^4$.

9.4. Pre dané lineárne podpriestory S, T vektorového priestoru $\mathbb{R}^4$ nájdite v každom z nasledujúcich prípadov nejaké bázy lineárnych podpriestorov $S \cap T$, $S + T \subseteq \mathbb{R}^4$ a určte ich dimenzie:

(a) $S = \big[(1,1,1,1)^{\mathsf{T}},(2,0,0,3)^{\mathsf{T}}\big]$, $T = \big[(3,2,1,0)^{\mathsf{T}},(6,3,2,4)^{\mathsf{T}}\big]$;

(b) $S = [\boldsymbol{e}_1 - \boldsymbol{e}_2 + 2\boldsymbol{e}_3 + \boldsymbol{e}_4]$, $T = [(1,0,2,0)^{\mathsf{T}},(2,-1,4,1)^{\mathsf{T}},(4,-1,8,1)^{\mathsf{T}}]$;

(c) $S = [\boldsymbol{e}_1 + \boldsymbol{e}_2, \boldsymbol{e}_2 + \boldsymbol{e}_3, \boldsymbol{e}_3 + \boldsymbol{e}_4]$, $T = \big[(1,0,1,0)^{\mathsf{T}},(0,1,0,1)^{\mathsf{T}},(1,0,0,1)^{\mathsf{T}}\big]$.

9.5. Pre dané afinné podpriestory M, N vektorového priestoru $\mathbb{R}^4$ nájdite v každom z nasledujúcich prípadov všeobecné aj parametrické rovnice ich prieniku $M \cap N$ a určte ich vzájomnú polohu ako aj dimenziu spojenia $M \sqcup N$:

(a) $M = \ell\big((1,0,2,0)^{\mathsf{T}},(0,2,0,1)^{\mathsf{T}}\big)$, $N = (4,1,7,2)^{\mathsf{T}} + \big[(1,2,3,4)^{\mathsf{T}},(0,1,2,3)^{\mathsf{T}},(0,0,1,2)^{\mathsf{T}}\big]$;

(b) $M = (0,0,1,-1)^{\mathsf{T}} + \big[(1,2,2,1)^{\mathsf{T}},(2,1,1,2)^{\mathsf{T}}\big]$, $N = \big[(0,1,1,0)^{\mathsf{T}},(1,0,0,1)^{\mathsf{T}},(1,-1,0,0)^{\mathsf{T}}\big]$;

(c) $M = [\boldsymbol{e}_1, \boldsymbol{e}_2]$, $N = (1,1,1,1)^{\mathsf{T}} + [\boldsymbol{e}_3, \boldsymbol{e}_4]$.

9.6. Afinné podpriestory M, N vektorového priestoru $\mathbb{R}^5$ sú dané všeobecnými rovnicami. V každom z nasledujúcich prípadov zistite ich vzájomnú polohu, napíšte parametrické rovnice ich prieniku aj spojenia a určte dimenzie afinných oboch podpriestorov $M \cap N$, $M \sqcup N$:

(a) $M\colon\ x_1 + 2x_2 = 0,\ x_2 - 3x_3 = 1,\ 4x_3 - x_4 + x_5 = 2,$

$N\colon\ 2x_1 + 3x_3 - x_5 = 5,\ x_2 + x_4 = 0,\ x_1 - 2x_3 = 7;$

(b) $M\colon\ x_1 + x_2 + x_3 + x_4 + x_5 = 5,\ x_1 - x_2 + x_3 = 1,\ x_2 + 2x_4 - x_5 = 2,\ x_3 = 1,$

$N\colon\ 3x_1 + x_2 + 2x_3 + 5x_4 - x_5 = 10,\ x_1 + x_3 + 2x_4 - x_5 = 1;$

(c) $M\colon\ x_1 = 1,\ x_2 = 2,\ x_3 = 3,$

$N\colon\ x_1 + x_2 + x_3 = 6,\ x_1 - x_2 + x_3 = 2.$

9.7. Jeden z afinných podpriestorov M, N vektorového priestoru $\mathbb{R}^4$ je daný všeobecnými rovnicami a druhý parametricky. V každom z nasledujúcich prípadov zistite ich vzájomnú polohu, napíšte parametrické rovnice ich prieniku $M \cap N$ a všeobecné rovnice ich spojenia $M \sqcup N$ a určte dimenzie afinných podpriestorov $M \cap N$, $M \sqcup N$:

(a) $M\colon\ x + y + z = 0,\ x - 2y + z = 4,\ x - u = 0,\qquad N = \big[(1,0,-1,0)^{\mathsf{T}}\big];$

(b) $M\colon\ 2x - y + 2z = 3,\ x + z - 2u = 0,\qquad N = \big[(1,4,1,1)^{\mathsf{T}},(1,0,1,0)^{\mathsf{T}}\big];$

(c) $M\colon\ x + y = 0,\ x + z = 1,\ x + u = 2,\qquad N = \ell\big((1,1,2,2)^{\mathsf{T}},(2,2,3,2)^{\mathsf{T}}\big).$

9.8. Nech V je vektorový priestor nad poľom K a S je jeho lineárny podpriestor. Označme V/S množinu všetkých afinných podpriestorov M priestoru V takých, že $\operatorname{Dir} M = S$. Dokážte postupne nasledujúce tvrdenia:

(a) Pre každý prvok $\boldsymbol{x} \in V$ existuje práve jeden podpriestor $X \in V/S$ taký, že $\boldsymbol{x} \in X$. Inými slovami, množina V/S tvorí *rozklad* množiny V a triedou rozkladu, do ktorej patrí prvok $\boldsymbol{x} \in V$, je afinný podpriestor $\boldsymbol{x} + S \in V/S$ (pozri paragraf 0.6).

(b) Prvky $\boldsymbol{x}, \boldsymbol{y} \in V$ patria do tej istej triedy rozkladu V/S práve vtedy, keď $\boldsymbol{x} - \boldsymbol{y} \in S$. Inak povedané, vzťahom $\boldsymbol{x} \equiv_S \boldsymbol{y} \Leftrightarrow \boldsymbol{x} - \boldsymbol{y} \in S$ je definovaná ekvivalencia na množine V prislúchajúca k rozkladu V/S, teda $V/S = V/{\equiv_S}$.

(c) Nech $\boldsymbol{x}_1 \equiv_S \boldsymbol{x}_2$, $\boldsymbol{y}_1 \equiv_S \boldsymbol{y}_2$ a $c \in K$. Potom tiež $\boldsymbol{x}_1 + \boldsymbol{y}_1 \equiv_S \boldsymbol{x}_2 + \boldsymbol{y}_2$ a $c\boldsymbol{x}_1 \equiv_S c\boldsymbol{x}_2$.

(d) Nech $X = \boldsymbol{x} + S$, $Y = \boldsymbol{y} + S$ patria do V/S. Vďaka (c) sú predpismi $X + Y = (\boldsymbol{x} + \boldsymbol{y}) + S$, $cX = c\boldsymbol{x} + S$ korektne definované operácie na množine V/S (to znamená, že výsledky týchto operácií nezávisia na reprezentantoch $\boldsymbol{x}$, $\boldsymbol{y}$ tried X resp. Y ale výlučne na podpriestoroch X, Y).

(e) Množina V/S s uvedenými operáciami súčtu a skalárneho násobku tvorí vektorový priestor nad poľom K; nazývame ho *faktorový priestor* vektorového priestoru V podľa podpriestoru S. Čo je nulou v tomto vektorovom priestore?

(f) Priradením $\boldsymbol{x} \mapsto \zeta_S(\boldsymbol{x}) = \boldsymbol{x} + S$ je definované surjektívne lineárne zobrazenie $\zeta_S \colon V \to V/S$ s jadrom $\operatorname{Ker} \zeta_S = S$.

(g) Nech U, V sú vektorové priestory nad poľom K a $\varphi \colon V \to U$ je lineárne zobrazenie. Priradením $\boldsymbol{x} + \operatorname{Ker} \varphi \mapsto \varphi(\boldsymbol{x})$ je *korektne* definované injektívne lineárne zobrazenie $V/\operatorname{Ker} \varphi \to U$. V dôsledku toho platí $V/\operatorname{Ker} \varphi \cong \operatorname{Im} \varphi$.

(h) Ak V je konečnorozmerný, tak $\dim V/S = \dim V - \dim S$, čo je ďalšia analógia medzi vlastnosťami dimenzie a logaritmu (porovnaj s tvrdením 5.4.3).

9.9. Nech V je vektorový priestor nad poľom K a S, T sú jeho lineárne podpriestory. Dokážte tzv. *kosoštvorcovú vetu o izomorfizme*:

$$(S + T)/T \cong S/(S \cap T).$$

(*Návod*: Dokážte, že priradením $(\boldsymbol{x} + \boldsymbol{y}) + T \mapsto \boldsymbol{x} + (S \cap T)$, kde $\boldsymbol{x} \in S$, $\boldsymbol{y} \in T$, je korektne definovaný lineárny izomorfizmus $(S + T)/T \to S/(S \cap T)$.)

9.10. Nech pole K má konečný počet prvkov q, V je vektorový priestor nad K konečnej dimenzie n a S je jeho lineárny podpriestor dimenzie k ($0 \le k \le n$). Dokážte postupne nasledujúce tvrdenia:

(a) Faktorový priestor V/S má práve q^{n-k} prvkov.

(b) Počet všetkých k-rozmerných *afinných* podpriestorov priestoru V je práve $q^{n-k}\binom{n}{k}_q$, kde $\binom{n}{k}_q$ je q-binomický koeficient (pozri cvičenia 5.18, 5.17).

10. Determinanty

V tejto kapitole zavedieme *determinanty* štvorcových matíc ľubovoľného rozmeru $n \times n$ nad pevným poľom K, preskúmame ich základné vlastnosti a naučíme sa ich počítať. Taktiež si ukážeme niekoľko príkladov ich využitia.

Čitateľ sa pravdepodobne už na strednej škole stretol s determinantmi reálnych matíc rozmerov 2×2 a 3×3. Možno tiež vie previesť výpočet determinantov vyšších rádov na výpočet determinantov nižších rádov pomocou ich rozvoja podľa nejakého riadku alebo stĺpca. So všeobecnou definíciou determinantu sa však asi dosiaľ nestretol. Ako čoskoro uvidíme, nie je to nijako priezračná definícia a na prvý pohľad určite nepôsobí „prirodzeným" dojmom. Keďže nechceme, aby táto definícia „spadla z neba", náš výklad začneme pomerne dlhým úvodom, ktorý má poslúžiť ako jej motivácia.

10.1 Orientovaný objem a multilineárne alternujúce funkcie

Na začiatok si položme prirodzenú otázku: Ako vyzerajú vzorce pre plošný obsah rovnobežníka v rovine $\mathbb{R}^2$, ktorého dve susedné strany tvoria vektory $\boldsymbol{u} = (u_1, u_2)^\mathsf{T}$, $\boldsymbol{v} = (v_1, v_2)^\mathsf{T}$, resp. pre objem rovnobežnostena v priestore $\mathbb{R}^3$, ktorého tri susedné hrany tvoria vektory $\boldsymbol{u} = (u_1, u_2, u_3)^\mathsf{T}$, $\boldsymbol{v} = (v_1, v_2, v_3)^\mathsf{T}$, $\boldsymbol{w} = (w_1, w_2, w_3)^\mathsf{T}$?

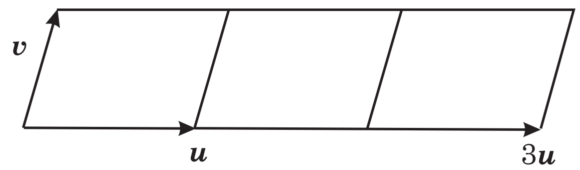
Vzorce, ktoré by vyjadrovali príslušný obsah alebo objem len pomocou súradníc vektorov $\boldsymbol{u}$, $\boldsymbol{v}$ resp. $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$, asi len tak z rukáva nevysypeme, môžeme sa však pokúsiť ich odvodiť. Najschodnejšia cesta vedie cez ujasnenie si vlastností, ktorým by mali takéto vzorce vyhovovať. Uvidíme, že tieto vlastnosti už jednoznačne (až na voľbu jednotkového obsahu či objemu) určujú hľadané vzorce nielen v rovine či v trojrozmernom priestore, ale možno ich bezprostredne zovšeobecniť na n-rozmerné vektorové priestory K^n nad ľubovoľným poľom K, hoci tu pojem „n-rozmerného objemu" stráca svoj názorný geometrický význam.

Označme teda $P(X)$ obsah rovinného útvaru X. Zrejme $P(X)$ je vždy nezáporné reálne číslo a pre zhodné útvary X, Y platí $P(X) = P(Y)$. Obsah je navyše *aditívny*, t. j. pre útvary X, Y také, že $P(X \cap Y) = 0$, platí $P(X \cup Y) = P(X) + P(Y)$. Konečne, $P(X) = 0$ pre ľubovoľnú úsečku X.

Obsah rovnobežníka $\{a\boldsymbol{u} + b\boldsymbol{v};\ a, b \in \langle 0, 1 \rangle\}$ určeného vektormi $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^2$ budeme značiť $P(\boldsymbol{u}, \boldsymbol{v})$. Z práve sformulovaných vlastností obsahu vyplývajú rovnosti

$$P(\boldsymbol{u},\boldsymbol{v}) = P(\boldsymbol{v},\boldsymbol{u}), \qquad\qquad P(c\boldsymbol{u},\boldsymbol{v}) = |c|P(\boldsymbol{u},\boldsymbol{v})$$

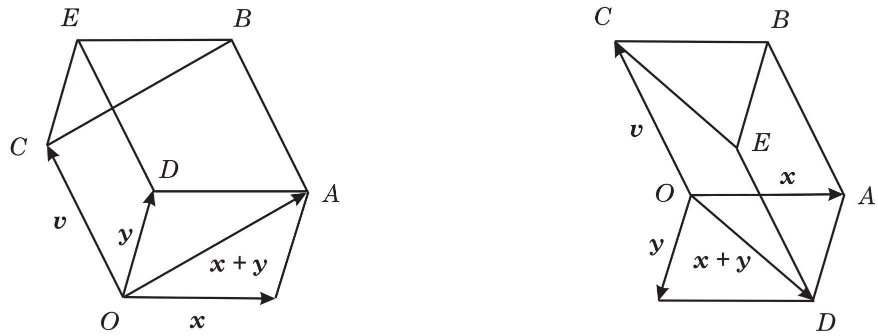
pre ľubovoľné $\boldsymbol{u},\boldsymbol{v} \in \mathbb{R}^2$, $c \in \mathbb{Z}$. Druhá vlastnosť sa nazýva *pozitívna homogenita* a pre $c = 3$ je znázornená na nasledujúcom obrázku.



Obr. 10.1. K pozitívnej homogenite obsahu vektorového rovnobežníka

Platnosť druhej rovnosti pre všetky $c \in \mathbb{Q}$ možno už z toho jednoducho dokázať (pozri cvičenie 10.1). S jej platnosťou pre všetky $c \in \mathbb{R}$ je to už trochu zložitejšie – zakladá sa na istých úvah o „spojitosti" obsahu –, a tak jej radšej uveríme bez dôkazu.

Pozrime sa teraz na ďalšie dva obrázky. (Podotýkame, že oba znázorňujú situáciu *v rovine*, teda pri pohľade na ne treba potlačiť priestorové videnie, ktoré sa nám mimovoľne otvára.)



Obr. 10.2. K aditivite obsahu vektorového rovnobežníka

V prvom prípade určujú vektory $\boldsymbol{x}+\boldsymbol{y}$, $\boldsymbol{v}$ rovnobežník $OABC$, vektory $\boldsymbol{y}$, $\boldsymbol{v}$ rovnobežník $ODEC$ a rovnobežník vektorov $\boldsymbol{x}$, $\boldsymbol{v}$ je zhodný s rovnobežníkom $DABE$. Zo zhodnosti trojuholníkov OAD, CBE potom na základe uvedených vlastností obsahu vyplýva rovnosť

$$P(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{v}) = P(\boldsymbol{x},\boldsymbol{v}) + P(\boldsymbol{y},\boldsymbol{v}).$$

V druhom prípade určujú vektory $\boldsymbol{x}$, $\boldsymbol{v}$ rovnobežník $OABC$, vektory $\boldsymbol{x}+\boldsymbol{y}$, $\boldsymbol{v}$ rovnobežník $ODEC$ a rovnobežník vektorov $\boldsymbol{y}$, $\boldsymbol{v}$ je zhodný s rovnobežníkom $DABE$. Zo zhodnosti trojuholníkov ODA, CEB vyplýva $P(\boldsymbol{x},\boldsymbol{v}) = P(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{v}) + P(\boldsymbol{y},\boldsymbol{v})$, teda

$$P(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{v}) = P(\boldsymbol{x},\boldsymbol{v}) - P(\boldsymbol{y},\boldsymbol{v}).$$

To je v porovnaní s prvým prípadom nepríjemné prekvapenie, určite by sme dali prednosť rovnakej formule. Všimnime si však , že „kratšie otočenie“ vektora $\boldsymbol{y}$ do vektora $\boldsymbol{v}$ je orientované proti „kratším otočeniam“ vektorov $\boldsymbol{x}$ aj $\boldsymbol{x}+\boldsymbol{y}$ do vektora $\boldsymbol{v}$. V druhom prípade by sa nám preto hodilo, aby obsah rovnobežníka určeného vektormi $\boldsymbol{y}$, $\boldsymbol{v}$ mal z toho dôvodu opačné znamienko ako obsahy rovnobežníkov prislúchajúcich vektorom $\boldsymbol{x}$, $\boldsymbol{v}$ resp. $\boldsymbol{x}+\boldsymbol{y}$, $\boldsymbol{v}$. Tento cieľ možno dosiahnuť, ak namiesto plošného obsahu vektorových rovnobežníkov budeme uvažovať ich *orientovaný plošný obsah*, ktorý mení znamienko zámenou poradia dvoch vektorov, teda môže nadobúdať aj záporné hodnoty. Pôvodný nezáporný plošný obsah potom dostaneme ako absolútnu hodnotu orientovaného obsahu. Tento prístup nám navyše umožní zbaviť sa absolútnej hodnoty v rovnosti $P(c\boldsymbol{u},\boldsymbol{v}) = |c|P(\boldsymbol{u},\boldsymbol{v})$.

Podobnými úvahami, ktoré by si však vyžiadali trochu zložitejšie obrázky, tentokrát znázorňujúce naozaj priestorové situácie, by sme mohli dospieť i k potrebe skúmať *orientovaný objem* rovnobežnostena $\{a\boldsymbol{u}+b\boldsymbol{v}+c\boldsymbol{w};\ a,b,c \in \langle 0,1\rangle\}$ určeného vektormi $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ v trojrozmernom priestore $\mathbb{R}^3$, prípadne *n-rozmerný objem* rovnobežnostena $\{a_1\boldsymbol{u}_1 + \ldots + a_n\boldsymbol{u}_n;\ a_1,\ldots,a_n \in \langle 0,1\rangle\}$ určeného vektormi $\boldsymbol{u}_1,\ldots,\boldsymbol{u}_n$ v n-rozmernom priestore $\mathbb{R}^n$ (pre $n > 3$ však bez možnosti sprostredkovať si geometrický vhľad obrázkami).

Pre čitateľa, ktorý sa už stretol s *vektorovým súčinom* v $\mathbb{R}^3$, poznamenajme, že orientovaný n-rozmerný objem sa správa do značnej miery podobne. Vektorový súčin $\boldsymbol{u}\times\boldsymbol{v}$ dvoch vektorov $\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^3$, je vektor kolmý na rovinu $[\boldsymbol{u},\boldsymbol{v}]$, ktorého dĺžka sa rovná plošnému obsahu rovnobežníka vektorov $\boldsymbol{u}$, $\boldsymbol{v}$ a orientácia je daná pravidlom pravej ruky (ak položíme dlaň pravej ruky malíčkom na vektor $\boldsymbol{u}$ tak, že zakrivené prsty smerujú k vektoru $\boldsymbol{v}$ po oblúku zodpovedajúcom uhlu $\leq 180°$, vztýčený palec ukazuje smer aj orientáciu vektora $\boldsymbol{u}\times\boldsymbol{v}$). Z toho dôvodu $\boldsymbol{u}\times\boldsymbol{v} = -(\boldsymbol{v}\times\boldsymbol{u})$.

Ak nahradíme reálne čísla ľubovoľným poľom K, vykonané úvahy nás privádzajú k nasledujúcim definíciám. Nech V je vektorový priestor nad poľom K a $1 \leq n \in \mathbb{N}$. Hovoríme, že zobrazenie $F\colon V^n \to K$ je

(a) *n-lineárne* alebo tiež *multilineárne*, ak pre každé $1 \leq j \leq n$ a ľubovoľné vektory $\boldsymbol{u}_1,\ldots,\boldsymbol{u}_{j-1},\boldsymbol{u}_{j+1},\ldots,\boldsymbol{u}_n \in V$ priradenie

$$\boldsymbol{x} \mapsto F(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_{j-1},\boldsymbol{x},\boldsymbol{u}_{j+1},\ldots,\boldsymbol{u}_n)$$

definuje lineárne zobrazenie $V \to K$, t. j. pre všetky $\boldsymbol{x},\boldsymbol{y}\in V$, $a,b\in K$ platí

$$\begin{aligned}&F(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_{j-1},a\boldsymbol{x}+b\boldsymbol{y},\boldsymbol{u}_{j+1},\ldots,\boldsymbol{u}_n)\\ &\quad = aF(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_{j-1},\boldsymbol{x},\boldsymbol{u}_{j+1},\ldots,\boldsymbol{u}_n) + bF(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_{j-1},\boldsymbol{y},\boldsymbol{u}_{j+1},\ldots,\boldsymbol{u}_n);\end{aligned}$$

(b) *antisymetrické*, ak pre všetky $1 \leq i < j \leq n$ a ľubovoľné vektory $\boldsymbol{u}_1,\ldots,\boldsymbol{u}_n \in V$ platí

$$F(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_i,\ldots,\boldsymbol{u}_j,\ldots,\boldsymbol{u}_n) = -F(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_j,\ldots,\boldsymbol{u}_i,\ldots,\boldsymbol{u}_n).$$

Inak povedané, $F\colon V^n \to K$ je n-lineárne, ak dosadením ľubovoľných $n-1$ pevných vektorov na akékoľvek miesta do F dostaneme lineárne zobrazenie vo zvyšnej voľnej premennej; F je antisymetrické, ak zámenou poradia ľubovoľných dvoch argumentov v F sa hodnota výsledku zmení na opačnú.

Cieľom našich úvah teda bolo čitateľa presvedčiť, že n-rozmerný orientovaný objem v $\mathbb{R}^n$ je multilineárna antisymetrická funkcia

$$\underbrace{\mathbb{R}^n \times \ldots \times \mathbb{R}^n}_{n\text{-krát}} \to \mathbb{R}.$$

Ukazuje sa však, že antisymetriu možno nahradiť zdanlivo slabšou, geometricky názornou podmienkou, motivovanou očividným vzťahom $P(\boldsymbol{u},\boldsymbol{u}) = 0$ pre obsah degenerovaného vektorového rovnobežníka. Hovoríme, že zobrazenie $F\colon V^n \to K$ je

(c) *alternujúce*, ak pre ľubovoľné $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in V$ z podmienky $\boldsymbol{u}_i = \boldsymbol{u}_j$ pre nejaké $1 \le i < j \le n$ vyplýva

$$F(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) = 0.$$

Ukážeme si, že uvedené tri vlastnosti spolu tesne súvisia. Najprv ale pripomeňme, že pole K má charakteristiku 2, ak v ňom platí $1 + 1 = 0$, čo je ekvivalentné s podmienkou $(\forall\, a \in K)(a = -a)$. Príkladom je pole $\mathbb{Z}_2$ (pozri paragraf 1.2). Ak $\operatorname{char} K \neq 2$, tak $(\forall\, a \in K)(a = -a \Rightarrow a = 0)$.

10.1.1. Lema. *Nech V je vektorový priestor nad poľom K a $F\colon V^n \to K$ je ľubovoľné zobrazenie.*

(a) Ak $\operatorname{char} K \neq 2$ a F je antisymetrické, tak F je alternujúce.

(b) Ak F je multilineárne a alternujúce, tak F je antisymetrické.

Dôkaz. (a) sme už vlastne dokázali v úvahe predchádzajúcej túto lemu.

(b) Nech F je multilineárne a alternujúce. Položme $\boldsymbol{x} = \boldsymbol{u}_i$, $\boldsymbol{y} = \boldsymbol{u}_j$ a zafixujme zvyšné z vektorov $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$. Potom $G(\boldsymbol{x},\boldsymbol{y}) = F(\boldsymbol{u}_1, \ldots, \boldsymbol{x}, \ldots, \boldsymbol{y}, \ldots, \boldsymbol{u}_n)$ je bilineárne (t. j. 2-lineárne) alternujúce zobrazenie $V^2 \to K$. Stačí dokázať, že G je antisymentrické. Vďaka uvedeným vlastnostiam platí

$$\begin{aligned} G(\boldsymbol{x},\boldsymbol{y}) + G(\boldsymbol{y},\boldsymbol{x}) &= G(\boldsymbol{x},\boldsymbol{x}) + G(\boldsymbol{x},\boldsymbol{y}) + G(\boldsymbol{y},\boldsymbol{x}) + G(\boldsymbol{y},\boldsymbol{y}) \\ &= G(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{x}+\boldsymbol{y}) = 0, \end{aligned}$$

teda $G(\boldsymbol{x},\boldsymbol{y}) = -G(\boldsymbol{y},\boldsymbol{x})$.

Pre multilineárne zobrazenie tak alternácia implikuje antisymetriu, kým opačná implikácia platí len za dodatočného predpokladu $\operatorname{char} K \neq 2$ (no, na druhej strane, aj bez multilinearity).

10.1.2. Lema. *Nech $F\colon V^n \to K$, $\sigma\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ sú ľubovoľné zobrazenia a $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in V$.*

(a) Ak F je antisymetrické a σ je permutácia, tak

$$F(\boldsymbol{u}_{\sigma(1)}, \dots, \boldsymbol{u}_{\sigma(n)}) = (-1)^{|\sigma|} F(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n).$$

(b) Ak F je alternujúce a σ nie je permutácia, tak

$$F(\boldsymbol{u}_{\sigma(1)}, \dots, \boldsymbol{u}_{\sigma(n)}) = 0.$$

Dôkaz. (a) Stačí si uvedomiť, že $|\sigma|$ označuje najmenší počet traspozícií (t. j. výmien poradia dvojíc), z ktorých možno zložiť permutáciu σ (pozri paragraf 0.5).

(b) Ak σ nie je permutácia, tak $\sigma(i) = \sigma(j)$, preto tiež $\boldsymbol{u}_{\sigma(i)} = \boldsymbol{u}_{\sigma(j)}$, pre nejaké $1 \leq i < j \leq n$. Označme $\boldsymbol{v}_k = \boldsymbol{u}_{\sigma(k)}$ pre $1 \leq k \leq n$. Potom $\boldsymbol{v}_i = \boldsymbol{v}_j$, a v dôsledku alternácie $F(\boldsymbol{v}_1, \dots, \boldsymbol{v}_i, \dots, \boldsymbol{v}_j, \dots, \boldsymbol{v}_n) = 0$.

Zaznamenajme teraz niektoré základné vlastnosti multilineárnych alternujúcich (teda automaticky aj antisymetrických) zobrazení, ktoré budeme sústavne využívať.

10.1.3. Lema. *Nech $F\colon V^n \to K$ je multilineárne alternujúce zobraznie. Potom pre ľubovoľné $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n \in V$ platí:*
(a) Pripočítaním skalárneho násobku nejakého z vektorov k inému vektoru sa hodnota $F(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ nezmení, t. j. pre ľubovoľné $c \in K$ a $i, j \leq n$, $i \neq j$ platí

$$F(\boldsymbol{v}_1, \dots, \boldsymbol{v}_i, \dots, \boldsymbol{v}_j + c\boldsymbol{v}_i, \dots, \boldsymbol{v}_n) = F(\boldsymbol{v}_1, \dots, \boldsymbol{v}_i, \dots, \boldsymbol{v}_j, \dots, \boldsymbol{v}_n).$$

(b) Ak sú vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ lineárne závislé, tak $F(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) = 0$.

Dôkaz. (a) Nech napr. $i < j$. Priamym výpočtom s použitím multilinearity a alternácie dostávame

$$\begin{aligned} &F(\boldsymbol{v}_1, \dots, \boldsymbol{v}_i, \dots, \boldsymbol{v}_j + c\boldsymbol{v}_i, \dots, \boldsymbol{v}_n) \\ &\qquad = F(\boldsymbol{v}_1, \dots, \boldsymbol{v}_i, \dots, \boldsymbol{v}_j, \dots, \boldsymbol{v}_n) + cF(\boldsymbol{v}_1, \dots, \boldsymbol{v}_i, \dots, \boldsymbol{v}_i, \dots, \boldsymbol{v}_n) \\ &\qquad = F(\boldsymbol{v}_1, \dots, \boldsymbol{v}_i, \dots, \boldsymbol{v}_j, \dots, \boldsymbol{v}_n) \end{aligned}$$

(b) Ak sú vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ lineárne závislé, tak niektorý z nich, povedzme $\boldsymbol{v}_k$, je lineárnou kombináciou ostatných, teda $\boldsymbol{v}_k = \sum_{i \neq k} c_i \boldsymbol{v}_i$ pre vhodné skaláry c_i. Z multilinearity a alternácie F potom vyplýva

$$F(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \dots, \boldsymbol{v}_n) = \sum_{i \neq k} c_i F(\boldsymbol{v}_1, \dots, \boldsymbol{v}_i, \dots, \boldsymbol{v}_n) = 0,$$

lebo v každom z uvedených výrazov $F(\boldsymbol{v}_1, \dots, \boldsymbol{v}_i, \dots, \boldsymbol{v}_n)$ sa vektor $\boldsymbol{v}_i$ vyskytuje ako argument na i-tom aj na k-tom mieste.

Pozrime sa teraz bližšie, ako vyzerajú všetky bilineárne alternujúce zobrazenia $F\colon K^2 \times K^2 \to K$ nad poľom K. Zvoľme ľubovoľné vektory $\boldsymbol{u} = u_1\boldsymbol{e}_1 + u_2\boldsymbol{e}_2$ a $\boldsymbol{v} = v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2$ z K^2. Ak dvakrát po sebe využijeme bilinearitu a na záver alternáciu a antisymetriu F, postupne dostaneme

$$\begin{aligned}
F(\boldsymbol{u},\boldsymbol{v}) &= F(u_1\boldsymbol{e}_1 + u_2\boldsymbol{e}_2,\, \boldsymbol{v}) = u_1F(\boldsymbol{e}_1,\boldsymbol{v}) + u_2F(\boldsymbol{e}_2,\boldsymbol{v})\\
&= u_1F(\boldsymbol{e}_1,\, v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2) + u_2F(\boldsymbol{e}_2,\, v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2)\\
&= u_1v_1F(\boldsymbol{e}_1,\boldsymbol{e}_1) + u_1v_2F(\boldsymbol{e}_1,\boldsymbol{e}_2) + u_2v_1F(\boldsymbol{e}_2,\boldsymbol{e}_1) + u_2v_2F(\boldsymbol{e}_2,\boldsymbol{e}_2)\\
&= F(\boldsymbol{e}_1,\boldsymbol{e}_2)(u_1v_2 - u_2v_1) = F(\boldsymbol{e}_1,\boldsymbol{e}_2)\begin{vmatrix} u_1 & v_1\\ u_2 & v_2\end{vmatrix},
\end{aligned}$$

kde výraz

$$\begin{vmatrix} u_1 & v_1\\ u_2 & v_2\end{vmatrix} = u_1v_2 - u_2v_1$$

čitateľ už iste pozná ako determinant matice $(\boldsymbol{u},\boldsymbol{v}) = \begin{pmatrix} u_1 & v_1\\ u_2 & v_2\end{pmatrix} \in K^{2\times 2}$.

Podobným spôsobom možno odvodiť aj tvar ľubovoľnej n-lineárnej alternujúcej funkcie $F\colon K^{n\times n} \to K$ (i teraz, ako obyčajne, prirodzene stotožňujeme n-tú karteziánsku mocninu $(K^n)^n$ stĺpcového vektorového priestoru $V = K^n$ s priestorom matíc $K^{n\times n}$). Nech $\boldsymbol{A} = (a_{ij}) \in K^{n\times n}$ je matica so stĺpcami

$$\boldsymbol{s}_j(\boldsymbol{A}) = a_{1j}\boldsymbol{e}_1 + \ldots + a_{nj}\boldsymbol{e}_n = \sum_{i=1}^{n} a_{ij}\boldsymbol{e}_i.$$

S využitím n-linearity F pre každý z n stĺpcov matice $\boldsymbol{A}$ možno výraz $F(\boldsymbol{A})$ postupne roznásobiť, čím dostaneme súčet n^n členov tvaru

$$a_{\sigma(1)\,1}\ldots a_{\sigma(n)\,n}F(\boldsymbol{e}_{\sigma(1)},\ldots,\boldsymbol{e}_{\sigma(n)}),$$

z ktorých každý zodpovedá jednému zobrazeniu σ množiny $\{1,\ldots,n\}$ do seba. Podľa lemy 10.1.2 sčítance prislúchajúce zobrazeniam $\sigma \notin \mathcal{S}_n$ sú všetky rovné 0 a pre $\sigma \in \mathcal{S}_n$ platí

$$F(\boldsymbol{e}_{\sigma(1)},\ldots,\boldsymbol{e}_{\sigma(n)}) = (-1)^{|\sigma|}F(\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n).$$

Záverom tak dostávame

$$\begin{aligned}
F(\boldsymbol{A}) &= F(\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n)\sum_{\sigma\in\mathcal{S}_n}(-1)^{|\sigma|}a_{\sigma(1)\,1}\ldots a_{\sigma(n)\,n}\\
&= F(\boldsymbol{I}_n)\sum_{\sigma\in\mathcal{S}_n}(-1)^{|\sigma|}a_{\sigma(1)\,1}\ldots a_{\sigma(n)\,n},
\end{aligned}$$

kde príslušná suma obsahuje $n!$ sčítancov, jeden pre každú permutáciu $\sigma \in \mathcal{S}_n$.

10.2 Definícia a základné vlastnosti determinantu

Determinantom štvorcovej matice $\boldsymbol{A} = (a_{ij}) \in K^{n\times n}$ nazývame výraz

$$\det \boldsymbol{A} = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \sum_{\sigma\in\mathcal{S}_n} (-1)^{|\sigma|} a_{\sigma(1)\,1} \dots a_{\sigma(n)\,n}.$$

Ak nehrozí zámena s absolútnou hodnotou, používame tiež označenie $|\boldsymbol{A}|$. Determinant štvorcovej matice rádu n budeme nazývať *determinant rádu n.*

Špeciálne pre maticu $(a_{ij}) \in K^{3\times 3}$ dostávame vzorec

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\ &- a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23}, \end{aligned}$$

známy ako *Sarrusovo pravidlo.*

Skôr ako kuriozitu poznamenajme, že uvedená definícia zahŕňa aj prípad $n = 0$: pre (jedinú) prázdnu maticu $\boldsymbol{I}_0 = (\;) \in K^{0\times 0}$ dáva $\det \boldsymbol{I}_0 = \det(\;) = 1$.

Nasledujúce dve vlastnosti determinantov dokážeme ako dôsledky našej definície.

10.2.1. Tvrdenie. *Determinant transponovanej matice sa rovná determinantu pôvodnej matice, t. j.*

$$\det \boldsymbol{A}^{\mathsf{T}} = \det \boldsymbol{A}$$

pre ľubovoľnú $\boldsymbol{A} \in K^{n\times n}$.

Dôkaz. Podľa definícií transponovanej matice a determinantu

$$\det \boldsymbol{A}^{\mathsf{T}} = \sum_{\sigma\in\mathcal{S}_n} (-1)^{|\sigma|} a_{1\sigma(1)} \dots a_{n\sigma(n)}.$$

Keďže pre $\sigma \in \mathcal{S}_n$ platí $i = \sigma(j) \Leftrightarrow j = \sigma^{-1}(i)$, zoradením činiteľov v súčine $a_{1\sigma(1)} \dots a_{n\sigma(n)}$ podľa druhého indexu tento nadobudne tvar $a_{\sigma^{-1}(1)\,1} \dots a_{\sigma^{-1}(n)\,n}$. Pritom priradenie $\sigma \mapsto \sigma^{-1}$ je bijekcia $\mathcal{S}_n \to \mathcal{S}_n$. Navyše, ak $\sigma = \tau_1 \circ \ldots \circ \tau_k$ je rozklad permutácie σ na transpozície, tak σ^{-1} je kompozícia tých istých transpozícií v opačnom poradí, preto $|\sigma| = \left|\sigma^{-1}\right|$. V dôsledku toho zámenou sumácie cez $\sigma \in \mathcal{S}_n$ za sumáciu cez $\sigma^{-1} = \varrho \in \mathcal{S}_n$ dostávame

$$\det \boldsymbol{A}^{\mathsf{T}} = \sum_{\varrho\in\mathcal{S}_n} (-1)^{|\varrho|} a_{\varrho(1)\,1} \dots a_{\varrho(n)\,n} = \det \boldsymbol{A}.$$

Vďaka práve dokázanému tvrdeniu si všetky výsledky o determinantoch matíc zachovajú svoju platnosť, ak v nich každý výskyt slova „stĺpec“ nahradíme slovom „riadok“ a naopak. Tento princíp zámeny riadkov a stĺpcov budeme často využívať.

10.2.2. Tvrdenie. *Nech* $1 \leq m < n$ *a* $\boldsymbol{A} \in K^{n\times n}$ *je bloková matica tvaru*

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{D} \end{pmatrix},$$

kde $\boldsymbol{B} \in K^{m\times m}$, $\boldsymbol{C} \in K^{m\times(n-m)}$ *a* $\boldsymbol{D} \in K^{(n-m)\times(n-m)}$. *Potom*

$$\det \boldsymbol{A} = \det \boldsymbol{B} \cdot \det \boldsymbol{D}.$$

Dôkaz. Z uvedeného blokového tvaru matice $\boldsymbol{A}$ vyplýva

$$a_{ij} = \begin{cases} b_{ij}, & \text{ak } 1 \leq i, j \leq m, \\ c_{i\,j-m}, & \text{ak } 1 \leq i \leq m < j \leq n, \\ 0, & \text{ak } 1 \leq j \leq m < i \leq n, \\ d_{i-m\,j-m}, & \text{ak } m < i, j \leq n. \end{cases}$$

Označme $G = \{\sigma \in \mathcal{S}_n;\, (\forall\, j \leq m)(\sigma(j) \leq m)\}$. Potom pre $\sigma \in \mathcal{S}_n \smallsetminus G$ platí $a_{\sigma(1)\,1} \dots a_{\sigma(n)\,n} = 0$, teda do hodnoty determinantu matice $\boldsymbol{A}$ prispievajú len sčítance zodpovedajúce permutáciám $\sigma \in G$. Navyše, $(\forall\, \sigma \in G)(m < j \Rightarrow m < \sigma(j))$, takže pre $\sigma \in G$ možno definovať permutácie $\sigma' \in \mathcal{S}_m$ a $\sigma'' \in \mathcal{S}_{n-m}$ predpismi $\sigma'(j) = \sigma(j)$, ak $1 \leq j \leq m$, resp. $\sigma''(k) = \sigma(k+m) - m$, ak $1 \leq k \leq n-m$. Zrejme priradením $\sigma \mapsto (\sigma', \sigma'')$ je daná bijekcia $G \to \mathcal{S}_m \times \mathcal{S}_{n-m}$ a platí $|\sigma| = |\sigma'| + |\sigma''|$. Takže môžeme písať

$$\begin{aligned}
\det \boldsymbol{A} &= \sum_{\sigma\in G} (-1)^{|\sigma|} a_{\sigma(1)\,1} \dots a_{\sigma(m)\,m} a_{\sigma(m+1)\,m+1} \dots a_{\sigma(n)\,n} \\
&= \sum_{\sigma\in G} (-1)^{|\sigma'|+|\sigma''|} b_{\sigma'(1)\,1} \dots b_{\sigma'(m)\,m} d_{\sigma''(1)\,1} \dots d_{\sigma''(n-m)\,n-m} \\
&= \sum_{\sigma'\in\mathcal{S}_m} (-1)^{|\sigma'|} b_{\sigma'(1)\,1} \dots b_{\sigma'(m)\,m} \sum_{\sigma''\in\mathcal{S}_{n-m}} (-1)^{|\sigma''|} d_{\sigma''(1)\,1} \dots d_{\sigma''(n-m)\,n-m} \\
&= \det \boldsymbol{B} \cdot \det \boldsymbol{D}.
\end{aligned}$$

Na základe tvrdenia 10.2.1 teraz vieme, že $\det \boldsymbol{A} = \det \boldsymbol{B} \cdot \det \boldsymbol{D}$, aj keď sa nulový blok $\boldsymbol{0}$ nachádza nad a blok $\boldsymbol{C} \in K^{(n-m)\times m}$ pod diagonálou matice $\boldsymbol{A}$. Tvrdenie 10.2.2 možno taktiež zrejmým spôsobom zovšeobecniť na matice pozostávajúce z viacerých diagonálne zoradených štvorcových blokov, pod (nad) ktorými sú samé nuly. Spomeňme explicitne nasledujúce dva prípady:

(1) Ak $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k$ sú štvorcové matice, tak

$$\det \operatorname{diag}(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k) = \det \boldsymbol{A}_1 \cdot \ldots \cdot \det \boldsymbol{A}_k.$$

(2) Matica $\boldsymbol{A} \in K^{n\times n}$ sa nazýva *horná (dolná) trojuholníková matica*, ak $a_{ij} = 0$ pre $i < j$ (resp. pre $i > j$). Pre horné aj dolné trojuholníkové matice platí

$$\det \boldsymbol{A} = a_{11} \ldots a_{nn},$$

t. j. determinant takej matice je súčinom jej diagonálnych prvkov. Špeciálne to platí pre diagonálne matice.

10.3 Charakterizácia determinantu a regulárnych matíc

Úvahy z paragrafu 10.1 možno zhrnúť do nasledujúcej vety.

10.3.1. Veta. *Determinant rádu n je n-lineárna alternujúca funkcia $K^{n\times n} \to K$ stĺpcov matice. Navyše, pre každý skalár $c \in K$ existuje jediné multilineárne alternujúce zobrazenie $F\colon K^{n\times n} \to K$ stĺpcov matice také, že $F(\boldsymbol{I}_n) = c$. Toto F je dané predpisom*

$$F(\boldsymbol{A}) = c \det \boldsymbol{A}.$$

Determinant $\det\colon K^{n\times n} \to K$ je teda jednoznačne určený ako n-lineárna alternujúca funkcia stĺpcov matice taká, že

$$\det \boldsymbol{I}_n = \det(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) = 1.$$

Táto rovnosť zodpovedá prirodzenej voľbe jednotky orientovaného n-rozmerného objemu v K^n – je ňou orientovaný objem rovnobežnostena určeného vektormi $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ (v tomto poradí).

V paragrafe 10.1 sme vlastne dokázali, že každá n-lineárna alternujúca funkcia $F\colon K^{n\times n} \to K$ musí mať uvedený tvar, t. j. je skalárnym násobkom determinantu. Zostáva však overiť, že determinant, tak, ako sme ho definovali v paragrafe 10.2, je naozaj multilineárne alternujúce zobrazenie. Hoci tieto vlastnosti sú intutívne jasné z našej konštrukcie, pre ambicióznejšieho čitateľa podáme ich dôkaz vychádzajúci len z definície determinantu. Navyše sa tým náš výklad stane formálne nezávislým na motivačných úvahách o orientovanom objeme z prvej časti úvodného paragrafu 10.1.

Dôkaz vety 10.3.1 odložíme až do nasledujúceho paragrafu, kde nám poslúži ako vhodný úvod do ďalšieho okruhu otázok. Na tomto mieste však zaznamenáme dva bezprostredné dôsledky tejto charakterizačnej vety. Samozrejme, v jej dôkaze sa na ne nebudeme odvolávať.

10.3.2. Veta. *(Cauchy) Pre ľubovoľné matice* $\boldsymbol{A}$, $\boldsymbol{B} \in K^{n\times n}$ *platí*

$$\det(\boldsymbol{A}\cdot\boldsymbol{B}) = \det\boldsymbol{A}\cdot\det\boldsymbol{B};$$

t. j. determinant súčinu matíc sa rovná súčinu ich determinantov.

Dôkaz. Zvoľme pevne maticu $\boldsymbol{A} \in K^{n\times n}$ a definujme zobrazenie $F\colon K^{n\times n} \to K$ predpisom $F(\boldsymbol{B}) = \det(\boldsymbol{A}\cdot\boldsymbol{B})$ pre $\boldsymbol{B} \in K^{n\times n}$. Overíme, že F je n-lineárne alternujúce zobrazenie stĺpcov matice $\boldsymbol{B}$; označme ich $\boldsymbol{v}_1,\dots,\boldsymbol{v}_n$.

Najprv overíme, že F je alternujúce. Nech $1 \le i < j \le n$ a $\boldsymbol{B}$ je matica taká, že $\boldsymbol{v}_i = \boldsymbol{v}_j$ pre nejaké $i < j$. Potom aj $\boldsymbol{A}\cdot\boldsymbol{v}_i = \boldsymbol{A}\cdot\boldsymbol{v}_j$, a s využitím alternácie determinantu dostávame

$$\begin{aligned}F(\boldsymbol{B}) &= \det(\boldsymbol{A}\cdot(\boldsymbol{v}_1,\dots,\boldsymbol{v}_i,\dots,\boldsymbol{v}_j,\dots,\boldsymbol{v}_n))\\ &= \det(\boldsymbol{A}\cdot\boldsymbol{v}_1,\dots,\boldsymbol{A}\cdot\boldsymbol{v}_i,\dots,\boldsymbol{A}\cdot\boldsymbol{v}_j,\dots,\boldsymbol{A}\cdot\boldsymbol{v}_n) = 0.\end{aligned}$$

Teraz dokážeme multilinearitu F. Zafixujme stĺpce $\boldsymbol{v}_1,\dots,\boldsymbol{v}_{j-1},\boldsymbol{v}_{j+1},\dots,\boldsymbol{v}_n$ a na miesto j-teho stĺpca dosaďme vektor $a\boldsymbol{x}+b\boldsymbol{y}$. S využitím n-linearity determinantu nám vyjde

$$\begin{aligned}F(\boldsymbol{v}_1,\dots,a\boldsymbol{x}+b\boldsymbol{y},\dots,\boldsymbol{v}_n) &= \det(\boldsymbol{A}\cdot(\boldsymbol{v}_1,\dots,a\boldsymbol{x}+b\boldsymbol{y},\dots,\boldsymbol{v}_n))\\ &= \det(\boldsymbol{A}\cdot\boldsymbol{v}_1,\dots,\boldsymbol{A}\cdot(a\boldsymbol{x}+b\boldsymbol{y}),\dots,\boldsymbol{A}\cdot\boldsymbol{v}_n))\\ &= \det(\boldsymbol{A}\cdot\boldsymbol{v}_1,\dots,a(\boldsymbol{A}\cdot\boldsymbol{x})+b(\boldsymbol{A}\cdot\boldsymbol{y}),\dots,\boldsymbol{A}\cdot\boldsymbol{v}_n))\\ &= a\det(\boldsymbol{A}\cdot\boldsymbol{v}_1,\dots,\boldsymbol{A}\cdot\boldsymbol{x},\dots,\boldsymbol{A}\cdot\boldsymbol{v}_n)\\ &\qquad + b\det(\boldsymbol{A}\cdot\boldsymbol{v}_1,\dots,\boldsymbol{A}\cdot\boldsymbol{y},\dots,\boldsymbol{A}\cdot\boldsymbol{v}_n)\\ &= a\det(\boldsymbol{A}\cdot(\boldsymbol{v}_1,\dots,\boldsymbol{x},\dots,\boldsymbol{v}_n))\\ &\qquad + b\det(\boldsymbol{A}\cdot(\boldsymbol{v}_1,\dots,\boldsymbol{y},\dots,\boldsymbol{v}_n))\\ &= aF(\boldsymbol{v}_1,\dots,\boldsymbol{x},\dots,\boldsymbol{v}_n) + bF(\boldsymbol{v}_1,\dots,\boldsymbol{y},\dots,\boldsymbol{v}_n).\end{aligned}$$

Podľa vety 10.3.1 má F tvar $F(\boldsymbol{B}) = c\det\boldsymbol{B}$ pre jednoznačne určený skalár $c = F(\boldsymbol{I}_n) = \det(\boldsymbol{A}\cdot\boldsymbol{I}_n) = \det\boldsymbol{A}$.

10.3.3. Veta. *Štvorcová matica* $\boldsymbol{A} \in K^{n\times n}$ *je regulárna práve vtedy, keď* $\det\boldsymbol{A} \ne 0$. *V tom prípade*

$$\det\bigl(\boldsymbol{A}^{-1}\bigr) = (\det\boldsymbol{A})^{-1}.$$

Dôkaz. Ak $\boldsymbol{A}$ je singulárna, tak jej stĺpce sú lineárne závislé. Podľa lemy 10.1.3 (b) je $F(\boldsymbol{A}) = 0$ pre ľubovoľnú n-lineárnu alternujúcu funkciu $F\colon K^{n\times n} \to K$. Teda špeciálne $\det\boldsymbol{A} = 0$. Naopak, nech $\boldsymbol{A}$ je regulárna. Potom podľa vety 10.3.2,

$$\det\boldsymbol{A}\cdot\det\bigl(\boldsymbol{A}^{-1}\bigr) = \det\bigl(\boldsymbol{A}\cdot\boldsymbol{A}^{-1}\bigr) = \det\boldsymbol{I}_n = 1.$$

Preto $\det\boldsymbol{A} \ne 0$ a $\det\bigl(\boldsymbol{A}^{-1}\bigr) = (\det\boldsymbol{A})^{-1}$.

10.4 Laplaceov rozvoj determinantu

Náš výklad začneme sľúbeným dôkazom. Keďže pre $n = 0, 1$ niet čo dokazovať, aby sme sa vyhli rozpitvávaniu trivialít, budeme v celom paragrafe predpokladať, že $n \geq 2$.

Dôkaz vety 10.3.1 Najprv dokážeme, že determinant je alternujúca funkcia. Nech $\boldsymbol{A} \in K^{n\times n}$ je taká, že $\boldsymbol{s}_i(\boldsymbol{A}) = \boldsymbol{s}_j(\boldsymbol{A})$ pre nejaké $i < j$. Označme $\tau \in \mathcal{S}_n$ transpozíciu, ktorá vymieňa i a j (a ostatné prvky necháva namieste). Množina všetkých párnych permutácií množiny $\{1, \ldots, n\}$ sa zvykne značiť $\mathcal{A}_n$. Zrejme priradením $\sigma \mapsto \sigma \circ \tau$ je daná bijekcia $\mathcal{A}_n \to \mathcal{S}_n \smallsetminus \mathcal{A}_n$. Pre všetky $k, l \leq n$ platí $a_{kl} = a_{k\,\tau(l)}$, a keďže $\tau^{-1} = \tau$, pre každú permutáciu $\sigma \in \mathcal{S}_n$ máme

$$a_{(\sigma\circ\tau)(1)\,1} \ldots a_{(\sigma\circ\tau)(n)\,n} = a_{\sigma(1)\,\tau(1)} \ldots a_{\sigma(n)\,\tau(n)} = a_{\sigma(1)\,1} \ldots a_{\sigma(n)\,n}$$

S využitím toho môžeme počítať

$$\begin{aligned}
\det \boldsymbol{A} &= \sum_{\sigma\in\mathcal{S}_n} (-1)^{|\sigma|} a_{\sigma(1)\,1} \ldots a_{\sigma(n)\,n} \\
&= \sum_{\sigma\in\mathcal{A}_n} a_{\sigma(1)\,1} \ldots a_{\sigma(n)\,n} - \sum_{\sigma\in\mathcal{S}_n\smallsetminus\mathcal{A}_n} a_{\sigma(1)\,1} \ldots a_{\sigma(n)\,n} \\
&= \sum_{\sigma\in\mathcal{A}_n} a_{\sigma(1)\,1} \ldots a_{\sigma(n)\,n} - \sum_{\sigma\in\mathcal{A}_n} a_{(\sigma\circ\tau)(1)\,1} \ldots a_{(\sigma\circ\tau)(n)\,n} = 0.
\end{aligned}$$

Teraz dokážeme, že $\det \boldsymbol{A}$ je lineárnou funkciou j-teho stĺpca $(a_{1j}, \ldots, a_{nj})^{\mathsf{T}}$. Pre $i \leq n$ označme $\mathcal{S}_n(i, j) = \{\sigma \in \mathcal{S}_n;\ i = \sigma(j)\}$ a položme

$$\tilde{a}_{ij} = \sum_{\sigma\in\mathcal{S}_n(i,j)} (-1)^{|\sigma|} a_{\sigma(1)\,1} \ldots a_{\sigma(j-1)\,j-1} a_{\sigma(j+1)\,j+1} \ldots a_{\sigma(n)\,n}.$$

Potom zrejme

$$\det \boldsymbol{A} = \sum_{i=1}^{n} \tilde{a}_{ij} a_{ij} = (\tilde{a}_{1j}, \ldots, \tilde{a}_{nj}) \cdot (a_{1j}, \ldots, a_{nj})^{\mathsf{T}},$$

čo dokazuje spomínanú linearitu.

Na základe tvrdenia 10.2.1 platí aj „riadková verzia“ práve dokázanej vety 10.3.1 Špeciálne, determinant je takisto multilineárna alternujúca funkcia riadkov matice a (keďže $\sigma \in \mathcal{S}_n(i, j) \Leftrightarrow \sigma^{-1} \in \mathcal{S}_n(j, i)$) pre i-ty riadok $(a_{i1}, \ldots, a_{in})$ matice $\boldsymbol{A}$ jej determinant má rozvoj

$$\det \boldsymbol{A} = \sum_{j=1}^{n} a_{ij} \tilde{a}_{ij} = (a_{i1}, \ldots, a_{in}) \cdot (\tilde{a}_{i1}, \ldots, \tilde{a}_{in})^{\mathsf{T}}.$$

s rovnako definovanými koeficientmi $\tilde{a}_{ij}$.

Uvedený prvok $\tilde{a}_{ij}$ nazývame *algebraickým doplnkom* prvku a_{ij} v matici $\boldsymbol{A}$. Maticu $\widetilde{\boldsymbol{A}} = (\tilde{a}_{ij})_{n\times n}$ nazývame *maticou algebraických doplnkov* k matici $\boldsymbol{A}$.

10.4.1. Tvrdenie. *Nech $\boldsymbol{A}_{ij}$ označuje maticu rádu $n-1$, ktorá vznikne z matice $\boldsymbol{A} \in K^{n\times n}$ vynechaním i-teho riadku a j-teho stĺpca. Potom*

$$\tilde{a}_{ij} = (-1)^{i+j}|\boldsymbol{A}_{ij}|.$$

Dôkaz. Označme $\boldsymbol{B}$ maticu, ktorá vznikne nahradením j-teho stĺpca matice $\boldsymbol{A}$ stĺpcovým vektorom $\boldsymbol{e}_i \in K^n$. Zrejme $\tilde{a}_{ij} = |\boldsymbol{B}|$. Ak budeme v matici $\boldsymbol{B}$ postupne vymieňať stĺpce s indexmi j a $j+1$, ďalej $j+1$ a $j+2$, atď., až nakoniec $n-1$ a n, a potom riadky s indexmi i a $i+1$, ďalej $i+1$ a $i+2$, atď., až napokon $n-1$ a n, dostaneme maticu tvaru

$$\boldsymbol{C} = \begin{pmatrix} \boldsymbol{A}_{ij} & \boldsymbol{0} \\ \boldsymbol{b} & 1 \end{pmatrix},$$

kde $\boldsymbol{b}$ vznikne z i-teho riadku matice $\boldsymbol{A}$ vynechaním j-teho prvku a $\boldsymbol{0}$ je nulový stĺpec dĺžky $n-1$. Podľa tvrdenia 10.2.2 (a poznámky za jeho dôkazom), determinant tejto matice je $|\boldsymbol{A}_{ij}|$. Keďže determinant je alternujúca funkcia tak stĺpcov ako aj riadkov matice a všetkých výmien bolo dohromady $(n-j)+(n-i) = 2n-(i+j)$, platí

$$\tilde{a}_{ij} = |\boldsymbol{B}| = (-1)^{2n-(i+j)}|\boldsymbol{C}| = (-1)^{i+j}|\boldsymbol{A}_{ij}|.$$

Determinanty matíc, ktoré vzniknú vynechaním niektorých riadkov a rovnakého počtu sĺtpcov z matice $\boldsymbol{A} \in K^{n\times n}$, nazývame jej *minormi*, prípadne *subdeterminantmi* determinantu $|\boldsymbol{A}|$. Dosadením získaných hodnôt algebraických doplnkov do rozvoja determinantu rádu n podľa niektorého riadku resp. stĺpca tak dostávame jeho vyjadrenie pomocou subdeterminantov rádu $n-1$.

10.4.2. Veta. *Nech $\boldsymbol{A} \in K^{n\times n}$, $1 \le k, l \le n$. Potom*

$$|\boldsymbol{A}| = \sum_{j=1}^{n} (-1)^{k+j} a_{kj}\,|\boldsymbol{A}_{kj}| = \sum_{i=1}^{n} (-1)^{i+l} a_{il}\,|\boldsymbol{A}_{il}|.$$

Uvedené súčty nazývame *Laplaceovými rozvojmi* determinantu $|\boldsymbol{A}|$ – prvý podľa k-teho riadku, druhý podľa l-teho stĺpca.

10.5 Výpočet determinantu

Skôr než sa pustíme do výpočtov konkrétnych determinantov, skúsme si urobiť inventúru prostriedkov, ktoré máme nato k dispozícii, a posúdiť ich vhodnosť.

Asi sa zhodneme na tom, že výpočet determinantu rádu n podľa jeho definície, ako súčtu $n!$ súčinov po n činiteľoch, by bol značne ťažkopádny. Pokiaľ sme sa stretli len s prípadmi $n = 2$ alebo $n = 3$, nemusíme si to jasne uvedomiť. Avšak už $4! = 24$, $5! = 120$ a funkcia $n!$ veľmi rýchlo rastie. Preto je potrebné pouvažovať o nejakej inej metóde.

Keďže determinant je multilineárnou alternujúcou funkciou tak riadkov ako aj stĺpcov matice, ako najprirodzenejšia sa nám ponúka metóda úpravy matice na hornú prípadne dolnú trojuholníkovú maticu pomocou elementárnych riadkových i stĺpcových operácií. Ako sme už spomínali v poznámke (2) za dôkazom tvrdenia 10.2.2:

(0) Determinant trojuholníkovej matice sa rovná súčinu jej diagonálnych prvkov.

Pripomeňme si aj nasledujúce pravidlá z paragrafu 10.1 o vplyve ERO a ESO na determinant:

(1) Výmenou poradia dvoch riadkov alebo stĺpcov matice sa hodnota jej determinantu zmení na opačnú.

(2) Vynásobením nejakého riadku alebo stĺpca matice nenulovým skalárom $c \in K$ sa jej determinant zmení na c-násobok pôvodnej hodnoty.

(3) Pripočítaním skalárneho násobku nejakého riadku matice k jej inému riadku, resp. násobku nejakého jej stĺpca k inému stĺpcu sa hodnota jej determinantu nezmení.

Všimnite si, že len tretí typ menovaných úprav necháva determinant bezo zmeny! Poznamenajme, že úpravy typu (3) spolu s pravidlom (0) plne postačujú na výpočet akéhokoľvek determinantu. Bez pravidiel (1) a (2) sa možno kľudne zaobísť, občas nám však môžu pomôcť sprehľadniť situáciu, preto sa im nebudeme vyhýbať.

Často býva užitočné výslovne si uvedomiť nasledujúci dôsledok pravidiel (1)–(3):

(4) Ak matica obsahuje nulový riadok alebo stĺpec, prípadne dva rovnaké riadky alebo stĺpce, tak jej determinant je 0.

Mohlo by sa zdať, že sme akosi pozabudli na Laplaceov rozvoj. Táto metóda umožňuje previesť výpočet determinantu rádu n na výpočet n determinantov rádu $n - 1$, presnejšie na istú ich lineárnu kombináciu. Ak by sme dôsledne pokračovali ďalej, mohli by sme túto úlohu previesť na výpočet $n(n - 1)$ determinantov rádu $n - 2$, atď., až by sme napokon dostali $n!$ determinantov rádu 1. Ak si to dobre premyslíme, zistíme, že takýto výpočet by bol rovnako efektívny (či, lepšie povedané, neefektívny) ako výpočet determinantu priamo na základe jeho definície.

Jednako sa Laplaceovho rozvoja celkom nezriekame. Odporúčame ho však používať len vtedy, keď sú všetky prvky príslušného riadku či stĺpca, podľa ktorého determinant rozvíjame, až na jednu výnimku rovné nule. Vtedy vlastne nejde ani tak o rozvoj ako o zníženie rádu daného determinantu o 1 (bez nárastu počtu determinantov). Toto odporúčanie sformulujeme do nášho predposledného pravidla:

(5) Nech všetky prvky i-teho riadku prípadne j-teho stĺpca matice $\boldsymbol{A}$ s výnimkou prvku a_{ij} sú rovné 0. Potom

$$|\boldsymbol{A}| = (-1)^{i+j} a_{ij}\, |\boldsymbol{A}_{ij}|.$$

Všimnite si, že pravidlo (0) možno dostať ako dôsledok $(n-1)$-násobného použitia pravidla (5) a zrejmého faktu, že determinant matice (a) typu 1×1 je samotná hodnota a.

Ak si ešte uvedomíme, že determinanty rádu 2 možno najvýhodnejšie počítať priamo z definície:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}, \tag{6}$$

je to už naozaj všetko, čo potrebujeme vedieť na *efektívny* výpočet determinantu.

10.5.1. Príklad. Vypočítame determinant reálnej matice

$$\boldsymbol{A} = \begin{pmatrix} 2 & 2 & 1 & 1 & 1 \\ 5 & 6 & 3 & 4 & 5 \\ 7 & 5 & 3 & 5 & 7 \\ 13 & 10 & 3 & 8 & 13 \\ 7 & 2 & 1 & 1 & 6 \end{pmatrix}.$$

Najprv odpočítame prvý riadok od piateho a druhý od štvrtého. V matici, ktorú takto získame, odpočítame piaty stĺpec od prvého a štvrtý od druhého. Postupne tak dostaneme

$$|\boldsymbol{A}| = \begin{vmatrix} 2 & 2 & 1 & 1 & 1 \\ 5 & 6 & 3 & 4 & 5 \\ 7 & 5 & 3 & 5 & 7 \\ 8 & 4 & 0 & 4 & 8 \\ 5 & 0 & 0 & 0 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 & 5 \\ 0 & 0 & 3 & 5 & 7 \\ 0 & 0 & 0 & 4 & 8 \\ 0 & 0 & 0 & 0 & 5 \end{vmatrix} = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120.$$

Priznávame, že výpočet, ktorý sme práve predviedli je tak trochu podraz voči čitateľovi. Úpravy, ktoré sme pri ňom použili, boli totiž len opačným postupom, ktorým sme pri formulácii úlohy z vopred narafičenej výslednej hornej trojuholníkovej matice „uvarili“ zadanie. Napokon, tak je tomu s väčšinou úloh v učebniciach. No čitateľ, ktorého autor „nevpustil do kuchyne“, má len malú nádej toto optimálne riešenie nájsť. Teda aspoň pokiaľ je úloha dobre postavená. Predvedieme preto aj iné, „normálne“ riešenie, na aké má šancu prísť aj nezasvätený riešiteľ.

Najprv odpočítame tretí stĺpec od štvrtého aj piateho a jeho dvojnásobok od prvého aj druhého stĺpca. V ďalšom kroku determinant rozvinieme podľa prvého riadku:

$$|\boldsymbol{A}| = \begin{vmatrix} 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 3 & 1 & 2 \\ 1 & -1 & 3 & 2 & 4 \\ 7 & 4 & 3 & 5 & 10 \\ 5 & 0 & 1 & 0 & 5 \end{vmatrix} = \begin{vmatrix} -1 & 0 & 1 & 2 \\ 1 & -1 & 2 & 4 \\ 7 & 4 & 5 & 10 \\ 5 & 0 & 0 & 5 \end{vmatrix}.$$

Teraz odpočítame prvý stĺpec od posledného a získaný determinant rozvinieme podľa posledného riadku:

$$|\boldsymbol{A}| = \begin{vmatrix} -1 & 0 & 1 & 3 \\ 1 & -1 & 2 & 3 \\ 7 & 4 & 5 & 3 \\ 5 & 0 & 0 & 0 \end{vmatrix} = -5 \begin{vmatrix} 0 & 1 & 3 \\ -1 & 2 & 3 \\ 4 & 5 & 3 \end{vmatrix}.$$

Po odpočítaní trojnásobku druhého stĺpca od tretieho a rozvinutí podľa prvého riadku sme konečne v cieli:

$$|\boldsymbol{A}| = -5 \begin{vmatrix} 0 & 1 & 0 \\ -1 & 2 & -3 \\ 4 & 5 & -12 \end{vmatrix} = 5 \begin{vmatrix} -1 & -3 \\ 4 & -12 \end{vmatrix} = 5\,(12 + 3 \cdot 4) = 5 \cdot 24 = 120.$$

10.5.2. Príklad. Vypočítame tzv. *Vandermondov determinant* rádu n

$$\mathrm{VD}_n(x_1, x_2, \dots, x_n) = \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix}.$$

Odpočítaním prvého riadku od všetkých ostatných riadkov a následným rozvojom podľa prvého stĺpca dostaneme

$$\mathrm{VD}_n(x_1, x_2, \ldots, x_n) = \begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 0 & x_2 - x_1 & x_2^2 - x_1^2 & \ldots & x_2^{n-1} - x_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_n - x_1 & x_n^2 - x_1^2 & \ldots & x_n^{n-1} - x_1^{n-1} \end{vmatrix}$$

$$= \begin{vmatrix} x_2 - x_1 & x_2^2 - x_1^2 & \ldots & x_2^{n-1} - x_1^{n-1} \\ \vdots & \vdots & & \vdots \\ x_n - x_1 & x_n^2 - x_1^2 & \ldots & x_n^{n-1} - x_1^{n-1} \end{vmatrix}.$$

Odpočítajme teraz od každého stĺpca počnúc druhým x_1-násobok predchádzajúceho stĺpca. V determinante, ktorý získame, je na mieste (i, k), kde $1 \le i, k \le n-1$, prvok

$$\bigl(x_{i+1}^k - x_1^k\bigr) - x_1\bigl(x_{i+1}^{k-1} - x_1^{k-1}\bigr) = x_{i+1}^{k-1}(x_{i+1} - x_1).$$

Ak teda vyjmeme z i-teho riadku činiteľ $x_{i+1} - x_1$, postupne nám vyjde

$$\mathrm{VD}_n(x_1, x_2, \ldots, x_n) = \begin{vmatrix} x_2 - x_1 & x_2(x_2 - x_1) & \ldots & x_2^{n-2}(x_2 - x_1) \\ \vdots & \vdots & & \vdots \\ x_n - x_1 & x_n(x_n - x_1) & \ldots & x_n^{n-2}(x_n - x_1) \end{vmatrix}$$

$$= (x_2 - x_1) \ldots (x_n - x_1) \begin{vmatrix} 1 & x_2 & \ldots & x_2^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \ldots & x_n^{n-2} \end{vmatrix}$$

$$= (x_2 - x_1) \ldots (x_n - x_1)\ \mathrm{VD}_{n-1}(x_2, \ldots, x_n).$$

Teraz už aj bez počítania determinantov vidíme, že

$$\mathrm{VD}_{n-1}(x_2, \ldots, x_n) = (x_3 - x_2) \ldots (x_n - x_2)\ \mathrm{VD}_{n-2}(x_3, \ldots, x_n),$$

atď. Keďže zrejme $\mathrm{VD}_1(x_n) = 1$, dostávame výsledok

$$\mathrm{VD}_n(x_1, x_2, \ldots, x_n) = \prod_{1 \le i < j \le n} (x_j - x_i),$$

kde symbolom $\prod$ označujeme súčin príslušných činiteľov.

10.6 Inverzná matica a Cramerovo pravidlo

V tomto záverečnom paragrafe si predvedieme dva príklady využitia determinantov. Vyjadríme pomocou nich inverznú maticu k regulárnej štvorcovej matici a riešenie sústavy lineárnych rovníc s regulárnou štvorcovou maticou. Vopred poznamenajme, že tieto vyjadrenia sa na priame výpočty príliš nehodia. Na druhej strane tým, že vyjadrujú inverznú maticu a riešenie spomínanej sústavy v tvare prehľadných ucelených formúl, sú významné hlavne z teoretického hľadiska.

Nech $\boldsymbol{A} \in K^{n\times n}$ a $1 \leq i, k \leq n$ sú rôzne indexy. Označme $\boldsymbol{B}$ maticu, ktorá vznikne z matice $\boldsymbol{A}$ nahradením jej k-teho riadku i-tym riadkom. Potom matica $\boldsymbol{B}$ má (aspoň) dva riadky rovnaké, menovite i-tý a k-tý, preto $|\boldsymbol{B}| = 0$. Na druhej strane matice $\boldsymbol{A}$ a $\boldsymbol{B}$ sa líšia nanajvýš v k-tom riadku, preto $\boldsymbol{A}_{kj} = \boldsymbol{B}_{kj}$ pre každé $1 \leq j \leq n$. Z toho dôvodu sú algebraické doplnky zodpovedajúcich si prvkov k-tych riadkov oboch matíc rovnaké:

$$\tilde{b}_{kj} = (-1)^{k+j}|\boldsymbol{B}_{kj}| = (-1)^{k+j}|\boldsymbol{A}_{kj}| = \tilde{a}_{kj}.$$

Ak rozvinieme determinant matice $\boldsymbol{B}$ podľa jej k-teho riadku, dostaneme

$$\det \boldsymbol{B} = \sum_{j=1}^{n} b_{kj}\tilde{b}_{kj} = \sum_{j=1}^{n} a_{ij}\tilde{a}_{kj} = 0.$$

Spojenie tejto rovnosti s Laplaceovým rozvojom determinantu matice $\boldsymbol{A}$ podľa k-teho riadku dáva

$$\sum_{j=1}^{n} a_{ij}\tilde{a}_{kj} = \boldsymbol{r}_i(\boldsymbol{A}) \cdot \boldsymbol{r}_k\big(\widetilde{\boldsymbol{A}}\big)^{\mathsf{T}} = \boldsymbol{r}_i(\boldsymbol{A}) \cdot \boldsymbol{s}_k\big(\widetilde{\boldsymbol{A}}^{\mathsf{T}}\big) = \begin{cases} |\boldsymbol{A}|, & \text{ak } i = k, \\ 0, & \text{ak } i \neq k. \end{cases}$$

Inak povedané

$$\boldsymbol{A} \cdot \widetilde{\boldsymbol{A}}^{\mathsf{T}} = |\boldsymbol{A}|\, \boldsymbol{I}_n.$$

Inverznú maticu k regulárnej štvorcovej matici $\boldsymbol{A}$ potom dostaneme tak, že transponovanú maticu jej algebraických doplnkov vydelíme determinantom $|\boldsymbol{A}|$. Tým sme dokázali nasledujúcu vetu.

10.6.1. Veta. *Nech $\boldsymbol{A} \in K^{n\times n}$ je regulárna matica. Potom*

$$\boldsymbol{A}^{-1} = \frac{1}{|\boldsymbol{A}|}\widetilde{\boldsymbol{A}}^{\mathsf{T}}.$$

10.6.2. Príklad. Nájdeme inverznú maticu k reálnej matici

$$\boldsymbol{A} = \begin{pmatrix} 1 & -2 \\ 5 & -3 \end{pmatrix}.$$

Jej determinant a maticu algebraických doplnkov vypočítame ľahko:

$$|\boldsymbol{A}| = 1 \cdot (-3) - 5 \cdot (-2) = 7, \qquad \widetilde{\boldsymbol{A}} = \begin{pmatrix} -3 & -5 \\ 2 & 1 \end{pmatrix}.$$

Preto

$$\boldsymbol{A}^{-1} = \frac{1}{7}\begin{pmatrix} -3 & 2 \\ -5 & 1 \end{pmatrix}.$$

Poznamenajme však, že okrem štvorcových matíc rádu 2, kedy je to v podstate jedno, je výpočet inverznej matice pomocou ERO alebo ESO, ako sme ho popísali v paragrafe 7.4, podstatne výhodnejší než výpočet na základe vety 10.6.1. Už pre matice rádu 3 by sme na to potrebovali vypočítať jeden determinant rádu 3 a deväť determinantov rádu 2. Vo všeobecnom prípade by sme museli vypočítať jeden determinant rádu n a n^2 determinantov rádu $n-1$. Čitateľ sa pravdepodobne po prvýkrát stretol s determinantmi v súvislosti s riešením sústav lineárnych rovníc, v ktorých je počet rovníc a neznámych ten istý. Možno by si v prípadoch 2×2 a 3×3 ešte vedel spomenúť aj na príslušné vzorce. Takéto vzorce, známe ako *Cramerovo pravidlo*, však platia v ľubovoľnom rozmere $n \times n$.

10.6.3. Veta. *Nech $\boldsymbol{A} \in K^{n\times n}$ je regulárna matica, $\boldsymbol{b} \in K^n$ a pre $1 \le j \le n$ nech $\boldsymbol{A}_j^{\boldsymbol{b}}$ označuje maticu, ktorá vznikne z matice $\boldsymbol{A}$ nahradením jej j-teho stĺpca stĺpcovým vektorom $\boldsymbol{b}$. Potom sústava $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ má jediné riešenie*

$$\boldsymbol{x} = \left(\frac{|\boldsymbol{A}_1^{\boldsymbol{b}}|}{|\boldsymbol{A}|}, \frac{|\boldsymbol{A}_2^{\boldsymbol{b}}|}{|\boldsymbol{A}|}, \dots, \frac{|\boldsymbol{A}_n^{\boldsymbol{b}}|}{|\boldsymbol{A}|} \right)^{\mathsf{T}}.$$

Dôkaz. Podľa vety 7.3.6 má uvedená sústava jediné riešenie $\boldsymbol{x} = \boldsymbol{A}^{-1} \cdot \boldsymbol{b}$. Ak do tohto vyjadrenia dosadíme za $\boldsymbol{A}^{-1}$ z vety 10.6.1, pre j-tu zložku vektora $\boldsymbol{x}$ nám vyjde

$$x_j = \frac{1}{|\boldsymbol{A}|} \sum_{i=1}^{n} \tilde{a}_{ij} b_i = \frac{|\boldsymbol{A}_j^{\boldsymbol{b}}|}{|\boldsymbol{A}|},$$

lebo zodpovedajúce si prvky j-tych stĺpcov matíc $\boldsymbol{A}$ a $\boldsymbol{A}_j^{\boldsymbol{b}}$ majú rovnaké algebraické doplnky, takže uvedený súčet je Laplaceov rozvoj determinantu $|\boldsymbol{A}_j^{\boldsymbol{b}}|$ podľa j-teho stĺpca.

Varujeme však čitateľa pred používaním Cramerovho pravidla na riešenie konkrétnych sústav n lineárnych rovníc o n neznámych. Metóda úpravy rozšírenej matice sústavy na redukovaný stupňovitý tvar pomocou ERO je oveľa rýchlejšia a pohodlnejšia. Kým vyriešenie takej sústavy pomocou ERO vyžaduje úpravu jedinej matice typu $n \times (n+1)$, pri riešení Cramerovým pravidlom by sme museli pri výpočte determinantov upraviť $n+1$ matíc typu $n \times n$.

Na obranu determinantov však poznamenajme, že stretnutie s najdôležitejšími príkladmi ich využitia nás ešte len čaká. Popri tzv. Gramových determinantoch to bude najmä v súvislosti s charakteristickým polynómom a vlastnými hodnotami lineárnych transformácií a štvorcových matíc.

Cvičenia

10.1. Nech U, V sú vektorové priestory nad poľom $\mathbb{Q}$. Dokážte nasledujúce tvrdenia:

(a) Ak $F\colon V \to U$ je zobrazenie také, že pre všetky $c \in \mathbb{Z}$, $\boldsymbol{v} \in V$ platí $F(c\boldsymbol{v}) = cF(\boldsymbol{v})$, tak uvedená rovnosť platí aj pre všetky $c \in \mathbb{Q}$.

(b) Ak $F\colon V \to U$ je zobrazenie také, že pre všetky $c \in \mathbb{Z}$, $\boldsymbol{v} \in V$ platí $F(c\boldsymbol{v}) = |c|F(\boldsymbol{v})$, tak uvedená rovnosť platí aj pre všetky $c \in \mathbb{Q}$.

10.2. Nech U, $V_1, \dots, V_n$ sú vektorové priestory nad poľom K. Podobne ako v paragrafe 10.1 definujte pojem n-lineárneho zobrazenia $F\colon V_1 \times \ldots \times V_n \to U$. Označme $\mathcal{L}_n(V_1, \dots, V_n, U)$ množinu všetkých n-lineárnych zobrazení $F\colon V_1 \times \ldots \times V_n \to U$; ak $V_1 = \ldots = V_n = V$, tak miesto $\mathcal{L}_n(V_1, \dots, V_n, U)$ píšeme len $\mathcal{L}_n(V, U)$. Dokážte, že množina $\mathcal{L}_n(V_1, \dots, V_n, U)$ tvorí lineárny podpriestor vektorového priestoru $U^{V_1 \times \ldots \times V_n}$ všetkých zobrazení $V_1 \times \ldots \times V_n \to U$.

10.3. Nech U, V sú vektorové priestory nad poľom K a $n \geq 2$. Definujte pojmy symetrického, antisymetrického a alternujúceho zobrazenia $F\colon V^n \to U$. Dokážte postupne nasledujúce tvrdenia:

(a) Množiny $\mathrm{Sym}_n(V, U)$ všetkých symetrických zobrazení $V^n \to U$, $\mathrm{Asym}_n(V, U)$ všetkých antisymetrických zobrazení $V^n \to U$ a $\mathrm{Alt}_n(V, U)$ všetkých alternujúcich zobrazení $V^n \to K$ sú lineárne podpriestory vektorového priestoru U^{V^n} všetkých zobrazení $V^n \to U$.

(b) Ak $\mathrm{char}\, K = 2$, tak $\mathrm{Sym}_n(V, U) = \mathrm{Asym}_n(V, U)$.

(c) Ak $\mathrm{char}\, K \neq 2$, tak $\mathrm{Asym}_n(V, U) \subseteq \mathrm{Alt}_n(V, U)$ a $\mathrm{Sym}_n(V, U) \cap \mathrm{Asym}_n(V, U) = \{\mathbf{0}\}$, kde $\mathbf{0}$ tentokrát označuje identicky nulové zobrazenie $V^n \to U$.

(d) $\mathcal{L}_n(V, U) \cap \mathrm{Alt}_n(V, U) \subseteq \mathcal{L}_n(V, U) \cap \mathrm{Asym}_n(V, U)$.

(e) Ak $\mathrm{char}\, K \neq 2$, tak $\mathcal{L}_n(V, U) \cap \mathrm{Alt}_n(V, U) = \mathcal{L}_n(V, U) \cap \mathrm{Asym}_n(V, U)$.

(f) Nájdite príklad bilineárneho (t. j. 2-lineárneho) zobrazenia $\mathbb{Z}_2 \times \mathbb{Z}_2 \to \mathbb{Z}_2$, ktoré je symetrické (teda aj antisymetrické), no nie je alternujúce.

(g) Ak $\dim V = n$, tak $\dim\big(\mathcal{L}_n(V, K) \cap \mathrm{Alt}_n(V, K)\big) = 1$. (*Návod*: Dobre si uvedomte, čo vlastne hovorí druhá časť vety 10.3.1)

10.4. Dokážte Lemy 10.1.2 a 10.1.3 za všeobecnejších podmienok cvičenia 10.3, t. j. pre zobrazenia $F\colon V^n \to U$.

10.5. (a) Vypočítajte orientovaný aj neorientovaný plošný obsah rovnobežníka určeného vektormi $\boldsymbol{u} = (1, 5)^\mathsf{T}$, $\boldsymbol{v} = (3, -4)^\mathsf{T}$ v rovine $\mathbb{R}^2$.

(b) Vypočítajte orientovaný aj neorientovaný objem rovnobežnostena určeného vektormi $\boldsymbol{u} = (3, 2, -1)$, $\boldsymbol{v} = (5, 0, 4)$, $\boldsymbol{w} = (1, 1, 1)$ v priestore $\mathbb{R}^3$.

(c) Vypočítajte orientovaný aj neorientovaný štvorrozmerný objem štvorrozmerného „rovnobežnonadstena“ určeného vektormi $\boldsymbol{e}_1 + \boldsymbol{e}_2$, $\boldsymbol{e}_2 - 2\boldsymbol{e}_3$, $4\boldsymbol{e}_3 - 2\boldsymbol{e}_4$, $3\boldsymbol{e}_1 + \boldsymbol{e}_4$ v priestore $\mathbb{R}^4$.

10.6. Vypočítajte nasledujúce determinanty nad poľom $\mathbb{R}$:

(a) $\begin{vmatrix} 0 & 5 & 3 & -2 \\ 1 & 1 & 6 & -4 \\ 2 & 1 & -3 & 0 \\ 3 & 3 & 9 & -6 \end{vmatrix}$; (b) $\begin{vmatrix} 3 & 5 & 7 & 4 \\ 0 & 1 & 0 & -1 \\ 1 & 4 & 4 & 0 \\ 4 & 7 & 11 & 6 \end{vmatrix}$; (c) $\begin{vmatrix} 3 & 7 & 131 & -17 \\ 2 & 6 & -21 & 401 \\ 0 & 0 & 100 & 29 \\ 0 & 0 & 3 & 1 \end{vmatrix}$; (d) $\begin{vmatrix} 2 & -2 & 0 & 4 & 0 \\ 2 & 3 & 5 & 7 & 11 \\ 1 & -2 & 4 & 0 & 2 \\ 0 & 6 & 0 & 1 & -1 \\ 1 & -1 & 1 & 1 & -1 \end{vmatrix}$;

(e) $\begin{vmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 & 5 \\ 1 & 2 & 3 & 0 & 4 \\ 1 & -1 & 0 & 7 & 0 \\ 1 & 0 & 1 & 2 & 3 \end{vmatrix}$.

10.7. Rozhodnite, pre ktoré hodnoty parametra λ je uvedená matica regulárna resp. singulárna. Riešte nad každým z polí $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}_2$ a $\mathbb{Z}_7$ pre nasledujúce matice (ak vám prekáža, že niektorý prvok matice nepatrí do príslušného poľa $\mathbb{Z}_p$, nahraďte ho jeho zvyškom po delení číslom p):

(a) $\begin{pmatrix} 1 & \lambda & 0 \\ 0 & 1 & \lambda \\ \lambda & 0 & 1 \end{pmatrix}$; (b) $\begin{pmatrix} 1-\lambda & \lambda & 0 \\ \lambda & 2-\lambda & 0 \\ \lambda^3 & 3-\lambda^2 & \lambda \end{pmatrix}$; (c) $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & \lambda & 1 & \lambda \\ 0 & 1 & \lambda & 1 \\ \lambda & 1 & 0 & 1 \end{pmatrix}$; (d) $\begin{pmatrix} \lambda-1 & 1 & 0 & 4 \\ -2 & \lambda+1 & 5 & 3 \\ 0 & 0 & \lambda-1 & 0 \\ 0 & 0 & 5 & \lambda \end{pmatrix}$.

10.8. Pre ľubovoľnú maticu $\boldsymbol{A}$ rozmeru $n \times n$ nad poľom K a skalár $c \in K$ platí $\det(c\boldsymbol{A}) = c^n \det \boldsymbol{A}$, špeciálne $\det(-\boldsymbol{A}) = (-1)^n \det \boldsymbol{A}$. Dokážte.

10.9. Nech $n \geq 2$ a $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ sú štvorcové matice rádu n nad poľom K. Na príklade ukážte, že (okrem prípadu $\boldsymbol{B} = \boldsymbol{0}$ alebo $\boldsymbol{C} = \boldsymbol{0}$) pre z nich zloženú blokovú maticu neplatí nasledujúce „zovšeobecnenie“ pravidla na výpočet determinantu matíc rozmeru 2×2:

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = |\boldsymbol{A}|\,|\boldsymbol{D}| - |\boldsymbol{B}|\,|\boldsymbol{C}|.$$

10.10. Pomocou algebraických doplnkov nájdite explicitné vyjadrenie pre inverzné matice k nasledujúcim maticiam a sformulujte nutné a postačujúce podmienky ich existencie:

(a) $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$; (b) $\begin{pmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{pmatrix}$; (c) $\begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}$; (d) $\begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$; (e) $\begin{pmatrix} 0 & 0 & a \\ 0 & b & 0 \\ c & 0 & 0 \end{pmatrix}$.

Zovšeobecnite výsledky úloh (d), (e) na matice ľubovoľného rádu n.

10.11. Nájdite riešenia nasledujúcich sústav lineárnych rovníc nad poľom $\mathbb{R}$ pomocou Cramerovho pravidla:

(a) $x + y = 1$, $3x - 5y = 4$;

(b) $x - 4y = -2$, $7x + 2y = 1$;

(c) $x + y - z = 0$, $2x + 3z = 3$, $3x + y - 2z = -5$;

(d) $2x + y + 2z = 10$, $4x - y = -3$, $y + 2z = 0$;

(e) $x + 2y + 3z + 4u = 10$, $y + 2z + 3u = 6$, $z + 2u = 3$, $y - z + u = 1$.

10.12. Nech $\boldsymbol{A} \in K^{m \times n}$ je matica nad poľom K a $I \subseteq \{1, \ldots, m\}$, $J \subseteq \{1, \ldots, n\}$ sú dve množiny indexov s rovnakým počtom prvkov $k \leq \min(m, n)$. Označme $\boldsymbol{A}_{IJ} \in K^{k \times k}$ štvorcovú maticu rádu k tvorenú tými prvkami a_{ij} matice $\boldsymbol{A}$, pre ktoré $i \in I$ a $j \in J$. Determinanty takto vzniknutých matíc $\boldsymbol{A}_{IJ}$ nazývame *minormi matice* $\boldsymbol{A} \in K^{m \times n}$ rádu k. Jednoducho povedané, minory matice $\boldsymbol{A}$ sú determinanty štvorcových matíc, ktoré vzniknú vynechaním niektorých riadkov a stĺpcov matice $\boldsymbol{A}$. Minory prislúchajúce množinám indexov tvaru $I = J = \{1, \dots, k\}$ sa nazývajú *hlavné minory* matice $\boldsymbol{A}$.

Dokážte, že hodnosť matice $\boldsymbol{A}$ je najväčšie prirodzené číslo k, pre ktoré existuje nenulový minor rádu k matice $\boldsymbol{A}$. (*Návod*: Ukážte, že ak $\boldsymbol{r}_{i_1}(\boldsymbol{A}), \dots, \boldsymbol{r}_{i_k}(\boldsymbol{A})$ sú lineárne nezávislé riadky matice $\boldsymbol{A}$, tak existujú indexy stĺpcov $j_1, \dots, j_k$ také, že matica $\left(a_{i_p j_q}\right)_{k\times k}$ má hodnosť k.)

10.13. (a) Vypočítajte všetky minory reálnej matice $\left(\begin{smallmatrix} 1 & 1 & -1 \\ 3 & 0 & -2 \end{smallmatrix}\right)$.

(b) Vypočítajte všetky hlavné minory komplexnej matice $\left(\begin{smallmatrix} 2+\mathrm{i} & 0 & 3\mathrm{i} & 4 \\ 4\mathrm{i} & 3-2\mathrm{i} & 2 & 6-\mathrm{i} \\ 2-\mathrm{i} & 1+5\mathrm{i} & 0 & 2\mathrm{i} \end{smallmatrix}\right)$.

10.14. Nech $\boldsymbol{A} \in K^{n\times n}$ je štvorcová matica rádu n a I, J sú dve podmnožiny množiny $\{1, \dots, n\}$ s rovnakým počtom prvkov $k \le n$. *Algebraickým doplnkom minora* $|\boldsymbol{A}_{IJ}|$ matice $\boldsymbol{A}$ (pozri cvičenie 12) nazývame výraz $(-1)^{I+J}|\boldsymbol{A}_{I'J'}|$, kde I', J' označujú doplnky množín I resp. J v množine $\{1, \dots, n\}$ a $I + J = \sum_{i\in I} i + \sum_{j\in J} j$. Nech ďalej $\mathcal{P}_k(n)$ označuje množinu všetkých k-prvkových podmnožín množiny $\{1, \dots, n\}$ (zrejme $\#\,\mathcal{P}_k(n) = \binom{n}{k}$).

Dokážte nasledujúce zovšeobecnenia Laplaceovho rozvoja determinantu z vety 10.4.2:

(a) *Laplaceov rozvoj determinantu podľa vybraných riadkov.* Pre ľubovoľnú množinu $I \in \mathcal{P}_k(n)$ vybraných (indexov) riadkov matice $\boldsymbol{A}$ platí

$$|\boldsymbol{A}| = \sum_{J\in\mathcal{P}_k(n)} (-1)^{I+J}\,|\boldsymbol{A}_{IJ}|\,|\boldsymbol{A}_{I'J'}|\,.$$

(b) *Laplaceov rozvoj determinantu podľa vybraných stĺpcov.* Pre ľubovoľnú množinu $J \in \mathcal{P}_k(n)$ vybraných (indexov) stĺpcov matice $\boldsymbol{A}$ platí

$$|\boldsymbol{A}| = \sum_{I\in\mathcal{P}_k(n)} (-1)^{I+J}\,|\boldsymbol{A}_{IJ}|\,|\boldsymbol{A}_{I'J'}|\,.$$

10.15. (a) Vypočítajte algebraické doplnky minorov rádu 2 reálnej matice $\left(\begin{smallmatrix} 1 & 1 & 3 & 0 \\ -1 & 0 & -2 & 1 \\ 4 & 6 & 4 & -2 \\ 3 & 7 & 2 & 0 \end{smallmatrix}\right)$.

(b) Rozviňte determinant matice z úlohy (a) podľa prvého a druhého riadku. Výsledok prorovnajte s hodnotou determinatu získanou priamym výpočtom.

10.16. Nech K je pole, $x_0, x_1, \dots, x_n$ sú navzájom rôzne a $y_0, y_1, \dots, y_n$ sú ľubovoľné prvky z K.

(a) Dokážte, že potom existuje práve jeden polynóm $f(x) = a_0 + a_1x + \ldots + a_nx^n \in K^{(n)}[x]$ taký, že $f(x_0) = y_0$, $f(x_1) = y_1$, …, $f(x_n) = y_n$; $f(x)$ sa nazýva *interpolačný polynóm* konečnej funkcie $x_i \mapsto y_i$, kde $0 \le i \le n$. (*Návod*: Rozpíšte uvedené rovnosti do sústavy lineárnych rovníc v neznámych $a_0, a_1, \dots, a_n$ a využite Vandermondov determinant.)

(b) Odvoďte tzv. *Lagrangeov tvar interpolačného polynómu*

$$f(x) = \sum_{i=0}^{n} y_i \prod_{j\ne i} \frac{x - x_j}{x_i - x_j}\,.$$

(c) Vyjadrite koeficienty interpolačného polynómu $f(x)$ pomocou Cramerovho pravidla.

(d) Nech K je pole reálnych čísel $\mathbb{R}$. V prípade *ekvidistančného delenia* x_0, $x_1 = x_0 + d$, $x_2 = x_0 + 2d$, …, $x_n = x_0 + nd$ s krokom $d > 0$, nájdite explicitné vzorce pre koeficienty interpolačného polynómu, ak $0 \le n \le 3$.

10.17. S využitím výsledkov predchádzajúceho cvičenia riešte nasledujúce úlohy:

(a) Nájdite kvadratické polynómy $f_0(x), f_1(x), f_2(x) \in \mathbb{R}^{(2)}[x]$, ktorých grafy prechádzajú bodmi $(-1,0)$, $(1,0)$, $(3,-2)$, resp. $(-1,0)$, $(0,-1)$, $(2,5)$, resp. $(0,-1)$, $(2,5)$, $(3,-2)$.

(b) Nájdite kubické polynómy $g_1(x), g_2(x) \in \mathbb{R}^{(3)}[x]$, ktorých grafy prechádzajú bodmi $(0,-1)$, $(1,0)$, $(2,5)$, $(3,-2)$, resp. $(-1,0)$, $(0,-1)$, $(1,0)$, $(2,5)$.

(c) Nájdite bikvadratický polynóm $h(x) \in \mathbb{R}^{(4)}[x]$, ktorého graf prechádza bodmi $(-1,0)$, $(0,-1)$, $(1,0)$, $(2,5)$, $(3,-2)$.

(d) Načrtnite grafy funkcií f_0, f_1, f_2, g_1, g_2 a h na intervale $\langle -2, 4\rangle$ a porovnajte ich.

10.18. *Permanentom* štvorcovej matice $\boldsymbol{A} = (a_{ij})_{n\times n}$ nad poľom K nazývame výraz

$$\operatorname{per} \boldsymbol{A} = \sum_{\sigma \in \mathcal{S}_n} a_{\sigma(1)\,1} \ldots a_{\sigma(n)\,n}.$$

Formálne teda permanent dostaneme vynechaním znamienkových koeficientov $\operatorname{sgn}\sigma = (-1)^{|\sigma|}$ v definícii determinantu. Dokážte postupne nasledujúce tvrdenia:

(a) Ak maticu chápeme ako riadok jej stĺpcov, tak permanent je n-lineárne symetrické zobrazenie $K^{n\times n} \to K$ (pozri cvičenie 10.3).

(b) Pre ľubovoľnú maticu $\boldsymbol{A} \in K^{n\times n}$ platí $\operatorname{per} \boldsymbol{A}^{\mathsf{T}} = \operatorname{per} \boldsymbol{A}$.

(c) Ak $\operatorname{char} K = 2$, tak pre každú maticu $\boldsymbol{A} \in K^{n\times n}$ platí $\operatorname{per} \boldsymbol{A} = \det \boldsymbol{A}$.

10.19. (a) Odvoďte všeobecný vzorec na výpočet permanentu matíc rozmeru 2×2.

(b) Odvoďte vzorec analogický Sarrusovmu pravidlu na výpočet permanentu matíc rozmeru 3×3. [1]

(c) Vypočítajte permanenty reálnych matíc $\boldsymbol{A} = \begin{pmatrix} 1 & 3 \\ 2 & -6 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 2 & 4 & 0 \\ 3 & 7 & -1 \\ 1 & 1 & 1 \end{pmatrix}$.

(c) Odvoďte z výsledkov úlohy (a), že permanent regulárnej štvorcovej matice sa môže rovnať nule, kým permanent singulárnej štvorcovej matice nemusí byť 0.

10.20. Nech $X = \{x_1, \ldots, x_n\}$ je n-prvková množina a (X, H) je orientovaný graf s množinou vrcholov X (pozri cvičenie 2.7). Hovoríme, že *permutácia* $\sigma \in \mathcal{S}_n$ *žije na grafe* (X, H), ak pre každé $i \le n$ platí $(x_i, x_{\sigma(i)}) \in H$.

(a) Nech $\boldsymbol{H} \in \mathbb{R}^{n\times n}$ je incidenčná matica grafu (X, H). Potom počet všetkých permutácií $\sigma \in \mathcal{S}_n$, ktoré žijú na grafe (X, H), je práve $\operatorname{per} \boldsymbol{H}$. Dokážte.

(b) Pre každý z grafov z obrázku 2.1 vypočítajte permanent jeho incidenčnej matice a určte počet permutácií, ktoré na ňom žijú (pokúste sa riešiť obe úlohy pre daný graf v optimálnom poradí).

[1] Stojí za poznámku, že – na rozdiel od determinantu – nepoznáme nijaký jednoduchý, rýchly algoritmus na výpočet permanentu štvorcových matíc všeobecného rádu n nad poľami charakteristiky $\neq 2$. Dokonca máme dobré dôvody domnievať sa, že taký algoritmus neexistuje.

Časť II

Bilineárne formy a geometria

11. Bilineárne a kvadratické formy

V predchádzajúcej kapitole sme sa stretli s multilineárnymi zobrazeniami, ktorých príkladmi boli práve determinanty. V tejto kapitole podrobnejšie preskúmame tzv. *bilineárne zobrazenia*, ktoré sú najjednoduchším netriviálnym prípadom takýchto zobrazení. Z nich sa sústredíme hlavne na *bilineárne formy*, t. j. na bilineárne zobrazenia s hodnotami v príslušnom poli. Od bilineárnych foriem je už len krok k tzv. *kvadratickým formám*. Týmto krokom, prísne vzaté, vystúpime zo sveta objektov lineárnych (t. j. „prvého stupňa“) a dostaneme sa do sveta objektov kvadratických (t. j. „druhého stupňa“). Zatiaľ sa však sústredíme len na algebraickú stránku celej veci. Geometrické aspekty bilineárnych a kvadratických foriem preskúmame až neskôr, keď si na tento účel zadovážime aj ďalšie účinné nástroje.

V celej kapitole K označuje nejaké pevné no inak ľubovoľné pole. Ak sa o tom osobitne nezmienime, pod vektorovým priestorom budeme rozumieť vektorový priestor nad poľom K. Väčšina výsledkov tejto kapitoly však platí len za dodatočnej podmienky, že charakteristika nášho poľa K sa nerovná dvom. Na rozdiel od kapitol 8 a 10, kde nám tento predpoklad slúžil len na zjednodušenie formulácií niektorých definícií, tvrdení a dôkazov, v tejto kapitole už hrá podstatnú úlohu.

11.1 Bilineárne zobrazenia a bilineárne formy

Nech U, V, W sú vektorové priestory nad poľom K. Hovoríme, že $F : U \times V \to W$ je *bilineárne zobrazenie*, ak pre všetky $\boldsymbol{x}, \boldsymbol{x}_1, \boldsymbol{x}_2 \in U$, $\boldsymbol{y}, \boldsymbol{y}_1, \boldsymbol{y}_2 \in V$ a $c_1, c_2 \in K$ platí

$$\begin{aligned} F(\boldsymbol{x}, c_1\boldsymbol{y}_1 + c_2\boldsymbol{y}_2) &= c_1F(\boldsymbol{x}, \boldsymbol{y}_1) + c_2F(\boldsymbol{x}, \boldsymbol{y}_2), \\ F(c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2, \boldsymbol{y}) &= c_1F(\boldsymbol{x}_1, \boldsymbol{y}) + c_2F(\boldsymbol{x}_2, \boldsymbol{y}). \end{aligned}$$

Inak povedané, zobrazenie $F\colon U \times V \to W$ je bilineárne práve vtedy, keď pre ľubovoľné pevné $\boldsymbol{x} \in U$ je priradením $\boldsymbol{y} \mapsto F(\boldsymbol{x}, \boldsymbol{y})$ definované lineárne zobrazenie $V \to W$ a pre ľubovoľné pevné $\boldsymbol{y} \in V$ je priradením $\boldsymbol{x} \mapsto F(\boldsymbol{x}, \boldsymbol{y})$ definované lineárne zobrazenie $U \to W$.

11.1.1. Príklad. (a) Príklad 6.1.4 vlastne hovorí, že pre pevné m, n, p je predpisom $F(\boldsymbol{A}, \boldsymbol{B}) = \boldsymbol{A} \cdot \boldsymbol{B}$ dané bilineárne zobrazenie $K^{m\times n} \times K^{n\times p} \to K^{m\times p}$. Teda násobenie je bilineárne zobrazenie medzi vektorovými priestormi matíc príslušných rozmerov.

(b) Pre ľubovoľné vektorové priestory U, V je predpisom $F(\varphi, \boldsymbol{x}) = \varphi(\boldsymbol{x})$ definované bilineárne zobrazenie $\mathcal{L}(V, U) \times V \to U$. To znamená, že na dosadenie argumentu do funkcie sa možno dívať ako na bilineárne zobrazenie na príslušných vektorových priestoroch (pozri paragraf 6.5). Najdôležitejším prípadom takéhoto bilineárneho zobrazenia je tzv. *párovanie* alebo *dualita* $(\eta, \boldsymbol{x}) \mapsto \eta(\boldsymbol{x}) : V^* \times V \to K$ medzi priestorom V a jeho duálom V^*.

(c) Pre ľubovoľné vektorové priestory U, V, W je predpisom $F(\varphi, \psi) = \varphi \circ \psi$ definované bilineárne zobrazenie $\mathcal{L}(V, U) \times \mathcal{L}(W, V) \to \mathcal{L}(W, U)$. Teda kompozíciu možno chápať ako bilineárne zobrazenie na uvedených vektorových priestoroch lineárnych zobrazení.

Pod *bilineárnou formou* na vektorových priestoroch U, V rozumieme ľubovoľné bilineárne zobrazenie $F \colon U \times V \to K$.

Nech U, V sú konečnorozmerné vektorové priestory s bázami $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$ resp. $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$. Potom pre ľubovoľnú maticu $\boldsymbol{A} = (a_{ij}) \in K^{m \times n}$ je predpisom

$$F(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot (\boldsymbol{y})_{\boldsymbol{\beta}} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_i y_j,$$

kde $(\boldsymbol{x})_{\boldsymbol{\alpha}} = (x_1, \ldots, x_m)^{\mathsf{T}}$, $(\boldsymbol{y})_{\boldsymbol{\beta}} = (y_1, \ldots, y_n)^{\mathsf{T}}$ sú súradnice vektorov $\boldsymbol{x} \in U$, $\boldsymbol{y} \in V$ v príslušných bázach, definovaná bilineárna forma $F \colon U \times V \to K$ (presvedčte sa o tom). Ukážeme, že tiež naopak, každá bilineárna forma na vektorových priestoroch U, V má uvedený tvar pre jednoznačne určenú maticu $\boldsymbol{A} \in K^{m \times n}$.

Maticou bilineárnej formy $F \colon U \times V \to K$ vzhľadom na bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ nazývame maticu

$$[F]_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \big(F(\boldsymbol{u}_i, \boldsymbol{v}_j)\big) \in K^{m \times n},$$

ktorá je tvorená hodnotami formy F na dvojiciach vektorov báz $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$.

11.1.2. Tvrdenie. *Nech U, V sú konečnorozmerné vektorové priestory s bázami $\boldsymbol{\alpha}$, resp. $\boldsymbol{\beta}$ a $F \colon U \times V \to K$ je bilineárna forma. Potom pre všetky $\boldsymbol{x} \in U$, $\boldsymbol{y} \in V$ platí*

$$F(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot [F]_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot (\boldsymbol{y})_{\boldsymbol{\beta}}$$

a $\boldsymbol{A} = [F]_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ je jediná matica s touto vlastnosťou.

Dôkaz. Nech $(\boldsymbol{x})_{\boldsymbol{\alpha}} = (x_1, \ldots, x_m)^{\mathsf{T}}$, $(\boldsymbol{y})_{\boldsymbol{\beta}} = (y_1, \ldots, y_n)^{\mathsf{T}}$ sú súradnice vektorov $\boldsymbol{x} \in U$, $\boldsymbol{y} \in V$ v príslušných bázach $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$, $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$. S použitím bilinearity F dostávame

$$\begin{aligned} F(\boldsymbol{x}, \boldsymbol{y}) &= F(x_1 \boldsymbol{u}_1 + \ldots + x_m \boldsymbol{u}_m,\ y_1 \boldsymbol{v}_1 + \ldots + y_n \boldsymbol{v}_n) \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j F(\boldsymbol{u}_i, \boldsymbol{v}_j) = (\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot [F]_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot (\boldsymbol{y})_{\boldsymbol{\beta}}. \end{aligned}$$

Zostáva ukázať, že pre ľubovoľnú maticu $\boldsymbol{A} = (a_{ij}) \in K^{m\times n}$ platí

$$\big(\forall \boldsymbol{x} \in U\big)\big(\forall \boldsymbol{y} \in V\big)\big(F(\boldsymbol{x},\boldsymbol{y}) = (\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot (\boldsymbol{y})_{\boldsymbol{\beta}}\big) \ \Rightarrow\ \boldsymbol{A} = [F]_{\boldsymbol{\alpha},\boldsymbol{\beta}}.$$

Za uvedeného predpokladu voľbou $\boldsymbol{x} = \boldsymbol{u}_i$, $\boldsymbol{y} = \boldsymbol{v}_j$ dostávame

$$F(\boldsymbol{u}_i, \boldsymbol{v}_j) = (u_i)_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot (\boldsymbol{v}_j)_{\boldsymbol{\beta}} = \boldsymbol{e}_i^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \boldsymbol{e}_j = a_{ij}.$$

Teda špeciálne každá bilineárna forma $F\colon K^m \times K^n \to K$ na (stĺpcových) vektorových priestoroch K^m, K^n má tvar

$$F(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{x}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \boldsymbol{y} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}x_iy_j,$$

kde $\boldsymbol{A} = [F]_{\boldsymbol{\varepsilon}^{(m)},\boldsymbol{\varepsilon}^{(n)}}$ je matica formy F vzhľadom na kanonické bázy $\boldsymbol{\varepsilon}^{(m)}$, $\boldsymbol{\varepsilon}^{(n)}$. Teraz preskúmame ako závisí matica bilineárnej formy $F\colon U \times V \to K$ na bázach priestorov U, V, presnejšie, ako sa mení v závislosti od zmien týchto báz.

11.1.3. Tvrdenie. *Nech V_1, V_2 sú konečnorozmerné vektorové priestory nad poľom K, $\boldsymbol{\alpha}_1$, $\boldsymbol{\beta}_1$ sú dve bázy priestoru V_1, $\boldsymbol{\alpha}_2$, $\boldsymbol{\beta}_2$ sú dve bázy priestoru V_2 a $F\colon V_1 \times V_2 \to K$ je bilineárna forma. Potom*

$$[F]_{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2} = (\boldsymbol{P}_{\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1})^{\mathsf{T}} \cdot [F]_{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2} \cdot \boldsymbol{P}_{\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2}.$$

Dôkaz. Označme $\boldsymbol{A} = [F]_{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2}$, $\boldsymbol{B} = [F]_{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2}$ matice formy F v príslušných bázach. Pre ľubovoľné vektory $\boldsymbol{x} \in V_1$, $\boldsymbol{y} \in V_2$ platí

$$\begin{aligned}
(\boldsymbol{x})_{\boldsymbol{\beta}_1}^{\mathsf{T}} \cdot \boldsymbol{B} \cdot (\boldsymbol{y})_{\boldsymbol{\beta}_2} = F(\boldsymbol{x},\boldsymbol{y}) &= (\boldsymbol{x})_{\boldsymbol{\alpha}_1}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot (\boldsymbol{y})_{\boldsymbol{\alpha}_2} \\
&= \big(\boldsymbol{P}_{\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1} \cdot (\boldsymbol{x})_{\boldsymbol{\beta}_1}\big)^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \big(\boldsymbol{P}_{\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2} \cdot (\boldsymbol{y})_{\boldsymbol{\beta}_2}\big) \\
&= (\boldsymbol{x})_{\boldsymbol{\beta}_1}^{\mathsf{T}} \cdot \big((\boldsymbol{P}_{\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1})^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \boldsymbol{P}_{\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2}\big) \cdot (\boldsymbol{y})_{\boldsymbol{\beta}_2}.
\end{aligned}$$

Požadovaná rovnosť

$$\boldsymbol{B} = (\boldsymbol{P}_{\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1})^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \boldsymbol{P}_{\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2}$$

vyplýva z jednoznačnosti matice bilineárnej formy F vzhľadom na bázy $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ dokázanej v tvrdení 11.1.2.

11.1.4. Príklad. Nech $F : K^m \times K^n \to K$ je bilineárna forma a $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ sú nejaké bázy priestorov K^m, resp. K^n. Označme $\boldsymbol{A} = [F]_{\boldsymbol{\alpha},\boldsymbol{\beta}}$, $\boldsymbol{M} = [F]_{\boldsymbol{\varepsilon}^{(m)},\boldsymbol{\varepsilon}^{(n)}}$

matice formy F vzhľadom na bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, resp. vzhľadom na kanonické bázy $\boldsymbol{\varepsilon}^{(m)}$, $\boldsymbol{\varepsilon}^{(n)}$. Podľa práve dokázaného tvrdenia platia rovnosti

$$\boldsymbol{A} = \left(\boldsymbol{P}_{\boldsymbol{\varepsilon}^{(m)},\boldsymbol{\alpha}}\right)^{\mathsf{T}} \cdot \boldsymbol{M} \cdot \boldsymbol{P}_{\boldsymbol{\varepsilon}^{(n)},\boldsymbol{\beta}} = \boldsymbol{\alpha}^{\mathsf{T}} \cdot \boldsymbol{M} \cdot \boldsymbol{\beta},$$

$$\boldsymbol{M} = \left(\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\varepsilon}^{(m)}}\right)^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\varepsilon}^{(n)}} = \left(\boldsymbol{\alpha}^{-1}\right)^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \boldsymbol{\beta}^{-1},$$

umožňujúce (po stotožnení každej bázy s maticou tvorenou jej stĺpcami) priamy výpočet jednej z matíc $\boldsymbol{A}$, $\boldsymbol{M}$ na základe znalosti príslušných báz a druhej z nich.

Tvrdenie 11.1.3 a príklad 11.1.4 nás priamo nabádajú k porovnaniu s vetou 7.5.1 a príkladom 7.5.2. Analógia s lineárnymi zobrazeniami a ich maticami však siaha ešte ďalej a zahŕňa aj vety 7.5.3 a 7.5.4.

11.1.5. Tvrdenie. *Nech U je m-rozmerný a V je n-rozmerný vektorový priestor nad poľom K. Potom pre ľubovoľné matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{m\times n}$ nasledujúce podmienky sú ekvivalentné:*

(i) $\boldsymbol{A}$, $\boldsymbol{B}$ sú maticami tej istej bilineárnej formy $F\colon U \times V \to K$ vzhľadom na nejaké dve (možno no nie nutne rôzne) dvojice báz priestorov U, V;

(ii) existujú regulárne matice $\boldsymbol{P} \in K^{m\times m}$, $\boldsymbol{Q} \in K^{n\times n}$ také, že $\boldsymbol{B} = \boldsymbol{P}\cdot\boldsymbol{A}\cdot\boldsymbol{Q}$;

(iii) $h(\boldsymbol{A}) = h(\boldsymbol{B})$.

Dôkaz. Podľa tvrdenia 7.2.4 je štvorcová matica regulárna práve vtedy, keď k nej transponovaná matica je regulárna. Ekvivalencia (i) $\Leftrightarrow$ (ii) je tak priamym dôsledkom tvrdena 11.1.3 a vety 7.5.3. Ekvivalencia (ii) $\Leftrightarrow$ (iii) je už obsiahnutá vo vete 7.5.3.

Na základe uvedeného tvrdenia možno korektne definovať *hodnosť* $h(F)$ *bilineárnej formy* F na konečnorozmerných vektorových priestoroch ako hodnosť jej matice vzhľadom na ľubovoľnú dvojicu báz. Je totiž zrejmé, že táto hodnota na voľbe príslušných báz nezávisí.

11.1.6. Dôsledok. *Pre každú bilineárnu formu $F\colon U\times V \to K$ na konečnorozmerných vektorových priestoroch nad poľom K možno zvoliť bázu $\boldsymbol{\alpha}$ priestoru U a bázu $\boldsymbol{\beta}$ priestoru V tak, že F má vzhľadom na bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ maticu v blokovom tvare*

$$[F]_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \begin{pmatrix} \boldsymbol{I}_h & \boldsymbol{0}_{h,n-h} \\ \boldsymbol{0}_{m-h,h} & \boldsymbol{0}_{m-h,n-h} \end{pmatrix},$$

kde $m = \dim U$, $n = \dim V$ a $h = h(F)$.

Otázkou, ako možno k danej bilineárnej forme F nájsť také bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, sa tu nebudeme zaoberať. Pre zvedavého čitateľa podávame stručný návod v cvičeniach. Za istých okolností sa s ňou však ešte stretneme. No v tejto chvíli obrátime svoju pozornosť trochu iným smerom.

Ak $F\colon U \times V \to K$ je ľubovoľná bilineárna forma, tak pre každé $\boldsymbol{x} \in U$ je predpisom $\varphi_{\boldsymbol{x}}(\boldsymbol{y}) = F(\boldsymbol{x}, \boldsymbol{y})$ definovaný lineárny funkcionál $\varphi_{\boldsymbol{x}} : V \to K$, t. j. prvok duálu $V^* = \mathcal{L}(V, K)$ vektorového priestoru V (pozri paragraf 6.5). Ak položíme $F^\times(\boldsymbol{x}) = \varphi_{\boldsymbol{x}}$, je tým definované zobrazenie $F^\times : U \to V^*$. Bude nás zaujímať, pre aké F má *každý* lineárny funkcionál $\varphi \in V^*$ tvar $\varphi = F^\times(\boldsymbol{x})$ pre nejaké, prípadne pre jediné $\boldsymbol{x} \in U$.

11.1.7. Veta. *(a) Nech U, V sú vektorové priestory a $F : U \times V \to K$ je bilineárna forma. Potom $F^\times : U \to V^*$ je lineárne zobrazenie.*

(b) Ak U, V sú konečnorozmerné, tak $h(F^\times) = h(F)$. V dôsledku toho $F^\times$ je injektívne práve vtedy, keď $h(F) = \dim U$, a $F^\times$ je surjektívne práve vtedy, keď $h(F) = \dim V$.

Dôkaz. S využitím linearity F v prvej premennej možno podmienku (a) overiť priamym výpočtom, ktorý prenechávame ako cvičenie čitateľovi.

(b) Položme $\dim V = n$, zvoľme nejaké bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ priestorov U resp. V a označme $\boldsymbol{A} = [F]_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ maticu formy F vzhľadom na ne. Potom

$$\operatorname{Ker} F^\times = \big\{\boldsymbol{x} \in U;\ (\forall \boldsymbol{y} \in V)((\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot (\boldsymbol{y})_{\boldsymbol{\beta}} = 0\big\} = \{\boldsymbol{x} \in U;\ (\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot \boldsymbol{A} = \boldsymbol{0}\},$$

lebo pre ľubovoľné $\boldsymbol{x} \in U$ je predpisom $\boldsymbol{y} \mapsto (\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot (\boldsymbol{y})_{\boldsymbol{\beta}}$ určené lineárne zobrazenie $V \to K$, ktoré má vzhľadom na bázy $\boldsymbol{\beta}$ vo V a (1) v K maticu $(\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot \boldsymbol{A}$, a matica lineárneho zobrazenia v daných bázach je určená jednoznačne. Keďže $(\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot \boldsymbol{A} = \big(\boldsymbol{A}^{\mathsf{T}} \cdot (\boldsymbol{x})_{\boldsymbol{\alpha}}\big)^{\mathsf{T}}$ a $\boldsymbol{x} \mapsto (\boldsymbol{x})_{\boldsymbol{\alpha}}$ je lineárny izomorfizmus $U \to K^n$, z tvrdenia 9.2.2 vyplýva

$$\dim \operatorname{Ker} F^\times = \dim \mathcal{R}\big(\boldsymbol{A}^{\mathsf{T}}\big) = n - h(\boldsymbol{A}) = n - h(F).$$

Podľa vety 6.2.3 o dimenzii jadra a obrazu z toho dostávame

$$h(F^\times) = \dim \operatorname{Im} F^\times = n - \dim \operatorname{Ker} F^\times = h(F).$$

Zvyšok je už triviálnym dôsledkom tejto rovnosti, vety 6.2.2 a tvrdenia 6.5.3

11.1.8. Dôsledok. *Nech U, V sú konečnorozmerné vektorové priestory rovnakej dimenzie. Potom pre ľubovoľnú bilineárnu formu $F : U \times V \to K$ nasledujúce podmienky sú ekvivalentné:*

(i) $F^\times : U \to V^*$ je lineárny izomorfizmus;

(ii) $h(F) = \dim U = \dim V$;

(iii) pre ľubovoľné (pre nejaké) bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ priestorov U resp. V je $[F]_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ regulárna matica.

Bilineárna forma $F : V \times V \to K$ sa nazýva *regulárna*, ak spĺňa niektorú (teda všetky) z podmienok posledného dôsledku; v opačnom prípade sa F nazýva *singulárna*.

11.2 Symetrické bilineárne formy a kvadratické formy

V tomto paragrafe sa budeme zaoberať výlučne bilineárnymi formami, v ktorých prvá aj druhá premenná prebieha ten istý vektorový priestor V, t. j. bilineárnymi formami tvaru $F\colon V \times V \to K$. Budeme ich nazývať *bilineárnymi formami na vektorovom priestore* V.

Bilineárna forma $F : V^2 \to K$ sa nazýva *symetrická*, ak pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí

$$F(\boldsymbol{x}, \boldsymbol{y}) = F(\boldsymbol{y}, \boldsymbol{x}).$$

Pre istotu ešte pripomíname, že bilineárna forma $F : V^2 \to K$ sa nazýva *antisymetrická* (pozri paragraf 10.1), ak pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí

$$F(\boldsymbol{x}, \boldsymbol{y}) = -F(\boldsymbol{y}, \boldsymbol{x}).$$

11.2.1. Tvrdenie. *Nech F je bilineárna forma na vektorovom priestore V nad poľom K a* $\operatorname{char} K \neq 2$. *Potom F možno rozložiť na súčet*

$$F = F_0 + F_1$$

pre jednoznačne určené bilineárne formy F_0, F_1 na V, pričom F_0 je symetrická a F_1 je antisymetrická.

Dôkaz Keďže $\operatorname{char} K \neq 2$, v K platí $2 = 1 + 1 \neq 0$, teda existuje prvok $\frac{1}{2} = 2^{-1} \in K$. Pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ položme

$$F_0(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2}\big(F(\boldsymbol{x}, \boldsymbol{y}) + F(\boldsymbol{y}, \boldsymbol{x})\big),$$
$$F_1(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2}\big(F(\boldsymbol{x}, \boldsymbol{y}) - F(\boldsymbol{y}, \boldsymbol{x})\big).$$

Jednoduchými priamymi výpočtami, ktoré prenechávame čitateľovi, možno overiť, že F_0 aj F_1 sú bilineárne formy na V. Na prvý pohľad vidno, že F_0 je symetrická, F_1 je antisymetrická a pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí

$$F(\boldsymbol{x}, \boldsymbol{y}) = F_0(\boldsymbol{x}, \boldsymbol{y}) + F_1(\boldsymbol{x}, \boldsymbol{y}).$$

Zostáva overiť jednoznačnosť. Uvedený tvar foriem F_0, F_1 je však bezprostredným dôsledkom poslednej rovnosti a rovnosti

$$F(\boldsymbol{y}, \boldsymbol{x}) = F_0(\boldsymbol{x}, \boldsymbol{y}) - F_1(\boldsymbol{x}, \boldsymbol{y}),$$

ktorá z nej vyplýva na základe symetrie formy F_0 a antisymetrie formy F_1.

Ak F je bilineárna forma na konečnorozmernom vektorovom priestore V s bázou $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$, tak pod *maticou formy* F vzhľadom na túto bázu budeme rozumieť jej maticu vzhľadom na dvojicu báz $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}$; značíme ju $[F]_{\boldsymbol{\alpha}}$. Teda

$$[F]_{\boldsymbol{\alpha}} = [F]_{\boldsymbol{\alpha},\boldsymbol{\alpha}} = \big(F(\boldsymbol{u}_i, \boldsymbol{u}_j)\big)_{n\times n}.$$

Toto obmedzenie v porovnaní so všeobecnou definíciou z predošlého paragrafu je prirodzené – naopak, značne umelo by pôsobilo vyjadrovať súradnice prvej a druhej premennej v F, hoci ležia v tom istom priestore V, vzhľadom na rôzne bázy. Matice bilineárnych foriem na vektorovom priestore V vzhľadom na dvojice rôznych báz $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ vo V preto odteraz vylúčime z ďalších úvah. Jedna z výhod takéhoto prístupu spočíva v nasledujúcom zrejmom tvrdení, ktoré by však bez spomínaného obmedzenia neplatilo.

Pripomeňme, že štvorcová matica $\boldsymbol{A} = (a_{ij})_{n\times n}$ sa nazýva *symetrická*, ak $\boldsymbol{A} = \boldsymbol{A}^{\mathsf{T}}$, t. j. ak pre všetky $i, j \le n$ platí $a_{ij} = a_{ji}$; podobne, $\boldsymbol{A}$ sa nazýva *antisymetrická*, ak $\boldsymbol{A} = -\boldsymbol{A}^{\mathsf{T}}$, t. j. ak pre všetky $i, j \le n$ platí $a_{ij} = -a_{ji}$.

11.2.2. Tvrdenie. *Nech $\boldsymbol{\alpha}$ je ľubovoľná báza konečnorozmerného vektorového priestoru V a $F : V^2 \to K$ je bilineárna forma na V. Potom*

(a) F je symetrická práve vtedy, keď jej matica $[F]_{\boldsymbol{\alpha}}$ je symetrická;

(b) F je antisymetrická práve vtedy, keď jej matica $[F]_{\boldsymbol{\alpha}}$ je antisymetrická.

Násobenie v poli možno chápať ako bilineárnu formu $F : K^2 \to K$, kde $F(a, b) = ab$ pre $a, b \in K$. Stotožnením prvej a druhej premennej dostávame zobrazenie $q \colon K \to K$, kde $q(a) = F(a, a) = a^2$, t. j. „a-kvadrát“. Zovšeobecnením tohto postupu dospejeme k pojmu kvadratickej formy.

Zobrazenie $q : V \to K$ vektorového priestoru V do poľa K sa nazýva *kvadratická forma* na V, ak existuje bilineárna forma $F : V^2 \to K$ taká, že pre všetky $\boldsymbol{x} \in V$ platí

$$q(\boldsymbol{x}) = F(\boldsymbol{x}, \boldsymbol{x}).$$

Hovoríme tiež, že bilineárna forma F *indukuje* kvadratickú formu q.

Vo všeobecnosti existuje k danej kvadratickej forme $q : V \to K$ mnoho bilineárnych foriem $F : V^2 \to K$, pre ktoré platí uvedená rovnosť. Ak je totiž $F : V^2 \to K$ nejaká bilineárna forma a $G : V^2 \to K$ je ľubovoľná antisymetrická bilineárna forma, tak, aspoň pokiaľ char $K \ne 2$,

$$F(\boldsymbol{x}, \boldsymbol{x}) = F(\boldsymbol{x}, \boldsymbol{x}) + G(\boldsymbol{x}, \boldsymbol{x}),$$

lebo zrejme $G(\boldsymbol{x}, \boldsymbol{x}) = 0$. Špeciálne v označení tvrdenia 11.2.1 platí

$$F(\boldsymbol{x}, \boldsymbol{x}) = F_0(\boldsymbol{x}, \boldsymbol{x}).$$

To nás privádza na myšlienku pokúsiť sa odstrániť spomínanú nejednoznačnosť dodatočnou požiadavkou symetrie príslušnej bilineárnej formy.

Polárnou formou kvadratickej formy $q : V \to K$ nazývame *symetrickú* bilineárnu formu $F : V^2 \to K$, ktorá indukuje formu q.

11.2.3. Tvrdenie. *Nech q je kvadratická forma na vektorovom priestore V nad poľom K, pričom* $\operatorname{char} K \neq 2$. *Potom existuje jediná symetrická bilineárna forma $F : V^2 \to K$ taká, že*

$$q(\boldsymbol{x}) = F(\boldsymbol{x}, \boldsymbol{x})$$

pre všetky $\boldsymbol{x} \in V$.

Dôkaz. Ak F je ľubovoľná bilineárna forma na V, ktorá indukuje q, tak z tvrdenia 11.2.1 a pred chvíľou učinenej poznámky vyplýva, že hľadaná polárna forma ku q je daná vzťahom $F_0(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2}(F(\boldsymbol{x}, \boldsymbol{y}) + F(\boldsymbol{y}, \boldsymbol{x}))$.

Požadovaná jednoznačnosť polárnej formy je bezprostredným dôsledkom nasledujúceho tvrdenia, ktoré nám poskytuje ďalšiu cennú informáciu o vzťahu medzi ňou a indukovanou kvadratickou formou.

11.2.4. Tvrdenie. *Nech* $\operatorname{char} K \neq 2$, *$q : V \to K$ je kvadratická forma a $F \colon V^2 \to K$ je jej polárna forma. Potom pre ľubovoľné $\boldsymbol{x}, \boldsymbol{y} \in V$ platí*

$$\begin{aligned} F(\boldsymbol{x}, \boldsymbol{y}) &= \frac{1}{2}\big(q(\boldsymbol{x}+\boldsymbol{y}) - q(\boldsymbol{x}) - q(\boldsymbol{y})\big) \\ &= \frac{1}{2}\big(q(\boldsymbol{x}) + q(\boldsymbol{y}) - q(\boldsymbol{x}-\boldsymbol{y})\big) \\ &= \frac{1}{4}\big(q(\boldsymbol{x}+\boldsymbol{y}) - q(\boldsymbol{x}-\boldsymbol{y})\big) \end{aligned}$$

Dôkaz. Najprv si uvedomme, že $\operatorname{char} K \neq 2$ má za dôsledok $4 = 2 \cdot 2 \neq 0$. Každú z rovností

$$\begin{aligned} F(\boldsymbol{x}, \boldsymbol{y}) &= \frac{1}{2}\big(F(\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}) - F(\boldsymbol{x}, \boldsymbol{x}) - F(\boldsymbol{y}, \boldsymbol{y})\big) \\ &= \frac{1}{2}\big(F(\boldsymbol{x}, \boldsymbol{x}) + F(\boldsymbol{y}, \boldsymbol{y}) - F(\boldsymbol{x}-\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y})\big) \\ &= \frac{1}{4}\big(F(\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}) - F(\boldsymbol{x}-\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y})\big) \end{aligned}$$

možno teraz na základe bilinearity formy F overiť priamym výpočtom, ktorý prenechávame čitateľovi.

Uvedené rovnosti nám tak dávajú k dispozícii hneď tri rôzne formuly, z ktorých každá umožňuje zrekonštruovať polárnu formu na základe znalosti jej kvadratickej formy. Ešte si všimnite, že tieto rovnosti sú len jednoduchým zovšeobecnením známych vzorcov

$$ab = \frac{1}{2}\big((a+b)^2 - a^2 - b^2\big) = \frac{1}{2}\big(a^2 + b^2 - (a-b)^2\big) = \frac{1}{4}\big((a+b)^2 - (a-b)^2\big)$$

platných pre ľubovoľné $a, b \in K$.

Maticou kvadratickej formy $q : V \to K$ na konečnorozmernom vektorovom priestore V nad poľom charakteristiky rôznej od dvoch vzhľadom na bázu $\boldsymbol{\alpha}$ nazývame maticu jej polárnej formy vzhľadom na túto bázu a značíme ju $[q]_{\boldsymbol{\alpha}}$. Matica $[q]_{\boldsymbol{\alpha}}$ je touto požiadavkou jednoznačne určená a je to vždy symetrická matica. *Hodnosťou kvadratickej formy* potom nazývame hodnosť jej matice vzhľadom na akúkoľvek bázu a značíme ju $h(q)$. Zrejme hodnosť $h(q)$ nezávisí od voľby bázy a rovná sa hodnosti $h(F)$ príslušnej polárnej formy. Kvadratická forma sa nazýva *regulárna*, resp. *singulárna*, ak má príslušnú vlastnosť jej polárna forma.

Poznámka. Upozorňujeme, že ani jeden z výsledkov uvádzaných v tvrdeniach 11.2.1 a 11.2.3 (t. j. ani existencia ani jednoznačnosť) nie je splnený vo vektorových priestoroch nad poľom charakteristiky 2. Dokonca jednu a tú istú symetrickú bilineárnu, resp. kvadratickú formu možno zadať maticami rôznej hodnosti. Príklady možno nájsť v cvičeniach. V prípade tvrdenia 11.2.4 nedávajú uvedené vzorce nad poľom K charakteristiky 2 vôbec žiadny zmysel.

11.3 Diagonalizácia kvadratických foriem

Ak $F : V^2 \to K$ je ľubovoľná bilineárna forma na konečnorozmernom vektorovom priestore V a $\boldsymbol{A} = (a_{ij})_{n\times n}$ je jej matica vzhľadom na nejakú bázu $\boldsymbol{\alpha}$ priestoru V, tak pre indukovanú kvadratickú formu $q : V \to K$ a všetky $\boldsymbol{x} \in V$ platí

$$\begin{aligned} q(\boldsymbol{x}) = F(\boldsymbol{x}, \boldsymbol{x}) = (\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot (\boldsymbol{x})_{\boldsymbol{\alpha}} &= \sum_{i,j=1}^{n} a_{ij} x_i x_j \\ &= \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1\le i<j\le n} (a_{ij} + a_{ji})\, x_i x_j, \end{aligned}$$

kde $(\boldsymbol{x})_{\boldsymbol{\alpha}} = (x_1, \ldots, x_n)^{\mathsf{T}}$ sú príslušné súradnice. Ak F je navyše symetrická, tak $\boldsymbol{A}$ je priamo (symetrická) matica formy q v báze $\boldsymbol{\alpha}$ a uvedený výraz prejde na tvar

$$q(\boldsymbol{x}) = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{1\le i<j\le n} a_{ij} x_i x_j.$$

V tomto paragrafe si ukážeme, že voľbou vhodnej bázy, t. j. zavedením „nových súradníc“, sa možno zbaviť všetkých sčítancov $a_{ij} x_i x_j$ obsahujúcich zmiešané členy a upraviť tak celý výraz na *diagonálny tvar*

$$q(\boldsymbol{x}) = \sum_{i=1}^{n} a_{ii} x_i^2.$$

Príslušný výpočet možno uskutočniť pomocou tzv. *Lagrangeovej metódy*, ktorá používa dva typy úprav: jednak metódu *doplnenia na štvorec*, známu už zo strednej školy, jednak substitúciu

$$x_i x_j = \frac{1}{4}(x_i + x_j)^2 - \frac{1}{4}(x_i - x_j)^2,$$

ku ktorej sa musíme uchýliť zakaždým, keď pre $i \neq j$ je $a_{ij} \neq 0$, ale $a_{ii} = a_{jj} = 0$. Ako naznačuje posledná identita, Lagrangeova metóda funguje len pre polia charakteristiky rôznej od dvoch. Bez ďalšieho komentára si celý postup ozrejmíme na príklade.

11.3.1. Príklad. Kvadratická forma $q : \mathbb{R}^4 \to \mathbb{R}$ je pre $\boldsymbol{x} = (x_1, x_2, x_3, x_4)^{\mathsf{T}} \in \mathbb{R}^4$ daná predpisom

$$q(\boldsymbol{x}) = -2x_2^2 + x_1x_2 + 2x_2x_3 - 3x_3x_4 = \boldsymbol{x}^{\mathsf{T}} \cdot \begin{pmatrix} 0 & 1/2 & 0 & 0 \\ 1/2 & -2 & 1 & 0 \\ 0 & 1 & 0 & -3/2 \\ 0 & 0 & -3/2 & 0 \end{pmatrix} \cdot \boldsymbol{x}\,.$$

Všetky zmiešané členy obsahujúce x_2 pripojíme k členu $-2x_2^2$ a doplníme na štvorec:

$$\begin{aligned} q(\boldsymbol{x}) &= -2\left(x_2^2 - \frac{1}{2}x_1x_2 - x_2x_3\right) - 3x_3x_4 \\ &= -2\left(\left(x_2 - \frac{1}{4}x_1 - \frac{1}{2}x_3\right)^2 - \frac{1}{16}x_1^2 - \frac{1}{4}x_3^2 - \frac{1}{4}x_1x_3\right) - 3x_3x_4 \\ &= -\frac{1}{8}(x_1 - 4x_2 + 2x_3)^2 + \frac{1}{8}x_1^2 + \frac{1}{2}x_3^2 + \frac{1}{2}x_1x_3 - 3x_3x_4\,. \end{aligned}$$

Teraz pripojíme zmiešané členy obsahujúce x_1 k členu $\frac{1}{8}x_1^2$ a doplníme na štvorec. Dostaneme

$$\begin{aligned} q(\boldsymbol{x}) &= -\frac{1}{8}(x_1 - 4x_2 + 2x_3)^2 + \frac{1}{8}(x_1^2 + 4x_3^2 + 4x_1x_3) - 3x_3x_4 \\ &= -\frac{1}{8}(x_1 - 4x_2 + 2x_3)^2 + \frac{1}{8}(x_1 + 2x_3)^2 - 3x_3x_4\,. \end{aligned}$$

Po použití spomínanej substitúcie konečne máme

$$q(\boldsymbol{x}) = -\frac{1}{8}(x_1 - 4x_2 + 2x_3)^2 + \frac{1}{8}(x_1 + 2x_3)^2 - \frac{3}{4}(x_3 + x_4)^2 + \frac{3}{4}(x_3 - x_4)^2.$$

Ak si na $\mathbb{R}^4$ zavedieme nové súradnice $\boldsymbol{y} = (y_1, y_2, y_3, y_4)^{\mathsf{T}}$, kde

$$\begin{aligned} y_1 &= x_1 - 4x_2 + 2x_3 \\ y_2 &= x_1 + 2x_3 \\ y_3 &= x_3 + x_4 \\ y_4 &= x_3 - x_4, \end{aligned}$$

tak pôvodná kvadratická forma q v nich nadobudne diagonálny tvar

$$q(\boldsymbol{x}) = q'(\boldsymbol{y}) = -\frac{1}{8}y_1^2 + \frac{1}{8}y_2^2 - \frac{3}{4}y_3^2 + \frac{3}{4}y_4^2 .$$

Pritom matica

$$\boldsymbol{Q} = \begin{pmatrix} 1 & -4 & 2 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix},$$

ktorá zabezpečuje zmenu súradníc $\boldsymbol{y} = \boldsymbol{Q}\cdot\boldsymbol{x}$ je maticou prechodu z kanonickej bázy $\boldsymbol{\varepsilon}$ do takej bázy $\boldsymbol{\alpha}$ priestoru $\mathbb{R}^4$, v ktorej pre všetky $\boldsymbol{x} \in \mathbb{R}^4$ platí $(\boldsymbol{x})_{\boldsymbol{\alpha}} = \boldsymbol{y} = \boldsymbol{Q}\cdot\boldsymbol{x}$. To znamená, že $\boldsymbol{Q} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\varepsilon}} = \boldsymbol{\alpha}^{-1}$. Teda $\boldsymbol{\alpha} = \boldsymbol{Q}^{-1}$, presnejšie, hľadaná báza $\boldsymbol{\alpha}$ je tvorená stĺpcami matice

$$\boldsymbol{Q}^{-1} = \begin{pmatrix} 0 & 1 & -1 & -1 \\ -1/4 & 1/4 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & -1/2 \end{pmatrix}.$$

11.3.2. Príklad. Ešte si všimnime veľkú mieru nejednoznačnosti diagonálneho tvaru a príslušnej transformácie súradníc. Napr. kvadratickú formu

$$q(x,y) = 10x^2 + 5y^2 - 2xy$$

na $\mathbb{R}^2$ možno pri troche postrehu jednoducho upraviť na diagonálny tvar

$$q(x,y) = (x^2 + 4y^2 + 4xy) + (9x^2 + y^2 - 6xy) = (x+2y)^2 + (3x-y)^2,$$

zatiaľ čo pri dôslednom dodržiavaní predchádzajúceho receptu dostaneme

$$\begin{aligned} q(\boldsymbol{x}) &= 10\left(x^2 - \frac{1}{5}xy\right) + 5y^2 \\ &= 10\left(\left(x - \frac{1}{10}y\right)^2 - \frac{1}{100}y^2\right) + 5y^2 \\ &= \frac{1}{10}(10x-y)^2 + \frac{49}{10}y^2. \end{aligned}$$

Podobne, ako sme sa pri riešení sústav lineárnych rovníc vyhli manipulácii s premennými a celú úlohu sme previedli na úpravu istej matice pomocou elementárnych riadkových operácii, aj teraz bude našim cieľom upraviť danú kvadratickú či symetrickú bilineárnu formu na diagonálny tvar a zároveň nájsť príslušnú bázu len vhodnými vhodnými jej matice. Najprv si však ujasníme, ako vplýva zmena bázy na maticu bilineárnej resp. kvadratickej formy na konečnorozmerom priestore V. Ako špeciálny prípad tvrdenia 11.1.3 okamžite dostávame

11.3.3. Tvrdenie. *Nech $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ sú dve bázy konečnorozmerného vektorového priestoru V a $F : V^2 \to K$ je bilineárna forma na V. Potom pre matice $\boldsymbol{A} = [F]_{\boldsymbol{\alpha}}$, $\boldsymbol{B} = [F]_{\boldsymbol{\beta}}$ formy F v týchto bázach platí*

$$\boldsymbol{B} = (\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}})^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}.$$

Štvorcové matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{n\times n}$ sa nazývajú *kongruentné*, označenie $\boldsymbol{A} \equiv \boldsymbol{B}$, ak existuje regulárna matica $\boldsymbol{P} \in K^{n\times n}$ taká, že

$$\boldsymbol{B} = \boldsymbol{P}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \boldsymbol{P}.$$

Zrejme kongruentné matice majú rovnakú hodnosť. Čitateľ by si mal taktiež sám overiť, že pre ľubovoľné matice $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C} \in K^{n\times n}$ platí

$$\begin{gathered} \boldsymbol{A} \equiv \boldsymbol{A}, \\ \boldsymbol{A} \equiv \boldsymbol{B} \;\Rightarrow\; \boldsymbol{B} \equiv \boldsymbol{A}, \\ \boldsymbol{A} \equiv \boldsymbol{B} \;\&\; \boldsymbol{B} \equiv \boldsymbol{C} \;\Rightarrow\; \boldsymbol{A} \equiv \boldsymbol{C}. \end{gathered}$$

Inak povedané, vzťah kongruencie je *reflexívny*, *symetrický* a *tranzitívny*, t. j. je vzťahom *ekvivalencie* na množine $K^{n\times n}$.

Ďalej nás bude zaujímať len kongruencia symetrických matíc; v celej všeobecnosti sa preto týmto vzťahom viac zaoberať nebudeme.

Jednoducho možno nahliadnuť (prípadne overiť) platnosť implikácie

$$\boldsymbol{A} \equiv \boldsymbol{B} \;\&\; \boldsymbol{B} = \boldsymbol{B}^{\mathsf{T}} \;\Rightarrow\; \boldsymbol{A} = \boldsymbol{A}^{\mathsf{T}},$$

t. j. ak je jedna z dvoch kongruentných matíc symetrická, tak je symetrická aj druhá z nich.

Keďže každá regulárna matica jej maticou prechodu medzi vhodnou dvojicou báz, tvrdenie 11.3.3 má za bezprostredný dôsledok nasledujúcu vetu.

11.3.4. Veta. *Nech V je n-rozmerný vektorový priestor nad poľom K charakteristiky $\neq 2$. Potom pre ľubovoľné symetrické matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{n\times n}$ nasledujúce podmienky sú ekvivalentné:*

(i) *$\boldsymbol{A}$, $\boldsymbol{B}$ sú maticami tej istej symetrickej bilineárnej formy $F\colon V \times V \to K$ vzhľadom na nejaké dve (možno no nie nutne rôzne) bázy priestoru V;*

(ii) *$\boldsymbol{A}$, $\boldsymbol{B}$ sú maticami tej istej kvadratickej formy $q\colon V \to K$ vzhľadom na nejaké dve (možno no nie nutne rôzne) bázy priestoru V;*

(iii) $\boldsymbol{A} \equiv \boldsymbol{B}$.

Poznamenajme, že ekvivalencia (i) $\Leftrightarrow$ (iii) platí aj nad poľami charakteristiky 2.

Konečne môžeme systematicky pristúpiť k otázke diagonalizácie symetrických bilineárnych resp. kvadratických foriem. Obe zredukujeme na úpravy príslušnej symetrickej matice na diagonálnu maticu pomocou úprav zachovávajúcich kongruenciu.

Najprv si všimnime, čo to vlastne znamená, že matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{n\times n}$ sú kongruentné prostredníctvom regulárnej matice $\boldsymbol{P} \in K^{n\times n}$. Maticu $\boldsymbol{P}$ možno upraviť na jednotkovú maticu $\boldsymbol{I}_n$ pomocou elementárnych *stĺpcových* úprav, ktorým zodpovedá rozklad matice $\boldsymbol{P}$ na súčin elementárnych matíc

$$\boldsymbol{P} = \boldsymbol{E}_1 \cdot \ldots \cdot \boldsymbol{E}_k.$$

Potom

$$\begin{aligned}\boldsymbol{B} = \boldsymbol{P}^\mathsf{T} \cdot \boldsymbol{A} \cdot \boldsymbol{P} &= (\boldsymbol{E}_1 \cdot \ldots \cdot \boldsymbol{E}_k)^\mathsf{T} \cdot \boldsymbol{A} \cdot (\boldsymbol{E}_1 \cdot \ldots \cdot \boldsymbol{E}_k)\\ &= \boldsymbol{E}_k^\mathsf{T} \cdot \ldots \cdot \boldsymbol{E}_1^\mathsf{T} \cdot \boldsymbol{A} \cdot \boldsymbol{E}_1 \cdot \ldots \cdot \boldsymbol{E}_k\end{aligned}$$

Nech teraz $\boldsymbol{C} \in K^{n\times n}$ je ľubovoľná matica a $\boldsymbol{E} \in K^{n\times n}$ je elementárna matica, zodpovedajúca nejakej ESO. Uvedomme si, v akom vzťahu je matica $\boldsymbol{E}^\mathsf{T} \cdot \boldsymbol{C} \cdot \boldsymbol{E}$ k pôvodnej matici $\boldsymbol{C}$. Samozrejme, že sú kongruentné, no nielen to. Matica $\boldsymbol{E}^\mathsf{T}$ totiž zodpovedá „rovnakej“ ERO, ako bola ESO reprezentovaná maticou $\boldsymbol{E}$. Presnejšie,

(a) ak $\boldsymbol{E}$ zodpovedala výmene i-teho a j-teho stĺpca, tak $\boldsymbol{E}^\mathsf{T}$ zodpovedá výmene i-teho a j-teho riadku;

(b) ak $\boldsymbol{E}$ zodpovedala vynásobeniu i-teho stĺpca nenulovým skalárom $c \in K$, tak $\boldsymbol{E}^\mathsf{T}$ zodpovedá vynásobeniu i-teho riadku tým istým skalárom c;

(c) ak $\boldsymbol{E}$ zodpovedala pripočítaniu c-násobku i-teho stĺpca k j-temu stĺpcu, tak $\boldsymbol{E}^\mathsf{T}$ zodpovedá pripočítaniu c-násobku i-teho riadku k j-temu riadku.

Matica $\boldsymbol{E}^\mathsf{T} \cdot \boldsymbol{C} \cdot \boldsymbol{E}$ teda vznikne z matice $\boldsymbol{C}$ pomocou dvojice symetricky združených elementárnych úprav – jednej ESO a jednej ERO. Navyše, keďže vďaka asociatívnosti násobenia platí $\boldsymbol{E}^\mathsf{T} \cdot (\boldsymbol{C} \cdot \boldsymbol{E}) = (\boldsymbol{E}^\mathsf{T} \cdot \boldsymbol{C}) \cdot \boldsymbol{E}$, je jedno, v akom poradí obe úpravy vykonáme.

Úprava symetrickej matice $\boldsymbol{A} \in K^{n\times n}$ na s ňou kongruentnú diagonálnu maticu $\boldsymbol{B} \in K^{n\times n}$ sa teda realizuje v konečnej postupnosti krokov, z ktorých každý pozostáva z jednej ESO a s ňou symetricky združenej ERO. Postupnosťou len príslušných ESO vykonaných na jednotkovej matici $\boldsymbol{I}_n$ získame tiež príslušnú maticu prechodu $\boldsymbol{P}$ takú, že platí $\boldsymbol{B} = \boldsymbol{P}^\mathsf{T} \cdot \boldsymbol{A} \cdot \boldsymbol{P}$. Celý postup možno stručne zachytiť pomocou schémy

$$\boldsymbol{A} \xrightarrow[\text{ERO}]{\text{ESO}} \boldsymbol{B}$$

$$\boldsymbol{I}_n \xrightarrow{\text{ESO}} \boldsymbol{P}$$

Ak teda $\boldsymbol{A}$ bola maticou nejakej symetrickej bilineárnej formy, prípadne kvadratickej formy v báze $\boldsymbol{\alpha}$, tak $\boldsymbol{B}$ je maticou tejto formy v báze $\boldsymbol{\beta} = \boldsymbol{\alpha} \cdot \boldsymbol{P}$ (t. j.

$\boldsymbol{P} = \boldsymbol{P}_{\boldsymbol{\alpha\beta}}$). Špeciálne, ak $\boldsymbol{A}$ je maticou príslušnej formy na (stĺpcovom) vektorovom priestore K^n v kanonickej báze $\boldsymbol{\alpha} = \boldsymbol{\varepsilon}^{(n)}$, tak $\boldsymbol{\beta} = \boldsymbol{P}$, t.j. vektory novej bázy $\boldsymbol{\beta}$ sú priamo stĺpce matice $\boldsymbol{P}$.

Zostáva dokázať, že uvedený postup vždy vedie k cieľu.

11.3.5. Veta. *Nech* char $K \neq 2$ *a* V *je konečnorozmerný vektorový priestor nad poľom* K. *Potom*

(a) *každá symetrická matica* $\boldsymbol{A} \in K^{n\times n}$ *je kongruentná s nejakou diagonálnou maticou;*

(b) *každá symetrická bilineárna forma* $F : V^2 \to K$ *má vo vhodnej báze priestoru* V *diagonálnu maticu;*

(c) *každá kvadratická forma* $q : V \to K$ *má vo vhodnej báze priestoru* V *diagonálnu maticu.*

Dôkaz. Stačí dokázať len tvrdenie (a), tvrdenia (b), (c) sú už jeho bezprostrednými dôsledkami. Popíšeme algoritmus, ktorý nám hovorí, aké ESO a ERO treba postupne aplikovať na maticu $\boldsymbol{A}$. V celom postupe sa používajú dva typy úprav:

(1) Nech $i \leq n$ je najmenší index taký, že $a_{ii} \neq 0$. Potom postupne pre každé $j \leq n$ také, že $j \neq i$ a $a_{ij} = a_{ji} \neq 0$, pripočítame k j-temu stĺpcu matice $\left(-\frac{a_{ij}}{a_{ii}}\right)$-násobok i-teho stĺpca a v takto získanej matici pripočítame $\left(-\frac{a_{ji}}{a_{ii}}\right)$-násobok i-teho riadku k j-temu riadku. Inak povedané, pomocou diagonálneho prvku $a_{ii} \neq 0$ vynulujeme všetky ostatné nenulové prvky i-teho riadku aj stĺpca.

(2) Nech pre každé $i \leq n$ platí: ak $a_{ii} \neq 0$, tak $a_{ij} = a_{ji} = 0$ pre každé $j \neq i$. Nech $k \leq n$ je najmenší index taký, že $a_{kk} = 0$, ale k-ty riadok nie je identicky nulový. Nech ďalej $j \leq n$ je najmenší index taký, že $a_{kj} = a_{jk} \neq 0$. Potom ku k-temu stĺpcu matice pripočítame jej j-ty stĺpec a ku k-temu riadku takto získanej matice pripočítame jej j-ty riadok. (Výsledná matica má na mieste (k,k) prvok $a_{kj} + a_{jk} = 2a_{jk} \neq 0$.)

Úpravy typu (1) majú prednosť, t.j. vykonávame ich tak dlho, ako je to len možné alebo kým nezískame diagonálnu maticu. V opačnom prípade aplikujeme jednu úpravu typu (2). Po nej nikdy nedostaneme diagonálnu maticu a vždy možno aplikovať úpravu typu (1). Takto postupujeme, kým sa to len dá. Ak už nemožno aplikovať žiadnu z úprav (1), (2), znamená to, že sme dospeli k diagonálnej matici. Keďže po každej úprave typu (1) pribudnú aspoň dva nulové prvky mimo diagonály a prípadné nenulové prvky mimo diagonály, ktoré pribudli v dôsledku nej alebo jednej úpravy typu (2), budú vynulované ďalšími úpravami typu (1), celý postup nevyhnutne skončí po konečnom počte krokov.

Okrem úprav (1), (2) možno použiť ďalšie dva typy úprav, bez ktorých sa síce možno zaobísť, no s ich pomocou možno docieliť „krajší“ tvar diagonálnej matice formy prípadne matice prechodu.

(3) Výmena i-teho a j-teho stĺpca matice a vzápätí aj jej i-teho a j-teho riadku.

(4) Vynásobenie i-teho stĺpca a vzápätí aj i-teho riadku matice ľubovoľným nenulovým skalárom $c \in K$ (diagonálny prvok a_{ii} sa tým zmení na $c^2 a_{ii}$).

Význam úprav (3) a (4) je jasný: (3) zodpovedá zámene poradia súradníc $x_i \leftrightarrow x_j$ a (4) substitúcii x_i/c miesto x_i. Čitateľ by si mal sám premyslieť, ako zodpovedá úprava typu (1) doplneniu na štvorec a úprava typu (2) (spolu s následnou úpravou typu (1)) substitúcii súčtu štvorcov miesto $x_i x_j$. Aby sme mu uľahčli premýšľanie, upravíme práve opísanou metódou maticu kvadratickej formy z príkladu 11.3.1 na s ňou kongruentný diagonálny tvar a zároveň nájdeme príslušnú bázu.

11.3.6. Príklad. Budeme upravovať symetrickú maticu

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1/2 & 0 & 0 \\ 1/2 & -2 & 1 & 0 \\ 0 & 1 & 0 & -3/2 \\ 0 & 0 & -3/2 & 0 \end{pmatrix} \in \mathbb{R}^{4\times 4}$$

na s ňou kongruentný diagonálny tvar. Začneme úpravami typu (1). Najprv vynulujeme pomocou prvku $a_{22} = -2$ zvyšné nenulové prvky druhého stĺpca a riadku. Za tým účelom pripočítame $\frac{1}{4}$-násobok druhého stĺpca k prvému stĺpcu a vzápätí vykonáme rovnakú operáciu s riadkami. Potom pripočítame $\frac{1}{2}$-násobok druhého stĺpca k tretiemu stĺpcu a to isté urobíme s riadkami. Postupne dostaneme

$$\boldsymbol{A} \equiv \begin{pmatrix} 1/8 & 0 & 1/4 & 0 \\ 0 & -2 & 1 & 0 \\ 1/4 & 1 & 0 & -3/2 \\ 0 & 0 & -3/2 & 0 \end{pmatrix} \equiv \begin{pmatrix} 1/8 & 0 & 1/4 & 0 \\ 0 & -2 & 0 & 0 \\ 1/4 & 0 & 1/2 & -3/2 \\ 0 & 0 & -3/2 & 0 \end{pmatrix}.$$

Vykonaním príslušných ESO na jednotkovej matici dostaneme

$$\boldsymbol{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \wr \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/4 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \wr \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/4 & 1 & 1/2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

pričom znakom $\wr$ označujeme *stĺpcovú* ekvivalenciu matíc. Ďalej vynulujeme pomocou prvku $a_{11} = \frac{1}{8}$ zvyšné nenulové prvky prvého stĺpca a riadku, t. j.

odpočítame dvojnásobok prvého stĺpca od tretieho a to isté urobíme s riadkami. Príslušnú stĺpcovú operáciu vykonáme aj na matici získanej z $\boldsymbol{I}_4$. Vyjdú nám matice

$$\boldsymbol{A} \equiv \begin{pmatrix} 1/8 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & -3/2 \\ 0 & 0 & -3/2 & 0 \end{pmatrix} \quad \text{a} \quad \boldsymbol{I}_4 \wr \begin{pmatrix} 1 & 0 & -2 & 0 \\ 1/4 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Keďže úpravu typu (1) nemožno viac aplikovať, použijeme úpravu typu (2). Pripočítame štvrtý stĺpec k tretiemu a to isté aj pre riadky. Potom už opäť môžeme použiť úpravu typu (1). Odpočítame $\frac{1}{2}$-násobok tretieho stĺpca od štvrtého a to isté urobíme s riadkami. Po vykonaní príslušných ESO na matici získanej z $\boldsymbol{I}_4$ dostaneme výsledok

$$\boldsymbol{A} \equiv \begin{pmatrix} 1/8 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -3 & -3/2 \\ 0 & 0 & -3/2 & 0 \end{pmatrix} \equiv \begin{pmatrix} 1/8 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 3/4 \end{pmatrix} = \boldsymbol{B},$$

$$\boldsymbol{I}_4 \wr \begin{pmatrix} 1 & 0 & -2 & 0 \\ 1/4 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \wr \begin{pmatrix} 1 & 0 & -2 & 1 \\ 1/4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1/2 \\ 0 & 0 & 1 & 1/2 \end{pmatrix} = \boldsymbol{P}.$$

Vidíme, že napriek úpornej snahe pridŕžať sa maticových úprav zodpovedajúcich úpravam kvadratickej formy z príkladu 11.3.1, nám obe matice vyšli mierne odlišné. Úplnú zhodu možno dosiahnuť práve „kozmetickými“ úpravami (3) a (4). Najprv prehodíme prvý a druhý prvok na diagonále. Tomu zodpovedá výmena prvého a druhého stĺpca v matici prechodu $\boldsymbol{P}$. Potom vynásobíme prvý diagonálny prvok skalárom $\frac{1}{16} = \left(-\frac{1}{4}\right)^2$, tretí skalárom $\frac{1}{4} = \left(\frac{1}{2}\right)^2$ a štvrtý skalárom $1 = (-1)^2$. Tomu zodpovedá vynásobenie prvého stĺpca príslušnej matice prechodu skalárom $-\frac{1}{4}$, tretieho skalárom $\frac{1}{2}$ a štvrtého skalárom -1. Teda

$$\boldsymbol{A} \equiv \boldsymbol{B} \equiv \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 1/8 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 3/4 \end{pmatrix} \equiv \begin{pmatrix} -1/8 & 0 & 0 & 0 \\ 0 & 1/8 & 0 & 0 \\ 0 & 0 & -3/4 & 0 \\ 0 & 0 & 0 & 3/4 \end{pmatrix},$$

$$\boldsymbol{I}_4 \wr \boldsymbol{P} \wr \begin{pmatrix} 0 & 1 & -2 & 1 \\ 1 & 1/4 & 0 & 0 \\ 0 & 0 & 1 & -1/2 \\ 0 & 0 & 1 & 1/2 \end{pmatrix} \wr \begin{pmatrix} 0 & 1 & -1 & -1 \\ -1/4 & 1/4 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & -1/2 \end{pmatrix} = \boldsymbol{Q}^{-1}.$$

„Našinec“ by asi dal prednosť tvarom bez zlomkov a „s menej mínusmi“

$$\boldsymbol{A} \equiv \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \qquad \text{a} \qquad \boldsymbol{I}_4 \wr \begin{pmatrix} 0 & 4 & -2 & 2 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Rozmyslite si, akými úpravami ich možno získať. Keďže však inverznou maticou k poslednej matici prechodu je matica

$$\begin{pmatrix} -1/4 & 1 & -1/2 & 0 \\ 1/4 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & -1/2 & 1/2 \end{pmatrix},$$

cenou za „pekný“ tvar diagonálnej matice formy a matice prechodu (t. j. príslušnej bázy) sú zlomky a „viac mínusov“ v matici transformácie súradníc, čo sa prejaví vo vnútri zátvoriek v algebraickom vyjadrení pôvodnej formy:

$$\begin{aligned} q(\boldsymbol{x}) &= -2x_2^2 + x_1x_2 + 2x_2x_3 - 3x_3x_4 \\ &= -\frac{1}{8}(x_1 - 4x_2 + 2x_3)^2 + \frac{1}{8}(x_1 + 2x_3)^2 - \frac{3}{4}(x_3 + x_4)^2 + \frac{3}{4}(x_3 - x_4)^2 \\ &= -2\left(-\frac{1}{4}x_1 + x_2 - \frac{1}{2}x_3\right)^2 + 2\left(\frac{1}{4}x_1 + \frac{1}{2}x_3\right)^2 \\ &\quad - 3\left(\frac{1}{2}x_3 + \frac{1}{2}x_4\right)^2 + 3\left(-\frac{1}{2}x_3 + \frac{1}{2}x_4\right)^2. \end{aligned}$$

Je už len otázkou vkusu, čomu dáme prednosť.

Pre poučenie si všimnite, že v prípade, ak je matica formy v diagonálnom tvare, možno stĺpce matice prechodu „beztrestne“ násobiť skalárom -1. Vynásobenie skalárom $(-1)^2 = 1$ sa totiž na diagonále neprejaví.

Celkom na záver ešte poznamenajme, že každá kvadratická forma na *riadkovom* vektorovom priestore K^n má tvar

$$q(\boldsymbol{x}) = \boldsymbol{x} \cdot \boldsymbol{A} \cdot \boldsymbol{x}^{\mathsf{T}}$$

pre nejakú maticu $\boldsymbol{A} \in K^{n \times n}$. Diagonalizácia (symetrickej matice) takejto formy (ak char $K \neq 2$) funguje rovnako ako v stĺpcovom priestore K^n, s jediným rozdielom, že príslušnú bázu, vzhľadom na ktorú má q diagonálnu maticu, dostaneme pomocou zodpovedajúcich *riadkových* úprav na jednotkovej matici; presnejšie, táto báza je tvorená *riadkami* takto získanej výslednej matice prechodu.

Cvičenia

11.1. (a) Predpokladajme, že bilineárna forma $F : K^m \times K^n \to K$ má vzhľadom na kanonické bázy $\boldsymbol{\varepsilon}^{(m)}$, $\boldsymbol{\varepsilon}^{(n)}$ maticu $\boldsymbol{A} \in K^{m\times n}$. Navrhnite a vysvetlite postup podľa schémy

$$\boldsymbol{A} \xrightarrow[\text{ESO}]{\text{ERO}} \boldsymbol{D}$$
$$\boldsymbol{I}_m \xrightarrow{\text{ERO}} \boldsymbol{P}$$
$$\boldsymbol{I}_n \xrightarrow[\text{ESO}]{} \boldsymbol{Q}$$

tak, aby ste ním získali bázy $\boldsymbol{\alpha} = \boldsymbol{P}^{\mathsf{T}}$, $\boldsymbol{\beta} = \boldsymbol{Q}$ priestorov K^m, resp. K^n, vzhľadom na ktoré má matica F blokovo diagonálny tvar $\boldsymbol{D} = \operatorname{diag}(\boldsymbol{I}_h, \boldsymbol{0})$, kde $h = h(F)$.

(b) Upravte podľa navrhnutého postupu matice $\boldsymbol{A} = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 0 & -2 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 1 & 4 & 0 \\ 3 & 2 & -2 \\ -2 & 2 & 2 \end{pmatrix}$ a nájdite príslušné bázy.

11.2. Dokážte časť (a) vety 11.1.7.

11.3. Nech $F : U \times V \to K$ je bilineárna forma na vektorových priestoroch U a V. Potom množiny

$$\operatorname{LR}(F) = \big\{\boldsymbol{u} \in U;\ (\forall \boldsymbol{v} \in V)(F(\boldsymbol{u},\boldsymbol{v}) = 0)\big\},$$
$$\operatorname{PR}(F) = \big\{\boldsymbol{v} \in V;\ (\forall \boldsymbol{u} \in U)(F(\boldsymbol{u},\boldsymbol{v}) = 0)\big\}$$

nazývame *ľavým* resp. *pravým radikálom* formy F.

(a) Dokážte, že $\operatorname{LR}(F)$, $\operatorname{PR}(F)$ sú lineárne podpriestory priestorov U resp. V.

(b) Ako súvisí $\operatorname{LR}(F)$ s lineárnym zobrazením $F^{\times}$ z vety 11.1.7? Nájdite lineárne zobrazenie, ktoré rovnakým spôsobom súvisí s $\operatorname{PR}(F)$.

(c) Dokážte, že pre dimenzie uvedených radikálov platí: $\dim \operatorname{LR}(F) = \dim U - h(F)$, $\dim \operatorname{PR}(F) = \dim V - h(F)$.

(d) S využitím (c) zdôvodnite platnosť časti (b) vety 11.1.7 a dôsledku 11.1.8.

11.4. Nech $F\colon V \times V \to K$ je regulárna bilineárna forma na konečnorozmernom vektorovom priestore V. Podľa dôsledku 11.1.8 má každý lineárny funkcionál $\varphi : V \to K$ tvar $\varphi(\boldsymbol{x}) = F(\boldsymbol{u}, \boldsymbol{x})$ pre jednoznačne určený vektor $\boldsymbol{u} \in V$. Ako súvisí matica $(\varphi)_{1,\boldsymbol{\alpha}}$ funkcionálu φ v báze $\boldsymbol{\alpha}$ priestoru V s maticou $\boldsymbol{A} = [F]_{\boldsymbol{\alpha}}$ a súradnicami $(\boldsymbol{u})_{\boldsymbol{\alpha}}$? Riešte rovnakú úlohu pre reprezentáciu lineárnych funkcionálov $\varphi : V \to K$ v tvare $\varphi(\boldsymbol{x}) = F(\boldsymbol{x}, \boldsymbol{v})$ pre $\boldsymbol{v} \in V$.

11.5. Nech F je symetrická bilineárna forma na vektorovom priestore V a q je ňou indukovaná kvadratická forma. Potom pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí $F(\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y}) = q(\boldsymbol{x}) - q(\boldsymbol{y})$. Dokážte. Ktorý elementárny vzorec zovšeobecňuje uvedená identita?

11.6. Napíšte maticu každej z kvadratických foriem

(a) $q(x, y, z) = x^2 + y^2 - 3z^2 + 4xy - 2xz + 6yz$ pre $(x, y, z) \in \mathbb{R}^3$,

(b) $r(x, y, z) = -2x^2 + (1/2)y^2 - 4xz + 2yz$ pre $(x, y, z) \in \mathbb{R}^3$,

(c) $s(\boldsymbol{x}) = 2x_1x_2 - 2x_1x_3 + 4x_2x_4$ pre $\boldsymbol{x} = (x_1, x_2, x_3, x_4)^{\mathsf{T}} \in \mathbb{R}^4$.

Nájdite diagonálny tvar každej z nich a príslušnú bázu (maticu prechodu) jednak doplnením na úplné štvorce, jednak úpravou jej matice. Výsledky zakaždým porovnajte a zdôvodnite prípadné rozdiely.

11.7. Napíšte polárnu formu kvadratickej formy $q(x,y,z) = x^2+2y^2+3z^2+4xy+5xz+6yz$

(a) nad poľom $\mathbb{R}$;

(b) nad poľom $\mathbb{Z}_7$.

V oboch prípadoch ju upravte na diagonálny tvar a nájdite príslušnú bázu.

11.8. Presvedčte sa, že symetrické matice $\boldsymbol{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ a $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ nad poľom $\mathbb{Z}_2$ indukujú tú istú kvadratickú formu na priestore $\mathbb{Z}_2^2$. Nájdite ďalšie príklady podobnej nejednoznačnosti.

11.9. Dokážte, že kvadratickú formu $q(x,y) = xy$ na vektorovom priestore $\mathbb{Z}_2^2$ nemožno indukovať nijakou symetrickou bilineárnou formou $F : \mathbb{Z}_2^2 \times \mathbb{Z}_2^2 \to \mathbb{Z}_2$.

11.10. Teraz zovšeobecníme príklady z cvičení 11.8 a 11.9. Nech V je konečnorozmerný vektorový priestor nad poľom K charakteristiky 2 a $F : V^2 \to K$ je bilineárna forma. Dokážte nasledujúce tvrdenia:

(a) F je symetrická práve vtedy, keď F je antisymetrická.

(b) Ak F je symetrická, tak kvadratická forma $q(\boldsymbol{x}) = F(\boldsymbol{x},\boldsymbol{x})$ má diagonálny tvar (t. j. neobsahuje zmiešané členy) pri vyjadrení súradníc vektora $\boldsymbol{x} \in V$ vzhľadom na ľubovoľnú bázu priestoru V.

(c) Na základe (b) zdôvodnite, že ak kvadratická forma $q : V \to K$ má v niektorej báze priestoru V tvar, v ktorom sa vyskytujú zmiešané členy, tak q nie je indukovaná žiadnou symetrickou bilineárnou formou. Uveďte príklady takých kvadratických foriem q.

11.11. Funkciu $\Gamma\colon \mathbb{R}^4 \to \mathbb{R}$, danú predpisom $\Gamma(x_1,x_2,y_1,y_2) = x_1y_1 - x_1y_2 + 2x_2y_2$, možno chápať dvoma spôsobmi:

(1) ako kvadratickú formu $\Gamma\colon \mathbb{R}^4 \to \mathbb{R}$ takú, že $\Gamma(\boldsymbol{a}) = \Gamma(x_1,x_2,y_1,y_2)$ pre $\boldsymbol{a} = (x_1,x_2,y_1,y_2) \in \mathbb{R}^4$;

(2) ako bilineárnu formu $\Gamma\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ takú, že $\Gamma(\boldsymbol{x},\boldsymbol{y}) = \Gamma(x_1,x_2,y_1,y_2)$ pre $\boldsymbol{x} = (x_1,x_2), \boldsymbol{y} = (y_1,y_2) \in \mathbb{R}^2$.

(a) V prípade (1) napíšte polárnu formu $F(\boldsymbol{a},\boldsymbol{b})$ kvadratickej formy Γ pre $\boldsymbol{a} = (x_1,x_2,y_1,y_2)$, $\boldsymbol{b} = (u_1,u_2,v_1,v_2) \in \mathbb{R}^4$.

(b) V prípade (2) napíšte kvadratickú formu $q(\boldsymbol{x}) = \Gamma(\boldsymbol{x},\boldsymbol{x})$ indukovanú bilineárnou formou Γ a polárnu formu F kvadratickej formy q. Porovnajte F a Γ.

(c) Ktoré kvadratické formy $\Gamma\colon K^{2n} \to K$ možno takýmto spôsobom považovať zároveň za bilineárne formy $\Gamma\colon K^n \times K^n \to K$? Za akých podmienok je bilineárna forma Γ symetrická resp. antisymetrická?

12. Bilineárne a kvadratické formy nad poľom $\mathbb{R}$

V tejto kapitole budeme pokračovať v štúdiu bilineárnych a kvadratických foriem. Obmedzíme sa však na bilineárne a kvadratické formy na vektorových priestoroch nad poľom reálnych čísel. Tento zvláštny prípad je zároveň najdôležitejší z hľadiska aplikácií lineárnej algebry v geometrii a matematickej analýze. Ako príklad toho si v poslednom paragrafe predvedieme využitie reálnych kvadratických foriem pri hľadaní a klasifikácii extrémov a sedlových bodov funkcií viac premenných.

Pole $\mathbb{R}$ má charakteristiku ∞, teda $\operatorname{char}\mathbb{R} \neq 2$. To nám umožňuje plne využiť všetky výsledky predchádzajúcej kapitoly. Množina reálnych čísel je však popri štruktúre poľa vybavená aj reláciou usporiadania, ktorá je vhodne zladená so sčítaním a násobením na $\mathbb{R}$. Navyše, pre $a \in \mathbb{R}$ platí $a \geq 0$ práve vtedy, keď existuje nejaké $b \in \mathbb{R}$ také, že $a = b^2$. Práve táto vlastnosť nám umožní hlbšie objasniť a jemnejšie klasifikovať štruktúru bilineárnych a kvadratických foriem nad $\mathbb{R}$.

12.1 Signatúra

Nech $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ je symetrická matica. Podľa vety 11.3.5 $\boldsymbol{A}$ je kongruentná s nejakou diagonálnou maticou $\boldsymbol{B} \in \mathbb{R}^{n\times n}$. Potom $\boldsymbol{A}$ a $\boldsymbol{B}$ majú rovnakú hodnosť $h(\boldsymbol{A}) = h(\boldsymbol{B})$, ktorá sa zrejme rovná počtu nenulových prvkov na diagonále matice $\boldsymbol{B}$. Popri hodnosti sú však i počty kladných a záporných prvkov na diagonále matice $\boldsymbol{B}$ invariantmi, spoločnými pre navzájom kongruentné matice.

Pre ľubovoľnú *diagonálnu* maticu $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ položme

$$\sigma(\boldsymbol{A}) = (s_+, s_-, s_0),$$

kde s_+ je počet kladných, s_- počet záporných a s_0 počet nulových prvkov na diagonále matice $\boldsymbol{A}$. Usporiadanú trojicu $\sigma(\boldsymbol{A}) = (s_+, s_-, s_0)$ nazývame *signatúrou matice* $\boldsymbol{A}$.

Všimnite si, že tri zložky signatúry $\sigma(\boldsymbol{A})$ nie sú nezávislé. Platí totiž

$$s_+ + s_- = h(\boldsymbol{A}) \qquad \text{a} \qquad s_+ + s_- + s_0 = n,$$

to znamená, že pri znalosti rozmeru n a hodnosti $h(\boldsymbol{A})$ je signatúra jednoznačne určená už jedným z čísel s_+, s_-. Z tohto dôvodu niektorí autori definujú signatúru len ako počet s_+.

Teraz už môžeme sformulovať tvrdenie o invariantnosti signatúry presnejšie.

12.1.1. Veta. (Sylvestrov zákon zotrvačnosti) *Nech* $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$ *sú diagonálne matice. Potom*

$$\boldsymbol{A} \equiv \boldsymbol{B} \;\Rightarrow\; \sigma(\boldsymbol{A}) = \sigma(\boldsymbol{B}).$$

Dôkaz. Nech $\boldsymbol{A} \equiv \boldsymbol{B}$. Označme $h = h(\boldsymbol{A}) = h(\boldsymbol{B})$ hodnosť oboch matíc a k, l počty kladných diagonálnych prvkov v maticiach $\boldsymbol{A}$ resp. $\boldsymbol{B}$. Potom $\sigma(\boldsymbol{A}) = (k, h-k, n-h)$ a $\sigma(\boldsymbol{B}) = (l, h-l, n-h)$, takže stačí dokázať rovnosť $k = l$. Keďže poradie prvkov na diagonále možno prehadzovať pri zachovaní vzťahu kongruencie, môžeme si dovoliť predpokladať, že naše matice majú tvar

$$\boldsymbol{A} = \operatorname{diag}(a_1, \dots, a_k, -a_{k+1}, \dots, -a_h, 0, \dots, 0),$$

$$\boldsymbol{B} = \operatorname{diag}(b_1, \dots, b_l, -b_{l+1}, \dots, -b_h, 0, \dots, 0),$$

kde $a_i > 0$, $b_i > 0$ pre $i \leq h$. Keďže $\boldsymbol{A} \equiv \boldsymbol{B}$, existuje regulárna matica $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ taká, že $\boldsymbol{B} = \boldsymbol{P}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \boldsymbol{P}$. Jej stĺpce tvoria bázu $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ vektorového priestoru $\mathbb{R}^n$. Potom $\boldsymbol{A}$ je maticou kvadratickej formy

$$q(\boldsymbol{x}) = \boldsymbol{x}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \boldsymbol{x}$$

na $\mathbb{R}^n$ v báze $\boldsymbol{\varepsilon}$, zatiaľ čo $\boldsymbol{B}$ je jej maticou v báze $\boldsymbol{\beta}$ (pozri tvrdenie 11.3.3). Označme

$$S = [\boldsymbol{e}_1, \dots, \boldsymbol{e}_k], \qquad T = [\boldsymbol{v}_{l+1}, \dots, \boldsymbol{v}_n]$$

lineárne podpriestory v $\mathbb{R}^n$. Pre každý nenulový vektor $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \ldots + x_k\boldsymbol{e}_k \in S$ platí

$$q(\boldsymbol{x}) = \boldsymbol{x}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \boldsymbol{x} = a_1x_1^2 + \ldots + a_kx_k^2 > 0.$$

Podobne, pre každý vektor $\boldsymbol{y} = y_{l+1}\boldsymbol{v}_{l+1} + \ldots + y_n\boldsymbol{v}_n \in T$ platí

$$q(\boldsymbol{y}) = (\boldsymbol{y})_{\boldsymbol{\beta}}^{\mathsf{T}} \cdot \boldsymbol{B} \cdot (\boldsymbol{y})_{\boldsymbol{\beta}} = -b_{l+1}y_{l+1}^2 - \ldots - b_hy_h^2 \leq 0.$$

Z toho vyplýva, že $S \cap T = \{\boldsymbol{0}\}$, teda

$$\dim(S+T) = \dim S + \dim T = k + (n-l).$$

Keďže $S + T \subseteq \mathbb{R}^n$, zrejme $\dim(S+T) \leq n$. Z nerovnosti $k + n - l \leq n$ okamžite vyplýva $k \leq l$. Zo symetrie vzťahu kongruencie dostávame tiež $l \leq k$.

Práve dokázaná veta umožňuje korektne rozšíriť definíciu signatúry z diagonálnych matíc na všetky symetrické matice, a taktiež na symetrické bilineárne a kvadratické formy.

Signatúrou symetrickej matice $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, označenie $\sigma(\boldsymbol{A})$, rozumieme signatúru ľubovoľnej s ňou kongruentnej diagonálnej matice $\boldsymbol{B} \in \mathbb{R}^{n\times n}$. *Signatúrou symetrickej bilineárnej formy* $F : V^2 \to \mathbb{R}$ na konečnorozmernom reálnom vektorovom priestore V, označenie $\sigma(F)$, rozumieme signatúru jej matice vzhľadom na ľubovoľnú bázu vo V. Konečne *signatúrou kvadratickej formy* $q : V \to \mathbb{R}$ na konečnorozmernom vektorovom priestore nad $\mathbb{R}$, označenie $\sigma(q)$, rozumieme signatúru jej polárnej formy.

Všimnite si, že pre symetrickú bilneárnu aj pre kvadratickú formu sa príslušná signatúra rovná signatúre nejakej jej diagonálnej matice. Sylvestrov zákon zotrvačnosti spolu s vetou 11.3.4 nám zaručujú, že ľubovoľné dve diagonálne matice zodpovedajúce danej forme vzhľadom na rôzne bázy, v ktorých má táto forma diagonálnu maticu, majú rovnakú signatúru.

Každú reálnu symetrickú maticu $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ možno upraviť na s ňou kongruentnú diagonálnu maticu. Tú zasa možno zmenou poradia prvkov na diagonále upraviť na s ňou kongruentnú maticu tvaru

$$\operatorname{diag}(d_1, \ldots, d_k, -d_{k+1}, \ldots, -d_{k+l}, \underbrace{0, \ldots, 0}_{m \text{ krát}}),$$

kde $\sigma(\boldsymbol{A}) = (k, l, m)$ a $d_i > 0$ pre $i \le k + l$. Ak teraz pre každé $i \le k + l$ vynásobime i-ty stĺpec aj riadok skalárom $1/\sqrt{d_i}$, vyjde nám matica v blokovo diagonálnom tvare

$$\boldsymbol{A} \equiv \operatorname{diag}(\boldsymbol{I}_k, -\boldsymbol{I}_l, \mathbf{0}_{m,m}).$$

Spojením tejto úvahy so Sylvestrovým zákonom dostávame

12.1.2. Veta. *Nech $\boldsymbol{A}$, $\boldsymbol{B} \in \mathbb{R}^{n\times n}$ sú ľubovoľné symetrické matice. Potom*

$$\boldsymbol{A} \equiv \boldsymbol{B} \;\Leftrightarrow\; \sigma(\boldsymbol{A}) = \sigma(\boldsymbol{B}).$$

12.1.3. Dôsledok. *Nech V je vektorový priestor nad $\mathbb{R}$ konečnej dimenzie n. Potom*

(a) každá symetrická bilineárna forma $F : V^2 \to \mathbb{R}$ má vo vhodnej báze $\boldsymbol{\alpha}$ priestoru V blokovo diagonálnu maticu tvaru

$$[F]_{\boldsymbol{\alpha}} = \operatorname{diag}(\boldsymbol{I}_k, -\boldsymbol{I}_l, \mathbf{0}_{m,m}),$$

kde $\sigma(F) = (k, l, m)$;

(b) každá kvadratická forma $q : V \to \mathbb{R}$ má vo vhodnej báze $\boldsymbol{\alpha}$ priestoru V diagonálny tvar

$$q(\boldsymbol{x}) = x_1^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_{k+l}^2,$$

kde $\sigma(q) = (k, l, n - k - l)$ a $(\boldsymbol{x})_{\boldsymbol{\alpha}} = (x_1, \ldots, x_n)^{\mathsf{T}}$.

Pre porovnanie ešte uvedieme pár poznámok o symetrických bilineárnych a kvadratických formách na vektorových priestoroch nad poľom $\mathbb{C}$ všetkých komplexných čísel a poľom $\mathbb{Q}$ všetkých racionálnych čísel.

Keďže $\operatorname{char}\mathbb{C} = \operatorname{char}\mathbb{Q} = \infty \neq 2$, každú symetrickú maticu $\boldsymbol{A}$ typu $n \times n$ nad jedným i druhým poľom možno upraviť na s ňou kongruentnú diagonálnu maticu
$$\boldsymbol{A} \equiv \operatorname{diag}(d_1, \dots, d_h, 0, \dots, 0),$$
kde $h = h(\boldsymbol{A})$ a $d_j \neq 0$ pre $j \leq h$.

V komplexnom prípade si každý z prvkov d_j môžeme vyjadriť v goniometrickom tvare
$$d_j = r_j(\cos\alpha_j + \mathrm{i}\sin\alpha_j),$$
kde $r_j = |d_j| > 0$ a $0 \leq \alpha_j < 2\pi$. Ak pre každé $j \leq h$ vynásobíme j-ty stĺpec i riadok skalárom
$$c_j = \frac{1}{\sqrt{r_j}}\left(\cos\frac{\alpha_j}{2} - \mathrm{i}\sin\frac{\alpha_j}{2}\right),$$
pre ktorý platí $c_j^2 d_j = 1$, dostaneme
$$\boldsymbol{A} \equiv \operatorname{diag}(\boldsymbol{I}_h, \mathbf{0}_{m,m}),$$
kde $m = n - h$. Z toho vidíme, že nad poľom $\mathbb{C}$ sa nič podobné rozdeleniu nenulových prvkov na kladné a záporné nekoná – všetky nenulové prvky na diagonále sú rovnocenné a možno ich nahradiť jednotkou. Jediným invariantom, ktorý jednoznačne určuje kongruenciu symetrických matíc ako aj kanonický tvar matíc symetrických bilineárnych i kvadratických foriem nad $\mathbb{C}$, je ich hodnosť, ktorá tak plne preberá úlohu signatúry v reálnom prípade. Nasledujúce tvrdenie je upresnením a zhrnutím našich úvah.

12.1.4. Tvrdenie. *(a) Nech $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$ sú symetrické matice. Potom $\boldsymbol{A} \equiv \boldsymbol{B}$ práve vtedy, keď $h(\boldsymbol{A}) = h(\boldsymbol{B})$.*

(b) Nech V je konečnorozmerný vektorový priestor nad $\mathbb{C}$, a $F : V^2 \to \mathbb{C}$ je symetrická bilineárna forma. Potom F má vzhľadom na nejakú bázu $\boldsymbol{\alpha}$ priestoru V maticu v blokovo diagonálnom tvare $[F]_{\boldsymbol{\alpha}} = \operatorname{diag}(\boldsymbol{I}_h, \mathbf{0}_{m,m})$, kde $h = h(F)$ a $m = \dim V - h$.

Kým situácia nad $\mathbb{C}$ je podstatne jednoduchšia než nad $\mathbb{R}$ a dokázali sme ju úplne popísať, nad $\mathbb{Q}$ si tak ľahko poradiť nevieme. Základný problém tkvie v tom, že nie všetky kladné racionálne čísla majú racionálne druhé odmocniny. Tak už pre matice rozmeru 1×1 máme napr. $(2) \not\equiv (1)$ v dôsledku iracionality čísla $\sqrt{2}$. Aby to však nebolo také jednoduché, v rozmere 2×2 napr. platí
$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \not\equiv \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$
(Presvedčte sa o tom!) Systematickému štúdiu kongruencie racionálnych symetrických matíc, ktoré už nie je čiste záležitosťou lineárnej algebry ale aj teórie čísel, sa v tomto kurze viac venovať nebudeme.

12.2 Definitnosť

Nech V je vektorový priestor nad poľom $\mathbb{R}$. Kvadratická forma $q : V \to \mathbb{R}$ sa nazýva
(a) *kladne definitná*, ak $q(\boldsymbol{x}) > 0$ pre každé $\mathbf{0} \neq \boldsymbol{x} \in V$;
(b) *kladne semidefinitná*, ak $q(\boldsymbol{x}) \geq 0$ pre každé $\boldsymbol{x} \in V$;
(c) *záporne definitná*, ak $q(\boldsymbol{x}) < 0$ pre každé $\mathbf{0} \neq \boldsymbol{x} \in V$;
(d) *záporne semidefinitná*, ak $q(\boldsymbol{x}) \leq 0$ pre každé $\boldsymbol{x} \in V$;
(e) *indefinitná*, ak existujú $\boldsymbol{x}, \boldsymbol{y} \in V$ také, že $q(\boldsymbol{x}) < 0 < q(\boldsymbol{y})$.
Rovnakú klasifikáciu zavádzame aj pre symetrické bilineárne formy $F : V^2 \to \mathbb{R}$ – F má príslušnú vlastnosť definitnosti práve vtedy, keď ňou indukovaná kvadratická forma $q(\boldsymbol{x}) = F(\boldsymbol{x}, \boldsymbol{x})$ má túto vlastnosť. Podobne, symetrická matica $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ má príslušnú vlastnosť definitnosti práve vtedy, keď túto vlastnosť má kvadratická forma $q(\boldsymbol{x}) = \boldsymbol{x}^\mathsf{T} \cdot \boldsymbol{A} \cdot \boldsymbol{x}$ na priestore $\mathbb{R}^n$.

Na začiatok zaznamenáme niekoľko jednoduchých pozorovaní: (a) $\Rightarrow$ (b), (c) $\Rightarrow$ (d), no každá z podmienok (a), (c), (e) vylučuje zvyšné dve. Dokonca (e) vylučuje každú z podmienok (b), (d). Podmienky (b), (d) sa vzájomne nevylučujú, no jediná kvadratická forma, ktorá je zároveň kladne aj záporne semidefinitná, je forma identicky rovná nule na V. V dimenzii $n = 1$ je to však jediná (kladne alebo záporne) semidefinitná forma. V dimenzii $n = 1$ takisto neexistujú nijaké indefinitné formy.

Nasledujúce očividné tvrdenie poskytuje úplný popis vlastností definitnosti a regularity kvadratických foriem (a zároveň aj symetrických bilineárnych foriem a symetrických matíc) prostredníctvom ich signatúry.

12.2.1. Tvrdenie. *Nech V je n-rozmerý vektorový priestor nad poľom $\mathbb{R}$ a $q : V \to \mathbb{R}$ je kvadratická forma so signatúrou $\sigma(q) = (s_+, s_-, s_0)$. Potom*
(a) q je kladne definitná práve vtedy, keď $\sigma(q) = (n, 0, 0)$;
(b) q je kladne semidefinitná práve vtedy, keď $\sigma(q) = (h(q), 0, n - h(q))$;
(c) q je záporne definitná práve vtedy, keď $\sigma(q) = (0, n, 0)$;
(d) q je záporne semidefinitná práve vtedy, keď $\sigma(q) = (0, h(q), n - h(q))$;
(e) q je indefinitná práve vtedy, keď $s_+ \geq 1$ a $s_- \geq 1$;
(f) q je regulárna práve vtedy, keď $s_0 = 0$.

Časť (a) predchádzajúceho tvrdenia v spojení s dôsledkom 12.1.3 okamžite dáva

12.2.2. Dôsledok. *Symetrická matica $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ je kladne definitná práve vtedy, keď existuje regulárna matica $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ taká, že*

$$\boldsymbol{A} = \boldsymbol{P}^\mathsf{T} \cdot \boldsymbol{P}.$$

Tvrdenie 12.2.1 nám spolu s algoritmom z dôkazu vety 11.3.5 (prípadne Lagrangeovou metódou) dáva priamy návod na zistenie charakteru definitnosti nejakej formy či matice. Tak napríklad kvadratická forma z príkladov 11.3.1, 11.3.6 má signatúru $(2,2,0)$, teda je indefinitná. Kvadratická forma z príkladu 11.3.2 má signatúru $(2,0,0)$, čiže je kladne definitná.

Niekedy však môže byť užitočné, ak dokážeme určiť charakter definitnosti nejakej symetrickej matice (a tým pádom aj ňou určenej kvadratickej či bilineárnej formy) priamo, t. j. bez jej predchádzajúcej úpravy na s ňou kongruentný diagonálny tvar. Za tým účelom najprv zavedieme istú modifikáciu úprav typu (1) z dôkazu vety 11.3.5 – nazveme ich úpravami typu

(1^+) Nech $i \le n$ je najmenší index taký, že $a_{ii} \neq 0$. Potom postupne pre každé $j \le n$ také, že $j > i$ a $a_{ij} = a_{ji} \neq 0$, pripočítame k j-temu stĺpcu matice $\big(-\frac{a_{ij}}{a_{ii}}\big)$-násobok i-teho stĺpca a v takto získanej matici pripočítame $\big(-\frac{a_{ji}}{a_{ii}}\big)$-násobok i-teho riadku k j-temu riadku. Inak povedané, pomocou diagonálneho prvku $a_{ii} \neq 0$ vynulujeme všetky nenulové prvky i-teho riadku aj stĺpca, ktoré ležia *napravo* resp. *nadol* od prvku a_{ii}.

Uvedomme si, že matica prechodu, ktorá vznikne vykonaním ESO, zodpovedajúcich nejakým úpravám typu (1^+) na jednotkovej matici, je vždy horná trojuholníková matica s jednotkami na diagonále. Navyše, súčin dvoch matíc takéhoto tvaru má tiež takýto tvar (presvedčte sa o tom).

Ak $\boldsymbol{A} = (a_{ij})_{n\times n}$ je matica nad ľubovoľným poľom K a $1 \le k \le n$, tak pre potreby zvyšku tohto paragrafu $\boldsymbol{A}_k$ označuje maticu tvorenú ľavým horným rohom rozmeru $k \times k$ matice $\boldsymbol{A}$. Teda

$$\boldsymbol{A}_1 = (a_{11}), \quad \boldsymbol{A}_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \dots, \quad \boldsymbol{A}_n = \boldsymbol{A}.$$

Determinanty matíc $\boldsymbol{A}_k$ pre $1 \le k \le n$ nazývame *hlavnými minormi* matice $\boldsymbol{A}$.

12.2.3. Veta. *(Jacobi) Nech K je ľubovoľné pole a $\mathbf{0} \neq \boldsymbol{A} \in K^{n\times n}$ je symetrická matica hodnosti h. Potom nasledujúce podmienky sú ekvivalentné:*

(i) matica $\boldsymbol{A}_h$ je regulárna a maticu $\boldsymbol{A}$ možno upraviť na s ňou kongruentný diagonálny tvar výlučne pomocou úprav typu (1^+);

(ii) $|\boldsymbol{A}_k| \neq 0$ pre každé $1 \le k \le h$, a platí

$$\boldsymbol{A} \equiv \operatorname{diag}\left(|\boldsymbol{A}_1|, \frac{|\boldsymbol{A}_2|}{|\boldsymbol{A}_1|}, \dots, \frac{|\boldsymbol{A}_h|}{|\boldsymbol{A}_{h-1}|}, 0, \dots, 0\right);$$

(iii) $|\boldsymbol{A}_k| \neq 0$ pre každé $1 \le k \le h$;

(iv) $|\boldsymbol{A}_k| \neq 0$ pre každé $1 \le k \le h$ a maticu $\boldsymbol{A}$ možno upraviť na s ňou kongruentný diagonálny tvar

$$\operatorname{diag}\left(|\boldsymbol{A}_1|, \frac{|\boldsymbol{A}_2|}{|\boldsymbol{A}_1|}, \dots, \frac{|\boldsymbol{A}_h|}{|\boldsymbol{A}_{h-1}|}, 0, \dots, 0\right)$$

výlučne pomocou úprav typu (1^+).

Dôkaz. (i) $\Rightarrow$ (ii): Nech platí (i). Podľa poznámky, predchádzajúcej dokazovanú vetu existuje horná trojuholníková matica $\boldsymbol{P} \in K^{n\times n}$ s jednotkami na diagonále taká, že matica $\boldsymbol{B} = \boldsymbol{P}^\mathsf{T} \cdot \boldsymbol{A} \cdot \boldsymbol{P}$ je v diagonálnom tvare. Zrejme i každá z matíc $\boldsymbol{P}_k$, $1 \le k \le n$, tvorená ľavým horným rohom rozmeru $k \times k$ matice $\boldsymbol{P}$, je v takomto tvare. Označme $\boldsymbol{B} = \operatorname{diag}(b_1, \dots, b_n)$. Prenechávame čitateľovi, aby sa sám presvedčil, že potom pre každé $k \le n$ platí

$$\boldsymbol{B}_k = \operatorname{diag}(b_1, \dots, b_k) = \boldsymbol{P}_k^\mathsf{T} \cdot \boldsymbol{A}_k \cdot \boldsymbol{P}_k.$$

Determinant každej z matíc $\boldsymbol{P}_k$ je súčin jej diagonálnych prvkov, čiže $|\boldsymbol{P}_k| = 1$. Preto

$$|\boldsymbol{B}_k| = b_1 \dots b_k = |\boldsymbol{A}_k|$$

pre každé $k \le n$. Keďže matica $\boldsymbol{A}_h$ je regulárna, platí

$$|\boldsymbol{A}_h| = |\boldsymbol{B}_h| = b_1 \dots b_h \ne 0,$$

teda $b_k \ne 0$ pre všetky $k \le h$. Z jednotlivých rovností $|\boldsymbol{A}_1| = b_1$, $|\boldsymbol{A}_2| = b_1 b_2$, …, $|\boldsymbol{A}_h| = b_1 b_2 \dots b_h$ už priamo vyplýva nenulovosť všetkých hlavných minorov $|\boldsymbol{A}_k|$ pre $k \le h$, ako aj rovnosti $b_1 = |\boldsymbol{A}_1|$, $b_2 = |\boldsymbol{A}_2|/|\boldsymbol{A}_1|$, …, $b_h = |\boldsymbol{A}_h|/|\boldsymbol{A}_{h-1}|$. Keďže $h(\boldsymbol{B}) = h(\boldsymbol{A}) = h$, pre $h \le k \le n$ platí $b_k = 0$.

(ii) $\Rightarrow$ (iii) platí triviálne.

(iii) $\Rightarrow$ (iv): Nech platí (iii). Dokážeme, že maticu $\boldsymbol{A}$ možno upraviť na diagonálny tvar

$$\boldsymbol{A} \equiv \operatorname{diag}\left(|\boldsymbol{A}_1|, \frac{|\boldsymbol{A}_2|}{|\boldsymbol{A}_1|}, \dots, \frac{|\boldsymbol{A}_h|}{|\boldsymbol{A}_{h-1}|}, 0, \dots, 0\right)$$

len pomocou úprav typu (1^+). Nakoľko platí $a_{11} = |\boldsymbol{A}_1| \ne 0$, pomocou a_{11} možno vynulovať všetky ostatné prvky prvého riadku aj stĺpca matice $\boldsymbol{A}$. Keďže ležia napravo resp. nadol od a_{11}, ide o úpravu typu (1^+). Dostaneme tak maticu v blokovo diagonálnom tvare

$$\boldsymbol{A} \equiv \operatorname{diag}\left(a_{11}, \boldsymbol{C}^{(1)}\right),$$

kde $\boldsymbol{C}^{(1)} = \left(c_{ij}^{(1)}\right)$ je matica rozmeru $(n-1)\times(n-1)$ nad K. Vzhľadom na charakter vykonaných stĺpcových a riadkových úprav platí

$$\boldsymbol{A}_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \equiv \begin{pmatrix} a_{11} & 0 \\ 0 & c_{11}^{(1)} \end{pmatrix},$$

a taktiež $|\boldsymbol{A}_2| = a_{11} c_{11}^{(1)}$, čiže

$$c_{11}^{(1)} = \frac{|\boldsymbol{A}_2|}{a_{11}} = \frac{|\boldsymbol{A}_2|}{|\boldsymbol{A}_1|} \ne 0.$$

Pomocou prvku $c_{11}^{(1)} \neq 0$ možno teraz vynulovať všetky ostatné prvky prvého riadku aj stĺpca matice $\boldsymbol{C}^{(1)}$. Opäť ide o úpravu typu (1^+) na matici $\operatorname{diag}\Big(a_{11}, \boldsymbol{C}^{(1)}\Big)$. Dostaneme tak maticu v blokovo diagonálnom tvare

$$\boldsymbol{A} \equiv \operatorname{diag}\big(a_{11}, \boldsymbol{C}^{(1)}\big) \equiv \operatorname{diag}\left(|\boldsymbol{A}_1|, \frac{|\boldsymbol{A}_2|}{|\boldsymbol{A}_1|}, \boldsymbol{C}^{(2)}\right),$$

kde $\boldsymbol{C}^{(2)} = \big(c_{ij}^{(2)}\big) \in K^{(n-2)\times(n-2)}$. Rovnakou úvahou ako v predošlom prípade dospejeme k záveru, že

$$c_{11}^{(2)} = \frac{|\boldsymbol{A}_3|}{|\boldsymbol{A}_2|} \neq 0.$$

Takto môžeme pokračovať tak dlho, až kým nedospejeme k blokovo diagonálnemu tvaru

$$\boldsymbol{A} \equiv \operatorname{diag}\left(|\boldsymbol{A}_1|, \frac{|\boldsymbol{A}_2|}{|\boldsymbol{A}_1|}, \dots, \frac{|\boldsymbol{A}_h|}{|\boldsymbol{A}_{h-1}|}, \boldsymbol{C}^{(h)}\right),$$

kde $\boldsymbol{C}^{(h)} \in K^{(n-h)\times(n-h)}$. Vzhľadom nato, že obe matice majú hodnosť h, (pokiaľ $h < n$) matica $\boldsymbol{C}^{(h)}$ je identicky nulová.

(iv) $\Rightarrow$ (i) je opäť triválne.

12.2.4. Veta. (*Sylvestrovo kritérium*) *Nech $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ je symetrická matica. Potom*

(a) $\boldsymbol{A}$ je kladne definitná práve vtedy, keď $|\boldsymbol{A}_k| > 0$ pre všetky $1 \leq k \leq n$;

(b) $\boldsymbol{A}$ je záporne definitná práve vtedy, keď $(-1)^k|\boldsymbol{A}_k| > 0$ pre všetky $1 \leq k \leq n$.

Dôkaz. (a) Nech $\boldsymbol{A}$ je kladne definitná. Podľa dôsledku 12.2.2 existuje regulárna matica $\boldsymbol{P} \in \mathbb{R}^{n\times n}$, taká, že $\boldsymbol{A} = \boldsymbol{P}^\mathsf{T} \cdot \boldsymbol{P}$. Keďže $|\boldsymbol{P}| \neq 0$, odtiaľ už priamo vyplýva

$$|\boldsymbol{A}| = |\boldsymbol{P}^\mathsf{T}| \cdot |\boldsymbol{P}| = |\boldsymbol{P}|^2 > 0.$$

Pre každé $1 \leq k \leq n$ označme $S_k = [\boldsymbol{e}_1, \dots, \boldsymbol{e}_k]$ lineárny podpriestor v $\mathbb{R}^n$ generovaný prvými k vektormi kanonickej bázy $\boldsymbol{\varepsilon}$ a $q_k = q \restriction S_k$ zúženie kvadratickej formy $q(\boldsymbol{x}) = \boldsymbol{x}^\mathsf{T} \cdot \boldsymbol{A} \cdot \boldsymbol{x}$ na podpriestor S_k. Zrejme každé q_k je kladne definitná kvadratická forma, ktorá má v báze $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_k)$ podpriestoru S_k maticu $\boldsymbol{A}_k$. Takže každá z matíc $\boldsymbol{A}_k$, $1 \leq k \leq n$, je kladne definitná. Podľa prvej časti dôkazu z toho vyplýva $|\boldsymbol{A}_k| > 0$.

Nech naopak všetky hlavné minory $|\boldsymbol{A}_k|$, $1 \leq k \leq n$, sú kladné. Potom $h(\boldsymbol{A}) = n$ a podľa Jacobiho vety platí

$$\boldsymbol{A} \equiv \operatorname{diag}\left(|\boldsymbol{A}_1|, \frac{|\boldsymbol{A}_2|}{|\boldsymbol{A}_1|}, \dots, \frac{|\boldsymbol{A}_n|}{|\boldsymbol{A}_{n-1}|}\right).$$

Keďže všetky diagonálne prvky v poslednej matici sú kladné, $\boldsymbol{A}$ je kladne definitná.

(b) vyplýva z (a) na základe faktu, že $\boldsymbol{A}$ je záporne definitná práve vtedy, keď $-\boldsymbol{A}$ je kladne definitná, a rovnosti $|-\boldsymbol{A}_k| = (-1)^k|\boldsymbol{A}_k|$ splnenej pre všetky $k \leq n$.

12.3 Extrémy funkcií viac premenných

V tomto paragrafe si predvedieme, ako nám čerstvo nadobunuté poznatky o kvadratických formách môžu pomôcť pri štúdiu funkcií viac premenných, presnejšie pri hľadaní extrémov a sedlových bodov takýchto funkcií a všeobecnejšie pri klasifikácii ich stacionárnych bodov.

Najprv si zopakujme, ako si počíname v prípade funkcie jednej premennej. Pre jednoduchosť sa obmedzíme len na „dostatočne hladké“ funkcie.

Reálnu funkciu jednej premennej $f : A \to \mathbb{R}$ definovanú na nejakej otvorenej množine [1] $A \subseteq \mathbb{R}$ nazveme pre účely tohto paragrafu *dostatočne hladkou*, ak f má na celej množine A konečnú a spojitú prvú i druhú deriváciu. Z matematickej analýzy si pripomeňme si, že pre takúto funkciu f máme v každom bode $a \in A$ k dispozícii Taylorov rozvoj

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \theta(x)(x-a)^2$$

pre x z istého okolia [2] $N \subseteq A$ bodu a, pričom funkcia $\theta : N \to \mathbb{R}$ je spojitá a vyhovuje podmienke $\theta(a) = 0$, teda absolútna hodnota zvyšku $\theta(x)(x-a)^2$ je v dosť malom okolí $M \subseteq N$ bodu a v porovnaní s ostatnými členmi uvedeného rozvoja zanedbateľne malá. Pre x z tohto malého okolia bodu a teda môžeme písať

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2.$$

Ak $f'(a) \neq 0$, tak lineárny člen $f'(a)(x-a)$ mení v bode a znamienko a v dostatočne malom okolí bodu a prevažuje nad kvadratickým členom $\frac{1}{2}f''(a)(x-a)^2$. Preto dostatočne hladká funkcia (dokonca už funkcia s konečnou a spojitou prvou deriváciou) môže nadobúdať na otvorenej množine extrémy len bodoch a, pre ktoré platí $f'(a) = 0$; hovoríme im *stacionárne* alebo tiež *kritické body* funkcie f. Ak $a \in A$ je stacionárny bod, tak uvedený Taylorov rozvoj má v tomto bode tvar

$$f(x) = f(a) + \frac{1}{2}f''(a)(x-a)^2 + \theta(x)(x-a)^2 \approx f(a) + \frac{1}{2}f''(a)(x-a)^2$$

[1] Množina $A \subseteq \mathbb{R}$ sa nazýva *otvorená*, ak pre každé $a \in A$ existuje $\varepsilon > 0$ také, že celý otvorený interval $(a-\varepsilon, a+\varepsilon)$ je obsiahnutý v množine A.

[2] Množina $N \subseteq \mathbb{R}$ sa nazýva *okolím* bodu $a \in \mathbb{R}$, ak existuje $\varepsilon > 0$ také, že $(a-\varepsilon, a+\varepsilon) \subseteq N$.

pre $x \in M$.

Ak $f''(a) > 0$, tak $f(a) < f(x)$ pre všetky $x \neq a$ z nejakého okolia $L \subseteq M$ bodu a, teda f má v bode a *ostré lokálne minimum.*

Ak $f''(a) < 0$, tak $f(a) > f(x)$ pre všetky $x \neq a$ z nejakého okolia $L \subseteq M$ bodu a, teda f má v bode a *ostré lokálne maximum.*

Ak $f''(a) = 0$, tak v bode a sa može diať v podstate „čokoľvek", presnejšie, len na základe prvej a druhej derivácie nevieme určiť, či f má v bode a extrém, ani charakter prípadného extrému. Ak f má aj derivácie vyších rádov, ich znalosť nám môže pomôcť. No tým sa už zaoberať nebudeme.

Podobne si budeme počínať aj pri skúmaní funkcií viac premenných. Namiesto „obyčajnej" derivácie však musíme uvažovať derivácie podľa rôznych premenných funkcie f – hovoríme im *parciálne derivácie.* Pre čitateľa, ktorý sa s parciálnymi deriváciami dosiaľ nestretol, poznamenávame, že parciálnu deriváciu $\partial f/\partial x_i$ funkcie f podľa premennej x_i dostaneme tak, že f jednoducho derivujeme podľa x_i ako funkciu jednej premennej, pričom všetky ostatné premenné považujeme za konštanty. Druhú parciálnu deriváciu funkcie f, najprv podľa premennej x_j a potom podľa premennej x_i, značíme $\partial^2 f/\partial x_i \partial x_j$. Namiesto $\partial^2 f/\partial x_i \partial x_i$ píšeme $\partial^2 f/\partial x_i^2$. Pre jednoduchosť sa opäť obmedzíme len na „dostatočne hladké" funkcie.

Reálnu funkciu n premenných definovanú na otvorenej množine [3] $A \subseteq \mathbb{R}^n$ nazveme pre účely tohto paragrafu *dostatočne hladkou*, ak f má na celej množine A konečné a spojité všetky parciálne derivácie prvého i druhého rádu.

Nech teda $f : A \to \mathbb{R}$ je dostatočne hladká funkcia definovaná na otvorenej množine $A \subseteq \mathbb{R}^n$ a $\boldsymbol{a} \in A$. *Prvou (totálnou) deriváciou* alebo tiež *gradientom* funkcie f v bode $\boldsymbol{a}$ nazývame vektor

$$f'(\boldsymbol{a}) = \mathrm{grad} f(\boldsymbol{a}) = \nabla f(\boldsymbol{a}) = \left(\frac{\partial f(\boldsymbol{a})}{\partial x_1}, \ldots, \frac{\partial f(\boldsymbol{a})}{\partial x_n} \right),$$

ktorého zložky tvoria prvé parciálne derivácie funkcie f v bode $\boldsymbol{a}$ (znak ∇ čítame ako *nabla*). *Druhou (totálnou) deriváciou* alebo tiež *Hesseho maticou* funkcie f v bode $\boldsymbol{a}$ nazývame maticu

$$f''(\boldsymbol{a}) = \left(\frac{\partial^2 f(\boldsymbol{a})}{\partial x_i \partial x_j} \right)_{n \times n},$$

tvorenú všetkými druhými parciálnymi deriváciami funkcie f v bode $\boldsymbol{a}$. V diferenciálnom počte funkcií viac premenných sa dokazuje tzv. *Clairautova-Schwartzova veta*, podľa ktorej zo spojitosti druhých parciálnych derivácií

[3] Množina $A \subseteq \mathbb{R}^n$ sa nazýva *otvorená*, ak pre každé $\boldsymbol{a} \in A$ existuje $\varepsilon > 0$ také, že celá otvorená guľa $B(\boldsymbol{a}, \varepsilon) = \left\{ \boldsymbol{x} \in \mathbb{R}^n \,;\, (x_1 - a_1)^2 + \ldots + (x_n - a_n)^2 < \varepsilon^2 \right\}$ je obsiahnutá v množine A.

na množine A vyplýva
$$\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j} = \frac{\partial^2 f(\boldsymbol{x})}{\partial x_j \partial x_i}$$
pre všetky $i, j \leq n$ a $\boldsymbol{x} \in A$. To znamená, že Hesseho matica $f''(\boldsymbol{a})$ dostatočne hladkej funkcie f je symetrická, teda je maticou kvadratickej formy
$$q(\boldsymbol{v}) = \boldsymbol{v} \cdot f(\boldsymbol{a}) \cdot \boldsymbol{v}^{\mathsf{T}} = \boldsymbol{v} \cdot \left(\frac{\partial^2 f(\boldsymbol{a})}{\partial x_i \partial x_j} \right) \cdot \boldsymbol{v}^{\mathsf{T}}$$
na (riadkovom) vektorovom priestore $\mathbb{R}^n$ vzhľadom na kanonickú bázu $\boldsymbol{\varepsilon}^{(n)}$. Ukážeme si, že práve signatúra druhej derivácie $f''(\boldsymbol{a})$ rozhodujúcim (a v prípade jej regularity dokonca jednoznačným) spôsobom určuje chovanie funkcie v jej kritických bodoch.

Podobne ako v prípade jednej premennej, aj dostatočne hladkú funkciu viac premenných možno v každom bode $\boldsymbol{a} = (a_1, \ldots, a_n) \in A$ písať v tvare Taylorovho rozvoja
$$f(\boldsymbol{x}) = f(\boldsymbol{a}) + f'(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a})^{\mathsf{T}} + \frac{1}{2}\,(\boldsymbol{x} - \boldsymbol{a}) \cdot f''(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a})^{\mathsf{T}} + (\boldsymbol{x} - \boldsymbol{a}) \cdot \boldsymbol{\Theta}(\boldsymbol{x}) \cdot (\boldsymbol{x} - \boldsymbol{a})^{\mathsf{T}}$$
pre $\boldsymbol{x} = (x_1, \ldots, x_n)$ z istého okolia [4] $N \subseteq A$ bodu $\boldsymbol{a}$, kde $\boldsymbol{\Theta}(\boldsymbol{x}) = \big(\theta_{ij}(\boldsymbol{x})\big)_{n \times n}$ je matica, ktorej zložky tvoria hodnoty spojitých funkcií $\theta_{ij} : N \to \mathbb{R}$ v bode $\boldsymbol{x}$. Tieto funkcie navyše vyhovujú podmienke $\theta_{ij}(\boldsymbol{a}) = 0$, teda absolútna hodnota zvyšku $(\boldsymbol{x} - \boldsymbol{a}) \cdot \boldsymbol{\Theta}(\boldsymbol{x}) \cdot (\boldsymbol{x} - \boldsymbol{a})^{\mathsf{T}}$ je v dosť malom okolí $M \subseteq N$ bodu $\boldsymbol{a}$ v porovnaní s ostatnými členmi uvedeného rozvoja zanedbateľne malá. Pre $\boldsymbol{x}$ z tohto malého okolia bodu $\boldsymbol{a}$ teda môžeme písať
$$f(\boldsymbol{x}) \approx f(\boldsymbol{a}) + f'(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a})^{\mathsf{T}} + \frac{1}{2}\,(\boldsymbol{x} - \boldsymbol{a}) \cdot f''(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a})^{\mathsf{T}}.$$

Ak $\frac{\partial f(\boldsymbol{a})}{\partial x_i} \neq 0$, tak zložka $\frac{\partial f(\boldsymbol{a})}{\partial x_i}(x_i - a_i)$ lineárneho člena $f'(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a})^{\mathsf{T}}$ mení v bode $\boldsymbol{a}$ znamienko a v dostatočne malom okolí bodu $\boldsymbol{a}$ takéto zložky prevažujú nad kvadratickým členom $\frac{1}{2}(\boldsymbol{x} - \boldsymbol{a}) \cdot f''(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a})^{\mathsf{T}}$. Preto dostatočne hladká funkcia (dokonca už funkcia s konečnými a spojitými prvými parciálnymi deriváciami) môže na otvorenej množine nadobúdať extrémy len v bodoch $\boldsymbol{a}$, pre ktoré platí $f'(\boldsymbol{a}) = \boldsymbol{0}$, t. j. všetky parciálne derivácie $\partial f / \partial x_i$ v bode $\boldsymbol{a}$ sa rovnajú nule. Takýmto bodom hovoríme *stacionárne* alebo tiež *kritické body* funkcie f. Ak $\boldsymbol{a} \in A$ je stacionárny bod, tak uvedený Taylorov rozvoj má v tomto bode tvar
$$\begin{aligned} f(\boldsymbol{x}) &= f(\boldsymbol{a}) + \frac{1}{2}\,(\boldsymbol{x} - \boldsymbol{a}) \cdot f''(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a})^{\mathsf{T}} + (\boldsymbol{x} - \boldsymbol{a}) \cdot \boldsymbol{\Theta}(\boldsymbol{x}) \cdot (\boldsymbol{x} - \boldsymbol{a})^{\mathsf{T}} \\ &\approx f(\boldsymbol{a}) + \frac{1}{2}\,(\boldsymbol{x} - \boldsymbol{a}) \cdot f''(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a})^{\mathsf{T}} \end{aligned}$$

[4] Množina $N \subseteq \mathbb{R}^n$ sa nazýva *okolím* bodu $\boldsymbol{a} \in \mathbb{R}^n$, ak existuje $\varepsilon > 0$ také, že $B(\boldsymbol{a}, \varepsilon) \subseteq N$.

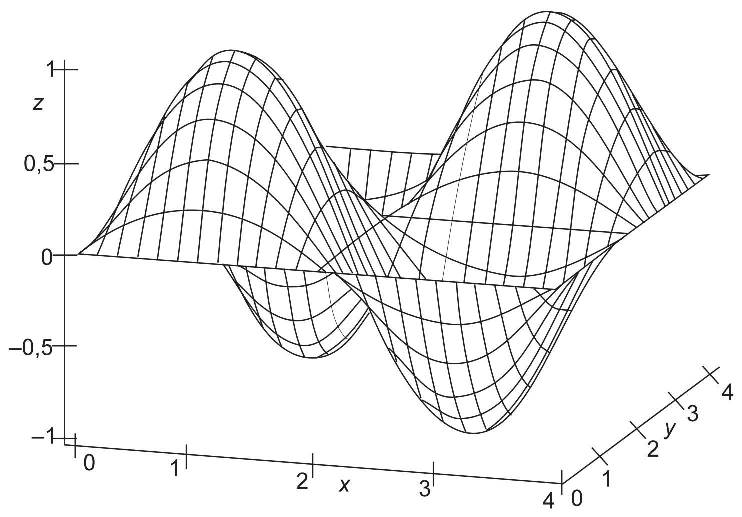
pre všetky $\boldsymbol{x}$ z okolia M bodu $\boldsymbol{a}$.

Ak Hesseho matica $f''(a)$ je kladne definitná, tak pre všetky $\boldsymbol{x} \neq \boldsymbol{a}$ z nejakého okolia $L \subseteq M$ bodu $\boldsymbol{a}$ platí $(\boldsymbol{x}-\boldsymbol{a}) \cdot f''(\boldsymbol{a}) \cdot (\boldsymbol{x}-\boldsymbol{a})^{\mathsf{T}} > 0$ a $f(\boldsymbol{a}) < f(\boldsymbol{x})$, teda f má v bode $\boldsymbol{a}$ *ostré lokálne minimum.*

Ak Hesseho matica $f''(\boldsymbol{a})$ je záporne definitná, tak pre všetky $\boldsymbol{x} \neq \boldsymbol{a}$ z nejakého okolia $L \subseteq M$ bodu $\boldsymbol{a}$ platí $(\boldsymbol{x}-\boldsymbol{a}) \cdot f''(\boldsymbol{a}) \cdot (\boldsymbol{x}-\boldsymbol{a})^{\mathsf{T}} < 0$ a $f(\boldsymbol{a}) > f(\boldsymbol{x})$, teda f má v bode $\boldsymbol{a}$ *ostré lokálne maximum.*

Ak Hesseho matica $f''(\boldsymbol{a})$ je indefinitná, tak existujú vektory $\boldsymbol{u}$, $\boldsymbol{v}$ a číslo $\varepsilon > 0$ také, že $\boldsymbol{u} \cdot f''(\boldsymbol{a}) \cdot \boldsymbol{u}^{\mathsf{T}} < 0 < \boldsymbol{v} \cdot f(\boldsymbol{a}) \cdot \boldsymbol{v}^{\mathsf{T}}$, obe úsečky $X = \{\boldsymbol{a}+t\boldsymbol{u};\ |t| \leq \varepsilon\}$, $Y = \{\boldsymbol{a}+t\boldsymbol{v};\ |t| \leq \varepsilon\}$ sú celé obsiahnuté v okolí M a pre všetky body $\boldsymbol{x} \in X$, $\boldsymbol{y} \in Y$ rôzne od $\boldsymbol{a}$ platí $f(\boldsymbol{x}) < f(\boldsymbol{a}) < f(\boldsymbol{y})$. Hovoríme, že f má v bode $\boldsymbol{a}$ *sedlový bod* alebo krátko *sedlo.* Tento prípad zrejme nemôže nastať pre funkcie jednej premennej. Samozrejme, v sedlovom bode funkcia nenadobúda extrém.

Ak Hesseho matica $f''(\boldsymbol{a})$ je nenulová, singulárna a semidefinitná, tak náš čitateľ asi očakáva, že f v bode $\boldsymbol{a}$ nadobudne *neostré lokálne minimum* (pre kladne semidefinitnú maticu) alebo *neostré lokálne maximum* (pre záporne semidefinitnú maticu). No nie vždy je tomu tak. Podobne ako v prípade nulovej druhej derivácie u funkcií jednej premennej, ak matica $f''(\boldsymbol{a})$ je singulárna a (kladne alebo záporne) semidefinitná, tak – bez ohľadu nato, či je nulová alebo nie, – v bode $\boldsymbol{a}$ sa môže udiať prakticky „čokoľvek“. Funkcia v tomto bode môže, ale nemusí mať extrém alebo sedlo. Niečo však predsa len môžeme povedať: ak uvedená matica je nenulová a kladne semidefinitná, tak f nemá v bode $\boldsymbol{a}$ (ostré ani neostré) lokálne maximum; ak je nenulová a



Obr. 12.1. Graf funkcie $z = f(x,y) = \sin\frac{\pi x}{2} \sin\frac{\pi y}{2}$

záporne semidefinitná, tak f nemá v bode $\boldsymbol{a}$ (ostré ani neostré) lokálne minimum. Často však môžeme podobné otázky rozhodnúť preskúmaním Hesseho matice $f''(\boldsymbol{x})$ pre body $\boldsymbol{x}$ v nejakom blízkom okolí bodu $\boldsymbol{a}$ – pozri cvičenie 12.9.

Ak platí $f''(\boldsymbol{a}) = \mathbf{0}$ a f má aj derivácie vyšších rádov, tak na ich základe môžeme niekedy povedať viac – pozri paragraf 33.2.

12.3.1. Príklad. (*Kopčeky a jamky*) Preskúmame funkciu $f : \mathbb{R}^2 \to \mathbb{R}$, danú predpisom $f(x, y) = \sin\frac{\pi x}{2}\sin\frac{\pi y}{2}$. Jej prvé parciálne derivácie sú

$$\frac{\partial f}{\partial x} = \frac{\pi}{2}\cos\frac{\pi x}{2}\sin\frac{\pi y}{2}, \qquad \frac{\partial f}{\partial y} = \frac{\pi}{2}\sin\frac{\pi x}{2}\cos\frac{\pi y}{2}.$$

Keďže sinus a kosinus nejakého čísla sa súčasne nemôžu rovnať nule, (x, y) je stacionárnym bodom funkcie f práve vtedy, keď $\cos\frac{\pi x}{2} = \cos\frac{\pi y}{2} = 0$ alebo $\sin\frac{\pi x}{2} = \sin\frac{\pi y}{2} = 0$. Teda stacionárnymi bodmi funkcie f sú práve všetky mrežové body roviny tvaru (m, n), kde m, n sú celé čísla rovnakej parity (t. j. $m - n$ je párne číslo).

Vypočítame i druhé parciálne derivácie

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial y^2} = -\frac{\pi^2}{4}\sin\frac{\pi x}{2}\sin\frac{\pi y}{2}, \qquad \frac{\partial^2 f}{\partial x\partial y} = \frac{\partial^2 f}{\partial y\partial x} = \frac{\pi^2}{4}\cos\frac{\pi x}{2}\cos\frac{\pi y}{2}.$$

Vidíme, že funkcia f je dostatočne hladká na celej rovine $\mathbb{R}^2$. Jej druhá totálna derivácia v stacionárnych bodoch má tvar

$$f''(m, n) = \frac{\pi^2}{4}\begin{pmatrix} -\sin\frac{m\pi}{2}\sin\frac{n\pi}{2} & \cos\frac{m\pi}{2}\cos\frac{n\pi}{2} \\ \cos\frac{m\pi}{2}\cos\frac{n\pi}{2} & -\sin\frac{m\pi}{2}\sin\frac{n\pi}{2} \end{pmatrix}.$$

Pre $m = 2k$, $n = 2l$ obe párne máme

$$f''(m, n) = \frac{\pi^2}{4}\begin{pmatrix} 0 & (-1)^{\frac{m+n}{2}} \\ (-1)^{\frac{m+n}{2}} & 0 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

V každom z týchto bodoch je Hesseho matica indefinitná, teda funkcia f má v bodoch tvaru $(2k, 2l)$, kde $k, l \in \mathbb{Z}$, sedlá. Hodnota funkcie vo všetkých sedlových bodoch je 0.

Pre $m = 2k + 1$, $n = 2l + 1$ obe nepárne máme

$$f''(m, n) = \frac{\pi^2}{4}\begin{pmatrix} (-1)^{\frac{m+n}{2}} & 0 \\ 0 & (-1)^{\frac{m+n}{2}} \end{pmatrix} \equiv (-1)^{k+l+1}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Táto matica je záporne definitná, ak $k + l$ je párne číslo, a kladne definitná pre nepárne $k + l$.

Funkcia f teda nadobúda ostré lokálne maximá v bodoch $(2k+1, 2l+1)$, kde k, l sú celé čísla a $k+l$ je párne. Hodnota všetkých týchto maxím je 1.

V bodoch $(2k+1, 2l+1)$, kde $k, l \in \mathbb{Z}$ a $k+l$ je nepárne číslo, nadobúda f ostré lokálne minimá, ktorých hodnota je vždy -1.

Na obrázku je graf funkcie f na časti definičného oboru $\langle 0, 4\rangle \times \langle 0, 4\rangle$. Vidno na ňom maximá v bodoch $(1, 1)$, $(3, 3)$, minimá v bodoch $(1, 3)$, $(3, 1)$ a sedlo v bode $(2, 2)$.

Nakoľko všetky typy kritických bodov, ktorých charakter možno určiť na základe druhej derivácie funkcie, sa nám podarilo ilustrovať na jedinom príklade, v ďalších ukážkach sa zameriame na funkcie so singulárnymi semidefinitnými druhými deriváciami v kritických bodoch. Keďže v takom prípade nám toho naša teória mnoho nepovie, zvolíme si funkcie, pri ktorých nám charakter kritických bodov bude jasný z názoru, ako napokon aj v prvom príklade. Najprv jeden príklad, kde všetko dopadne podľa očakávania.

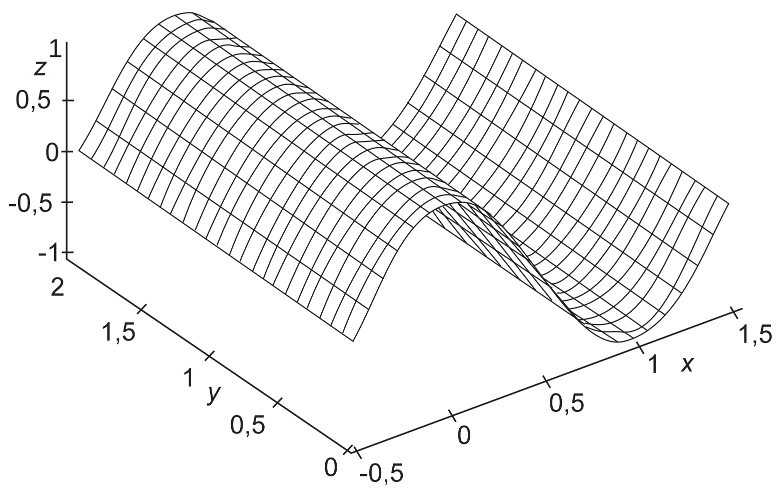
12.3.2. Príklad. (*Vlnitý plech*) Je daná funkcia $g : \mathbb{R}^2 \to \mathbb{R}$, kde $g(x, y) = \cos(\pi x)$. Jej prvé parciálne derivácie sú

$$\frac{\partial g}{\partial x} = -\pi \sin(\pi x), \qquad \frac{\partial g}{\partial y} = 0.$$

Stacionárne body funkcie g teda vytvárajú systém ekvidistantných rovnobežných priamok (s rovnicami) $x = k$, kde k je ľubovoľné celé číslo.

Druhé parciálne derivácie funkcie g sú

$$\frac{\partial^2 g}{\partial x^2} = -\pi^2 \cos(\pi x), \qquad \frac{\partial^2 g}{\partial y^2} = \frac{\partial^2 g}{\partial x \partial y} = \frac{\partial^2 g}{\partial y \partial x} = 0,$$



Obr. 12.2. Graf funkcie $z = g(x, y) = \cos(\pi x)$

takže funkcia g je dostatočne hladká. Druhá totálna derivácia v stacionárnych bodoch teda vyzerá takto

$$g''(k,y) = -\pi^2 \begin{pmatrix} \cos(k\pi) & 0 \\ 0 & 0 \end{pmatrix} = -\pi^2 \begin{pmatrix} (-1)^k & 0 \\ 0 & 0 \end{pmatrix}.$$

Pre k párne dostávame

$$g''(k,y) \equiv \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix},$$

čo je záporne semidefinitná singulárna matica. Podobne, pre k nepárne máme

$$g''(k,y) \equiv \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

čo je kladne semidefinitná singulárna matica. V takom prípade nám naša teória neposkytuje nijaké závery. Z grafu funkcie na časti $\left\langle -\frac{3}{2}, \frac{3}{2} \right\rangle \times \langle 0, 2 \rangle$ definičného oboru a z periodičnosti funkcie $\cos(\pi x)$ však možno usúdiť, že pre párne $k \in \mathbb{Z}$ nadobúda funkcia g na priamkach $x = k$ neostré lokálne maximum hodnoty 1 a pre nepárne $k \in \mathbb{Z}$ nadobúda g na priamkach $x = k$ neostré lokálne minimum hodnoty -1.

Na záver si predvedieme, čo všetko sa v kritickom bode môže ešte stať.

12.3.3. Príklad. (*Kreslo, sieť a sedlo*) Uvažujme reálne funkcie dvoch premenných

$$h_1(x,y) = x^2 + y^3, \qquad h_2(x,y) = x^2 + y^4, \qquad h_3(x,y) = x^2 - y^4.$$

Ľahko možno nahliadnuť, že všetky majú jediný a ten istý stacionárny bod $(0,0)$, v ktorom nadobúdajú tú istú hodnotu 0. Taktiež majú v tomto bode rovnakú druhú totálnu deriváciu

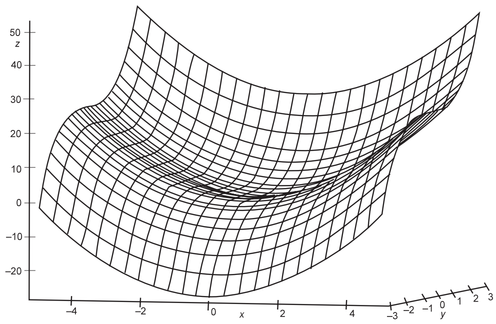
$$h_1''(0,0) = h_2''(0,0) = h_3''(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix},$$

(presvedčte sa o tom sami). Takže Hesseho matica je nenulová, singulárna a kladne semidefinitná. Z obrázkov grafov funkcií v istých okoliach bodu $(0,0)$ však vidno, že

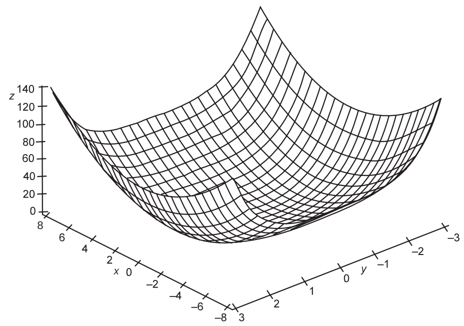
(1) h_1 nemá v bode $(0,0)$ extrém ani sedlo – v tomto bode sa totiž pretína krivka $z = x^2$, $y = 0$, pre ktorú je $(0,0)$ bodom ostrého lokálneho minima, s krivkou $x = 0$, $z = y^3$, ktorá má v bode $(0,0)$ inflexný bod;

(2) h_2 má v bode $(0,0)$ ostré lokálne minimum;

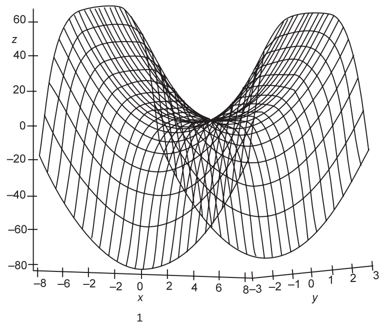
(3) h_3 má v bode $(0,0)$ sedlo.



Obr. 12.3. Graf funkcie $z = h_1(x, y) = x^2 + y^3$



Obr. 12.4. Graf funkcie
$z = h_2(x, y) = x^2 + y^4$



Obr. 12.5. Graf funkcie
$z = h_3(x, y) = x^2 - y^4$

Cvičenia

12.1. Napíšte matice nasledujúcich kvadratických foriem, upravte ich na diagonálny tvar a určte ich signatúry:

(a) $p(x) = 3(x_1^2 + x_2^2 + x_3^2) - 2(x_1x_2 + x_1x_3 + x_2x_3)$ pre $\boldsymbol{x} = (x_1, x_2, x_3)^\mathsf{T} \in \mathbb{R}^3$;

(b) $q(\boldsymbol{x}) = x_1x_2 - 2x_1x_3 + 3x_2x_3$ pre $\boldsymbol{x} = (x_1, x_2, x_3)^\mathsf{T} \in \mathbb{R}^3$;

(c) $r(\boldsymbol{x}) = x_1x_2 + x_1x_3 + x_2x_3$ pre $\boldsymbol{x} = (x_1, x_2, x_3)^\mathsf{T} \in \mathbb{R}^3$;

(d) $s(\boldsymbol{x}) = -2x_1^2 + x_2^2 - x_3^2 + 5x_4^2 - 2x_1x_2 - 4x_1x_3 - 2x_1x_4 - 2x_2x_3 2x_2x_4 4x_3x_4$ pre $\boldsymbol{x} = (x_1, x_2, x_3, x_4)^\mathsf{T} \in \mathbb{R}^4$;

(e) $t(\boldsymbol{x}) = -x_1^2 + x_2^2 + 8x_3^2 + 2x_4^2 + 4x_1x_3 - 2x_2x_4 + 10x_3x_4$ pre $\boldsymbol{x} = (x_1, x_2, x_3, x_4)^\mathsf{T} \in \mathbb{R}^4$.

12.2. Vypočítajte hlavné minory matíc kvadratických foriem z cvičenia 12.1. Pokúste sa určiť definitnosť každej z uvedených foriem na základe Sylvestrovho kritéria. Pre ktoré z uvedených foriem možno na základe Jacobiho vety 12.2.3 určiť aj signatúru?

12.3. Riešte rovnakú úlohu ako v cvičení 12.2 pre kvadratické formy z cvičení 11.6 a 11.7.

12.4. (a) Dokážte, že matica $2\boldsymbol{I}_2$ je kongruentná s maticou $\boldsymbol{I}_2$ nad poľom $\mathbb{Q}$.
(b) Dokážte, že matica $\operatorname{diag}(2, 1)$ nie je kongrunetná s maticou $\boldsymbol{I}_2$ nad poľom $\mathbb{Q}$.

12.5. Nájdite stacionárne body nasledujúcich funkcií dvoch premenných a klasifikujte ich pomocou Hesseho matice:

(a) $f_1(x,y) = x^3 - 3y^2 + 3xy - 3y$, (b) $f_2(x,y) = -\frac{2}{3}x^3 + y^2 + 4xy + 3y$,

(c) $g(x,y) = xy$, (d) $h_1(x,y) = x^2y^2(x-y+1)$,

(e) $h_2(x,y) = x^2(y+1) + 2y^3 + 5y^2$, (f) $h_3(x,y) = x^2(y+1) + y^2(x-1)$,

(g) $h_4(x,y) = x^2(y+1) + y^2(y-1)$, (h) $h_5(x,y) = (x^2+y^2)(y+1)$.

12.6. Pomocou Sylvestrovho kritéria zdôvodnite nasledujúce pravidlá pre klasifikáciu stacionárnych bodov funkcie dvoch premenných.
Nech $A \subseteq \mathbb{R}$ je otvorená množina a $\boldsymbol{a} \in A \subseteq \mathbb{R}^2$ je stacionárny bod funkcie $f : A \to \mathbb{R}$. Predpokladajme, že f je dostatočne hladká a determinant Hesseho matice funkcie f v bode $\boldsymbol{a}$ je nenulový, t. j. $\det f''(\boldsymbol{a}) \neq 0$. Potom platí:

(a) ak $\det f''(\boldsymbol{a}) > 0$ a $(\partial^2 f/\partial x^2)(\boldsymbol{a}) > 0$ alebo $(\partial^2 f/\partial y^2)(\boldsymbol{a}) > 0$, tak funkcia f má v bode $\boldsymbol{a}$ ostré lokálne minimum;

(b) ak $\det f''(\boldsymbol{a}) > 0$ a $(\partial^2 f/\partial x^2)(\boldsymbol{a}) < 0$ alebo $(\partial^2 f/\partial y^2)(\boldsymbol{a}) < 0$, tak funkcia f má v bode $\boldsymbol{a}$ ostré lokálne minimum;

(c) ak $\det f''(\boldsymbol{a}) < 0$, tak tak funkcia f má v bode $\boldsymbol{a}$ sedlo.

Len tak mimochodom si všimnite, že ak $\det f''(\boldsymbol{a}) > 0$, tak obe druhé parciálne derivácie $(\partial^2 f/\partial x^2)(\boldsymbol{a})$, $(\partial^2 f/\partial y^2)(\boldsymbol{a})$ sú nevyhnutne nenulové a majú rovnaké znamienko. Vysvetlite prečo.

12.7. S využitím cvičenia 12.6 preskúmajte ešte raz stacionárne body funkcií z cvičenia 12.5.

12.8. Nájdite stacionárne body nasledujúcich funkcií troch premenných a klasifikujte ich pomocou Hesseho matice:

(a) $f(x,y,z) = 3x^2 + x^3 + y^2 + xy^2 + z^3 - 3z$,

(b) $g(x,y,z) = (x+y+z)\,\mathrm{e}^{-(x^2+y^2+z^2)}$,

(c) $h(x,y,z) = \frac{x}{y+z} + \frac{y}{x+z} + \frac{z}{x+y}$.

12.9. Nasledujúce kritérium nám umožňuje klasifikovať aj niektoré stacionárne body funkcie, v ktorých je jej druhá totálna derivácia singulárna.
Predpokladajme, že $f\colon A \to \mathbb{R}$ je dostatočne hladká funkcia definovaná na otvorenej množine $A \subseteq \mathbb{R}^n$ a $\boldsymbol{a} \in A$ je jej stacionárny bod. Zdôvodnite nasledujúce tvrdenia:

(a) Ak existuje také okolie $N \subseteq A$ bodu $\boldsymbol{a}$, že pre všetky $\boldsymbol{x} \in N$, $\boldsymbol{x} \neq \boldsymbol{a}$ je druhá totálna derivácia f"$(\boldsymbol{x})$ funkcie f v bode $\boldsymbol{x}$ kladne (záporne) definitná, tak f má v bode $\boldsymbol{a}$ ostré lokálne minimum (maximum).

(b) Ak existuje také okolie $N \subseteq A$ bodu $\boldsymbol{a}$, že pre všetky $\boldsymbol{x} \in N$, $\boldsymbol{x} \neq \boldsymbol{a}$ je druhá totálna derivácia f"$(\boldsymbol{x})$ funkcie f v bode $\boldsymbol{x}$ kladne (záporne) semidefinitná, tak f má v bode $\boldsymbol{a}$ lokálne minimum (maximum).

(c) Ak existuje také okolie $N \subseteq A$ bodu $\boldsymbol{a}$, že pre všetky $\boldsymbol{x} \in N$, $\boldsymbol{x} \neq \boldsymbol{a}$ je druhá totálna derivácia f"$(\boldsymbol{x})$ funkcie f v bode $\boldsymbol{x}$ indefinitná, tak f má v bode $\boldsymbol{a}$ sedlo.

12.10. S využitím cvičenia 12.9 vysvetlite chovanie funkcií g z príkladu 12.3.2 a h_1, h_2, h_3 z príkladu 12.3.3 v ich kritických bodoch. Taktiež opätovne preskúmajte tie kritické body funkcií z predošlých cvičení, ktoré ste nedokázali klasifikovať podľa Hesseho matice v príslušnom bode.

12.11. Nájdite stacionárne body nasledujúcich funkcií dvoch premenných a klasifikujte ich pomocou Hesseho matice (ak treba) v okolí týchto bodov:

(a) $f_1(x,y) = (x+y-1)^3 + (x-y+1)^2$,

(b) $f_2(x,y) = (x+y-1)^3 + (x-y+1)^3$,

(c) $g_1(x,y) = x^2y^2$, (d) $g_2(x,y) = xy(xy-1)$,

(e) $h_1(x,y) = x^2\,\mathrm{e}^y$, (f) $h_2(x,y) = x\,\mathrm{e}^{-y^2}$.

12.12. Nájdite a klasifikujte stacionárne body funkcií dvoch premenných

(a) $f(x,y) = x^4(x-1) + ay(y-1)$, (b) $g(x,y) = x^4(x-1) + ay^2(y-1)$,

(c) $h(x,y) = x^4(x-1) + ay^3(y-1)$, (d) $k(x,y) = x^4(x-1) + ay^2(y-1)^2$.

Zakaždým urobte diskusiu vzhľadom na parameter $a \in \mathbb{R}$.

13. Euklidovské priestory

Naše štúdium vektorových priestorov sa doteraz nieslo prevažne v algebraickom duchu a bolo vedené takmer výlučne algebraickými prostriedkami. Geometria bola v tomto poňatí zredukovaná väčšinou len na otázky rovno- či rôznobežnosti a pretínania lineárnych a afinných podpriestorov, t. j. na tzv. *štruktúru incidencie.* Popri tom však geometrický názor bol pre nás dôležitým zdrojom prvotných motivácií alebo dodatočných ilustrácií mnohých pojmov. To bolo umožnené hlavne tým, že hoci sme sa zaoberali vektorovými priestormi nad ľubovoľným poľom, ako typické príklady sme si pod nimi väčšinou predstavovali vektorové priestory malej dimenzie nad poľom reálnych čísel a vektory v nich ako orientované úsečky (s počiatkom v bode $\mathbf{0}$), ktoré majú určitú dĺžku, smer a orientáciu. Inak povedané, lineárnu algebru sme často vedome a ešte častejšie podvedome zasadzovali do rámca elementárnej geometrie roviny alebo priestoru.

Prísne vzaté nám však len samotná štruktúra vektorového priestoru (ani keby sme sa obmedzili iba na prípad poľa $\mathbb{R}$) neumožňuje vôbec hovoriť o dĺžke – takýto pojem v doterajšom kontexte nemá žiadny zmysel. Pritom práve dĺžka, a spolu s ňou tiež uhol sú základnými kvantitatívnymi veličinami elementárnej euklidovskej geometrie. Naplnenie lineárnej algebry geometrickým obsahom teda v prvom rade vyžaduje dať práve týmto pojmom určitý dobre zakotvený význam, ktorý – hoci by sa neopieral len o náš geometrický názor – bol by s ním v dobrej zhode. V tejto kapitole sa o to pokúsime vo vektorových priestoroch nad poľom $\mathbb{R}$. Ukazuje sa, že celú základnú geometrickú štruktúru, vrátane dĺžok a uhlov, možno odvodiť z jedinej kladne definitnej symetrickej bilineárnej formy na takomto priestore.

13.1 Skalárny súčin

Skalárnym alebo tiež *vnútorným súčinom* na reálnom vektorovom priestore V rozumieme ľubovoľnú kladne definitnú, symetrickú bilineárnu formu na V. Hodnotu tejto formy na vektoroch $\boldsymbol{x}, \boldsymbol{y} \in V$ budeme značiť $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$.

Nezávisle na znalosti uvedených pojmov možno skalárny súčin na V definovať ako binárnu operáciu $V \times V \to \mathbb{R}$, ktorá každej dvojici $(\boldsymbol{x}, \boldsymbol{y})$ vektorov z V priradí reálne číslo $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ také, že pre všetky $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x}_1, \boldsymbol{x}_2 \in V$ a ľubovoľné $c \in \mathbb{R}$ platí:

$$\begin{aligned}
\langle \boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y} \rangle &= \langle \boldsymbol{x}_1, \boldsymbol{y} \rangle + \langle \boldsymbol{x}_2, \boldsymbol{y} \rangle && \text{(aditivita)},\\
\langle c\boldsymbol{x}, \boldsymbol{y} \rangle &= c\langle \boldsymbol{x}, \boldsymbol{y} \rangle && \text{(homogenita)},\\
\langle \boldsymbol{x}, \boldsymbol{y} \rangle &= \langle \boldsymbol{y}, \boldsymbol{x} \rangle && \text{(symetria)},\\
\boldsymbol{x} \neq \mathbf{0} \ &\Rightarrow\ \langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0 && \text{(kladná definitnosť)}.
\end{aligned}$$

Spojenie aditivity a homogenity skalárneho súčinu dáva jeho linearitu ako funkcie prvej premennej (pri pevnej druhej premennej). Vďaka symetrii z toho vyplýva aj linearita skalárneho súčinu ako funkcie druhej premennej (pri pevnej prvej premennej), t. j. rovnosti

$$\begin{aligned} \langle \boldsymbol{x}, \boldsymbol{y}_1 + \boldsymbol{y}_2 \rangle &= \langle \boldsymbol{x}, \boldsymbol{y}_2 \rangle + \langle \boldsymbol{x}, \boldsymbol{y}_2 \rangle, \\ \langle \boldsymbol{x}, c\boldsymbol{y} \rangle &= c\langle \boldsymbol{x}, \boldsymbol{y} \rangle, \end{aligned}$$

pre všetky $\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2 \in V$ a $c \in \mathbb{R}$. Z (bi)linearity takisto vyplýva nasledujúci podrobnejší rozpis podmienky kladnej definitnosti

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0 \ \& \ \big(\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}\big)$$

pre každé $\boldsymbol{x} \in V$. Prvá časť tejto podmienky nám umožňuje definovať *normu* alebo *dĺžku vektora* $\boldsymbol{x}$ rovnosťou

$$\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}.$$

Výraz $\|\boldsymbol{x}\|^2$ treba zatiaľ chápať len ako iné označenie pre kvadratickú formu $\langle \boldsymbol{x}, \boldsymbol{x} \rangle$ indukovanú skalárnym súčinom. O oprávnenosti názvu „dĺžka“ ako aj o ďalších vlastnostiach normy podrobnejšie pojednáme až v paragrafe 13.3.

Euklidovským priestorom nazývame ľubovoľný *konečnorozmerný* reálny vektorový priestor so skalárnym súčinom. Hoci sa v tejto kapitole hodláme sústrediť práve na euklidovské priestory, všetky pojmy a výsledky, v ktorých konečnosť dimenzie nehrá podstatnú úlohu, sa budeme snažiť formulovať tak, aby zahŕňali všetky (teda i nekonečnorozmerné) priestory so skalárnym súčinom.

13.1.1. Príklad. Náš čitateľ sa už na strednej škole v rámci analytickej geometrie, prípadne v rámci fyziky, asi stretol so skalárnym súčinom $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1y_1 + x_2y_2$ v rovine $\mathbb{R}^2$ a so skalárnym súčinom $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1y_1 + x_2y_2 + x_3y_3$ v priestore $\mathbb{R}^3$. Ľahko sa možno presvedčiť, že rovnaká formulka funguje pre každé n, t. j. pre $\boldsymbol{x} = (x_1, \dots, x_n)^\mathsf{T}$, $\boldsymbol{y} = (y_1, \dots, y_n)^\mathsf{T}$ je predpisom

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^\mathsf{T} \cdot \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i$$

definovaný skalárny súčin na stĺpcovom vektorovom priestore $\mathbb{R}^n$. V prípade riadkového priestoru $\mathbb{R}^n$ máme

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x} \cdot \boldsymbol{y}^\mathsf{T} = \sum_{i=1}^{n} x_i y_i$$

pre $\boldsymbol{x} = (x_1, \dots, x_n)$, $\boldsymbol{y} = (y_1, \dots, y_n)$. Takýto skalárny súčin budeme nazývať *štandardným skalárnym súčinom* na $\mathbb{R}^n$. Štandardný skalárny súčin vektorov $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ (či už ide o riadkové alebo stĺpcové vektory) sa obvykle značí $\boldsymbol{x} \cdot \boldsymbol{y}$. Dĺžka vektora $\boldsymbol{x}$ vzhľadom na štandardný skalárny súčin je

$$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}.$$

V rámci analytickej geometrie sa pre nenulové vektory $\boldsymbol{x}$, $\boldsymbol{y}$ dokazuje známy vzťah

$$\boldsymbol{x} \cdot \boldsymbol{y} = \|\boldsymbol{x}\| \, \|\boldsymbol{y}\| \cos \alpha,$$

ktorý zväzuje štandardný skalárny súčin v $\mathbb{R}^2$ či v $\mathbb{R}^3$ s dĺžkou príslušných vektorov a nimi zvieraným uhlom α.

13.1.2. Príklad. Nech V označuje vektorový priestor $\mathcal{C}\langle a, b\rangle$ všetkých *spojitých* reálnych funkcií definovaných na uzavretom intervale $\langle a, b\rangle$, kde $a < b$ sú reálne čísla, prípadne jeho ľubovoľný lineárny podpriestor. Pre $f, g \in V$ položme

$$\langle f, g\rangle = \int_a^b f(x)g(x) \, \mathrm{d}x.$$

Z komutatívnosti násobenia v $\mathbb{R}$ a aditivity a homogenity integrálu vyplýva, že $\langle f, g\rangle$ je symetrická bilineárna forma na V (podrobne si premyslite ako). Na dôkaz kladnej definitnosti si stačí uvedomiť, že $f(x)^2 \geq 0$ pre každé $x \in \langle a, b\rangle$ a pre $f \neq \mathbf{0}$ (t. j. f nie identicky rovné nule) zo spojitosti funkcie f (teda aj f^2) vyplýva existencia nejakého netriviálneho uzavretého podintervalu $\langle a_1, b_1\rangle \subseteq \langle a, b\rangle$ takého, že $f(x)^2 > 0$ pre všetky $x \in \langle a_1, b_1\rangle$. Keďže f^2 na $\langle a_1, b_1\rangle$ nadobúda minimum $m = \min_{a_1 \leq x \leq b_1} f(x)^2 > 0$, máme

$$\langle f, f\rangle = \int_a^b f(x)^2 \, \mathrm{d}x \geq \int_{a_1}^{b_1} f(x)^2 \, \mathrm{d}x \geq (b_1 - a_1)m > 0.$$

Teda predpisom $(f, g) \mapsto \langle f, g\rangle$ je definovaný skalárny súčin na $\mathcal{C}\langle a, b\rangle$ ako aj na jeho ľubovoľnom lineárnom podpriestore, napr. na priestoroch polynómov $\mathbb{R}[x]$, $\mathbb{R}^{(n)}[x]$, $n \in \mathbb{N}$, uvažovaných ako spojité funkcie na $\langle a, b\rangle$. V prípade priestorov $\mathcal{C}\langle a, b\rangle$, či $\mathbb{R}[x]$ ide o skalárny súčin na *nekonečnorozmerných* vektorových priestoroch. Norma spojitej funkcie $f : \langle a, b\rangle \to \mathbb{R}$ potom je

$$\|f\| = \sqrt{\langle f, f\rangle} = \left(\int_a^b f(x)^2 \, \mathrm{d}x \right)^{1/2}.$$

13.2 Gramova matica a Cauchyho-Schwartzova nerovnosť

Nech $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ je ľubovoľná usporiadaná k-tica vektorov vo vektorovom priestore V so skalárnym súčinom. Takmer všetky podstatné informácie o týchto vektoroch sú ukryté v tzv. *Gramovej matici*

$$\boldsymbol{G}(\boldsymbol{\alpha}) = \boldsymbol{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k) = \big(\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle\big)_{k \times k}$$

vektorov $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$. Determinant Gramovej matice

$$\det \boldsymbol{G}(\boldsymbol{\alpha}) = |\boldsymbol{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)| = \begin{vmatrix} \langle \boldsymbol{u}_1, \boldsymbol{u}_1 \rangle & \dots & \langle \boldsymbol{u}_1, \boldsymbol{u}_k \rangle \\ \vdots & \ddots & \vdots \\ \langle \boldsymbol{u}_k, \boldsymbol{u}_1 \rangle & \dots & \langle \boldsymbol{u}_k, \boldsymbol{u}_k \rangle \end{vmatrix}$$

sa nazýva *Gramovým determinantom* vektorov $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$.

13.2.1. Tvrdenie. *Nech $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ sú ľubovoľné vektory vo vektorovom priestore V so skalárnym súčinom. Potom*
(a) $\boldsymbol{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ je kladne semidefinitná symetrická matica;
(b) vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ sú lineárne nezávislé práve vtedy, keď matica $\boldsymbol{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ je kladne definitná.

Dôkaz. Označme $\boldsymbol{G} = \boldsymbol{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$.

(a) Symetria matice $\boldsymbol{G}$ je priamym dôsledkom symetrie skalárneho súčinu. Zostáva dokázať, že pre ľubovoľný vektor $\boldsymbol{c} = (c_1, \dots, c_k) \in \mathbb{R}^k$ platí $\boldsymbol{c} \cdot \boldsymbol{G} \cdot \boldsymbol{c}^\mathsf{T} \geq 0$. Položme $\boldsymbol{v} = c_1 \boldsymbol{u}_1 + \ldots + c_k \boldsymbol{u}_k$. Potom z bilinearity a kladnej definitnosti skalárneho súčinu vyplýva

$$\boldsymbol{c} \cdot \boldsymbol{G} \cdot \boldsymbol{c}^\mathsf{T} = \sum_{i=1}^{k} \sum_{j=1}^{k} c_i c_j \langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = \langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0.$$

(b) Ak $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ sú lineárne nezávislé, tak tvoria bázu lineárneho podpriestoru $S = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_k] \subseteq V$. Zúženie skalárneho súčinu $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ na podpriestor S je skalárny súčin (t. j. kladne definitná symetrická bilineárna forma) na S. $\boldsymbol{G}$ je maticou tejto formy vzhľadom na bázu $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$, teda je to kladne definitná matica.

Ak $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ sú lineárne závislé, tak v $\mathbb{R}^n$ existuje vektor $\boldsymbol{c} = (c_1, \dots, c_k) \neq \boldsymbol{0}$ taký, že $\boldsymbol{v} = c_1 \boldsymbol{u}_1 + \ldots + c_k \boldsymbol{u}_k = \boldsymbol{0}$. Potom podľa (a)

$$\boldsymbol{c} \cdot \boldsymbol{G} \cdot \boldsymbol{c}^\mathsf{T} = \langle \boldsymbol{v}, \boldsymbol{v} \rangle = \langle \boldsymbol{0}, \boldsymbol{0} \rangle = 0,$$

teda $\boldsymbol{G}$ nie je kladne definitná.

Práve dokázané tvrdenie má spolu s vetou 12.2.4 nasledujúci dôsledok.

13.2.2. Dôsledok. *Pre ľubovoľné* $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k \in V$ *platí*

$$|\boldsymbol{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)| \geq 0.$$

Pritom $|\boldsymbol{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)| = 0$ *práve vtedy, keď vektory* $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ *sú lineárne závislé.*

Špeciálne pre ľubovoľné dva vektory $\boldsymbol{u}, \boldsymbol{v} \in V$ platí

$$|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})| = \begin{vmatrix} \langle \boldsymbol{u}, \boldsymbol{u} \rangle & \langle \boldsymbol{u}, \boldsymbol{v} \rangle \\ \langle \boldsymbol{v}, \boldsymbol{u} \rangle & \langle \boldsymbol{v}, \boldsymbol{v} \rangle \end{vmatrix} = \langle \boldsymbol{u}, \boldsymbol{u} \rangle \langle \boldsymbol{v}, \boldsymbol{v} \rangle - \langle \boldsymbol{u}, \boldsymbol{v} \rangle \langle \boldsymbol{v}, \boldsymbol{u} \rangle = \|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2 - \langle \boldsymbol{u}, \boldsymbol{v} \rangle^2 \geq 0,$$

pričom rovnosť nastane práve vtedy, keď vektory $\boldsymbol{u}$, $\boldsymbol{v}$ sú lineárne závislé. Tým sme dokázali nasledujúci vzťah, známy ako *Cauchyho-Schwartzova nerovnosť.*

13.2.3. Veta. *Pre ľubovoľné vektory* $\boldsymbol{u}, \boldsymbol{v} \in V$ *v priestore* V *so skalárnym súčinom platí*

$$|\langle \boldsymbol{u}, \boldsymbol{v} \rangle| \leq \|\boldsymbol{u}\| \, \|\boldsymbol{v}\| \, ,$$

pričom rovnosť nastane práve vtedy, keď vektory $\boldsymbol{u}$, $\boldsymbol{v}$ *sú lineárne závislé.*

13.3 Dĺžka vektora a uhol dvoch vektorov

Normou na reálnom vektorovom priestore V rozumieme ľubovoľné zobrazenie $V \to \mathbb{R}$, ktoré vektoru $\boldsymbol{x} \in V$ priradí reálne číslo $\|\boldsymbol{x}\|$, nazývané *normou* alebo tiež *dĺžkou vektora* $\boldsymbol{x}$, také, že pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ a ľubovoľné $c \in \mathbb{R}$ platí

$$\begin{aligned} \|\boldsymbol{x} + \boldsymbol{y}\| &\leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\| && \text{(trojuholníková nerovnosť)}, \\ \|c\boldsymbol{x}\| &= |c| \, \|\boldsymbol{x}\| && \text{(pozitívna homogenita)}, \\ \|\boldsymbol{x}\| = 0 \ &\Rightarrow \ \boldsymbol{x} = \boldsymbol{0} && \text{(oddeliteľnosť)}. \end{aligned}$$

Z uvedených podmienok vyplýva nezápornosť normy, t. j. $\|\boldsymbol{x}\| \geq 0$ pre každé $\boldsymbol{x} \in V$. Keďže vďaka pozitívnej homogenite platí $\|\boldsymbol{0}\| = 0$ a $\|-\boldsymbol{x}\| = \|\boldsymbol{x}\|$, s použitím trojuholníkovej nerovnosti naozaj dostávame

$$\|\boldsymbol{x}\| = \frac{1}{2}\big(\|\boldsymbol{x}\| + \|-\boldsymbol{x}\|\big) \geq \frac{1}{2} \, \|\boldsymbol{x} - \boldsymbol{x}\| = \frac{1}{2} \, \|\boldsymbol{0}\| = 0.$$

Navyše, vďaka oddeliteľnosti máme $\|\boldsymbol{x}\| > 0$ pre každé $\boldsymbol{0} \neq \boldsymbol{x} \in V$; inak povedané, každý nenulový vektor môžeme pomocou jeho normy „oddeliť" od nulového vektora.

Reálny vektorový priestor s normou nazývame *normovaný priestor.* Intuitívne sa na normovaný priestor dívame ako na vektorový priestor, v ktorom možno merať dĺžky vektorov. Tri definujúce podmienky pre normu zaručujú, že takéto meranie dĺžok, t. j. priradenie $\boldsymbol{x} \mapsto \|\boldsymbol{x}\|$, má rozumné vlastnosti, aké od dĺžok očakávame.

Vzdialenosťou bodov $\boldsymbol{x}$, $\boldsymbol{y}$ vo vektorovom priestore V s normou $\|\cdot\|$ nazývame dĺžku vektora $\boldsymbol{x}-\boldsymbol{y}$, t. j. číslo $\|\boldsymbol{x}-\boldsymbol{y}\|$. Pomocou vzdialeností bodov možno trojuholníkovú nerovnosť vyjadriť iným, ekvivalentným spôsobom, ktorý vari ešte názornejšie osvetľuje jej názov:

$$\|\boldsymbol{x}-\boldsymbol{y}\|+\|\boldsymbol{y}-\boldsymbol{z}\| \geq \|\boldsymbol{x}-\boldsymbol{z}\|$$

pre všetky $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$.

Všeobecnou problematikou normovaných priestorov sa v tomto kurze nebudeme zaoberať. Obmedzíme sa len na normy, ktoré pochádzajú zo skalárnych súčinov.

13.3.1. Tvrdenie. *Nech V je reálny vektorový priestor so skalárnym súčinom. Potom rovnosťou*

$$\|\boldsymbol{x}\|=\sqrt{\langle\boldsymbol{x}, \boldsymbol{x}\rangle}$$

je definovaná norma na V.

Dôkaz. Zvoľme $\boldsymbol{x}, \boldsymbol{y} \in V$. S použitím bilinearity a symetrie skalárneho súčinu a Cauchyho-Schwartzovej nerovnosti (veta 13.2.3) dostávame

$$\begin{aligned}\|\boldsymbol{x}+\boldsymbol{y}\|^2 &= \langle\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}\rangle=\langle\boldsymbol{x}, \boldsymbol{x}\rangle+\langle\boldsymbol{x}, \boldsymbol{y}\rangle+\langle\boldsymbol{y}, \boldsymbol{x}\rangle+\langle\boldsymbol{y}, \boldsymbol{y}\rangle \\ &= \|\boldsymbol{x}\|^2+2\langle\boldsymbol{x}, \boldsymbol{y}\rangle+\|\boldsymbol{y}\|^2 \leq\|\boldsymbol{x}\|^2+2\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|+\|\boldsymbol{y}\|^2=\big(\|\boldsymbol{x}\|+\|\boldsymbol{y}\|\big)^2.\end{aligned}$$

To dokazuje trojuholníkovú nerovnosť. Jednoduchý dôkaz ďalších dvoch podmienok prenechávame čitateľovi.

Z Cauchyho-Schwartzovej nerovnosti vyplýva, že pre ľubovoľné *nenulové* vektory $\boldsymbol{x}$, $\boldsymbol{y}$ vo vektorovom priestore so skalárnym súčinom platí

$$-1 \leq \frac{\langle\boldsymbol{x}, \boldsymbol{y}\rangle}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|} \leq 1.$$

Preto existuje jediné reálne číslo α také, že $0 \leq \alpha \leq \pi$ a

$$\cos \alpha=\frac{\langle\boldsymbol{x}, \boldsymbol{y}\rangle}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|}.$$

Číslo α nazývame *uhlom* alebo tiež *odchýlkou vektorov* $\boldsymbol{x}$, $\boldsymbol{y}$ a značíme ho $\alpha=\sphericalangle(\boldsymbol{x}, \boldsymbol{y})$. Zo symetrie skalárneho súčinu vyplýva $\sphericalangle(\boldsymbol{x}, \boldsymbol{y})=\sphericalangle(\boldsymbol{y}, \boldsymbol{x})$, to znamená, že ide o *neorientovaný uhol.*

Pri takejto definícii uhla dvoch nenulových vektorov zostáva vzťah

$$\langle\boldsymbol{x}, \boldsymbol{y}\rangle=\|\boldsymbol{x}\|\,\|\boldsymbol{y}\| \cos \sphericalangle(\boldsymbol{x}, \boldsymbol{y}),$$

platný pre štandardný skalárny súčin v $\mathbb{R}^2$ a $\mathbb{R}^3$, zachovaný v ľubovoľnom priestore so skalárnym súčinom. Treba si však uvedomiť, že z logického hľadiska postupujeme obrátene ako v stredoškolskej analytickej geometrii. Tam totiž vopred vieme (či aspoň sa tak tvárime), čo sú to dĺžky a uhly vektorov, a pomocou nich definujeme skalárny súčin uvedeným vzťahom, prípadne rovnosťou $\boldsymbol{x} \cdot \boldsymbol{y} = \sum x_i y_i$, kedy musíme uvedený vzťah dokázať. Pri našom postupe vychádzame z pojmu skalárneho súčinu a pomocou neho definujeme dĺžky a uhly vektorov tak, že platnosť klasických pomočiek analytickej geometrie zostáva zachovaná, ba dokonca ju rozširujeme na vektorové priestory ľubovoľnej konečnej i nekonečnej dimenzie vybavené skalárnym súčinom.

Hovoríme, že vektory $\boldsymbol{x}, \boldsymbol{y} \in V$ sú (navzájom) *kolmé* alebo tiež *ortogonálne*, označenie $\boldsymbol{x} \perp \boldsymbol{y}$, ak $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$.

Teraz uvedieme niekoľko bezprostredných dôsledkov našich definícií. Ich jednoduché dôkazy prenechávame ako cvičenie čitateľovi.

13.3.2. Tvrdenie. *Nech V je vektorový priestor so skalárnym súčinom. Potom pre ľubovoľné nenulové vektory $\boldsymbol{x}, \boldsymbol{y} \in V$ platí:*

(a) $\sphericalangle(\boldsymbol{x}, \boldsymbol{y}) = 0 \iff (\exists\, c \in \mathbb{R})(c > 0 \;\&\; \boldsymbol{x} = c\boldsymbol{y})$;

(b) $\sphericalangle(\boldsymbol{x}, \boldsymbol{y}) = \pi \iff (\exists\, c \in \mathbb{R})(c < 0 \;\&\; \boldsymbol{x} = c\boldsymbol{y})$;

(c) $\sphericalangle(\boldsymbol{x}, \boldsymbol{y}) = \pi/2 \iff \boldsymbol{x} \perp \boldsymbol{y}$;

(d) $\sphericalangle(-\boldsymbol{x}, -\boldsymbol{y}) = \sphericalangle(\boldsymbol{x}, \boldsymbol{y}), \quad \sphericalangle(-\boldsymbol{x}, \boldsymbol{y}) = \sphericalangle(\boldsymbol{x}, -\boldsymbol{y}) = \pi - \sphericalangle(\boldsymbol{x}, \boldsymbol{y})$.

Pokračujeme štyrmi poučkami klasickej analytickej geometrie.

13.3.3. Tvrdenie. *Nech V je vektorový priestor so skalárnym súčinom. Potom pre ľubovoľné nenulové vektory $\boldsymbol{x}, \boldsymbol{y} \in V$ platí:*

(a) (kosinusová veta)

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2\,\|\boldsymbol{x}\|\,\|\boldsymbol{y}\| \cos\sphericalangle(\boldsymbol{x}, \boldsymbol{y}),$$

$$\|\boldsymbol{x} - \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 - 2\,\|\boldsymbol{x}\|\,\|\boldsymbol{y}\| \cos\sphericalangle(\boldsymbol{x}, \boldsymbol{y});$$

(b) (Pytagorova veta)

$$\boldsymbol{x} \perp \boldsymbol{y} \Rightarrow \|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x} - \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2;$$

(c) (pravidlo rovnobežníka)

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 + \|\boldsymbol{x} - \boldsymbol{y}\|^2 = 2\big(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\big);$$

(d) (uhlopriečky kosoštvorca sú na seba kolmé)

$$\|\boldsymbol{x}\| = \|\boldsymbol{y}\| \Rightarrow \boldsymbol{x} + \boldsymbol{y} \perp \boldsymbol{x} - \boldsymbol{y}.$$

(Samozrejme, tvrdenia (b), (c), (d) platia aj bez predpokladu nenulovosti vektorov $\boldsymbol{x}$, $\boldsymbol{y}$ – vidno to i z nášho dôkazu.)

Dôkaz. (a) Ako sme ukázali v dôkaze tvrdenia 13.3.1, pre ľubovoľné $\boldsymbol{x}, \boldsymbol{y} \in V$ máme

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle.$$

Z toho už okamžite vyplýva prvá podoba kosinusovej vety; jej druhú podobu dostaneme z prvej na základe 13.3.2 (d).

(b) Pytagorova veta je zvláštnym prípadom kosinusovej vety

(c) Pravidlo rovnobežníka dostaneme sčítaním oboch verzií kosinusovej vety.

(d) priamo vyplýva z identity $\langle \boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y}\rangle = \|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2$.

13.4 Ortogonálne a ortonormálne bázy

Usporiadaná k-tica $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ vektorov z vektorového priestoru so skalárnym súčinom V sa nazýva *ortogonálna*, ak $\boldsymbol{u}_i \perp \boldsymbol{u}_j$ pre všetky $1 \le i < j \le k$. Voľne tiež hovoríme, že *vektory* $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ sú (navzájom) *ortogonálne* alebo *kolmé*. Usporiadaná k-tica $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ sa nazýva *ortonormálna*, ak je ortogonálna a $\|\boldsymbol{u}_i\| = 1$ pre všetky $i \le k$. Taktiež hovoríme, že vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ tvoria *ortonormálny systém*. Podobne možno definovať pojmy ortogonálnosti a ortonormálnosti aj pre nekonečné postupnosti $(\boldsymbol{u}_k)_{k=0}^{\infty}$ vektorov z V, prípadne pre množiny $X \subseteq V$.

Nasledujúce tvrdenie je bezprostredným dôsledkom našich definícií.

13.4.1. Tvrdenie. *Nech V je vektorový priestor so skalárnym súčinom. Potom pre ľubovoľnú k-ticu $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_k) \in V^k$ platí:*

(a) $\boldsymbol{\alpha}$ je ortogonálna práve vtedy, keď jej Gramova matica $\boldsymbol{G}(\boldsymbol{\alpha})$ je diagonálna;

(b) $\boldsymbol{\alpha}$ je ortonormálna práve vtedy, keď $\boldsymbol{G}(\boldsymbol{\alpha}) = \boldsymbol{I}_k$.

Z posledného tvrdenia a dôsledku 13.2.2 priamo vyplýva

13.4.2. Dôsledok. *Ak $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k \in V$ sú navzájom kolmé nenulové vektory, špeciálne, ak $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ tvoria ortonormálny systém, tak sú lineárne nezávislé.*

V priestore $\mathbb{R}^n$ so štandardným skalárnym súčinom kanonická báza $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ je zrejme ortonormálna a v zhode s tvrdením 13.4.1 (b) možno ľahko nahliadnuť rovnosť $\boldsymbol{G}(\boldsymbol{\varepsilon}) = \boldsymbol{I}_n$. Ortogonálne a ortonormálne bázy však existujú v ľubovoľných euklidovských priestoroch.

13.4.3. Veta. *Každý euklidovský priestor má ortonormálnu bázu.*

Dôkaz Nech V je euklidovský priestor. Keďže skalárny súčin je kladne definitná, symetrická bilineárna forma na konečnorozmernom reálnom priestore V, na základe vety 11.3.5, dôsledku 12.1.3 a tvrdenia 12.2.1 existuje taká báza $\boldsymbol{\beta}$ priestoru V, vzhľadom na ktorú má tento súčin jednotkovú maticu. Inak povedané, $\boldsymbol{G}(\boldsymbol{\beta}) = \boldsymbol{I}_n$, čiže $\boldsymbol{\beta}$ je ortonormálna báza.

Nasledujúce tvrdenie zhŕňa niekoľko užitočných vlastností ortonormálnych báz v euklidovskom priestore. Podmienky (b) a (c), ktoré sú ekvivalentné, nesú spoločný názov *Parsevalova rovnosť*.[1]

13.4.4. Tvrdenie. *Nech* $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ *je ortonormálna báza euklidovského priestoru* V. *Potom pre ľubovoné vektory* $\boldsymbol{x}, \boldsymbol{y} \in V$ *platí:*

(a) $\boldsymbol{x} = \sum_{i=1}^{n} \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle \, \boldsymbol{u}_i$, *t. j.* $(\boldsymbol{x})_{\boldsymbol{\alpha}} = \big(\langle \boldsymbol{x}, \boldsymbol{u}_1 \rangle, \dots, \langle \boldsymbol{x}, \boldsymbol{u}_n \rangle\big)^{\mathsf{T}}$;

(b) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle \langle \boldsymbol{u}_i, \boldsymbol{y} \rangle$;

(c) $\|\boldsymbol{x}\|^2 = \sum_{i=1}^{n} \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle^2$.

Dôkaz. (a) Keďže $\boldsymbol{\alpha}$ je báza, $\boldsymbol{x}$ možno vyjadriť v tvare $\boldsymbol{x} = \sum_{i=1}^{n} c_i \boldsymbol{u}_i$ pre jednoznačne určené koeficienty $c_1, \dots, c_n$. Z ortonormálnosti vektorov $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ vyplýva

$$\langle \boldsymbol{x}, \boldsymbol{u}_j \rangle = \sum_{i=1}^{n} c_i \langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = c_j$$

pre každé $1 \le j \le n$.

(b) Ak $\boldsymbol{\alpha}$ je ortonormálna báza, tak matica skalárneho súčinu vzhľadom na bázu $\boldsymbol{\alpha}$ je $\boldsymbol{G}(\boldsymbol{\alpha}) = \boldsymbol{I}_n$. Preto $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = (\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot (\boldsymbol{y})_{\boldsymbol{\alpha}}$ pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$. Potrebný záver vyplýva z (a).

(c) je špeciálnym prípadom (b).

Pre istotu ešte si ešte raz zopakujme, čo vyplýva z tvrdenia 13.4.1, vety 13.4.3 a tvrdenia 13.4.4: V ľubovoľnom n-rozmernom euklidovskom priestore V skalárny súčin vektorov $\boldsymbol{x}, \boldsymbol{y} \in V$ splýva so štandardným skalárnym súčinom v $\mathbb{R}^n$

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i y_i = (\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot (\boldsymbol{y})_{\boldsymbol{\alpha}}$$

ich súradníc $(\boldsymbol{x})_{\boldsymbol{\alpha}} = (x_1, \dots, x_n)^{\mathsf{T}}$, $(\boldsymbol{y})_{\boldsymbol{\alpha}} = (y_1, \dots, y_n)^{\mathsf{T}}$ vzhľadom na ľubovoľnú ortonormálnu bázu $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ priestoru V. Podmienka (a) posledného tvrdenia nám navyše udáva explicitný tvar týchto súradníc: $x_i = \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle$, $y_i = \langle \boldsymbol{y}, \boldsymbol{u}_i \rangle$.

Teraz popíšeme algoritmus, ktorý umožňuje zostrojiť ortonormálne bázy, známy ako *Gramov-Schmidtov ortogonalizačný proces*, skrátene *GSO-proces*.

Nech V je vektorový priestor so skalárnym súčinom a $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n \in V$ sú lineárne nezávislé vektory. Chceme zostrojiť ortogonálne vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ tak, aby pre každé $k \le n$ platilo

$$[\boldsymbol{v}_1, \dots, \boldsymbol{v}_k] = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_k].$$

[1] Týmto názvom sa zvyknú označovať aj isté nekonečnorozmerné varianty uvedených rovností.

Ak položíme $\boldsymbol{v}_1 = \boldsymbol{u}_1$, tak samozrejme $[\boldsymbol{v}_1] = [\boldsymbol{u}_1]$. Vektor $\boldsymbol{v}_2 \in [\boldsymbol{u}_1, \boldsymbol{u}_2] = [\boldsymbol{v}_1, \boldsymbol{u}_2]$ budeme hľadať v tvare $\boldsymbol{v}_2 = a\boldsymbol{v}_1 + b\boldsymbol{u}_2$, kde $a, b \in \mathbb{R}$. Ak má však platiť $[\boldsymbol{v}_1, \boldsymbol{v}_2] = [\boldsymbol{u}_1, \boldsymbol{u}_2]$, musí byť $b \neq 0$. To najjednoduchšie dosiahneme voľbou $b = 1$. Navyše vektor $\boldsymbol{v}_2 = \boldsymbol{u}_2 + a\boldsymbol{v}_1$ má byť kolmý na vektor $\boldsymbol{v}_1$, t. j.

$$0 = \langle \boldsymbol{v}_2, \boldsymbol{v}_1 \rangle = \langle \boldsymbol{u}_2, \boldsymbol{v}_1 \rangle + a \langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle.$$

Odtiaľ dostávame

$$a = -\frac{\langle \boldsymbol{u}_2, \boldsymbol{v}_1 \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle}.$$

Predpokladajme, že $2 \leq k \leq n$ a už sme zostrojili ortogonálne vektory $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}$ také, že pre každé $1 \leq i \leq k-1$ platí $[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i] = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_i]$. Poučení prípadom $k = 2$ budeme vektor $\boldsymbol{v}_k \in [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}, \boldsymbol{u}_k] = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{u}_k]$ hľadať v tvare $\boldsymbol{v}_k = \boldsymbol{u}_k + a_1\boldsymbol{v}_1 + \cdots + a_{k-1}\boldsymbol{v}_{k-1}$, kde $a_1, \ldots, a_{k-1} \in \mathbb{R}$. Navyše vektor $\boldsymbol{v}_k$ musí byť kolmý na každý z vektorov $\boldsymbol{v}_i$ pre $1 \leq i \leq k-1$, čiže

$$0 = \langle \boldsymbol{v}_k, \boldsymbol{v}_i \rangle = \langle \boldsymbol{u}_k, \boldsymbol{v}_i \rangle + \sum_{j=1}^{k-1} a_j \langle \boldsymbol{v}_j, \boldsymbol{v}_i \rangle = \langle \boldsymbol{u}_k, \boldsymbol{v}_i \rangle + a_i \langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle,$$

lebo podľa nášho predpokladu $\langle \boldsymbol{v}_j, \boldsymbol{v}_i \rangle = 0$ pre $j \neq i$. Z toho dôvodu

$$a_i = -\frac{\langle \boldsymbol{u}_k, \boldsymbol{v}_i \rangle}{\langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle}.$$

Tým sme dokázali nasledujúcu vetu.

13.4.5. Veta. *Nech V je vektorový priestor so skalárnym súčinom a $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in V$ sú lineárne nezávislé vektory. Vektory $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in V$ definujeme rekurziou:*

$$\boldsymbol{v}_1 = \boldsymbol{u}_1 \qquad \text{a} \qquad \boldsymbol{v}_k = \boldsymbol{u}_k - \sum_{i=1}^{k-1} \frac{\langle \boldsymbol{u}_k, \boldsymbol{v}_i \rangle}{\langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle} \boldsymbol{v}_i,$$

pre $1 < k \leq n$. Potom $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ sú ortogonálne vektory a pre každé $1 \leq k \leq n$ platí

$$[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k] = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k].$$

Špeciálne, ak $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ je báza priestoru V, tak $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ je ortogonálna báza priestoru V; potom vektory $\|\boldsymbol{v}_1\|^{-1}\boldsymbol{v}_1, \ldots, \|\boldsymbol{v}_n\|^{-1}\boldsymbol{v}_n$ tvoria ortonormálnu bázu priestoru V.

Poznámka. Gramov-Schmidtov ortogonalizačný proces funguje aj pre nekonečné postupnosti $(\boldsymbol{u}_k)_{k=0}^{\infty}$ lineárne nezávislých vektorov v nekonečnorozmernom vektorovom priestore so skalárnym súčinom V. Rekurziou cez množinu všetkých prirodzených čísel

$$\boldsymbol{v}_0 = \boldsymbol{u}_0 \qquad \text{a} \qquad \boldsymbol{v}_k = \boldsymbol{u}_k - \sum_{i=0}^{k-1} \frac{\langle \boldsymbol{u}_k, \boldsymbol{v}_i \rangle}{\langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle} \boldsymbol{v}_i,$$

pre $k > 0$, je i v tomto prípade definovaná ortogonálna postupnosť $(\boldsymbol{v}_k)_{k=0}^{\infty}$ taká, že

$$[\boldsymbol{v}_0, \dots, \boldsymbol{v}_k] = [\boldsymbol{u}_0, \dots, \boldsymbol{u}_k]$$

pre každé $k \in \mathbb{N}$. Z toho už vyplýva lineárna nezávislosť postupnosti $(\boldsymbol{v}_k)_{k=0}^{\infty}$. Ak $(\boldsymbol{u}_k)_{k=0}^{\infty}$ bola bázou V, tak $(\boldsymbol{v}_k)_{k=0}^{\infty}$ je ortogonálna a $\left(\|\boldsymbol{v}_k\|^{-1}\boldsymbol{v}_k\right)_{k=0}^{\infty}$ ortonormálna báza vo V.

13.4.6. Príklad. Na základe Sylvestrovho kritéria (veta 12.2.4) ľahko nahliadneme, že symetrická matica

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 3 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

je kladne definitná. To znamená, že predpisom

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \boldsymbol{y}$$

je definovaný skalárny súčin na (stĺpcovom) priestore $\mathbb{R}^3$. Aplikáciou GSO-procesu na kanonickú bázu $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ zostrojíme ortogonálnu bázu $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ priestoru $\mathbb{R}^3$ s uvedeným skalárnym súčinom:

$$\begin{aligned}
\boldsymbol{v}_1 &= \boldsymbol{e}_1 = (1,0,0)^{\mathsf{T}}, \\
\boldsymbol{v}_2 &= \boldsymbol{e}_2 - \frac{\langle \boldsymbol{e}_2, \boldsymbol{v}_1 \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle}\boldsymbol{v}_1 = \boldsymbol{e}_2 - \frac{1}{2}\boldsymbol{v}_1 = (-1/2, 1, 0)^{\mathsf{T}}, \\
\boldsymbol{v}_3 &= \boldsymbol{e}_3 - \frac{\langle \boldsymbol{e}_3, \boldsymbol{v}_1 \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle}\boldsymbol{v}_1 - \frac{\langle \boldsymbol{e}_3, \boldsymbol{v}_2 \rangle}{\langle \boldsymbol{v}_2, \boldsymbol{v}_2 \rangle}\boldsymbol{v}_2 = \boldsymbol{e}_3 + \frac{1}{2}\boldsymbol{v}_1 - \frac{1}{3}\boldsymbol{v}_2 = (2/3, -1/3, 1)^{\mathsf{T}},
\end{aligned}$$

keďže $\langle \boldsymbol{e}_2, \boldsymbol{v}_1 \rangle = 1$, $\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle = 2$, $\langle \boldsymbol{e}_3, \boldsymbol{v}_1 \rangle = -1$, $\langle \boldsymbol{e}_3, \boldsymbol{v}_2 \rangle = 1/2$ a $\langle \boldsymbol{v}_2, \boldsymbol{v}_2 \rangle = 3/2$. Ak ešte dopočítame $\langle \boldsymbol{v}_3, \boldsymbol{v}_3 \rangle = 7/3$, dostaneme dĺžky jednotlivých vektorov $\|\boldsymbol{v}_1\| = \sqrt{2}$, $\|\boldsymbol{v}_2\| = \sqrt{3/2}$, $\|\boldsymbol{v}_3\| = \sqrt{7/3}$. Príslušná ortonormálna báza je potom tvorená vektormi

$$\frac{1}{\sqrt{2}}\boldsymbol{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \sqrt{\frac{2}{3}}\boldsymbol{v}_2 = \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \qquad \sqrt{\frac{3}{7}}\boldsymbol{v}_3 = \frac{1}{\sqrt{21}}\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}.$$

13.4.7. Príklad. Na vektorovom priestore $\mathbb{R}[x]$ všetkých reálnych polynómov v premennej x je daný skalárny súčin $\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)\,\mathrm{d}x$ (pozri príklad 13.1.2). Postupnosť $(x^n)_{n=0}^{\infty}$ všetkých mocnín premennej x tvorí bázu v $\mathbb{R}[x]$. GSO-procesom z nej zostrojíme ortogonálnu bázu $\left(L_n(x)\right)_{n=0}^{\infty}$ priestoru $\mathbb{R}[x]$ – jej prvky sa niekedy nazývajú *Legendreove polynómy.* Ak si uvedomíme, že

$$\langle x^m, x^n \rangle = \int_{-1}^{1} x^m x^n \,\mathrm{d}x = \left[\frac{x^{m+n+1}}{m+n+1}\right]_{-1}^{1} = \begin{cases} \frac{2}{m+n+1}, & \text{ak } m+n \text{ je párne,} \\ 0, & \text{ak } m+n \text{ je nepárne,} \end{cases}$$

postupne dostaneme

$$\begin{aligned}
L_0(x) &= 1,\\
L_1(x) &= x - \frac{\langle x, L_0\rangle}{\langle L_0, L_0\rangle}\, L_0(x) = x,\\
L_2(x) &= x^2 - \frac{\langle x^2, L_0\rangle}{\langle L_0, L_0\rangle}\, L_0(x) - \frac{\langle x^2, L_1\rangle}{\langle L_1, L_1\rangle}\, L_1(x) = x^2 - \frac{1}{3},\\
L_3(x) &= x^3 - \frac{\langle x^3, L_0\rangle}{\langle L_0, L_0\rangle}\, L_0(x) - \frac{\langle x^3, L_1\rangle}{\langle L_1, L_1\rangle}\, L_1(x) - \frac{\langle x^3, L_2\rangle}{\langle L_2, L_2\rangle}\, L_2(x) = x^3 - \frac{3}{5}\, x,
\end{aligned}$$

atď. Vidíme, že $L_n(x) = x^n - \dots$ je polynóm n-tého stupňa, ktorý pre párne n obsahuje len párne mocniny premennej x a pre nepárne n len nepárne mocniny x. S trochou námahy možno odvodiť explicitné vyjadrenie

$$L_n(x) = \frac{n!}{(2n)!}\, \frac{\mathrm{d}^n}{\mathrm{d}x^n}\, (x^2-1)^n.$$

V literatúre sa *Legendreovými polynómami* zvyčajne nazývajú násobky

$$P_n(x) = \frac{1}{2^n}\binom{2n}{n} L_n(x) = \frac{1}{2^n n!}\, \frac{\mathrm{d}^n}{\mathrm{d}x^n}\, (x^2-1)^n$$

polynómov $L_n(x)$, normované tak, aby pre každé n platilo $P_n(1) = 1$.

Gramovu-Schmidtovu ortogonalizáciu možno alternatívne vykonať len vhodnými úpravami istých matíc. Ak $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ je lineárne nezávislá n-tica vektorov v priestore so skalárnym súčinom V, tak podľa tvrdenia 13.2.1 (b) jej Gramova matica $\boldsymbol{G}(\boldsymbol{\alpha})$ je kladne definitná a je to matica skalárneho súčinu zúženého na lineárny podpriestor $S = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n] \subseteq V$ v báze $\boldsymbol{\alpha}$. Podľa vety 12.2.3 ju možno výlučne úpravami typu (1^+) upraviť na diagonálny tvar $\boldsymbol{D} = \operatorname{diag}(d_1, \dots, d_n)$ s kladnými prvkami na diagonále. Elementárne stĺpcové operácie zodpovedajúce týmto úpravám postupne vykonané na matici $\boldsymbol{I}_n$ nás privedú k hornej trojuholníkovej matici $\boldsymbol{P} = (p_{ij}) \in \mathbb{R}^{n\times n}$ s jednotkami na diagonále, t. j. $p_{ii} = 1$ a $p_{ij} = 0$ pre každé $i \le n$ a $j < i$. Ak položíme

$$\boldsymbol{\alpha} \cdot \boldsymbol{P} = \boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n),$$

tak $\boldsymbol{P} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$, t. j. $\boldsymbol{P}$ je maticou prechodu z bázy $\boldsymbol{\beta}$ do bázy $\boldsymbol{\alpha}$ (pozri paragraf 7.5). Potom $\boldsymbol{D}$ je maticou skalárneho súčinu zúženého na podpriestor S v báze $\boldsymbol{\beta}$, teda $\boldsymbol{G}(\boldsymbol{\beta}) = \boldsymbol{D}$, takže $\boldsymbol{\beta}$ je ortogonálna báza S. Navyše, vzhľadom na tvar matice $\boldsymbol{P}$, platí

$$\boldsymbol{v}_1 = \boldsymbol{u}_1 \qquad \text{a} \qquad \boldsymbol{v}_k = \boldsymbol{u}_k + \sum_{i=1}^{k-1} p_{ik}\boldsymbol{u}_i,$$

pre $1 < k \leq n$, z čoho možno ľahko nahliadnuť rovnosti

$$[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k] = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k]$$

pre všetky $k \leq n$. Taktiež normy vektorov $\boldsymbol{v}_k$ si možno prečítať priamo z matice $\boldsymbol{D}$. Platí totiž $d_k = \|\boldsymbol{v}_k\|^2$ pre každé $k \leq n$. Teda príslušná ortonormálna báza je tvorená vektormi $(1/\sqrt{d_1})\boldsymbol{v}_1, \ldots, (1/\sqrt{d_n})\boldsymbol{v}_n$. Tento posledný krok možno samozrejme realizovať úpravami typu (4) matice $\boldsymbol{G}(\boldsymbol{\beta})$ a príslušnými ESO na matici $\boldsymbol{P}$ (pozri paragraf 11.3).

Ak $V = \mathbb{R}^m$ s akýmkoľvek (t. j. nie nutne štandardným skalárnym súčinom), tak po stotožnení bázy $\boldsymbol{\alpha}$ s maticou, ktorej stĺpce sú vektory tejto bázy, možno k báze $\boldsymbol{\beta}$ dospieť priamo (t. j. bez matice $\boldsymbol{P}$), vykonaním príslušných ESO na matici $\boldsymbol{\alpha}$.

13.4.8. Príklad. Prepočítajme si ešte raz príklad 13.4.6 práve opísanou metódou. Uvedená matica $\boldsymbol{A}$ je zároveň Gramovou maticou $\boldsymbol{G}(\boldsymbol{\varepsilon})$. Najprv pomocou prvku na mieste $(1,1)$ vynulujeme ostatné prvky prvého riadku i stĺpca. Potom pomocou prvku na mieste $(2,2)$ vynulujeme prvky na miestach $(2,3)$ a $(3,2)$:

$$\boldsymbol{G}(\boldsymbol{\varepsilon}) = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 3 \end{pmatrix} \equiv \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3/2 & 1/2 \\ 0 & 1/2 & 5/2 \end{pmatrix} \equiv \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3/2 & 0 \\ 0 & 0 & 7/3 \end{pmatrix} = \boldsymbol{G}(\boldsymbol{\beta}).$$

Zrejme šlo o úpravy typu (1^+). Vykonaním príslušných ESO na jednotkovej matici dostaneme

$$\boldsymbol{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \wr \begin{pmatrix} 1 & -1/2 & 1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \wr \begin{pmatrix} 1 & -1/2 & 2/3 \\ 0 & 1 & -1/3 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{\beta}.$$

Čitateľ by si mal samostatne premyslieť detaily výpočtu. Vidíme, že výsledné matice sa presne zhodujú s výsledkom príkladu 13.4.6.

13.4.9. Príklad. Vráťme sa ešte k príkladu 13.4.7. Bez podrobnejšieho komentára zortogonalizujeme systém $(1, x, x^2, x^3)$ prvých štyroch mocnín x v $\mathbb{R}[x]$. Postupnými úpravami typu (1^+) Gramovej matice $\boldsymbol{G}(1, x, x^2, x^3)$ dostaneme

$$\boldsymbol{G}(1,x,x^2,x^3) = \begin{pmatrix} 2 & 0 & 2/3 & 0 \\ 0 & 2/3 & 0 & 2/5 \\ 2/3 & 0 & 2/5 & 0 \\ 0 & 2/5 & 0 & 2/7 \end{pmatrix} \equiv \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2/3 & 0 & 2/5 \\ 0 & 0 & 8/45 & 0 \\ 0 & 2/5 & 0 & 2/7 \end{pmatrix}$$

$$\equiv \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2/3 & 0 & 0 \\ 0 & 0 & 8/45 & 0 \\ 0 & 0 & 0 & 8/175 \end{pmatrix} = \boldsymbol{G}(L_0,L_1,L_2,L_3).$$

Príslušné ESO vykonané na jednotkovej matici dávajú

$$\boldsymbol{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \wr \begin{pmatrix} 1 & 0 & -1/3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \wr \begin{pmatrix} 1 & 0 & -1/3 & 0 \\ 0 & 1 & 0 & -3/5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \boldsymbol{P}.$$

Potom

$$(L_0,L_1,L_2,L_3) = \big(1,x,x^2,x^3\big)\cdot \boldsymbol{P} = \left(1,\ x,\ -\frac{1}{3}+x^2,\ -\frac{3}{5}\,x+x^3\right),$$

rovnako ako v príklade 13.4.7. Pohľad na maticu $\boldsymbol{G}(L_0,L_1,L_2,L_3)$ nám navyše prezradí normy polynómov $L_n(x)$ pre $n \le 3$:

$$\|L_0\| = \sqrt{2}, \qquad \|L_1\| = \sqrt{\frac{2}{3}}, \qquad \|L_2\| = \sqrt{\frac{8}{45}} = \frac{2}{3}\sqrt{\frac{2}{5}},$$
$$\|L_3\| = \sqrt{\frac{8}{175}} = \frac{2}{5}\sqrt{\frac{2}{7}}.$$

Pre normy Legendreových polynómov $P_n(x)$, $n \le 3$, z toho vyplýva

$$\|P_0\| = \sqrt{2}, \qquad \|P_1\| = \sqrt{\frac{2}{3}}, \qquad \|P_2\| = \sqrt{\frac{2}{5}}, \qquad \|P_3\| = \sqrt{\frac{2}{7}}.$$

13.4.10. Príklad. V euklidovskom priestore $\mathbb{R}^4$ so štandardným skalárnym súčinom je daný lineárny podpriestor S generovaný stĺpcami matice

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 1 & 2 \\ -1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

Nájdeme nejakú ortonormálnu bázu podpriestoru S. Zrejme stĺpce matice $\boldsymbol{\alpha}$ sú lineárne nezávislé, teda $\boldsymbol{\alpha}$ je bázou S. Jej Gramovu maticu $\boldsymbol{G}(\boldsymbol{\alpha}) = \boldsymbol{\alpha}^{\mathsf{T}} \cdot \boldsymbol{\alpha}$ upravíme pomocou úprav typu (1^+) na s ňou kongruentný diagonálny tvar

$$\boldsymbol{G}(\boldsymbol{\alpha}) = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 4 & 1 \\ 2 & 1 & 3 \end{pmatrix} \equiv \begin{pmatrix} 3 & 0 & 0 \\ 0 & 11/3 & 1/3 \\ 0 & 1/3 & 11/3 \end{pmatrix} \equiv \begin{pmatrix} 3 & 0 & 0 \\ 0 & 11/3 & 0 \\ 0 & 0 & 40/11 \end{pmatrix} = \boldsymbol{G}(\boldsymbol{\beta}).$$

Príslušnými ESO na matici $\boldsymbol{\alpha}$ dostaneme

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 1 & 2 \\ -1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \wr \begin{pmatrix} 0 & 1 & -1 \\ 1 & 2/3 & 4/3 \\ -1 & 4/3 & 2/3 \\ 1 & 2/3 & -2/3 \end{pmatrix} \wr \begin{pmatrix} 0 & 1 & -12/11 \\ 1 & 2/3 & 14/11 \\ -1 & 4/3 & 6/11 \\ 1 & 2/3 & -8/11 \end{pmatrix} = \boldsymbol{\beta}.$$

Ortonormálna báza podpriestoru S je potom tvorená stĺpcami matice $\boldsymbol{\beta}$ vynásobenými prevrátenými hodnotami druhých odmocnín príslušných diagonálnych prvkov matice $\boldsymbol{G}(\boldsymbol{\beta})$, teda vektormi

$$\frac{1}{\sqrt{3}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \qquad \sqrt{\frac{3}{11}} \begin{pmatrix} 1 \\ 2/3 \\ 4/3 \\ 2/3 \end{pmatrix} = \frac{1}{\sqrt{33}} \begin{pmatrix} 3 \\ 2 \\ 4 \\ 2 \end{pmatrix},$$

$$\sqrt{\frac{11}{40}} \begin{pmatrix} -12/11 \\ 14/11 \\ 6/11 \\ -8/11 \end{pmatrix} = \frac{1}{\sqrt{110}} \begin{pmatrix} -6 \\ 7 \\ 3 \\ -4 \end{pmatrix}.$$

13.5 Ortogonálne matice

Videli sme, že z vyjadrenia súradníc vektorov v euklidovských priestoroch vzhľadom na ortonormálne bázy vyplývajú dodatočné výhody, aké nám bázy, ktoré nespĺňajú túto podmienku, neposkytujú. Bude preto zaujímavé preskúmať, ako vyzerajú matice prechodu medzi takýmito bázami. Odpoveď na naznačenú otázku je prekvapivo jednoduchá.

Matica $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ sa nazýva *ortogonálna*, ak platí

$$\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A} = \boldsymbol{I}_n,$$

alebo, čo je to isté, $\boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathsf{T}}$. Uvedomme si, že prvá podmienka vlastne hovorí, že stĺpce matice $\boldsymbol{A}$ tvoria ortonormálnu bázu euklidovského priestoru $\mathbb{R}^n$ so štandardným skalárnym súčinom. Potom tiež platí $\boldsymbol{A} \cdot \boldsymbol{A}^{\mathsf{T}} = \boldsymbol{I}_n$, teda takisto riadky matice $\boldsymbol{A}$ tvoria ortonormálnu bázu v $\mathbb{R}^n$. Z tých dôvodov by bolo vari priliehavejšie nazývať takéto matice *ortonormálnymi.* Budeme sa však držať zaužívanej terminológie.

13.5.1. Veta. *Nech V je n-rozmerný euklidovský priestor, $\boldsymbol{\alpha}$ je ortonormálna a $\boldsymbol{\beta}$ je ľubovoľná báza priestoru V. Potom báza $\boldsymbol{\beta}$ je ortonormálna práve vtedy, keď matica prechodu $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ z bázy $\boldsymbol{\beta}$ do bázy $\boldsymbol{\alpha}$ je ortogonálna.*

Inak povedané, každá matica prechodu medzi ortonormálnymi bázami je ortogonálna, a tiež naopak, každá ortogonálna matica je maticou prechodu medzi ortonormálnymi bázami.
Dôkaz. Označme $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$, $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ a $\boldsymbol{P} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$. Podľa definície matice prechodu z paragrafu 7.5 a tvrdenia 13.4.4 (a) pre $i \le n$ platí

$$\boldsymbol{s}_i(\boldsymbol{P}) = (\boldsymbol{v}_i)_{\boldsymbol{\alpha}} = \big(\langle \boldsymbol{v}_i, \boldsymbol{u}_1\rangle, \dots, \langle \boldsymbol{v}_i, \boldsymbol{u}_n\rangle\big)^{\mathsf{T}},$$

teda $\boldsymbol{P} = \big(\langle \boldsymbol{v}_i, \boldsymbol{u}_j\rangle\big)_{n\times n}$. Keďže báza $\boldsymbol{\alpha}$ je ortonormálna, podľa tvrdenia 13.4.4 (b) má (i,k)-ty prvok matice $\boldsymbol{P}^{\mathsf{T}} \cdot \boldsymbol{P}$ tvar

$$\boldsymbol{s}_i(\boldsymbol{P})^{\mathsf{T}} \cdot \boldsymbol{s}_k(\boldsymbol{P}) = \sum_{j=1}^{n} \langle \boldsymbol{v}_i, \boldsymbol{u}_j\rangle \langle \boldsymbol{u}_j, \boldsymbol{v}_k\rangle = \langle \boldsymbol{v}_i, \boldsymbol{v}_k\rangle.$$

Teda matica $\boldsymbol{P}$ je ortogonálna, t. j. $\boldsymbol{P}^{\mathsf{T}} \cdot \boldsymbol{P} = \boldsymbol{I}_n$, práve vtedy, keď $\langle \boldsymbol{v}_i, \boldsymbol{v}_k\rangle = \delta_{ik}$ pre všetky $i, k \le n$, t. j. práve vtedy, keď $\boldsymbol{\beta}$ je ortonormálna báza.

Ortogonálne matice možno tiež charakterizovať ako matice, násobenie ktorými zachovávava štandardný skalárny súčin resp. euklidovskú dĺžku vektorov.

13.5.2. Veta. *Nech $\mathbb{R}^n$ je stĺpcový euklidovský priestor so štandardným skalárnym súčinom a $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. Potom nasledujúce podmienky sú ekvivalentné:*
(i) $\boldsymbol{A}$ je ortogonálna matica;
(ii) pre všetky $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ platí $\langle \boldsymbol{A} \cdot \boldsymbol{x}, \boldsymbol{A} \cdot \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$;
(iii) pre všetky $\boldsymbol{x} \in \mathbb{R}^n$ platí $\|\boldsymbol{A} \cdot \boldsymbol{x}\| = \|\boldsymbol{x}\|$.

Dôkaz. (i) ⇒ (ii) Nech $\boldsymbol{A}$ je ortogonálna a $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Potom

$$\langle \boldsymbol{A} \cdot \boldsymbol{x}, \boldsymbol{A} \cdot \boldsymbol{y}\rangle = (\boldsymbol{A} \cdot \boldsymbol{x})^{\mathsf{T}} \cdot (\boldsymbol{A} \cdot \boldsymbol{y}) = \boldsymbol{x}^{\mathsf{T}} \cdot \big(\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A}\big) \cdot \boldsymbol{y} = \boldsymbol{x}^{\mathsf{T}} \cdot \boldsymbol{I}_n \cdot \boldsymbol{y} = \langle \boldsymbol{x}, \boldsymbol{y}\rangle.$$

(ii) ⇒ (i) Z podmienky (ii) pre vektory $\boldsymbol{x} = \boldsymbol{e}_i$, $\boldsymbol{y} = \boldsymbol{e}_j$, kde $i, j \le n$, vyplýva

$$\boldsymbol{s}_i(\boldsymbol{A})^{\mathsf{T}} \cdot \boldsymbol{s}_j(\boldsymbol{A}) = \langle \boldsymbol{s}_i(\boldsymbol{A}), \boldsymbol{s}_j(\boldsymbol{A})\rangle = \langle \boldsymbol{A} \cdot \boldsymbol{e}_i, \boldsymbol{A} \cdot \boldsymbol{e}_j\rangle = \langle \boldsymbol{e}_i, \boldsymbol{e}_j\rangle = \delta_{ij},$$

teda $\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A} = \boldsymbol{I}_n$, čiže $\boldsymbol{A}$ je ortogonálna.

(ii) $\Rightarrow$ (iii) platí triviálne a (iii) $\Rightarrow$ (ii) je dôsledkom rovnosti

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \frac{1}{2}\big(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2\big).$$

Podrobnejší popis štruktúry ortogonálnych matíc nám umožnia až niektoré ďalšie poznatky o spektrálnych vlastnostiach matíc a lineárnych zobrazení, ktoré si začneme zadovažovať počnúc kapitolou 18. Ortogonálne matice rádu $n \leq 2$ však možno preskúmať celkom elementárnymi prostriedkami.

Čísla ± 1 sú zrejme jediné dve ortogonálne matice rozmeru 1×1. Prvý netriviálny prípad teda nastáva pre $n = 2$.

13.5.3. Veta. *Nech $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$. Potom $\boldsymbol{A}$ je ortogonálna práve vtedy, keď je maticou rotácie okolo počiatku, t. j.*

$$\boldsymbol{A} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \boldsymbol{R}_\alpha$$

pre nejaké $\alpha \in \mathbb{R}$, alebo maticou osovej súmernosti podľa osi prechádzajúcej počiatkom, t. j.

$$\boldsymbol{A} = \begin{pmatrix} \cos 2\beta & \sin 2\beta \\ \sin 2\beta & -\cos 2\beta \end{pmatrix} = \boldsymbol{S}_\beta$$

pre nejaké $\beta \in \mathbb{R}$.

Dôkaz. Nech $\boldsymbol{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Potom

$$\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2+c^2 & ab+cd \\ ba+dc & b^2+d^2 \end{pmatrix},$$

teda podmienka $\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A} = \boldsymbol{I}_2$, je ekvivalentná s rovnosťami

$$\begin{aligned} a^2 + c^2 &= 1, \\ ab + cd &= 0, \\ b^2 + d^2 &= 1. \end{aligned}$$

Prvá z nich je ekvivalentná s existenciou $\alpha \in \mathbb{R}$ takého, že $a = \cos\alpha$, $c = \sin\alpha$, a tretia s existenciou $\beta \in \mathbb{R}$ takého, že $b = \sin\beta$, $d = \cos\beta$. Podľa druhej rovnice musí pre ne platiť

$$\cos\alpha\sin\beta + \sin\alpha\cos\beta = \sin(\alpha+\beta) = 0,$$

t. j. $\alpha+\beta = k\pi$, kde $k \in \mathbb{Z}$. Pre k párne z toho dostávame $b = \sin\beta = -\sin\alpha$, $d = \cos\beta = \cos\alpha$, teda $\boldsymbol{A}$ má tvar

$$\boldsymbol{A} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \boldsymbol{R}_\alpha.$$

Pre k nepárne máme $b = \sin\beta = \sin\alpha$, $d = \cos\beta = -\cos\alpha$, čo vedie na maticu tvaru

$$\boldsymbol{A} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} = \boldsymbol{S}_{\alpha/2}.$$

Substitúciou (teraz už iného) $\beta = \alpha/2$ dostávame tvar uvedený v znení vety.

Cvičenia

13.1. Nech $\boldsymbol{A} \in \mathbb{R}^{n\times k}$ je ľubovoľná matica, ktorú stotožníme s usporiadanou k-ticou vektorov tvorenou jej stĺpcami. Potom v euklidovskom priestore $\mathbb{R}^n$ so štandardným skalárnym súčinom pre Gramovu maticu týchto vektorov platí $G(\boldsymbol{A}) = \boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A}$. Dokážte.

13.2. Nech $\boldsymbol{A} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$.

(a) Dokážte, že formulou $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^\mathsf{T} \cdot \boldsymbol{A} \cdot \boldsymbol{y}$ je definovaný skalárny súčin na vektorovom priestore $\mathbb{R}^3$.

(b) Vypočítajte dĺžky vektorov $\boldsymbol{e}_i$ bázy $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ a uhly medzi nimi vzhľadom na skalárny súčin daný maticou $\boldsymbol{A}$. Skúste potrebné hodnoty vyčítať priamo z matice $\boldsymbol{A}$.

(c) GSO-procesom aplikovaným na bázu $\boldsymbol{\varepsilon}$ nájdite ortogonálnu bázu $\boldsymbol{\beta} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ euklidovského priestoru $\mathbb{R}^3$.

(d) Vypočítajte normy vektorov bázy $\boldsymbol{\beta}$ a nájdite ortonormálnu bázu k báze $\boldsymbol{\beta}$.

13.3. Dokážte tzv. *Besselovu nerovnosť*: Ak $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ sú ortonormálne vektory v euklidovskom priestore V, tak $\langle \boldsymbol{x}, \boldsymbol{u}_1 \rangle^2 + \ldots + \langle \boldsymbol{x}, \boldsymbol{u}_k \rangle^2 \le \|\boldsymbol{x}\|^2$ pre každé $\boldsymbol{x} \in V$.

13.4. Dokážte Cauchyho-Schwartzovu nerovnosť priamo (t. j. bez použitia Gramovej matice a Sylvestrovho kritéria). (*Návod*: Uvažujte polynóm $f(\lambda) = \langle \lambda\boldsymbol{u} + \boldsymbol{v}, \lambda\boldsymbol{u} + \boldsymbol{v} \rangle$ v premennej λ a diskriminant kvadratickej rovnice $f(\lambda) = 0$.)

13.5. Dokážte dôsledok 13.4.2 bez použitia tvrdenia 13.4.1, len na základe definície lineárnej nezávislosti a ortogonálnosti vektorov $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$.

13.6. Pre $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ položme $\|\boldsymbol{x}\|_1 = \sum_{i=1}^n |x_i$, $\|\boldsymbol{x}\|_\infty = \max_{1\le i\le n} |x_i|$.

(a) Dokážte, že obe funkcie $\|\cdot\|_1$, $\|\cdot\|_\infty$ sú normy na vektorovom priestore $\mathbb{R}^n$.

(b) Nájdite príklady potvrdzujúce, že pre $n \ge 2$ ani jedna z noriem $\|\cdot\|_1$, $\|\cdot\|_\infty$ nespĺňa pravidlo rovnobežníka $\|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2 = 2\big(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\big)$.

(c) Odvoďte z toho, že pre $n \ge 2$ ani jedna z noriem $\|\cdot\|_1$, $\|\cdot\|_\infty$ nepochádza zo skalárneho súčinu.

13.7. Funkcia $f : \mathbb{R} \to \mathbb{R}$ sa nazýva *periodická* s periódou $T > 0$, ak pre každé $x \in \mathbb{R}$ platí $f(x+T) = f(x)$. Nech $a < b$ sú reálne čísla také, že $b - a = T$. Dokážte, že formulou $\langle f, g \rangle = \int_a^b f(x)g(x)\,\mathrm{d}x$ je definovaný skalárny súčin na vektorovom priestore všetkých spojitých periodických funkcií $f : \mathbb{R} \to \mathbb{R}$ s periódou T.

13.8. Uvažujte vektorový priestor všetkých periodických spojitých funkcií $f : \mathbb{R} \to \mathbb{R}$ s periódou $T = 2\pi$. Dokážte, že postupnosť funkcií $1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots$

v ňom tvorí ortogonálny systém vzhľadom na skalárny súčin $\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)\,\mathrm{d}x$. Vypočítajte normy týchto funkcií.

13.9. Dokážte nasledujúce tvrdenia.

(a) Ak $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ je ortogonálna matica, tak $\det \boldsymbol{A} = \pm 1$.

(b) Ak $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$ sú ortogonálne matice, tak aj matice $\boldsymbol{A}^T$, $\boldsymbol{A}^{-1}$ a $\boldsymbol{A}\cdot\boldsymbol{B}$ sú ortogonálne.

13.10. Doplňte dôkazy implikácií (ii) $\Rightarrow$ (iii) a (iii) $\Rightarrow$ (ii) vo vete 13.5.2.

13.11. Analýzou úvah o reprezentácii GSO-procesu maticovými úpravami (pozri text medzi príkladmi 13.4.7 a 13.4.8) dokážte nasledujúci špeciálny prípad tzv. *Iwasawovho rozkladu.*[2] Každú regulárnu maticu $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ možno rozložiť na súčin $\boldsymbol{A} = \boldsymbol{P}\cdot\boldsymbol{D}\cdot\boldsymbol{T}$, kde $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ je ortogonálna (t. j. jej stĺpce tvoria *ortonormálnu* bázu v $\mathbb{R}^n$), $\boldsymbol{D} \in \mathbb{R}^{n\times n}$ je diagonálna kladne definitná a $\boldsymbol{T} \in \mathbb{R}^{n\times n}$ je horná trojuholníková matica s jednotkami na diagonále.

13.12. Nájdite Iwasawove rozklady nasledujúcich matíc:

(a) $\boldsymbol{A} = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix}$, (b) $\boldsymbol{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,

(c) $\boldsymbol{C} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ -1 & 2 & -3 \end{pmatrix}$, (d) $\boldsymbol{D} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -3 \end{pmatrix}$.

13.13. (a) Zmodifikujte Gramov-Schmidtov ortogonalizačný proces tak, aby fungoval aj pre lineárne závislé postupnosti vektorov $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$. Ako vyzerá takto získaný k-ty vektor $\boldsymbol{v}_k$, ak $\boldsymbol{u}_k \in [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}]$?

(b) S využitím (a) dokážte reálnu verziu vety o *QR-rozklade*: Každú maticu $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, kde $m \geq n$, možno rozložiť na súčin $\boldsymbol{A} = \boldsymbol{Q}\cdot\boldsymbol{R}$, kde $\boldsymbol{Q} \in \mathbb{R}^{m\times m}$ je ortogonálna matica a $\boldsymbol{R} \in \mathbb{R}^{m\times n}$ je horná trojuholníková matica (t. j. $r_{ij} = 0$ pre $i > j$) s nulovými poslednými $m - n$ riadkami. (*Návod*: Ortogonalizujte stĺpce matice $\boldsymbol{A}$ a nejako si poraďte s nulovými stĺpcami.)

(c) Ak stĺpce matice $\boldsymbol{A}$ sú lineárne nezávislé, špeciálne, ak $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ je regulárna matica, tak jej QR-rozklad je určený jednoznačne za dodatočnej podmienky, že všetky diagonálne prvky matice $\boldsymbol{R}$ sú kladné. Dokážte.

13.14. Nájdite QR-rozklady nasledujúcich matíc a taktiež matíc z cvičenia 13.12:

(a) $\boldsymbol{A} = \begin{pmatrix} 1 & -2 \\ 1/3 & 7 \end{pmatrix}$, (b) $\boldsymbol{B} = \begin{pmatrix} 4 & 3 \\ -1 & 1 \end{pmatrix}$,

(c) $\boldsymbol{C} = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 3 & -2 \\ 1 & 1 & 0 \end{pmatrix}$, (d) $\boldsymbol{D} = \begin{pmatrix} 3 & 1 \\ -2 & 0 \\ -1 & 1 \end{pmatrix}$.

13.15. (a) Pre Legendreove polynómy $P_n(x)$ z príkladu 13.4.7 odvoďte rekurentný vzťah $P_0(x) = 1$, $P_1(x) = x$ a $P_{n+1}(x) = \frac{2n+1}{n+1}\, xP_n(x) - \frac{n}{n+1} P_{n-1}(x)$ pre $n \geq 1$.

(b) Dokážte explicitný vzorec $P_n(x) = \frac{1}{2^n n!}\,\frac{\mathrm{d}^n}{\mathrm{d}x^n}\,(x^2-1)^n$.

(c) Vypočitajte normy Legendreových polynómov $P_n(x)$.

13.16. (a) Dokážte, že rovnosťou $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,\mathrm{e}^{-x^2}\,\mathrm{d}x$ je definovaný skalárny súčin na istom vektorovom priestore spojitých funkcií $f\colon \mathbb{R} \to \mathbb{R}$ obsahujúcom všetky

[2] Všeobecný prípad sa týka rozkladu prvkov v istých typoch *Lieových grúp.*

polynómy (a okrem nich mnoho ďalších funkcií – o konvergenciu nevlastného integrálu sa postará faktor e^{-x^2}).

(b) Označme $(H_n)_{n=0}^{\infty}$ postupnosť funkcií, ktorú získame GSO-procesom z postupnosti polynómov $\big((2x)^n\big)_{n=0}^{\infty}$. Vypočítajte prvé štyri členy H_0, H_1, H_2, H_3 tejto postupnosti.

(c) Odvoďte rekurentný vzťah $H_0(x) = 1$, $H_1(x) = 2x$, $H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$ pre $n \geq 1$, a vypočítajte ďalšie dva členy H_4, H_5.

(d) Na základe (c) zdôvodnite, že funkcia H_n je polynóm n-tého stupňa, ktorý pre párne n obsahuje iba párne a pre nepárne n iba nepárne mocniny premennej x. Funkcie $H_n(x)$ nazývame *Hermitove polynómy*.

(e) Dokážte explicitný vzťah $H_n(x) = (-1)^n \mathrm{e}^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2}$ a odvoďte z neho podmienku $H_n'(x) = 2nH_{n-1}(x)$.

(f) Vychádzajúc z rovnosti pre Laplaceov integrál $\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\, \mathrm{d}x = \sqrt{\pi}$ (pozri cvičenie 14.20) overte, že pre normy Hermitových polynómov platí $\|H_n(x)\|^2 = 2^n n! \sqrt{\pi}$.

14. Ortogonálne projekcie a podpriestory

V tejto kapitole budeme pokračovať v štúdiu euklidovských priestorov s cieľom podať kvantitatívny popis vzájomnej polohy afinných podpriestorov v takomto priestore pomocou dvoch základných parametrov – ich *vzdialenosti* a *odchýlky (uhla)*. Našim hlavným nástrojom pri tom budú lineárne operátory *kolmého priemetu* vektorov do lineárnych podpriestorov, zvané tiež *ortogonálne projekcie*. V druhej časti kapitoly predvedieme tri aplikácie rozpracovaných pojmov a metód: využijeme ich jednak na zavedenie tzv. *sférických súradníc* v euklidovských priestoroch, jednak pri konštrukcii „najlepších približných riešení“ neriešiteľných sústav lineárnych rovníc a lineárnej regresii, a napokon sa oboznámime s formuláciu základných pojmov teórie pravdepodobnosti v jazyku euklidovských priestorov.

14.1 Ortokomplement a ortogonálna projekcia

Relácia ortogonality (kolmosti) vo vektorovom priestore so skalárnym súčinom má niekoľko zrejmých vlastností, ktoré tu zaznamenáme bez dôkazu.

14.1.1. Tvrdenie. *Nech V je vektorový priestor so skalárnym súčinom. Potom pre ľubovoľné $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$, $c, d \in \mathbb{R}$ platí:*
(a) $\boldsymbol{x} \perp \mathbf{0}$;
(b) $\boldsymbol{x} \perp \boldsymbol{x} \;\Leftrightarrow\; \boldsymbol{x} = \mathbf{0}$;
(c) $\boldsymbol{x} \perp \boldsymbol{y} \;\Leftrightarrow\; \boldsymbol{y} \perp \boldsymbol{x}$;
(d) $\big(\boldsymbol{x} \perp \boldsymbol{y} \;\&\; \boldsymbol{x} \perp \boldsymbol{z}\big) \;\Rightarrow\; \boldsymbol{x} \perp (c\boldsymbol{y} + d\boldsymbol{z})$.

Ortogonálnym doplnkom alebo tiež *ortokomplementom* ľubovoľnej množiny $X \subseteq V$ vo vektorovom priestore so skalárnym súčinom nazveme množinu

$$X^{\perp} = \{\boldsymbol{y} \in V;\ (\forall \boldsymbol{x} \in X)(\boldsymbol{x} \perp \boldsymbol{y})\}$$

všetkých vektorov $\boldsymbol{y} \in V$ kolmých na každý vektor $\boldsymbol{x} \in X$.

14.1.2. Tvrdenie. *Nech V je vektorový priestor so skalárnym súčinom. Potom pre ľubovoľné množiny $X, Y \subseteq V$ platí:*
(a) $\emptyset^{\perp} = \{\mathbf{0}\}^{\perp} = V,\ \ V^{\perp} = \{\mathbf{0}\}$;
(b) $X^{\perp} = [X]^{\perp} = [X^{\perp}]$;
(c) $X \subseteq Y \;\Rightarrow\; Y^{\perp} \subseteq X^{\perp}$;
(d) $X \subseteq X^{\perp\perp}$;

(e) $X^{\perp\perp\perp} = X^{\perp}$*;*
(f) $X \cap X^{\perp} = \{\mathbf{0}\}$*, ak* $\mathbf{0} \in X$*, a* $X \cap X^{\perp} = \emptyset$*, ak* $\mathbf{0} \notin X$*;*
(g) $(X \cup Y)^{\perp} = X^{\perp} \cap Y^{\perp}$*;*
(h) ak X*,* Y *sú navyše lineárne podpriestory, tak* $(X+Y)^{\perp} = X^{\perp} \cap Y^{\perp}$*.*

Dôkaz. Jednotlivé podmienky sú priamymi dôsledkami predošlého tvrdenia, ich jednoduché dôkazy preto prenechávame ako cvičenie čitateľovi. Na ukážku predvedieme, ako vyplýva (e) z podmienok (c) a (d). Podľa (d) platí $X \subseteq X^{\perp\perp}$, z čoho podľa (c) vyplýva $X^{\perp\perp\perp} \subseteq X^{\perp}$. Obrátená inklúzia $X^{\perp} \subseteq X^{\perp\perp\perp}$ je opäť dôsledkom (d).

Z podmienky (b) okrem iného vyplýva, že $X^{\perp}$ je lineárnym podpriestorom vo V pre *každú podmnožinu* $X \subseteq V$.

Nech $S \subseteq V$ je lineárny podpriestor priestoru so skalárnym súčinom V a $\boldsymbol{x} \in V$. Hovoríme, že vektor $\boldsymbol{z} \in S$ je *kolmým priemetom* alebo tiež *ortogonálnou projekciou* vektora $\boldsymbol{x}$ do podpriestoru S, ak $\boldsymbol{x} - \boldsymbol{z} \in S^{\perp}$. Tento vektor (ak existuje) budeme značiť $\boldsymbol{z} = \mathrm{pr}_S(\boldsymbol{x}) = \boldsymbol{x}_S$.

14.1.3. Veta. *Nech* V *je vektorový priestor so skalárnym súčinom,* $S \subseteq V$ *je jeho konečnorozmerný lineárny podpriestor a* $\boldsymbol{x} \in V$*. Potom*

(a) kolmý priemet vektora $\boldsymbol{x}$ *do podpriestoru* S *existuje a je jednoznačne určený rovnosťou*

$$\mathrm{pr}_S(\boldsymbol{x}) = \boldsymbol{x}_S = \sum_{i=1}^{k} \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i,$$

kde $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ *je ľubovoľná ortonormálna báza podpriestoru* S*;*

(b) pre ľubovoľný vektor $\boldsymbol{y} \in S$ *platí*

$$\|\boldsymbol{x} - \boldsymbol{x}_S\| \leq \|\boldsymbol{x} - \boldsymbol{y}\|,$$

pričom rovnosť nastane práve vtedy, keď $\boldsymbol{y} = \boldsymbol{x}_S$*;*

(c) ak $\boldsymbol{x} \neq \mathbf{0}$ a $S \neq \{\mathbf{0}\}$*, tak pre ľubovoľný vektor* $\mathbf{0} \neq \boldsymbol{y} \in S$ *platí*

$$\frac{\|\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|} \geq \frac{|\langle \boldsymbol{x}, \boldsymbol{y} \rangle|}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|},$$

pričom rovnosť nastane práve vtedy, keď vektory $\boldsymbol{x}_S$*,* $\boldsymbol{y}$ *sú lineárne závislé.*

Dôkaz (a) Nech $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ je ľubovoľná ortonormálna báza podpriestoru S. Ak má kolmý priemet vektora $\boldsymbol{x}$ do S existovať, musí mať tvar $\boldsymbol{x}_S = c_1\boldsymbol{u}_1 + \ldots + c_k\boldsymbol{u}_k$ pre nejaké $c_1, \ldots, c_k \in \mathbb{R}$. Podmienka $\boldsymbol{x} - \boldsymbol{x}_S \in S^{\perp}$ je podľa predchádzajúceho tvrdenia ekvivalentná s konjunkciou podmienok $\boldsymbol{x}-\boldsymbol{x}_S \perp \boldsymbol{u}_i$ pre každé $i \leq k$. Z toho vyplýva

$$0 = \langle \boldsymbol{x}-\boldsymbol{x}_S, \boldsymbol{u}_i\rangle = \langle \boldsymbol{x}, \boldsymbol{u}_i\rangle - \sum_{j=1}^{k} c_i\langle \boldsymbol{u}_j, \boldsymbol{u}_i\rangle = \langle \boldsymbol{x}, \boldsymbol{u}_i\rangle - c_i,$$

teda lineárna kombinácia $c_1\boldsymbol{u}_1+\ldots+c_k\boldsymbol{u}_k$ je kolmým priemetom $\boldsymbol{x}$ do S práve vtedy, keď $c_i = \langle \boldsymbol{x}, \boldsymbol{u}_i\rangle$, pre každé i.

(b) Nech $\boldsymbol{y} \in S$. Potom tiež $\boldsymbol{x}_S - \boldsymbol{y} \in S$, teda $(\boldsymbol{x}-\boldsymbol{x}_S) \perp (\boldsymbol{x}_S-\boldsymbol{y})$. Podľa Pytagorovej vety platí

$$\|\boldsymbol{x}-\boldsymbol{y}\|^2 = \|(\boldsymbol{x}-\boldsymbol{x}_S)+(\boldsymbol{x}_S-\boldsymbol{y})\|^2 = \|\boldsymbol{x}-\boldsymbol{x}_S\|^2 + \|\boldsymbol{x}_S-\boldsymbol{y}\|^2 \geq \|\boldsymbol{x}-\boldsymbol{x}_S\|^2,$$

pričom rovnosť nastane práve vtedy, keď $\boldsymbol{y} = \boldsymbol{x}_S$.

(c) Pre $\boldsymbol{y} \in S$ máme $(\boldsymbol{x}-\boldsymbol{x}_S) \perp \boldsymbol{y}$, preto $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}_S, \boldsymbol{y}\rangle + \langle \boldsymbol{x}-\boldsymbol{x}_S, \boldsymbol{y}\rangle = \langle \boldsymbol{x}_S, \boldsymbol{y}\rangle$. Ak $\boldsymbol{y} \neq \boldsymbol{0}$, tak s použitím Cauchyho-Schwartzovej nerovnosti z toho dostávame

$$\frac{|\langle \boldsymbol{x}, \boldsymbol{y}\rangle|}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|} = \frac{|\langle \boldsymbol{x}_S, \boldsymbol{y}\rangle|}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|} \leq \frac{\|\boldsymbol{x}_S\|\,\|\boldsymbol{y}\|}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|} = \frac{\|\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|}.$$

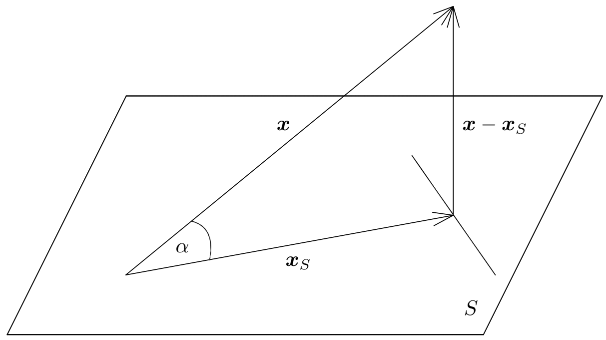
Rovnosť zrejme nastane práve vtedy, keď vektory $\boldsymbol{x}_S$, $\boldsymbol{y}$ sú lineárne závislé.

Podmienka (a) práve dokázanej vety má nasledujúci dôsledok.

14.1.4. Dôsledok. *Nech V je vektorový priestor so skalárnym súčinom a $S, T \subseteq V$ sú jeho konečnorozmerné lineárne podpriestory. Potom*
(a) $V = S \oplus S^{\perp}$, $S = S^{\perp\perp}$, $(S \cap T)^{\perp} = S^{\perp} + T^{\perp}$;
(b) $\mathrm{pr}_S : V \to V$ *je lineárny operátor;*
(c) $(\forall \boldsymbol{x} \in V)(\boldsymbol{x} \in S \;\Leftrightarrow\; \mathrm{pr}_S(\boldsymbol{x}) = \boldsymbol{x})$;
(d) $\mathrm{Im}\,\mathrm{pr}_S = S$ *a* $\mathrm{Ker}\,\mathrm{pr}_S = S^{\perp}$;
(e) $\boldsymbol{x}-\boldsymbol{x}_S$ *je kolmým priemetom vektora* $\boldsymbol{x}$ *do podpriestoru* $S^{\perp}$.

Lineárny operátor pr_S nazývame *ortogonálnou projekciou* na podpriestor S.

Poznámka. Ak V je euklidovský priestor, tak veta 14.1.3 a dôsledok 14.1.4 samozrejme platia pre *ľubovoľný* podpriestor $S \subseteq V$. Za istých okolností, ktorých rozbor však presahuje rámec lineárnej algebry (tzv. úplnosť priestoru V a uzavretosť podpriestoru S), spomínané výsledky platia aj pre nekonečnorozmerné podpriestory. V cvičeniach 14.4, 14.5 a 14.6 zakaždým uvedieme príklad vektorového priestoru so skalárnym súčinom V a jeho nekonečnorozmerného podpriestoru $S \neq V$ takého, že $S^{\perp} = \{\boldsymbol{0}\}$. Potom pre žiaden vektor $\boldsymbol{x} \in V \setminus S$ nemôže existovať jeho kolmý priemet do S; rovnako $S \oplus S^{\perp} = S \neq V$.



Obr. 14.1. Kolmý priemet vektora do lineárneho podpriestoru

Predpokladajme, že kolmý priemet vektora $\boldsymbol{x}$ do lineárneho podpriestoru S existuje. Vysvetlíme si, ako možno za tohto predpokladu definovať vzdialenosť aj odchýlku vektora $\boldsymbol{x}$ od každého z podpriestorov S, $S^{\perp}$. Situácia je znázornená na obrázku 14.1. Vektor $\boldsymbol{x}-\boldsymbol{x}_S$ je kolmý na každú priamku v podpriestore S, špeciálne trojuholník tvorený vektormi $\boldsymbol{x}$, $\boldsymbol{x}_S$, $\boldsymbol{x}-\boldsymbol{x}_S$ je pravouhlý, s pravým uhlom pri „konci“ vektora $\boldsymbol{x}_S$.

Podmienka (b) vety 14.1.3 nás oprávňuje nazvať dĺžku (anglicky *distance*) vektora $\boldsymbol{x}-\boldsymbol{x}_S$ *vzdialenosťou* vektora (prípadne bodu) $\boldsymbol{x}$ od podpriestoru S. Budeme ju značiť

$$\operatorname{dist}(\boldsymbol{x}, S) = \|\boldsymbol{x}-\boldsymbol{x}_S\| = \min\{\|\boldsymbol{x}-\boldsymbol{y}\|;\, \boldsymbol{y}\in S\}.$$

Vzhľadom na podmienku (e) dôsledku 14.1.4 je vzdialenosť vektora $\boldsymbol{x}$ od podpriestoru $S^{\perp}$ daná vzťahom

$$\operatorname{dist}\bigl(\boldsymbol{x}, S^{\perp}\bigr) = \|\boldsymbol{x}_S\|.$$

Podobne, keďže kosinus je na intervale $\langle 0,\pi\rangle$ klesajúca funkcia, podmienka (c) vety 14.1.3 nás oprávňuje nazvať výraz

$$\sphericalangle(\boldsymbol{x}, S) = \arccos\frac{\|\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|} = \min\{\sphericalangle(\boldsymbol{x},\boldsymbol{y});\boldsymbol{0}\neq\boldsymbol{y}\in S\}$$

odchýlkou vektor $\boldsymbol{x}\neq\boldsymbol{0}$ od podpriestoru $S\neq\{\boldsymbol{0}\}$, prípadne *uhlom* vektora $\boldsymbol{x}$ a podpriestoru S. Odchýlka $\sphericalangle(\boldsymbol{x}, S)$ je teda jednoznačne určená ako také reálne číslo $\alpha\in\langle 0,\pi/2\rangle$, pre ktoré platí

$$\cos\alpha = \frac{\|\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|},\quad \text{alebo ekvivalentne,}\quad \sin\alpha = \frac{\|\boldsymbol{x}-\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|}.$$

Zrejme opäť ide o *neorientovaný uhol.* Ak $\boldsymbol{x}_S\neq\boldsymbol{0}$, tak $\sphericalangle(\boldsymbol{x}, S) = \sphericalangle(\boldsymbol{x},\boldsymbol{x}_S)$; ak $\boldsymbol{x}_S=\boldsymbol{0}$, t. j. ak $\boldsymbol{x}\in S^{\perp}$, tak samozrejme $\sphericalangle(\boldsymbol{x}, S) = \pi/2$. Ešte si všimnite, že zatiaľ čo uhol dvoch vektorov nadobúda hodnoty z intervalu $\langle 0, \pi \rangle$, hodnoty, ktoré nadobúda uhol vektora a podpriestoru, sú obmedzené na interval $\langle 0, \pi/2 \rangle$. Vzhľadom na podmienku 14.1.3 (c), ak $S^\perp \neq \{\mathbf{0}\}$, tak odchýlka vektora $\boldsymbol{x} \neq \mathbf{0}$ od podpriestoru $S^\perp$ je daná vzťahom

$$\sphericalangle\big(\boldsymbol{x}, S^\perp\big) = \arccos \frac{\|\boldsymbol{x} - \boldsymbol{x}_S\|}{\|\boldsymbol{x}\|} = \arcsin \frac{\|\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|}.$$

Konečne časť (a) vety 14.1.3 nám dáva priamy návod, ako nájsť kolmý priemet vektora $\boldsymbol{x}$ do *konečnorozmerného* podpriestoru $S \subseteq V$, a tým aj vzdialenosti $\operatorname{dist}(\boldsymbol{x}, S)$, $\operatorname{dist}\big(\boldsymbol{x}, S^\perp\big)$ a odchýlky $\sphericalangle(\boldsymbol{x}, S)$, $\sphericalangle\big(\boldsymbol{x}, S^\perp\big)$ – potrebujeme však poznať aspoň jednu ortonormálnu bázu v S. Ak máme k dispozícii len nejakú „obyčajnú“ bázu podpriestoru S, je potrebné ju najprv ortonormalizovať, až potom možno použiť spomínaný vzorec. Diagonalizáciu Gramovej matice si však môžeme odpustiť, rovnaký výsledok sa totiž z nej dá získať aj priamo.

14.1.5. Tvrdenie. *Nech V je vektorový priestor so skalárnym súčinom, S je jeho konečnorozmerný lineárny podpriestor s bázou $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ a $\boldsymbol{x} \in V$. Potom pre $\boldsymbol{c} = (c_1, \dots, c_k)^\mathsf{T} \in \mathbb{R}^k$ platí $\boldsymbol{x}_S = c_1\boldsymbol{u}_1 + \dots + c_k\boldsymbol{u}_k$ práve vtedy, keď $\boldsymbol{c}$ je riešením sústavy lineárnych rovníc*

$$\boldsymbol{G}(\boldsymbol{\alpha}) \cdot \boldsymbol{c} = \langle \boldsymbol{x}, \boldsymbol{\alpha} \rangle^\mathsf{T},$$

kde $\langle \boldsymbol{x}, \boldsymbol{\alpha} \rangle$ označuje riadkový vektor $\big(\langle \boldsymbol{x}, \boldsymbol{u}_1 \rangle, \dots, \langle \boldsymbol{x}, \boldsymbol{u}_k \rangle\big) \in \mathbb{R}^k$.

Dôkaz. Pre $\boldsymbol{c} = (c_1, \dots, c_k)^\mathsf{T} \in \mathbb{R}^k$ platí $\boldsymbol{x}_S = c_1\boldsymbol{u}_1 + \dots + c_k\boldsymbol{u}_k$ práve vtedy, keď pre každé $i \leq k$ máme

$$0 = \langle \boldsymbol{x} - c_1\boldsymbol{u}_1 - \dots - c_k\boldsymbol{u}_k, \, \boldsymbol{u}_i \rangle = \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle - c_1\langle \boldsymbol{u}_1, \boldsymbol{u}_i \rangle - \dots - c_k\langle \boldsymbol{u}_k, \boldsymbol{u}_i \rangle.$$

Inými slovami, $\boldsymbol{c}$ musí vyhovovať sústave

$$\boldsymbol{G}(\boldsymbol{\alpha})^\mathsf{T} \cdot \boldsymbol{c} = \langle \boldsymbol{x}, \boldsymbol{\alpha} \rangle^\mathsf{T}.$$

Ale $\boldsymbol{G}(\boldsymbol{\alpha})^\mathsf{T} = \boldsymbol{G}(\boldsymbol{\alpha})$ vzhľadom na symetriu Gramovej matice. Vďaka regularite matice $\boldsymbol{G}(\boldsymbol{\alpha})$ ($\boldsymbol{\alpha}$ je báza S) má táto sústava jediné riešenie.

Ešte si všimnime, že rozšírená matica $\big(\boldsymbol{G}(\boldsymbol{\alpha}) \mid \langle \boldsymbol{x}, \boldsymbol{\alpha} \rangle^\mathsf{T}\big)$ uvedenej sústavy je vlastne Gramovou maticou $\boldsymbol{G}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{x})$ rádu $k+1$, z ktorej sme vynechali posledný riadok. Ak $\boldsymbol{\alpha}$ je ortonormálna báza, tak $\boldsymbol{G}(\boldsymbol{\alpha}) = \boldsymbol{I}_k$, t. j. príslušná sústava je už vo vyriešenom tvare $\boldsymbol{c} = \langle \boldsymbol{x}, \boldsymbol{\alpha} \rangle^\mathsf{T}$, presne v zhode s podmienkou (a) vety 14.1.3.

Poznámka. Tvrdenie zostáva bezo zmeny v platnosti, aj keď $\boldsymbol{\alpha}$ je ľubovoľný konečný (teda nie nutne lineárne nezávislý) systém generátorov v S. Jediné, čo sa zmení, je nejednoznačnosť *vyjadrenia* kolmého priemetu $\boldsymbol{x}_S = c_1\boldsymbol{u}_1 + \ldots + c_k\boldsymbol{u}_k$ ako lineárnej kombinácie vektorov systému $\boldsymbol{\alpha}$. Samotný kolmý priemet $\boldsymbol{x}_S$ je, samozrejme, určený jednoznačne. Rozmyslite si, prečo je to tak.

14.1.6. Príklad. V $\mathbb{R}^4$ so štandardným skalárnym súčinom je daný vektor $\boldsymbol{x} = (1,1,1,1)^\mathsf{T}$ a rovina $S = [\boldsymbol{u},\boldsymbol{v}]$, kde $\boldsymbol{u} = (0,-1,0,1)^\mathsf{T}$, $\boldsymbol{v} = (1,-2,1,-3)^\mathsf{T}$. Nájdeme kolmý priemet vektora $\boldsymbol{x}$ do roviny S a vypočítame vzdialenosť $\operatorname{dist}(\boldsymbol{x},S)$ a odchýlku $\sphericalangle(\boldsymbol{x},S)$. Kolmý priemet budeme hľadať v tvare $\boldsymbol{x}_S = c\boldsymbol{u} + d\boldsymbol{v}$, kde $(c,d)^\mathsf{T} \in \mathbb{R}^2$ vyhovuje sústave s rozšírenou maticou

$$\left(\begin{array}{cc|c} \langle\boldsymbol{u},\boldsymbol{u}\rangle & \langle\boldsymbol{v},\boldsymbol{u}\rangle & \langle\boldsymbol{x},\boldsymbol{u}\rangle \\ \langle\boldsymbol{u},\boldsymbol{v}\rangle & \langle\boldsymbol{v},\boldsymbol{v}\rangle & \langle\boldsymbol{x},\boldsymbol{v}\rangle \end{array}\right) = \left(\begin{array}{cc|c} 2 & -1 & 0 \\ -1 & 15 & -3 \end{array}\right).$$

Jej riešením dostaneme $c = -3/29$, $d = -6/29$, teda kolmý priemet vektora $\boldsymbol{x}$ do roviny $[\boldsymbol{u},\boldsymbol{v}]$ je

$$\boldsymbol{x}_S = (\boldsymbol{u},\boldsymbol{v}) \cdot \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -2 \\ 0 & 1 \\ 1 & -3 \end{pmatrix} \cdot \begin{pmatrix} -3/29 \\ -6/29 \end{pmatrix} = \frac{3}{29}\begin{pmatrix} -2 \\ 5 \\ -2 \\ 5 \end{pmatrix}.$$

Ako skúšku správnosti si overte rovnosti $\langle\boldsymbol{x}-\boldsymbol{x}_S,\boldsymbol{u}\rangle = \langle\boldsymbol{x}-\boldsymbol{x}_S,\boldsymbol{v}\rangle = 0$. Pre vzdialenosť resp. odchýlku $\boldsymbol{x}$ od S potom dostávame

$$\operatorname{dist}(\boldsymbol{x},S) = \|\boldsymbol{x}-\boldsymbol{x}_S\| = \left\|\frac{7}{29}(5,2,5,2)^\mathsf{T}\right\| = \frac{7}{29}\sqrt{58},$$

$$\sin\sphericalangle(\boldsymbol{x},S) = \frac{\|\boldsymbol{x}-\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|} = \frac{7}{2\cdot 29}\sqrt{58} = \frac{7}{\sqrt{58}}.$$

S požitím kalkulačky možno zistiť

$$\sphericalangle(\boldsymbol{x},S) = \arcsin\frac{7}{\sqrt{58}} \approx 1,1659 \text{ rad} \approx 66^\circ\, 48'\, 5''.$$

14.1.7. Príklad. Nech $\boldsymbol{A} \in R^{m\times n}$, pričom $m \geq n$ a $h(\boldsymbol{A}) = n$, t.j. stĺpce matice $\boldsymbol{A}$ sú lineárne nezávislé vektory v euklidovskom priestore $\mathbb{R}^m$ so štandardným skalárnym súčinom. Označme $S \subseteq \mathbb{R}^m$ lineárny podpriestor generovaný stĺpcami matice $\boldsymbol{A}$. Potom ortogonálna projekcia na podpriestor S je lineárny operátor $\mathrm{pr}_S : \mathbb{R}^m \to \mathbb{R}^m$. Nájdeme jeho maticu $\boldsymbol{B} = \left(\mathrm{pr}_S\right)_{\boldsymbol{\varepsilon},\boldsymbol{\varepsilon}} \in \mathbb{R}^{m\times m}$ vzhľadom na kanonickú ortonormálnu bázu $\boldsymbol{\varepsilon}$ priestoru $\mathbb{R}^m$.

Ak stotožníme maticu $\boldsymbol{A}$ s usporiadanou n-ticou jej stĺpcov, tak $\boldsymbol{A}$ je bázou S. Podľa tvrdenia 14.1.5 obraz $\boldsymbol{y} = \mathrm{pr}_S(\boldsymbol{x})$ vektora $\boldsymbol{x} \in \mathbb{R}^m$ dostaneme v tvare

$$\boldsymbol{y} = \boldsymbol{A}\cdot\boldsymbol{c},$$

kde $\boldsymbol{c} \in \mathbb{R}^n$ je (jediné) riešenie sústavy

$$\boldsymbol{G}(\boldsymbol{A}) \cdot \boldsymbol{c} = \langle \boldsymbol{x}, \boldsymbol{A} \rangle^{\mathsf{T}}.$$

Uvedomme si, že $\boldsymbol{G}(\boldsymbol{A}) = \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A}$ je regulárna matica. Ďalej platí $\langle \boldsymbol{x}, \boldsymbol{A} \rangle = \boldsymbol{x}^{\mathsf{T}} \cdot \boldsymbol{A}$, čiže $\langle \boldsymbol{x}, \boldsymbol{A} \rangle^{\mathsf{T}} = \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{x}$. Z toho dostávame

$$\begin{aligned} \boldsymbol{c} &= \boldsymbol{G}(\boldsymbol{A})^{-1} \cdot \langle \boldsymbol{x}, \boldsymbol{A} \rangle^{\mathsf{T}} = \left(\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A}\right)^{-1} \cdot \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{x}, \\ \boldsymbol{y} &= \boldsymbol{A} \cdot \boldsymbol{c} = \boldsymbol{A} \cdot \left(\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A}\right)^{-1} \cdot \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{x}. \end{aligned}$$

Teda hľadaná matica ortogonálnej projekcie pr_S je

$$\boldsymbol{B} = \left(\mathrm{pr}_S\right)_{\boldsymbol{\varepsilon},\boldsymbol{\varepsilon}} = \boldsymbol{A} \cdot \left(\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A}\right)^{-1} \cdot \boldsymbol{A}^{\mathsf{T}}.$$

14.2 Vzdialenosť dvoch afinných podpriestorov

Nech V je vektorový priestor so skalárnym súčinom a X, Y sú jeho dve neprázdne podmnožiny. *Vzdialenosťou množín* X, Y vo V nazývame číslo

$$\operatorname{dist}(X, Y) = \inf\{\|\boldsymbol{x} - \boldsymbol{y}\| \,;\, \boldsymbol{x} \in X \ \&\ \boldsymbol{y} \in Y\}.$$

Problematikou vzdialeností množín v plnom rozsahu sa tu zaoberať nebudeme. Obmedzíme sa len na *vzdialenosti konečnorozmerných afinných podpriestorov*. Úlohu prevedieme na určenie vzdialenosti vektora od lineárneho podpriestoru.

14.2.1. Lema. *Nech V je vektorový priestor so skalárnym súčinom a M, N sú jeho afinné podpriestory. Potom pre ľubovoľné body $\boldsymbol{p} \in M$, $\boldsymbol{q} \in M$ platí:*

$$\operatorname{dist}(M, N) = \operatorname{dist}(\boldsymbol{p} - \boldsymbol{q},\ \operatorname{Dir} M + \operatorname{Dir} N).$$

Dôkaz. Označme $S = \operatorname{Dir} M$, $T = \operatorname{Dir} N$ zamerania podpriestrov M, N. Potom $M = \boldsymbol{p} + S$, $N = \boldsymbol{q} + T$. Podľa definície vzdialenosti platí

$$\begin{aligned} \operatorname{dist}(M, N) &= \inf\left\{\|(\boldsymbol{p} + \boldsymbol{u}) - (\boldsymbol{q} + \boldsymbol{v})\| \,;\, \boldsymbol{u} \in S \ \&\ \boldsymbol{v} \in T\right\}, \\ \operatorname{dist}(\boldsymbol{p} - \boldsymbol{q},\ S + T) &= \inf\left\{\|(\boldsymbol{p} - \boldsymbol{q}) - (\boldsymbol{u} + \boldsymbol{v})\| \,;\, \boldsymbol{u} \in S \ \&\ \boldsymbol{v} \in T\right\}. \end{aligned}$$

Stačí teda overiť rovnosť množín na pravých stranách. Túto jednoduchú úlohu prenechávame ako cvičenie čitateľovi.

Hovoríme, že body $\boldsymbol{p} \in M$, $\boldsymbol{q} \in N$ tvoria *priečku* afinných podpriestorov M, N, ak

$$\operatorname{dist}(M, N) = \|\boldsymbol{p} - \boldsymbol{q}\|,$$

t. j. ak sa vzdialenosť podpriestorov M, N realizuje ako dĺžka vektora $\boldsymbol{p} - \boldsymbol{q}$.

14.2.2. Tvrdenie. *Nech M, N sú konečnorozmerné afinné podpriestory vektorového priestoru so skalárnym súčinom V. Potom*

(a) body $\boldsymbol{p} \in M$, $\boldsymbol{q} \in N$ tvoria priečku podpriestorov M, N práve vtedy, keď $\boldsymbol{p} - \boldsymbol{q} \in (\mathrm{Dir}\, M + \mathrm{Dir}\, N)^{\perp}$;

(b) pre ľubovoľné body $\boldsymbol{p} \in M$, $\boldsymbol{q} \in N$ a vektory $\boldsymbol{u} \in \mathrm{Dir}\, M$, $\boldsymbol{v} \in \mathrm{Dir}\, N$ platí: body $\boldsymbol{p} + \boldsymbol{u}$, $\boldsymbol{q} + \boldsymbol{v}$ tvoria priečku podpriestorov M, N práve vtedy, keď vektor $\boldsymbol{v} - \boldsymbol{u}$ je kolmým priemetom vektora $\boldsymbol{p} - \boldsymbol{q}$ do lineárneho podpriestoru $\mathrm{Dir}\, M + \mathrm{Dir}\, N$;

(c) existujú body $\boldsymbol{p} \in M$, $\boldsymbol{q} \in N$ tvoriace priečku podpriestorov M, N.

Dôkaz. (a) je bezprostredným dôsledkom vety 14.1.3 (b) a lemy 14.2.1;
(b) priamo vyplýva z (a). Konečne (c) dostaneme z (b) a vety 14.1.3 (a).

14.2.3. Dôsledok. *Pre konečnorozmerné afinné podpriestory M, $N \subseteq V$ vektorového priestoru so skalárnym súčinom platí $\mathrm{dist}(M, N) = 0$ práve vtedy, keď $M \cap N \neq \emptyset$.*

Podmienky 14.1.3 (a) a 14.2.2 (b) nám spolu s tvrdením 14.1.5 poskytujú priamy návod ako nájsť priečku a vzdialenosť ľubovoľných konečnorozmerných afinných podpriestorov. Ak $M = \boldsymbol{p} + [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m]$, $N = \boldsymbol{q} + [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]$ sú zadané parametricky, stačí nájsť jedno riešenie $\boldsymbol{c} = (c_1, \ldots, c_m, c_{m+1}, \ldots, c_{m+n})^{\mathsf{T}} \in \mathbb{R}^{m+n}$ sústavy

$$\boldsymbol{G}(\boldsymbol{\gamma}) \cdot \boldsymbol{c} = \langle \boldsymbol{p} - \boldsymbol{q}, \boldsymbol{\gamma} \rangle^{\mathsf{T}},$$

kde $\boldsymbol{\gamma} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$, a položiť

$$\boldsymbol{u} = c_1 \boldsymbol{u}_1 + \ldots c_m \boldsymbol{u}_m, \qquad \boldsymbol{v} = c_{m+1} \boldsymbol{v}_1 + \ldots + c_{m+n} \boldsymbol{v}_n.$$

Potom vektor $\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v} = \boldsymbol{\gamma} \cdot \boldsymbol{c}$ je kolmým priemetom vektora $\boldsymbol{p} - \boldsymbol{q}$ do lineárneho podpriestoru $\mathrm{Dir}\, M + \mathrm{Dir}\, N = [\boldsymbol{u}_1, \ldots \boldsymbol{u}_m, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]$ a priečka podpriestorov M, N je tvorená bodmi $\boldsymbol{p} - \boldsymbol{u}$, $\boldsymbol{q} + \boldsymbol{v}$. V dôsledku toho

$$\mathrm{dist}(M, N) = \|(\boldsymbol{p} - \boldsymbol{u}) - (\boldsymbol{q} + \boldsymbol{v})\| = \|\boldsymbol{p} - \boldsymbol{q} - \boldsymbol{w}\|.$$

Rozmyslite si, prečo – napriek formulácii tvrdenia 14.1.5 – si nemusíme robiť starosti s lineárnou nezávislosťou vektorov $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$.

14.2.4. Príklad. V euklidovskom priestore $\mathbb{R}^4$ so štandardným skalárnym súčinom nájdeme vzdialenosť rovín

$$M = (1, 1, 2, -2)^{\mathsf{T}} + [\boldsymbol{e}_1 + \boldsymbol{e}_2, \boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3],$$
$$N = (0, 0, 5, -1)^{\mathsf{T}} + [\boldsymbol{e}_2 + \boldsymbol{e}_4, \boldsymbol{e}_2 + \boldsymbol{e}_3 + \boldsymbol{e}_4].$$

Z príslušných skalárnych súčinov zostavíme (takmer Gramovu) rozšírenú maticu sústavy $\boldsymbol{G}(\boldsymbol{\gamma}) \cdot \boldsymbol{c} = \langle \boldsymbol{p} - \boldsymbol{q}, \boldsymbol{\gamma} \rangle^\mathsf{T}$ a upravíme ju na redukovaný stupňovitý tvar

$$\left(\begin{array}{cccc|c} 2 & 2 & 1 & 1 & 2 \\ 2 & 3 & 1 & 2 & -1 \\ 1 & 1 & 2 & 2 & 0 \\ 1 & 2 & 2 & 3 & -3 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 13/3 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 1 & 1 & -2/3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

So zreteľom na poslednú otázku si riešenie sústavy napíšeme vo všeobecnom tvare $\boldsymbol{c}_t = (13/3 + t, -3 - t, -2/3 - t, t)^\mathsf{T}$ s parametrom $t \in \mathbb{R}$. Položme

$$\begin{aligned} \boldsymbol{u}_t &= c_1(\boldsymbol{e}_1 + \boldsymbol{e}_2) + c_2(\boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3) = (4/3, 4/3, -3 - t, 0)^\mathsf{T}, \\ \boldsymbol{v}_t &= c_3(\boldsymbol{e}_2 + \boldsymbol{e}_4) + c_3(\boldsymbol{e}_2 + \boldsymbol{e}_3 + \boldsymbol{e}_4) = (0, -2/3, t, -2/3)^\mathsf{T} \end{aligned}$$

Potom pre každé $t \in \mathbb{R}$ dvojica bodov

$$\begin{aligned} \boldsymbol{p}_t &= (1, 1, 2, -2)^\mathsf{T} - \boldsymbol{u}_t = (-1/3, -1/3, 5 + t, -2)^\mathsf{T} \\ {}_t &= (0, 0, 5, -1)^\mathsf{T} + \boldsymbol{v}_t = (0, -2/3, 5 + t, -5/3)^\mathsf{T} \end{aligned}$$

tvorí priečku podpriestorov M, N. Vektory

$$\boldsymbol{u}_t + \boldsymbol{v}_t = (4/3, 2/3, -3, -2/3)^\mathsf{T}, \qquad \boldsymbol{p}_t - \boldsymbol{q}_t = \frac{1}{3}(-1, 1, 0, -1)^\mathsf{T}$$

však od parametra t nezávisia (a nie je to náhoda), rovnako ako vzdialenosť

$$\operatorname{dist}(M, N) = \|\boldsymbol{p}_t - \boldsymbol{q}_t\| = \frac{1}{\sqrt{3}}.$$

Pokiaľ by nás teda zaujímala len vzdialenosť podpriestorov M, N, prípadne by sme chceli nájsť len akúkoľvek ich jednu priečku, skutočne by stačilo použiť iba jedno riešenie $\boldsymbol{c}$ uvažovanej sústavy. Ešte si rozmyslite, ako zo získaných výsledkov vyplýva čiastočná rovnobežnosť podpriestorov M, N (pozri paragraf 8.4).

14.3 Odchýlka dvoch afinných podpriestorov

Odchýlku alebo *uhol* dvoch *netriviálnych konečnorozmerných afinných podpriestorov* vo vektorovom priestrore so skalárnym súčinom V značíme $\sphericalangle(M, N)$ a definujeme ju ako odchýlku $\sphericalangle(\operatorname{Dir} M, \operatorname{Dir} N)$ ich zameraní. Stačí teda povedať, čo rozumieme pod *odchýlkou* alebo *uhlom* $\sphericalangle(S, T)$ dvoch *netriválnych*

konečnorozmerných lineárnych podpriestorov S, $T \subseteq V$. Pre $S \subseteq T$ alebo $T \subseteq S$ položíme
$$\sphericalangle(S,T) = 0.$$
Ak $S \cap T = \{\mathbf{0}\}$, kladieme
$$\sphericalangle(S,T) = \inf\{\sphericalangle(\boldsymbol{x},\boldsymbol{y});\ \mathbf{0} \neq \boldsymbol{x} \in S \ \&\ \mathbf{0} \neq \boldsymbol{y} \in T\}.$$
Ak by sme takýmto spôsobom definovali odchýlku $\sphericalangle(S,T)$, aj keď $S \cap T \neq \{\mathbf{0}\}$, ľubovoľný spoločný nenulový vektor $\boldsymbol{x} \in S \cap T$ by sa postaral o to, aby platilo $\sphericalangle(S,T) = \sphericalangle(\boldsymbol{x},\boldsymbol{x}) = 0$, čo nevyzerá príliš rozumne. Preto pokiaľ $S \cap T \neq \{\mathbf{0}\}$, $S \not\subseteq T$ ani $T \not\subseteq S$, položíme
$$S_1 = S \cap (S \cap T)^\perp, \qquad T_1 = T \cap (S \cap T)^\perp.$$
Zrejme S_1, $T_1 \subseteq V$ sú netriválne lineárne podpriestory a $S_1 \cap T_1 = \{\mathbf{0}\}$ (za predpokladu $S \cap T = \{\mathbf{0}\}$ dokonca platí $S_1 = S$, $T_1 = T$). Preto môžeme konečne definovať
$$\sphericalangle(S,T) = \sphericalangle(S_1,T_1).$$
Takto definovaný uhol podpriestorov S, T je číslo z intervalu $\langle 0, \pi/2\rangle$ a platí preň $\sphericalangle(S,T) = \sphericalangle(T,S)$, teda je to *neorientovaný uhol.*

Je užitočné si uvedomiť, že na výpočet odchýlky dvoch podpriestorov stačí minimalizovať odchýlku vhodných vektorov z jedného podpriestoru od druhého z nich.

14.3.1. Tvrdenie. *Nech V je vektorový priestor so skalárnym súčinom a S, T sú jeho konečnorozmerné lineárne podpriestory, pričom $S \not\subseteq T$ ani $T \not\subseteq S$. Potom*
$$\sphericalangle(S,T) = \inf\big\{\sphericalangle(\boldsymbol{x},T);\ \mathbf{0} \neq \boldsymbol{x} \in S \cap (S \cap T)^\perp\big\}.$$

Dôkaz. Pre ľubovoľné $\boldsymbol{x} \neq \mathbf{0}$ je $\boldsymbol{x}_T \in T$ a $\boldsymbol{x} - \boldsymbol{x}_T \in T^\perp$. Pre $\boldsymbol{x} \in S \cap (S \cap T)^\perp$ však platí
$$\boldsymbol{x}_T = \boldsymbol{x} - (\boldsymbol{x} - \boldsymbol{x}_T) \in (S \cap T)^\perp + T^\perp = (S \cap T)^\perp,$$
lebo z inklúzie $S \cap T \subseteq T$ podľa podmienky 14.1.2 (c) vyplýva $T^\perp \subseteq (S \cap T)^\perp$. Teda $\boldsymbol{x}_T \in T \cap (S \cap T)^\perp$, preto $\sphericalangle(S,T)$ je menšia alebo rovná, ako výraz na pravej strane. Keďže arccos je klesajúca funkcia, opačná nerovnosť je dôsledkom rovnosti $\sphericalangle(\boldsymbol{x},T) = \arccos\big(\|\boldsymbol{x}_T\|\,\|\boldsymbol{x}\|^{-1}\big)$ a vety 14.1.3 (c).

Výpočet odchýlky dvoch všeobecných konečnorozmerných podpriestorov si teda vyžaduje minimalizovať hodnotu istého výrazu. To naznačuje možnosť jej výpočtu s použitím diferenciálneho počtu ako extrému funkcie viac premenných. Nebudeme sa však púšťať touto cestou, lebo neskôr hodláme predviesť elegantnejší spôsob vyjadrenia odchýlky. Prostriedkami, ktoré máme zatiaľ k dispozícii, dokážeme zrátať len odchýlku ľubovoľného netriviálneho konečnorozmerného podpriestoru od priamky alebo nadroviny, prípadne odchýlky podpriestorov, ktoré možno na tieto prípady previesť.

Ako vyplýva z vety 14.1.3 (c), odchýlka priamky $[\boldsymbol{x}]$, kde $\boldsymbol{x} \neq \mathbf{0}$, a konečnorozmerného lineárneho podpriestoru $S \neq \{\mathbf{0}\}$ je daná vzťahom

$$\sphericalangle\big([\boldsymbol{x}], S\big) = \sphericalangle(\boldsymbol{x}, S) = \arccos\frac{\|\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|} = \begin{cases} \sphericalangle(\boldsymbol{x}, \boldsymbol{x}_S), & \text{ak } \boldsymbol{x}_S \neq \mathbf{0}, \text{ t. j. ak } \boldsymbol{x} \notin S^{\perp}, \\ \pi/2, & \text{ak } \boldsymbol{x}_S = \mathbf{0}, \text{ t. j. ak } \boldsymbol{x} \in S^{\perp}. \end{cases}$$

Túto úlohu už vieme riešiť (pozri tvrdenie 14.1.5 a príklad 14.1.6).

14.3.2. Príklad. Vypočítame odchýlku rovín M, $N \subseteq \mathbb{R}^4$ z príkladu 14.2.4. Podľa definície $\sphericalangle(M, N) = \sphericalangle(S, T)$, kde $S = [\boldsymbol{e}_1 + \boldsymbol{e}_2, \boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3]$, $T = [\boldsymbol{e}_2 + \boldsymbol{e}_4, \boldsymbol{e}_2 + \boldsymbol{e}_3 + \boldsymbol{e}_4]$. Ľahko nahliadneme, že $S \cap T = [\boldsymbol{e}_3]$, teda $(S \cap T)^{\perp} = [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_4]$. V dôsledku toho

$$S_1 = S \cap (S \cap T)^{\perp} = [\boldsymbol{e}_1 + \boldsymbol{e}_2], \qquad T_1 = T \cap (S \cap T)^{\perp} = [\boldsymbol{e}_2 + \boldsymbol{e}_4].$$

Keďže $\langle \boldsymbol{e}_1 + \boldsymbol{e}_2, \boldsymbol{e}_2 + \boldsymbol{e}_4 \rangle = 1 \geq 0$ a $\|\boldsymbol{e}_1 + \boldsymbol{e}_2\| = \|\boldsymbol{e}_2 + \boldsymbol{e}_4\| = \sqrt{2}$,

$$\sphericalangle(M, N) = \sphericalangle(\boldsymbol{e}_1 + \boldsymbol{e}_2,\ \boldsymbol{e}_2 + \boldsymbol{e}_4) = \arccos\frac{1}{2} = \frac{\pi}{3} = 60^{\circ}.$$

Podľa dôsledku 14.1.4 (a) má každý $(n-1)$-rozmerný lineárny podpriestor S v n-rozmernom euklidovskom priestore V tvar $S = [\boldsymbol{a}]^{\perp}$ pre vhodný nenulový vektor $\boldsymbol{a} \in V$; potom každá nadrovina $N \subseteq V$ so zameraním S má tvar $N = \boldsymbol{p} + [\boldsymbol{a}]^{\perp}$ pre nejaké $\boldsymbol{p} \in N$. Vektor $\boldsymbol{a}$ sa nazýva *normála* alebo *normálový vektor* nadroviny N. Normála nadroviny je zrejme určená jednoznačne až na skalárny násobok. Výpočet odchýlky nadroviny a netriviálneho vlastného afinného podpriestoru možno previesť na výpočet odchýlky normály nadroviny a tohto podpriestoru. Tento prevod sa zakladá na nasledujúcom tvrdení.

14.3.3. Tvrdenie. *Nech S je netriviálny, vlastný lineárny podpriestor euklidovského priestoru V a $\mathbf{0} \neq \boldsymbol{a} \in V$. Potom*

$$\sphericalangle\big([\boldsymbol{a}]^{\perp}, S\big) = \frac{\pi}{2} - \sphericalangle(\boldsymbol{a}, S) = \sphericalangle\big(\boldsymbol{a}, S^{\perp}\big).$$

Dôkaz. Keďže druhá rovnosť je zrejmá, stačí dokázať prvú. Podmienky $\boldsymbol{a}_S = \mathbf{0}$, $\boldsymbol{a} \in S^{\perp}$ a $S \subseteq [\boldsymbol{a}]^{\perp}$ sú ekvivalentné. Podobne, $[\boldsymbol{a}]^{\perp} \subseteq S$ je ekvivalentné s podmienkou $S^{\perp} \subseteq [\boldsymbol{a}]$. V oboch prípadoch platí $\sphericalangle\big([\boldsymbol{a}]^{\perp}, S\big) = 0 = \sphericalangle\big(\boldsymbol{a}, S^{\perp}\big)$, $\sphericalangle(\boldsymbol{a}, S) = \pi/2$.

Nech teda $\boldsymbol{a}_S \neq \mathbf{0}$ a $[\boldsymbol{a}]$, $S^\perp$ ani $[\boldsymbol{a}]^\perp$, S nie sú vo vzťahu inklúzie. Označme $\boldsymbol{a}_{ST}$ kolmý priemet vektora $\boldsymbol{a}_S$ do podpriestoru $T = [\boldsymbol{a}]^\perp$. Postupne dokážeme rovnosti

$$\sphericalangle\big([\boldsymbol{a}]^\perp, S\big) = \arccos \frac{\|\boldsymbol{a}_{ST}\|}{\|\boldsymbol{a}_S\|} = \frac{\pi}{2} - \sphericalangle(\boldsymbol{a}, S).$$

Prvú z nich dokážme v dvoch krokoch. Najprv si uvedomme, že z dôsledku 14.1.4 (a) vyplýva $\big([\boldsymbol{a}]^\perp \cap S\big)^\perp = [\boldsymbol{a}] + S^\perp$, teda

$$\boldsymbol{a}_S = \boldsymbol{a} - (\boldsymbol{a} - \boldsymbol{a}_S) \in [\boldsymbol{a}] + S^\perp = \big([\boldsymbol{a}]^\perp \cap S\big)^\perp.$$

Takže $\sphericalangle\big([\boldsymbol{a}]^\perp, S\big) \leq \arccos\big(\|\boldsymbol{a}_{ST}\|\,\|\boldsymbol{a}_S\|^{-1}\big)$ na základe tvrdenia 14.3.1. Na dôkaz opačnej nerovnosti stačí podľa toho istého tvrdenia overiť

$$\frac{\|\boldsymbol{a}_{ST}\|}{\|\boldsymbol{a}_S\|} \geq \frac{\|\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|}$$

pre všetky $\mathbf{0} \neq \boldsymbol{x} \in [\boldsymbol{a}]^\perp \cap \big([\boldsymbol{a}]^\perp \cap S\big)^\perp$. Zrejme $\langle \boldsymbol{x}, \boldsymbol{a}_S\rangle = \langle \boldsymbol{x}_S, \boldsymbol{a}_S\rangle$, lebo $\boldsymbol{x} - \boldsymbol{x}_S \in S^\perp$. Keďže $\boldsymbol{x} \in [\boldsymbol{a}] + S^\perp$, vektory $\boldsymbol{x}_S$, $\boldsymbol{a}_S$ sú lineárne závislé (rozmyslite si prečo); preto $|\langle \boldsymbol{x}_S, \boldsymbol{a}_S\rangle| = \|\boldsymbol{x}_S\|\,\|\boldsymbol{a}_S\|$. Nakoľko $\boldsymbol{x} \in [\boldsymbol{a}]^\perp = T$, s použitím vety 14.1.3 (c) z toho vyplýva

$$\frac{\|\boldsymbol{x}_S\|}{\|\boldsymbol{x}\|} = \frac{|\langle \boldsymbol{x}_S, \boldsymbol{a}_S\rangle|}{\|\boldsymbol{x}\|\,\|\boldsymbol{a}_S\|} = \frac{|\langle \boldsymbol{x}, \boldsymbol{a}_S\rangle|}{\|\boldsymbol{x}\|\,\|\boldsymbol{a}_S\|} \leq \frac{\|\boldsymbol{a}_{ST}\|}{\|\boldsymbol{a}_S\|}.$$

Druhú rovnosť overíme priamym výpočtom

$$\begin{aligned}\frac{\pi}{2} - \sphericalangle(\boldsymbol{a}, S) &= \frac{\pi}{2} - \sphericalangle(\boldsymbol{a}_S, \boldsymbol{a}) = \frac{\pi}{2} - \sphericalangle\big(\boldsymbol{a}_S, [\boldsymbol{a}]\big)\\ &= \sphericalangle\big(\boldsymbol{a}_S, [\boldsymbol{a}]^\perp\big) = \arccos \frac{\|\boldsymbol{a}_{ST}\|}{\|\boldsymbol{a}_S\|}.\end{aligned}$$

Ako zvláštny prípad pre odchýlku dvoch nadrovín dostávame

14.3.4. Dôsledok. *Nech M, N sú dve nadroviny v euklidovskom priestore V s normálami $\boldsymbol{a}$, resp. $\boldsymbol{b}$. Potom*

$$\sphericalangle(M, N) = \sphericalangle(\boldsymbol{a}, [\boldsymbol{b}]) = \min\big\{\sphericalangle(\boldsymbol{a}, \boldsymbol{b}),\ \sphericalangle(\boldsymbol{a}, -\boldsymbol{b})\big\}.$$

V euklidovskom priestore $\mathbb{R}^n$ so štandardným skalárnym súčinom vystupuje normálový vektor danej nadroviny priamo v jej (všeobecnej) rovnici. Ak je totiž nadrovina M daná rovnicou

$$a_1 x_1 + \ldots + a_n x_n = b,$$

tak $\boldsymbol{a} = (a_1, \ldots, a_n)^\mathsf{T} \neq \mathbf{0}$ je jej normála a uvedenú rovnicu možno skrátene zapísať v tvare $\langle \boldsymbol{x}, \boldsymbol{a}\rangle = b$.

14.3.5. Príklad. V euklidovskom priestore V vypočítame odchýlku roviny $S = [\boldsymbol{u}, \boldsymbol{v}]$ a nadroviny $T = [\boldsymbol{a}]^\perp$. Podľa tvrdenia 14.3.3 platí

$$\sphericalangle(S,T) = \frac{\pi}{2} - \sphericalangle(\boldsymbol{a}, S) = \arcsin\frac{\|\boldsymbol{a}_S\|}{\|\boldsymbol{a}\|}.$$

Súradnice c, d kolmého priemetu $\boldsymbol{a}_S = c\boldsymbol{u} + d\boldsymbol{v}$ vzhľadom na bázu $(\boldsymbol{u}, \boldsymbol{v})$ podpriestoru S získame riešením sústavy

$$\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v}) \cdot \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} \langle \boldsymbol{a}, \boldsymbol{u} \rangle \\ \langle \boldsymbol{a}, \boldsymbol{v} \rangle \end{pmatrix}$$

podľa Cramerovho pravidla v tvare

$$c = \frac{\begin{vmatrix} \langle \boldsymbol{a}, \boldsymbol{u} \rangle & \langle \boldsymbol{v}, \boldsymbol{u} \rangle \\ \langle \boldsymbol{a}, \boldsymbol{v} \rangle & \langle \boldsymbol{v}, \boldsymbol{v} \rangle \end{vmatrix}}{|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})|}, \qquad d = \frac{\begin{vmatrix} \langle \boldsymbol{u}, \boldsymbol{u} \rangle & \langle \boldsymbol{a}, \boldsymbol{u} \rangle \\ \langle \boldsymbol{u}, \boldsymbol{v} \rangle & \langle \boldsymbol{a}, \boldsymbol{v} \rangle \end{vmatrix}}{|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})|}.$$

Spätné dosadenie do vzorca pre odchýlku $\sphericalangle(S,T)$ si odpustíme. Tieto vzorce možno zrejmým spôsobom zovšeobecniť aj pre odchýlku k-rozmerného podpriestoru S a priamky, resp. nadroviny. vo V. Pre $k \geq 3$ však výpočet pomocou Gramových determinantov už nie je výhodný.

14.3.6. Príklad. Podaktorý čitateľ si možno kladie otázku, prečo sme miesto tvrdenia 14.3.3 nedokázali silnejší výsledok, totiž rovnosť

$$\sphericalangle(S,T) + \sphericalangle\big(S^\perp, T\big) = \frac{\pi}{2},$$

pre *každý netriválny vlastný* (teda nie len jednorozmerný) lineárny podpriestor S a každý netriviálny lineárny podpriestor T euklidovského priestoru V. Uvedené tvrdenie, ako aj náš geometrický názor nám totiž napovedajú, že také niečo by malo platiť. O to bizarnejší sa nám preto bude zdať nasledujúci veľmi jednoduchý príklad, ktorý v ľubovoľnej dimenzii $n > 3$ našu domnienku vyvracia.

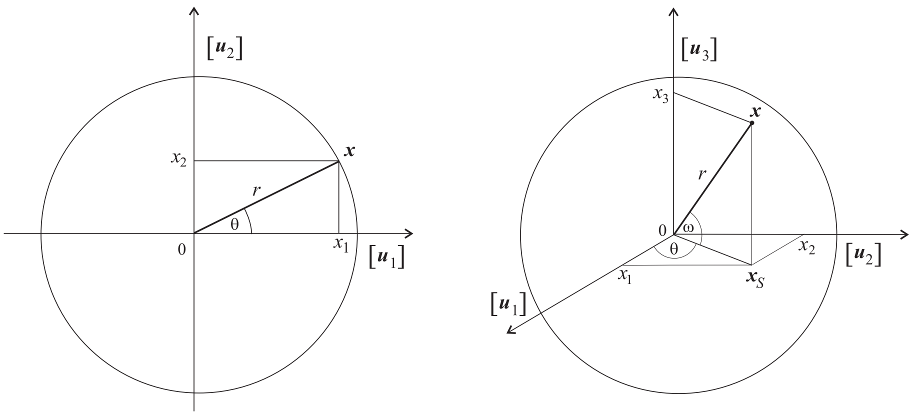
V euklidovskom priestore $\mathbb{R}^n$ so štandardným skalárnym súčinom sú dané lineárne podpriestory $S = [\boldsymbol{e}_1, \boldsymbol{e}_2]$, $T = [\boldsymbol{e}_1, \boldsymbol{e}_3]$. Zrejme $S \cap T = [\boldsymbol{e}_1]$, teda $S \cap [\boldsymbol{e}_1]^\perp = [\boldsymbol{e}_2]$ a $T \cap [\boldsymbol{e}_1]^\perp = [\boldsymbol{e}_3]$. V dôsledku toho

$$\sphericalangle(S,T) = \sphericalangle\big([\boldsymbol{e}_2], [\boldsymbol{e}_3]\big) = \frac{\pi}{2}.$$

Na druhej strane $S^\perp = [\boldsymbol{e}_1, \boldsymbol{e}_2]^\perp = [\boldsymbol{e}_3, \ldots, \boldsymbol{e}_n]$. Takže $S^\perp \cap T = [\boldsymbol{e}_3]$, z čoho dostávame $S^\perp \cap [\boldsymbol{e}_3]^\perp = [\boldsymbol{e}_4, \ldots, \boldsymbol{e}_n]$ a $T \cap [\boldsymbol{e}_3]^\perp = [\boldsymbol{e}_1]$. Konečne

$$\sphericalangle\big(S^\perp, T\big) = \sphericalangle\big([\boldsymbol{e}_4, \ldots, \boldsymbol{e}_n], [\boldsymbol{e}_1]\big) = \frac{\pi}{2}.$$

Teda hodnota súčtu $\sphericalangle(S,T) + \sphericalangle\big(S^\perp, T\big)$ je tentokrát π, a nie $\pi/2$, ako sme očakávali. Voľne povedané, podpriestor T je kolmý tak na podpriestor S, ako aj na jeho ortokomplement $S^\perp$. Samozrejme, platí tiež $\sphericalangle\big(S, S^\perp\big) = \pi/2$.



Obr. 14.2. Polárne súradnice v rovine $\mathbb{R}^2$ a sférické súradnice v priestore $\mathbb{R}^3$

14.4 Polárne a sférické súradnice

Ortonormálne bázy umožňujú zaviesť v euklidovských priestoroch aj iné typy súradníc, ako sme používali doteraz.

Polárne súradnice v rovine sú obdobou goniometrického vyjadrenia komplexných čísel. Ak $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \boldsymbol{u}_2)$ je nejaká ortonormálna báza dvojrozmerného euklidovského priestoru V, tak *polárne* (prípadne tiež *sférické*) *súradnice* vektora $\boldsymbol{x} \in V$, so súradnicami $(\boldsymbol{x})_{\boldsymbol{\alpha}} = (x_1, x_2)^\mathsf{T}$ vzhľadom na bázu $\boldsymbol{\alpha}$, tvorí usporiadaná dvojica (r, θ) reálnych čísel $r \geq 0$, $-\pi < \theta \leq \pi$ takých, že

$$
\begin{aligned}
x_1 &= r \cos \theta, \\
x_2 &= r \sin \theta.
\end{aligned}
$$

(Rovnako dobre by sme mohli vziať θ z intervalu $\langle 0, 2\pi)$.) Inak povedané, $r = \|\boldsymbol{x}\|$ je dĺžka vektora $\boldsymbol{x}$ a pre $\boldsymbol{x} \neq \boldsymbol{0}$ je θ *orientovaný uhol* vektorov $\boldsymbol{u}_1, \boldsymbol{x}$, t. j. uhol, o ktorý treba otočiť vektor $\boldsymbol{u}_1$, aby splynul s vektorom $r^{-1}\boldsymbol{x}$; pri otočení v kladnom zmysle (proti smeru hodinových ručičiek) je $\theta \geq 0$, v zápornom zmysle $\theta \leq 0$. Prípad $\boldsymbol{x} = \boldsymbol{0}$ je *singulárny*: jeho polárne súradnice tvorí každá usporiadaná dvojica tvaru $(0, \theta)$, kde $\theta \in (-\pi, \pi\rangle$.

Sférické súradnice v priestore majú názornú geografickú interpretáciu: Nájdeme ich tak, že koncovým bodom daného vektora $\boldsymbol{x}$ preložíme „glóbus“ so stredom v počiatku a polomerom $r = \|\boldsymbol{x}\|$, a určíme „zemepisnú dĺžku“ θ a „zemepisnú šírku“ ω tohto koncového bodu.

Presnejšie, ak $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$ je nejaká ortonormálna báza trojrozmerného euklidovského priestoru V, tak *sférické súradnice* vektora $\boldsymbol{x} \in V$ so súradnicami $(\boldsymbol{x})_{\boldsymbol{\alpha}} = (x_1, x_2, x_3)^\mathsf{T}$ vzhľadom na bázu $\boldsymbol{\alpha}$ tvorí usporiadaná trojica (r, θ, ω) reálnych čísel $r \geq 0$, $-\pi < \theta \leq \pi$, $-\frac{1}{2}\pi \leq \omega \leq \frac{1}{2}\pi$ takých, že

$$\begin{aligned} x_1 &= r\cos\theta\cos\omega, \\ x_2 &= r\sin\theta\cos\omega, \\ x_3 &= r\sin\omega. \end{aligned}$$

Teda $r = \|\boldsymbol{x}\|$ je opäť dĺžka vektora $\boldsymbol{x}$. Ak $\boldsymbol{x} \notin [\boldsymbol{u}_3]$, t. j ak $(x_1, x_2) \neq (0,0)$, tak θ je orientovaný uhol, ktorý zviera vektor $\boldsymbol{u}_1$ a kolmý priemet $\boldsymbol{x}_S$ vektora $\boldsymbol{x}$ do roviny $S = [\boldsymbol{u}_1, \boldsymbol{u}_2]$, a ω je orientovaný uhol, ktorý zviera tento kolmý priemet a pôvodný vektor $\boldsymbol{x}$.[1] Inak povedané, $(\|\boldsymbol{x}_S\|, \theta)$ sú polárne súradnice vektora $\boldsymbol{x}_S \in S$ vzhľadom na ortonormálnu bázu $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ roviny S. Priamka $[\boldsymbol{u}_3]$ je singulárna: ak $\mathbf{0} \neq \boldsymbol{x} \in [\boldsymbol{u}_3]$, tak $\theta \in (-\pi, \pi\rangle$ môže byť ľubovoľné a $\omega = \pm\frac{1}{2}\pi$, podľa toho, či $x_3 \gtrless 0$; ak $\boldsymbol{x} = \mathbf{0}$, tak aj $\omega \in \langle -\pi/2, \pi/2\rangle$ môže byť ľubovoľné.

Ak sa obmedzíme len na sférickú plochu s daným polomerom r a stredom v počiatku, stačí, samozrejme, sférické súradnice bodov na jej povrchu udávať v tvare dvojice „zemepisných súradníc“ (θ, ω).

Pozorný čitateľ v prechode od polárnych k sférickým súradniciam asi zahliadol všeobecnejšiu rekurentnú schému. Ak $n \geq 3$ a $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ je ortonormálna báza n-rozmerného euklidovského priestoru V, tak *sférické súradnice* vektora $\boldsymbol{x} \in V$ vzhľadom na bázu $\boldsymbol{\alpha}$ tvorí usporiadaná n-tica $(r, \theta, \omega_1, \ldots, \omega_{n-2})$ reálnych čísel $r \geq 0$, $-\pi < \theta \leq \pi$, $-\frac{1}{2}\pi \leq \omega_i \leq \frac{1}{2}\pi$ pre $1 \leq i \leq n-2$, takých, že

$$\begin{aligned} x_1 &= r\cos\theta\cos\omega_1\cos\omega_2\cos\omega_3\ldots\cos\omega_{n-2}, \\ x_2 &= r\sin\theta\cos\omega_1\cos\omega_2\cos\omega_3\ldots\cos\omega_{n-2}, \\ x_3 &= r\sin\omega_1\cos\omega_2\cos\omega_3\ldots\cos\omega_{n-2}, \\ x_4 &= r\sin\omega_2\cos\omega_3\ldots\cos\omega_{n-2}, \\ &\vdots \\ x_{n-1} &= r\sin\omega_{n-3}\cos\omega_{n-2}, \\ x_n &= r\sin\omega_{n-2}. \end{aligned}$$

Inak povedané, $r = \|\boldsymbol{x}\|$ a pre $\boldsymbol{x} \notin [\boldsymbol{u}_n]$ sú $(\|\boldsymbol{x}_S\|, \theta, \omega_1, \ldots, \omega_{n-3})$ sférické súradníce kolmého priemetu $\boldsymbol{x}_S$ vektora $\boldsymbol{x}$ do nadroviny $S = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-1}]$ vzhľadom na jej ortonormálnu bázu $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-1})$; ω_{n-2} je orientovaný uhol vektorov $\boldsymbol{x}_S$ a $\boldsymbol{x}$. Diskusiu singulárneho prípadu $\boldsymbol{x} \in [\boldsymbol{u}_n]$ prenechávame ako cvičenie čitateľovi. (Uvedomte si, že singulárny je vlastne celý $(n-2)$-rozmerný podpriestor $[\boldsymbol{u}_1, \boldsymbol{u}_2]^\perp$!)

Sférické súradnice možno s výhodou použiť pri riešení úloh so sférickou

[1] Niekedy sa v sférických súradniciach miesto uhla ω uvažuje uhol $\pi/2 - \omega$.

symetriou. Ako príklad môže poslúžiť výpočet n-rozmerných integrálov

$$\int \cdots \int f(x_1, \ldots, x_n)\,\mathrm{d}x_1 \ldots \mathrm{d}x_n$$

cez n-rozmernú guľu $B^{(n)}(R) = \{\boldsymbol{x} \in \mathbb{R}^n;\ \|\boldsymbol{x}\| \le R\}$ s polomerom R v euklidovskom priestore $\mathbb{R}^n$ so štandardným skalárnym súčinom.[2] Tie totiž možno substitúciou sférických súradníc previesť na integrály tvaru

$$\int \ldots \int F(r, \theta, \omega_1, \ldots, \omega_{n-2}) \left| \frac{\partial(x_1, \ldots, x_n)}{\partial(r, \theta, \omega_1, \ldots, \omega_{n-2})} \right| \mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\omega_1 \ldots \mathrm{d}\omega_{n-2}$$

cez karteziánsky súčin intervalov $\langle 0, R\rangle \times \langle -\pi, \pi\rangle \times \langle -\pi/2, \pi/2\rangle^{n-2}$, kde

$$F(r, \theta, \omega_1, \ldots, \omega_{n-2}) = f(x_1, \ldots, x_n),$$

pričom pôvodné súradnice $x_1, \ldots, x_n$ chápeme ako funkcie sférických súradníc r, θ, $\omega_1, \ldots, \omega_{n-2}$, a výraz $\frac{\partial(x_1,\ldots,x_n)}{\partial(r,\theta,\omega_1,\ldots,\omega_{n-2})}$ označuje tzv. *jakobián*, t. j. determinant *Jacobiho matice* sférickej transformácie súradníc

$$\begin{pmatrix} \frac{\partial x_1}{\partial r} & \frac{\partial x_1}{\partial \theta} & \frac{\partial x_1}{\partial \omega_1} & \cdots & \frac{\partial x_1}{\partial \omega_{n-2}} \\ \frac{\partial x_2}{\partial r} & \frac{\partial x_2}{\partial \theta} & \frac{\partial x_2}{\partial \omega_1} & \cdots & \frac{\partial x_2}{\partial \omega_{n-2}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial r} & \frac{\partial x_n}{\partial \theta} & \frac{\partial x_n}{\partial \omega_1} & \cdots & \frac{\partial x_n}{\partial \omega_{n-2}} \end{pmatrix}.$$

Priamym výpočtom, ktorý prenechávame ako cvičenie čitateľovi, sa možno presvedčiť, že jakobián sférických súradníc je

$$\frac{\partial(x_1, \ldots, x_n)}{\partial(r, \theta, \omega_1, \ldots, \omega_{n-2})} = r^{n-1} \cos\omega_1 \cos^2\omega_2 \ldots \cos^{n-2}\omega_{n-2}.$$

14.4.1. Príklad. Vypočítame objem štvorrozmernej gule $B^{(4)}(R)$ v euklidovskom priestore $\mathbb{R}^4$. Tento objem je daný štvorrozmerným integrálom

$$V_4(R) = \iiiint_{B^{(4)}(R)} \mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3\,\mathrm{d}x_4,$$

Po substitúcii sférických súradníc

$$\begin{aligned} x_1 &= r\cos\theta\cos\omega_1\cos\omega_2, \\ x_2 &= r\sin\theta\cos\omega_1\cos\omega_2, \\ x_3 &= r\sin\omega_1\cos\omega_2, \\ x_4 &= r\sin\omega_2 \end{aligned}$$

[2]Sledovanie ďalšieho textu a nasledujúceho príkladu si od čitateľa vyžaduje základné znalosti z integrálneho počtu funkcií viac premenných.

a s využitím *Fubiniho vety* prejde tento integrál na súčin štyroch jednoduchých integrálov

$$\begin{aligned}
V_4(R) &= \int_0^R \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} \int_{-\pi/2}^{\pi/2} \left| \frac{\partial(x_1, x_2, x_3, x_4)}{\partial(r, \theta, \omega_1, \omega_2)} \right| \mathrm{d}r \, \mathrm{d}\theta \, \mathrm{d}\omega_1 \, \mathrm{d}\omega_2 \\
&= \int_0^R r^3 \, \mathrm{d}r \int_{-\pi}^{\pi} \mathrm{d}\theta \int_{-\pi/2}^{\pi/2} \cos\omega_1 \, \mathrm{d}\omega_1 \int_{-\pi/2}^{\pi/2} \cos^2 \omega_2 \, \mathrm{d}\omega_2 \\
&= \left[\frac{r^4}{4} \right]_0^R \big[\theta\big]_{-\pi}^{\pi} \big[\sin\omega_1\big]_{-\pi/2}^{\pi/2} \left[\frac{2\omega_2 + \sin 2\omega_2}{4} \right]_{-\pi/2}^{\pi/2} \\
&= \frac{R^4}{4} \cdot 2\pi \cdot 2 \cdot \frac{\pi}{2} = \frac{\pi^2}{2} R^4.
\end{aligned}$$

Poznamenajme, že vychádzajúc zo známych hodnôt $V_1(R) = 2R$, $V_2(R) = \pi R^2$, možno objemy viacrozmených gúľ $B^{(n)}(R)$ vypočítať aj na základe rekurentného vzťahu

$$V_{n+2}(R) = 2\pi \int_0^R r V_n(r) \, \mathrm{d}r = \frac{2\pi}{n+2} \, R^2 V_n(R),$$

využívajúceho zovšeobecnenú integrálnu formulu pre objem rotačných telies.

14.5 Riešenie neriešiteľných sústav a lineárna regresia

V celom tomto paragrafe označujú m, n pevné kladné celé čísla. Stĺpcové vektorové priestory $\mathbb{R}^m$ a $\mathbb{R}^n$ sú vybavené štandardným skalárnym súčinom, takže to sú euklidovské priestory.

Nech $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$. Uvažujme sústavu lineárnych rovníc

$$\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$$

a označme $S = [\boldsymbol{s}_1(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A})]$ lineárny podpriestor v $\mathbb{R}^m$ generovaný stĺpcami matice $\boldsymbol{A}$. Podľa Frobeniovej vety naša sústava má nejaké riešenie $\boldsymbol{x} \in \mathbb{R}^n$ práve vtedy, keď $\boldsymbol{b} \in S$. Zložky riešenia $\boldsymbol{x} = (x_1, \dots, x_n)^{\mathsf{T}} \in \mathbb{R}^n$ sú potom koeficienty lineárnej kombinácie

$$x_1 \boldsymbol{s}_1(\boldsymbol{A}) + \ldots + x_n \boldsymbol{s}_n(\boldsymbol{A}) = \boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}.$$

No i v prípade, keď $\boldsymbol{b} \notin S$, t. j. riešenie sústavy neexistuje, sa môžeme pokúsiť nahradiť jej pravú stranu $\boldsymbol{b}$ čo najbližším vektorom podpriestoru S. Takto získaná nová sústava už má riešenie, ktoré môžeme právom považovať za najlepšie možné približné riešenie pôvodnej sústavy. Podľa vety 14.1.3 (b) je najbližší vektor podpriestoru S k vektoru $\boldsymbol{b}$ určený jednoznačne, a je to jeho kolmý priemet $\boldsymbol{b}_S$ do tohto podpriestoru. *Pseudoriešenie* neriešiteľnej (hovorí sa tiež *preurčenej*) sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ teda definujeme ako riešenie (tento raz už istotne riešiteľnej) sústavy

$$\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}_S.$$

Ak je pôvodná sústava riešiteľná, t. j. ak $\boldsymbol{b} \in S$, tak $\boldsymbol{b}_S = \boldsymbol{b}$ a obe sústavy splývajú, takže každé jej pseudoriešenie je priamo riešením pôvodnej sústavy.

14.5.1. Tvrdenie. *Nech $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$. Potom $\boldsymbol{x} \in \mathbb{R}^n$ je pseudoriešením sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ práve vtedy, keď $\boldsymbol{x}$ je riešením sústavy*

$$\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{b}$$

so štvorcovou maticou $\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A} \in \mathbb{R}^{n\times n}$ a ľavou stranou $\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{b} \in \mathbb{R}^n$.

Dôkaz. Najprv si uvedomme, že $\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A} = \boldsymbol{G}(\boldsymbol{s}_1(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A}))$ je Gramovou maticou prislúchajúcou stĺpcom matice $\boldsymbol{A}$. Podľa tvrdenia 14.1.5 a za ním nasledujúcej poznámky je lineárna kombinácia $x_1\boldsymbol{s}_1(\boldsymbol{A}) + \ldots + x_n\boldsymbol{s}_n(\boldsymbol{A}) = \boldsymbol{A}\cdot\boldsymbol{x}$ kolmým priemetom vektora $\boldsymbol{b}$ do podpriestoru $S = [\boldsymbol{s}_1(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A})]$, t. j. platí $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}_S$, práve vtedy, keď $\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{b}$.

Pseudoriešenie preurčenej sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ teda hľadáme ako riešenie zaručene riešiteľnej sústavy $\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{b}$. V typickom prípade má pôvodná sústava viac rovníc než neznámych, čiže $m > n$ a $\boldsymbol{A}$ je obdĺžniková matica, „vyššia ako širšia“. Potom je veľmi pravdepodobné, že štvorcová matica $\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A}$ rádu n (ako Gramova matica „malého“ počtu stĺpcových vektorov v euklidovskom priestore „veľkej“ dimenzie) je regulárna, teda k nej existuje inverzná matica $\left(\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A}\right)^{-1}$. V takom prípade je pseudoriešenie pôvodnej sústavy určené jednoznačne:

$$\boldsymbol{x} = \left(\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A}\right)^{-1} \cdot \boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{b}.$$

Samozrejme, ak $m = n$ a už samotná matica $\boldsymbol{A}$ je regulárna, dostávame

$$\left(\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A}\right)^{-1} \cdot \boldsymbol{A}^\mathsf{T} = \boldsymbol{A}^{-1}$$

a $\boldsymbol{x} = \boldsymbol{A}^{-1} \cdot \boldsymbol{b}$ je priamo jediným riešením pôvodnej sústavy.

V úlohách *lineárnej regresie* máme zadané hodnoty $y_1, \dots, y_m$ neznámej funkcie f v bodoch $x_1, \dots, x_m$ jej definičného oboru, získané väčšinou meraním. Funkciu f chceme aproximovať lineárnou kombináciou funkcií $f_1, \dots, f_n$, ktoré poznáme, či aspoň sú nám známe ich hodnoty $a_{ij} = f_j(x_i)$ v bodoch $x_1, \dots, x_m$. Zvyčajne je m podstatne väčšie ako n. V optimálnom prípade sa nám môže podariť zostrojiť funkciu $f = c_1 f_1 + \ldots + c_n f_n$ priamo ako lineárnu kombináciu funkcií f_j tak, aby f v bodoch x_i nadobúdala vopred predpísané hodnoty y_i, t. j.

$$y_i = f(x_i) = \sum_{j=1}^{n} c_j f_j(x_i) = \sum_{j=1}^{n} a_{ij} c_j.$$

Ak označíme $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{m\times n}$, $\boldsymbol{y} = (y_1, \dots, y_m)^\mathsf{T} \in \mathbb{R}^m$, $\boldsymbol{c} = (c_1, \dots, c_n)^\mathsf{T} \in \mathbb{R}^n$, vidíme, že vlastne hľadáme riešenie $\boldsymbol{c}$ sústavy

$$\boldsymbol{A} \cdot \boldsymbol{c} = \boldsymbol{y}.$$

Táto sústava je v typickom prípade preurčená, teda jej riešenie neexistuje. Úloha lineárnej regresie potom splýva s *metódou najmenších štvorcov* a spočíva v nájdení takých koeficientov c_j, ktoré minimalizujú výraz

$$\sum_{i=1}^{m} \Big(y_i - \sum_{j=1}^{n} a_{ij} c_j \Big)^2 = \|\boldsymbol{y} - \boldsymbol{A}\cdot\boldsymbol{c}\|^2 .$$

Toto minimum sa nadobúda pre $\boldsymbol{c}$ také, že $\boldsymbol{A}\cdot\boldsymbol{c} = \boldsymbol{y}_S$, kde S je podpriestor v $\mathbb{R}^m$ generovaný stĺpcami matice $\boldsymbol{A}$. Inak povedané, hľadanú lineárnu kombináciu dostaneme pre pseudoriešenie $\boldsymbol{c}$ sústavy $\boldsymbol{A}\cdot\boldsymbol{c} = \boldsymbol{y}$. Pre hodnoty pochádzajúce z rozumne postavených praktických úloh je takmer isté, že matica $\boldsymbol{A}^\mathsf{T}\cdot\boldsymbol{A}$ je regulárna. V takom prípade

$$\boldsymbol{c} = \big(\boldsymbol{A}^\mathsf{T}\cdot\boldsymbol{A}\big)^{-1}\cdot\boldsymbol{A}^\mathsf{T}\cdot\boldsymbol{y},$$

čiže hľadaná lineárna kombinácia $f = c_1 f_1 + \ldots + c_n f_n = (f_1, \dots, f_n)\cdot\boldsymbol{c}$ je určená jednoznačne.

Metódu lineárnej regresie, ako aj to, čo približne rozumieme pod „rozumne postavenou úlohou“, si ilustrujeme na jednom typickom a dôležitom príklade.

14.5.2. Príklad. V rovine $\mathbb{R}^2$ je daných $m \geq 2$ bodov $(x_1, y_1), \dots, (x_m, y_m)$, získaných meraním hodnôt nejakej neznámej funkcie f vo vybraných bodoch x_i jej definičného oboru. Túto funkciu hodláme aproximovať priamkou s rovnicou $y = a + bx$ tak, aby výraz $\sum_{i=1}^{m}(y_i - a - bx_i)^2$ bol minimálny. Ak si uvedomíme, že funkcia $y = a + bx$ je lineárnou kombináciou konštantnej funkcie $y = 1$ a identickej funkcie $y = x$, hneď vidíme, že ide o úlohu lineárnej regresie. Sústava

$$\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & \boldsymbol{x}_m \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

je okrem triviálneho prípadu, keď všetky body (x_i, y_i) ležia na jednej priamke, preurčená. Koeficienty a, b teda nájdeme ako pseudoriešenie tejto sústavy. Jej maticu si označíme $\boldsymbol{A}$. Jednoduchý výpočet dáva

$$\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A} = \begin{pmatrix} m & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix},$$

$$\det\bigl(\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A}\bigr) = m \sum_{i=1}^{m} x_i^2 - \left(\sum_{i=1}^{m} x_i\right)^2 = \sum_{i<j} (x_i - x_j)^2.$$

Teda $\det\bigl(\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A}\bigr) = 0$ práve vtedy, keď $x_1 = x_2 = \ldots = x_m$. Asi sa zhodneme na tom, že v takomto prípade by šlo, mierne povedané, o veľmi nerozumne postavenú úlohu. Naopak, na základe jej formulácie je prirodzené očakávať, že všetky hodnoty x_i budú navzájom rôzne. V čo len trochu rozumnom prípade teda dostávame jednoznačne určené koeficienty a, b v tvare

$$\begin{aligned}
\begin{pmatrix} a \\ b \end{pmatrix} = \bigl(\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A}\bigr)^{-1} \cdot \boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{y} &= \frac{1}{\sum_{i<j} (x_i - x_j)^2} \begin{pmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & m \end{pmatrix} \cdot \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix} \\
&= \frac{1}{\sum_{i<j} (x_i - x_j)^2} \begin{pmatrix} \sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i \\ m \sum x_i y_i - \sum x_i \sum y_i \end{pmatrix} \\
&= \frac{1}{\sum_{i<j} (x_i - x_j)^2} \begin{pmatrix} \sum_{i<j} (x_i - x_j)(x_i y_j - x_j y_i) \\ \sum_{i<j} (x_i - x_j)(y_i - y_j) \end{pmatrix}.
\end{aligned}$$

V krátkosti preskúmame aj duálnu situáciu. Ak má sústava $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ viac, t. j. nekonečne mnoho riešení (hovoríme tiež, že je *podurčená*), môže nám veľkosť riešenia (v zmysle euklidovskej normy) poslúžiť ako dodatočné kritérium výberu toho najvhodnejšieho (napr. najlacnejšieho) z nich. Hovoríme, že $\boldsymbol{x} \in \mathbb{R}^n$ je *minimálne riešenie* uvedenej sústavy, ak je jej riešením a pre každé riešenie $\boldsymbol{y}$ tejto sústavy platí $\|\boldsymbol{x}\| \leq \|\boldsymbol{y}\|$. Samozrejme, ak má sústava jediné riešenie, tak je to zároveň jej minimálne riešenie.

14.5.3. Tvrdenie. *Nech $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$. Potom $\boldsymbol{x} \in \mathbb{R}^n$ je minimálnym riešením sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ práve vtedy, keď má tvar $\boldsymbol{x} = \boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{z}$, kde $\boldsymbol{z} \in \mathbb{R}^m$ je riešením sústavy*

$$\boldsymbol{A} \cdot \boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{z} = \boldsymbol{b}$$

so štvorcovou maticou $\boldsymbol{A} \cdot \boldsymbol{A}^\mathsf{T} \in \mathbb{R}^{m \times m}$ a ľavou stranou $\boldsymbol{b} \in \mathbb{R}^m$.

Dôkaz. Pripomeňme si, že množina riešení $\mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b})$ sústavy $\boldsymbol{A}\cdot\boldsymbol{x}=\boldsymbol{b}$ tvorí afinný podpriestor v $\mathbb{R}^n$ a jeho zameraním je lineárny podpriestor $\mathcal{R}(\boldsymbol{A})$ všetkých riešení homogénnej sústavy $\boldsymbol{A}\cdot\boldsymbol{x}=\boldsymbol{0}$. Podľa definície vektor $\boldsymbol{x}\in\mathbb{R}^n$ je minimálnym riešením pôvodnej sústavy práve vtedy, keď body $\boldsymbol{0}$ a $\boldsymbol{x}$ tvoria priečku afinných podpriestorov $\{\boldsymbol{0}\}$ a $\mathcal{R}(\boldsymbol{A}\,|\,\boldsymbol{b})$, čo je podľa tvrdenia 14.2.2 (a) ekvivalentné s podmienkou $\boldsymbol{x}\in\mathcal{R}(\boldsymbol{A})^{\perp}$. Ak si uvedomíme, že $\mathcal{R}(\boldsymbol{A})=R^{\perp}$ pre lineárny podpriestor $R=\big[\boldsymbol{r}_1(\boldsymbol{A})^{\mathsf{T}},\dots,\boldsymbol{r}_m(\boldsymbol{A})^{\mathsf{T}}\big]$ v $\mathbb{R}^n$ generovaný stĺpcami $\boldsymbol{s}_i(\boldsymbol{A}^{\mathsf{T}})=\boldsymbol{r}_i(\boldsymbol{A})^{\mathsf{T}}$ matice $\boldsymbol{A}^{\mathsf{T}}$, a podľa dôsledku 14.1.4 (a) platí $\mathcal{R}(\boldsymbol{A})^{\perp}=R^{\perp\perp}=R$, vidíme, že $\boldsymbol{x}$ musí byť tvaru

$$\boldsymbol{x}=z_1\boldsymbol{r}_1(\boldsymbol{A})^{\mathsf{T}}+\ldots+z_m\boldsymbol{r}_m(\boldsymbol{A})^{\mathsf{T}}=\boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{z}$$

pre nejaké $\boldsymbol{z}\in\mathbb{R}^m$. Takýto vektor $\boldsymbol{x}$ je riešením pôvodnej sústavy práve vtedy, keď $\boldsymbol{A}\cdot\boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{z}=\boldsymbol{b}$.

Minimálne riešenie $\boldsymbol{x}$ podurčenej sústavy $\boldsymbol{A}\cdot\boldsymbol{x}=\boldsymbol{b}$ teda hľadáme tak, že najprv nájdeme nejaké (ľubovoľné) riešenie $\boldsymbol{z}$ sústavy $\boldsymbol{A}\cdot\boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{z}=\boldsymbol{b}$ a položíme $\boldsymbol{x}=\boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{z}$. V typickom prípade má pôvodná sústava menej rovníc než neznámych, čiže $m<n$ a $\boldsymbol{A}$ je obdĺžniková matica, „širšia ako vyššia“. Potom je veľmi pravdepodobné, že štvorcová matica $\boldsymbol{A}\cdot\boldsymbol{A}^{\mathsf{T}}$ rádu m (ako Gramova matica „malého“ počtu riadkových vektorov v (riadkovom) euklidovskom priestore „veľkej“ dimenzie) je regulárna, teda k nej existuje inverzná matica $\big(\boldsymbol{A}\cdot\boldsymbol{A}^{\mathsf{T}}\big)^{-1}$. V tom prípade je tak

$$\boldsymbol{z}=\big(\boldsymbol{A}\cdot\boldsymbol{A}^{\mathsf{T}}\big)^{-1}\cdot\boldsymbol{b},$$

ako aj minimálne riešenie pôvodnej sústavy

$$\boldsymbol{x}=\boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{z}=\boldsymbol{A}^{\mathsf{T}}\cdot\big(\boldsymbol{A}\cdot\boldsymbol{A}^{\mathsf{T}}\big)^{-1}\cdot\boldsymbol{b}$$

určené jednoznačne. Samozrejme, ak $m=n$ a už samotná $\boldsymbol{A}$ je regulárna, dostávame

$$\boldsymbol{A}^{\mathsf{T}}\cdot\big(\boldsymbol{A}\cdot\boldsymbol{A}^{\mathsf{T}}\big)^{-1}=\boldsymbol{A}^{-1}$$

a $\boldsymbol{x}=\boldsymbol{A}^{-1}\cdot\boldsymbol{b}$ je priamo jediným riešením pôvodnej sústavy. Oba prístupy možno kombinovať. Ak totiž sústava $\boldsymbol{A}\cdot\boldsymbol{x}=\boldsymbol{b}$ nemá riešenie, avšak má viac pseudoriešení, jej *minimálne pseudoriešenie* dostaneme v tvare

$$\boldsymbol{x}=\boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{A}\cdot\boldsymbol{z},$$

kde $\boldsymbol{z}$ je riešenie sústavy

$$\big(\boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{A}\big)^2\cdot\boldsymbol{z}=\boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{b}.$$

Detaily si premyslite samostatne. Na záver ešte poznamenajme, že pomocou podmienok tvrdení 14.5.1 a 14.5.3 možno definovať pojmy pseudoriešenia a minimálneho riešenia sústav lineárnych rovníc nad ľubovoľným poľom. Operácia transponovania matice totiž nie je nijako viazaná na štruktúru euklidovského priestoru. Vo všeobecnom prípade však uvedené pojmy strácajú svoj názorný geometrický význam. Neskôr, v paragrafe 25.2, uvidíme, že tento význam možno zachovať v prípade poľa komplexných čísel, ale toto zovšeobecnenie už bude trochu menej priamočiare, než sme naznačili.

14.6 Geometria pravdepodobnosti

V tomto paragrafe si predvedieme, ako možno základné pojmy teórie pravdepodobnosti formulovať v jazyku skalárneho súčinu a euklidovských priestorov. Kvôli jednoduchosti sa obmedzíme len na konečné pravdepodobnostné priestory. Bude to zároveň krátka a jednoduchá ukážka toho, čo sa zvykne nazývať „vnútornou jednotou matematiky“, prejavujúcou sa mnohorakými súvislosťami a hlbokou spätosťou jej jednotlivých disciplín. V tomto konkrétnom prípade dodáva geometrická interpretácia teórii pravdepodobnosti nový názorný rozmer.

Pravdepodobnostný priestor bude pre nás konečná množina X, vybavená funkciou $p : X \to \mathbb{R}$ takou, že

$$(\forall\, x \in X)(p(x) \geq 0) \qquad \text{a} \qquad \sum_{x \in X} p(x) = 1,$$

nazývanou *rozdelením pravdepodobnosti* alebo len stručne *pravdepodobnosťou* na X.[3] Typickým príkladom je *rovnomerné pravdepodobnostné rozdelenie* $p(x) = 1/n$ pre každé $x \in X$, kde $n = \#X$. Podmnožiny množiny X nazývame *javmi* a jej prvky (ktoré stotožňujeme s jednoprvkovými podmnožinami) *elementárnymi javmi. Pravdepodobnosťou javu* $A \subseteq X$ rozumieme číslo

$$P(A) = \sum_{x \in A} p(x).$$

Javy $A \subseteq X$ také, že $P(A) = 0$ nazývame *nemožnými.* Keďže nemožné javy sú už zo svojej definície tie, ktoré nemôžu nastať, lebo majú nulovú pravdepodobnosť, bez ujmy na všeobecnosti môžeme prijať zjednodušujúci predpoklad, že medzi prvkami množiny X sa nevyskytujú nemožné javy, t. j. platí $p(x) > 0$ pre každé $x \in X$ (v opačnom prípade môžeme nemožné

[3] V takýchto prípadoch zvykneme tiež hovoríme o *apriórnom* rozdelení pravdepodobnosti.

elementárne javy z množiny X jednoducho vylúčiť).[4] Inak povedané, $A = \emptyset$ je jediný nemožný jav $A \subseteq X$.

Náhodnou premennou na pravdepodobnostnom priestore X rozumieme ľubovoľnú funkciu $f : X \to \mathbb{R}$. Množina $\mathbb{R}^X$ všetkých náhodných premenných tvorí vektorový priestor nad poľom $\mathbb{R}$ (pozri odstavec 1.6.5). Zrejme $\dim \mathbb{R}^X = \# X$. Vďaka nášmu predpokladu, podľa ktorého $p(x) > 0$ pre každé $x \in X$, je rovnosťou

$$\langle f, g \rangle = \sum_{x \in X} f(x)g(x)p(x),$$

kde f, g sú náhodné premenné, definovaný skalárny súčin na $\mathbb{R}^X$. Priestor $\mathbb{R}^X$ všetkých náhodných premenných je tak euklidovským priestorom. Množina

$$C = \big\{ f \in \mathbb{R}^X ;\ (\forall\, x, y \in X)(f(x) = f(y)) \big\}$$

všetkých konštantných náhodných premenných tvorí jednorozmerný lineárny podpriestor priestoru $\mathbb{R}^X$. Keďže zobrazenie, ktoré reálnemu číslu a priradí konštantnú náhodnú premennú $f \in \mathbb{R}^X$ takú, že $f(x) = a$ pre každé $x \in X$, je lineárny izomorfizmus $\mathbb{R} \cong C$, môžeme si dovoliť stotožniť toto číslo s príslušnou konštantnou náhodnou premennou.

Nakoľko konštantná náhodná premenná 1 tvorí zrejme ortonormálnu bázu podpriestoru C, lineárny operátor pr_C ortogonálnej projekcie na podpriestor C má tvar

$$\mathsf{E}(f) = \mathrm{pr}_C(f) = \langle f, 1 \rangle = \sum_{x \in X} f(x)p(x)$$

pre $f \in \mathbb{R}^X$. Výraz $\mathsf{E}(f)$ nazývame *strednou* alebo aj *očakávanou hodnotou* náhodnej premennej f (označenie operátora $\mathsf{E} : \mathbb{R}^X \to \mathbb{R}^X$ pochádza z anglického slova *expectation*).

Ortokoplementom podpriestoru C je podľa dôsledku 14.1.4 (d) lineárny podpriestor

$$N = C^{\perp} = \operatorname{Ker} \mathsf{E} = \{ f \in \mathbb{R}^X ;\ \mathsf{E}(f) = 0 \},$$

tvorený všetkými náhodnými premennými s nulovou strednou hodnotou. Potom náhodná premenná $f - \mathsf{E}(f)$ je kolmým priemetom $f \in \mathbb{R}^X$ do nadroviny N.

Disperziou alebo tiež *rozptylom* náhodnej premennej $f \in \mathbb{R}^X$ nazývame výraz $\mathsf{D}(f) = \|f - \mathsf{E}(f)\|^2$. Zrejme platí

$$\sqrt{\mathsf{D}(f)} = \|f - \mathsf{E}(f)\| = \operatorname{dist}(f, C).$$

[4]Treba poznamenať, že v nekonečnom pravdepodobnostnom priestore X by už takýto predpoklad spôsobil nielen značnú ujmu na všeobecnosti ale i ďalšie vážne ťažkosti.

To znamená, že toto číslo, nazývané *stredná kvadratická odchýlka* prípadne *smerodajná odchýlka* náhodnej premennej f, udáva akúsi strednú mieru nekonštantnosti f, presnejšie, vzdialenosť f od jej očakávanej hodnoty.

Nech $f, g \in \mathbb{R}^X$ sú náhodné premenné s nenulovými disperziami. Ich *korelačný koeficient* definujeme vzťahom

$$\mathsf{R}(f,g) = \cos\sphericalangle(f - \mathsf{E}(f),\, g - \mathsf{E}(g)) = \frac{\langle f - \mathsf{E}(f),\, g - \mathsf{E}(g)\rangle}{\|f - \mathsf{E}(f)\|\,\|g - \mathsf{E}(g)\|}.$$

To znamená, že $\arccos \mathsf{R}(f,g)$ je uhol, ktorý zvierajú kolmé priemety náhodných premenných f, g do podpriestoru N náhodných premenných s nulovou strednou hodnotou. Špeciálne pre $f, g \in N$ je

$$\mathsf{R}(f,g) = \cos\sphericalangle(f,g) = \frac{\langle f, g\rangle}{\|f\|\,\|g\|}.$$

Niekoľko doplňujúcich informácií o „geometrii pravdepodobnosti“ môže čitateľ nájsť v cvičeniach.

Cvičenia

14.1. Doplňte dôkazy tvrdení 14.1.1 a 14.1.2 a dôsledku 14.1.4.

14.2. Nech S, T sú lineárne podpriestory vektorového priestoru V so skalárnym súčinom. Potom podľa podľa 14.1.2 (h) platí $(S+T)^\perp = S^\perp \cap T^\perp$. Dokážte, že za predpokladu $S = S^{\perp\perp}$, $T = T^{\perp\perp}$, špeciálne ak S, T sú konečnorozmerné platí aj duálna rovnosť $(S \cap T)^\perp = S^\perp + T^\perp$. Je uvedený predpoklad naozaj nevyhnutný?

14.3. Nech V je vektorový priestor so skalárnym súčinom. Dokážte nasledujúce tvrdenia:

(a) Kolmý priemet vektora $\boldsymbol{x} \in V$ do lineárneho podpriestoru $S \subseteq V$ existuje práve vtedy, keď $\boldsymbol{x} \in S + S^\perp$.

(b) Kolmý priemet do lineárneho podpriestoru $S \subseteq V$ existuje pre každé $\boldsymbol{x} \in V$ práve vtedy, keď $V = S + S^\perp$.

(c) Pre ľubovoľný lineárny podpriestor $S \subseteq V$ platí $S + S^\perp = V \;\Leftrightarrow\; S = S^{\perp\perp}$.

(d) Ak V je euklidovský priestor, tak pre ľubovoľnú podmnožinu $X \subseteq V$ platí $[X] = X^{\perp\perp}$.

14.4. (a) Dokážte, že množina ℓ_2 všetkých postupností $\boldsymbol{x} = (x_n)_{n=0}^\infty$ takých, že rad $\sum_{n=0}^\infty x_n^2$ konverguje, tvorí lineárny podpriestor vektorového priestoru $\mathbb{R}^\mathbb{N}$ všetkých postupností $\mathbb{N} \to \mathbb{R}$.

(b) Dokážte, že pre $\boldsymbol{x}, \boldsymbol{y} \in \ell_2$ rad $\sum_{n=0}^\infty x_n y_n$ konverguje. Odvoďte z toho, že formulou $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \sum_{n=0}^\infty x_n y_n$ je definovaný skalárny súčin na vektorovom priestore $V = \ell_2$.

(c) Označme $S = [\boldsymbol{e}_0, \boldsymbol{e}_1, ...]$ lineárny podpriestor priestoru ℓ_2, generovaný všetkými vektormi $\boldsymbol{e}_n$, kde $\boldsymbol{e}_n$ je postupnosť pozostávajúca zo samých núl s výnimkou n-tého člena $e_{nn} = 1$. Nájdite príklad postupnosti $\boldsymbol{a} \in \ell_2 \setminus S$. Potom $S \neq \ell_2$. Dokážte, že

$S^\perp = \{\mathbf{0}\}$.[5]

(d) Ako je to s rovnosťami $S + S^\perp = V$, $S = S^{\perp\perp}$ a $(S \cap T)^\perp = S^\perp + T^\perp$ pre $T = S^{\perp\perp}$?

14.5. Riešte rovnaké otázky ako v bodoch (c) a (d) predchádzajúceho cvičenia pre lineárny podpriestor $S = \mathbb{R}[x] = [1, x, x^2, \dots]$ všetkých polynomických funkcií vo vektorovom priestore $V = \mathcal{C}\langle 0, 1\rangle$ všetkých spojitých funkcií $\langle 0, 1\rangle \to \mathbb{R}$ so skalárnym súčinom $\langle f, g\rangle = \int_0^1 f(x)g(x)\,\mathrm{d}x$.

14.6. Riešte rovnaké otázky ako v bodoch (c) a (d) cvičenia 14.4 pre lineárny podpriestor $S = [1, \cos x, \sin x, \cos 2x, \sin 2x, \dots]$ vektorového priestoru V všetkých periodických spojitých funkcií $\mathbb{R} \to \mathbb{R}$ s periódou 2π. Skalárny súčin na ňom má tvar $\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)\,\mathrm{d}x$ (pozri cvičenie 13.8).

14.7. V každom z nasledujúcich prípadov nájdite kolmý priemet vektora $\boldsymbol{x} \in V$ do lineárneho podpriestoru $S \subseteq V$ a určte vzdialenosť $\operatorname{dist}(\boldsymbol{x}, S)$ a odchýlku $\sphericalangle(\boldsymbol{x}, S)$ (euklidovské priestory $\mathbb{R}^n$ sú vybavené štandardným skalárnym súčinom):

(a) $V = \mathbb{R}^3$, $S = [(2, -1, 0)]$, $\boldsymbol{x} = (1, 1, 1)$;

(b) $V = \mathbb{R}^3$, $S = [(1, -3, 5)]$, $\boldsymbol{x} = (1, 1, 1)$;

(c) $V = \mathbb{R}^3$, $S = [(2, -1, 0), (1, -3, 5)]$, $\boldsymbol{x} = (1, 1, 1)$;

(d) $V = \mathbb{R}^4$, $S = [(1, -2, 0, 3), (1, -3, 5, -2)]$, $\boldsymbol{x} = (1, 1, 1, 1)$;

(e) $V = \mathbb{R}^4$, $S = \{\boldsymbol{x} \in V;\ x_1 - x_2 - x_4 = 0\}$, $\boldsymbol{x} = (1, 1, 1, 1)$.

Porovnajte výsledky úloh (a), (b), (c), resp. (d), (e) a vysvetlite rozdiely prípadne zhodu.

14.8. Nech V je ľubovoľný vektorový priestor. Lineárny operátor $p : V \to V$ sa nazýva *projektor* alebo *projekcia*, ak $p \circ p = p$. Za predpokladu, že $p : V \to V$ je projekcia, dokážte nasledujúce tvrdenia:

(a) lineárny operátor $\mathrm{id}_V - p$ je tiež projekcia;

(b) $\operatorname{Im} p = \{\boldsymbol{x} \in V;\ \boldsymbol{x} = p(\boldsymbol{x})\} = \operatorname{Ker}(\mathrm{id}_V - p)$;

(c) $V = \operatorname{Im} p \oplus \operatorname{Ker} p$.

14.9. Nech V je vektorový priestor so skalárnym súčinom a $p : V \to V$ je projekcia. Hovoríme, že p je *ortogonálna projekcia*, ak pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí $p(\boldsymbol{y}) \perp \boldsymbol{x} - p(\boldsymbol{x})$. Dokážte nasledujúce tvrdenia:

(a) Ak $S \subseteq V$ je lineárny podpriestor taký, že $V = S \oplus S^\perp$, tak lineárny operátor $\mathrm{pr}_S : V \to V$ je ortogonálna projekcia.

(b) Ak $p : V \to V$ je ortogonálna projekcia, tak $p = \mathrm{pr}_S$ pre $S = \operatorname{Im} p$.

14.10. Nech $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n, \dots$ je ľubovoľná (konečná či nekonečná) postupnosť vektorov vo vektorovom priestore V so skalárnym súčinom. Postupnosť $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n, \dots$ definujeme rekurziou $\boldsymbol{v}_1 = \boldsymbol{u}_1$, $\boldsymbol{v}_{n+1} = \boldsymbol{u}_{n+1} - \mathrm{pr}_{S_n}(\boldsymbol{u}_{n+1})$, kde $S = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n] = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_n]$. Dokážte, že vektory $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n, \dots$ sú ortogonálne, a porovnajte túto konštrukciu s Gramovým-Schmidtovým ortogonalizačným procesom.

14.11. V nasledujúcich úlohách najdite nejakú priečku afinných podpriestorov M, N v euk-

[5] V tomto aj v nasledujúcich dvoch cvičeniach, ak rovnosť $S^\perp = \{\mathbf{0}\}$ neviete dokázať (vyžaduje to isté netriviálne poznatky z matematickej analýzy), berte ju ako východiskový postulát pre ďalšie časti.

lidovskom priestore $V = \mathbb{R}^n$ so štandardným skalárnym súčinom a vypočítajte vzdialenosť $\operatorname{dist}(M, N)$ a odchýlku $\sphericalangle(M, N)$:

(a) $V = \mathbb{R}^3$, $M = (1,1,1) + [(1,2,3)]$, $N = [(1,0,1)]$;

(b) $V = \mathbb{R}^3$, $M = (1,1,1) + [(1,2,3)]$, $N = [(0,1,1)]$;

(c) $V = \mathbb{R}^3$, $M = (1,1,1) + [(1,2,3)]$, $N = [(1,0,1),(0,1,1)]$;

(d) $V = \mathbb{R}^4$, $M = (1,1,1,1)+[(1,2,3,4),(1,-2,3,-4)]$, $N = [(0,1,0,1),(1,0,3,0)]$;

(e) $V = \mathbb{R}^4$, $M = (1,1,1,1)+[(1,2,3,4)]$, $N = [(0,1,0,1),(1,0,3,0),(1,-2,3,-4)]$.

Porovnajte výsledky úloh (a), (b), (c), resp. (d), (e) a vysvetlite rozdiely prípadne zhodu.

14.12. Nech M je nadrovina v euklidovskom priestore V daná rovnicou $\langle \boldsymbol{x}, \boldsymbol{a}\rangle = b$, kde $\boldsymbol{a} \in V$, $b \in \mathbb{R}$, a $\boldsymbol{p} \in V$ je bod. Odvoďte vzorec pre ich vzdialenosť $\operatorname{dist}(\boldsymbol{p}, M) = \|\langle \boldsymbol{p}, \boldsymbol{a}\rangle - b\rangle\| / \|\boldsymbol{a}\|$.

14.13. V trojrozmernom euklidovskom priestore je daný pravidelný štvorsten s hranami dĺžky 1.

(a) Vypočítajte vzdialenosť jeho vrchola od protiľahlej podstavy.

(b) Vyočítajte uhol, ktorý zvierajú (roviny, v ktorých ležia) jeho dve steny.

14.14. V n-rozmernom euklidovskom priestore je daná n-rozmerná (nad)kocka s hranami dĺžky 1. Vypočítajte dĺžku jej (nad)telesovej uhlopriečky a uhol, ktorý zviera so svojou $(n-1)$-rozmernou stenou (t. j. s nadrovinou, v ktorej táto stena leží).

14.15. Nech $S, T \subseteq V$ sú vlastné netriviálne lineárne podpriestory v euklidovskom priestore V. Dokážte nasledujúce rovnosti:
(a) $\sphericalangle(S, S^\perp) = \pi/2$; (b) $\sphericalangle(S, T^\perp) = \sphericalangle(S^\perp, T)$; (c) $\sphericalangle(S^\perp, T^\perp) = \sphericalangle(S, T)$.

14.16. V euklidovskom priestore $\mathbb{R}^5$ (so štandardným skalárnym súčinom) nájdite odchýlku lineárnych podpriestorov $S = [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]$, $T = [4\boldsymbol{e}_3 + 3\boldsymbol{e}_4, 5\boldsymbol{e}_4 + 4\boldsymbol{e}_5, 3\boldsymbol{e}_5 + 5\boldsymbol{e}_3]$.

14.17. Nájdite explicitné vyjadrenie pre n-rozmerný objem n-rozmernej gule s polomerom R pre $n = 5$, 6.

14.18. Podobne ako v prípade objemu preskúmajte $(n-1)$-rozmerný „plošný obsah“ $(n-1)$-rozmerného povrchu n-rozmernej gule s polomerom R výpočtom integrálu
$\int_{S^{(n-1)}(R)} \mathrm{d}x_1 \dots \mathrm{d}x_n$
cez množinu $S^{(n-1)}(R) = \{\boldsymbol{x} \in \mathbb{R}^n;\ \|\boldsymbol{x}\| = R\}$ s použitím sférických súradníc.

14.19. Nech $\alpha > -1/2$ je reálne číslo. S využitím polárnych resp. sférických súradníc vypočítajte určité integrály:

(a) $\int_{B^{(2)}(1)} (x^2+y^2)^\alpha \,\mathrm{d}x\,\mathrm{d}y$,
cez jednotkový kruh $B^{(2)}(1) = \{(x,y) \in \mathbb{R}^2;\ x^2+y^2 \le 1\}$,
resp. jednotkovú kružnicu $A = S^{(1)}(1) = \{(x,y) \in \mathbb{R}^2;\ x^2+y^2 = 1\}$;

(b) $\int_{B^{(3)}(1)} (x^2+y^2+z^2)^\alpha \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$,
cez jednotkovú guľu $B^{(3)}(1) = \{(x,y,z) \in \mathbb{R}^3;\ x^2+y^2+z^2 \le 1\}$,
resp. jednotkovú sféru $A = S^{(2)}(1) = \{(x,y,z) \in \mathbb{R}^3;\ x^2+y^2+z^2 = 1\}$.
Pokúste o zovšeobecnenie v $\mathbb{R}^n$. V čom je problém pre $\alpha \le -1/2$?

14.20. Vypočítajte hodnotu *Laplaceovho integrálu* [6] $\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x = \sqrt{\pi}$. (*Návod*: Overte rovnosť $\left(\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x\right)^2 = \iint_{\mathbb{R}^2} \mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y$ a dvojný integrál vpravo vypočítajte substitúciou pomocou polárnych súradníc.)

14.21. Nájdite pseudoriešenie sústavy m lineárnych rovníc s jednou neznámou $x = b_1$, $x = b_2, \ldots, x = b_m$. Pekný výsledok, čo vy na to?

14.22. Nájdite minimálne riešenie lineárnej rovnice o n neznámych $a_1x_1 + \ldots + a_nx_n = b$ za predpokladu, že $(a_1, \ldots, a_n) \neq \mathbf{0}$.

14.23. Za predpokladu, že $(a_1, a_2) \neq \mathbf{0}$, nájdite všetky pseudoriešenia sústavy dvoch rovníc o dvoch neznámych $a_1x_1 + a_2x_2 = b_1$, $a_1x_1 + a_2x_2 = b_2$ a určte medzi nimi to minimálne. Pokúste sa o zovšeobecnenie na m rovníc o n neznámych.

14.24. Metódou lineárnej regresie postupne aproximujte funkciu $f(x)$, z ktorej poznáte hodnoty $f(-3) = 4$, $f(-2) = 1$, $f(-1) = 0$, $f(0) = -2$, $f(1) = 0$, $f(2) = -3$, $f(3) = -5$, polynómami $f_1(x), \ldots, f_5(x)$ prvého stupňa až piateho stupňa. Načrtnite ich grafy na intervale $\langle -5, 5\rangle$ a porovnajte ich s grafom interpolačného polynómu $f_6(x)$ (pozri cvičenie 10.17).

14.25. Nech $\boldsymbol{A} \in \mathbb{R}^{m\times n}$.

(a) Predpokladajme, že $\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A}$ je regulárna matica, a položme $\boldsymbol{A}^\neg = \left(\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{A}\right)^{-1} \cdot \boldsymbol{A}^\mathsf{T}$.

(b) Predokladajme, že $\boldsymbol{A} \cdot \boldsymbol{A}^\mathsf{T}$ je regulárna matica, a položme $\boldsymbol{A}^\neg = \boldsymbol{A}^\mathsf{T} \cdot \left(\boldsymbol{A} \cdot \boldsymbol{A}^\mathsf{T}\right)^{-1}$.

(c) Dokážte, že v oboch prípadoch (a), (b) sú splnené rovnosti $\boldsymbol{A} \cdot \boldsymbol{A}^\neg \cdot \boldsymbol{A} = \boldsymbol{A}$, $\boldsymbol{A}^\neg \cdot \boldsymbol{A} \cdot \boldsymbol{A}^\neg = \boldsymbol{A}^\neg$ a obe matice $\boldsymbol{A} \cdot \boldsymbol{A}^\neg$, $\boldsymbol{A}^\neg \cdot \boldsymbol{A}$ sú symetrické.

(d) Dokážte, že pre každú maticu $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ existuje jednoznačne určená matica $\boldsymbol{A}^\neg \in \mathbb{R}^{n\times m}$, ktorá spĺňa podmienky časti (c). Matica $\boldsymbol{A}^\neg$ sa nazýva *Mooreovou-Penroseovou pseudoinverziou* matice $\boldsymbol{A}$ (Mooreovou-Penroseovou pseudoinverziou komplexných matíc sa budeme zaoberať v paragrafe 25.2).

(e) Sformulujte výsledky o pseudoriešeniach resp. minimálnych riešeniach sústav lineárnych rovníc pomocou Mooreovej-Penroseovej pseudoinverzie matice sústavy.

V nasledujúcich cvičeniach X je konečný pravdepodobnostný priestor s pravdepodobnosťou p, pričom $p(x) > 0$ pre každé $x \in X$.

14.26. Pre náhodné premenné $f, g \in \mathbb{R}^X$ dokážte rovnosti
(a) $\langle f, g\rangle = \mathsf{E}(fg)$, (b) $\|f\|^2 = \mathsf{E}(f^2)$, (c) $\mathsf{D}(f) = \mathsf{E}\big((f - \mathsf{E}(f))^2\big)$.

14.27. *Kovarianciou* náhodných premenných $f, g \in \mathbb{R}^X$ nazývame výraz $\mathrm{Cov}(f, g) = \langle f - \mathsf{E}(f), g - \mathsf{E}(g)\rangle$. Dokážte rovnosť $\mathrm{Cov}(f, g) = \mathsf{E}(fg) - \mathsf{E}(f)\mathsf{E}(g)$.

14.28. Hovoríme, že náhodné premenné $f, g \in \mathbb{R}^X$ sú *nekorelované*, ak $\mathrm{Cov}(f, g) = 0$. Dokážte nasledujúce tvrdenia:

(a) f, g sú nekorelované práve vtedy, keď $f - \mathsf{E}(f) \perp g - \mathsf{E}(g)$.

(b) Ak f, g majú nenulové disperzie, tak f, g sú nekorelované práve vtedy, keď $\mathsf{R}(f, g) = 0$.

(c) $\mathrm{Cov}(f, g)^2 \leq \mathsf{D}(f)\mathsf{D}(g)$, pričom rovnosť nastane práve vtedy, keď náhodné premenné $f - \mathsf{E}(f), g - \mathsf{E}(g)$ sú lineárne závislé. To je ďalej ekvivalentné s podmienkou $\mathsf{R}(f, g) = \pm 1$.

[6] Tento integrál sa zvykne tiež nazývať *Gaussovým* alebo *Poissonovým integrálom*.

14.29. Uvažujme pravdepodobnostný priestor $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ s rovnomerným pravdepodobnostným rozdelením $p(x) = 1/n$ pre každé $x \in \mathbb{Z}_n$. Pre prirodzené číslo a, $1 \leq a \leq n$, označme $f_a : \mathbb{Z}_n \to \mathbb{R}$ náhodnú premennú takú, že $f_a(x) = 1$, ak a delí x, a $f_a(x) = 0$ inak.

(a) Vypočítajte strednú hodnotu $\mathsf{E}(f_a)$ a disperziu $\mathsf{D}(f_a)$.

(b) Pre $a, b \in \{1, \ldots, n\}$ vypočítajte kovarianciu $\mathrm{Cov}(f_a, f_b)$ a (za predpokladu, že $a, b \neq 1$) aj koeficient korelácie $\mathsf{R}(f_a, f_b)$.

(c) Dokážte, že náhodné premenné f_a, f_b sú nekorelované práve vtedy, keď čísla a, b sú nesúdeliteľné a obe delia n alebo aspoň jedno z nich je 1.

15. Objem, orientácia a vektorový súčin

Naše zavedenie determinantov v kapitole 10 sme motivovali úvahami o n-rozmernom objeme a orientovanom objeme v priestore $\mathbb{R}^n$, a definíciu determinantu sme potom dostali prenesením formálnych vlastností orientovaného objemu do (stĺpcových) vektorových priestorov K^n nad ľubovoľným poľom K. Prísne vzaté však vo vektorových priestoroch, v ktorých nevieme povedať, ani čo je to dĺžka vektora, nedáva pojem objemu ani orientovaného objemu žiadny zmysel. Na druhej strane, vo vektorovom priestore V so skalárnym súčinom možno pre každé kladné celé číslo $k \leq \dim V$ zmysluplne definovať *k-rozmerný objem* ako aj *orientovaný k-rozmerný objem* k-rozmerného rovnobežnostena

$$\{a_1\boldsymbol{u}_1 + \cdots + a_k\boldsymbol{u}_k;\ a_1, \ldots, a_k \in \langle 0, 1\rangle\}$$

vytvoreného vektormi $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k \in V$ (o skutočne k-rozmerný rovnobežnosten ide samozrejme len vtedy, keď vektory $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ sú lineárne nezávislé, inak je to útvar nižšej dimenzie). Práve definícia takýchto objemov a vyjasnenie ich súvisu s determinantmi ako i s tzv. *vektorovým súčinom* bude náplňou tejto kapitoly. Objemami zložitejších útvarov sa tu zaoberať nebudeme – ich štúdium je predmetom *teórie miery a integrálu*, ktorá využíva viaceré podstatne hlbšie myšlienky a náročnejšie metódy, než sú tie, s ktorými sme sa doposiaľ zoznámili.

15.1 Objem

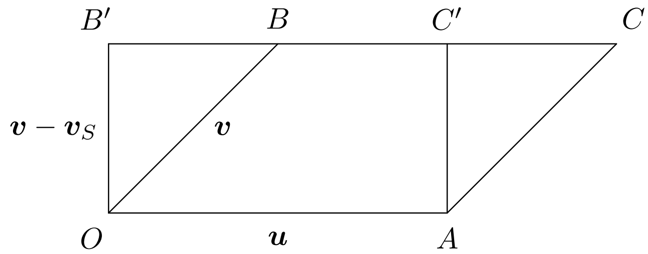
Pripomeňme si, ako počítame plošný obsah (t. j. dvojrozmerný objem) $\mathrm{vol}_2(\boldsymbol{u}, \boldsymbol{v})$ rovnobežníka vytvoreného vektormi $\boldsymbol{u}$, $\boldsymbol{v}$ v $\mathbb{R}^2$ alebo v $\mathbb{R}^3$ (označenie vol je z anglického *volume*). Jeden z vektorov, dajme tomu $\boldsymbol{v}$, rozložíme na súčet dvoch zložiek

$$\boldsymbol{v} = \boldsymbol{v}_S + (\boldsymbol{v} - \boldsymbol{v}_S),$$

kde $\boldsymbol{v}_S$ je kolmý priemet vektora $\boldsymbol{v}$ do podpriestoru $S = [\boldsymbol{u}]$ a zložka $\boldsymbol{v} - \boldsymbol{v}_S$ je kolmá na $\boldsymbol{u}$. Príslušný obsah potom dostaneme ako súčin dĺžok

$$\mathrm{vol}_2(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u}\|\,\|\boldsymbol{v} - \boldsymbol{v}_S\|\,.$$

Zo zhodnosti trojuholníkov OBB' a ACC' totiž vyplýva, že rovnobežník $OACB$ a obdĺžnik $OAC'B'$ majú rovnaký obsah. (Na obrázku vynechávame šípky vektorov.)



Obr. 15.1. K definfcii plošného obsahu

Podobne, trojrozmerný objem rovnobežnostena vytvoreného vektormi $\boldsymbol{u}, \boldsymbol{v}$, $\boldsymbol{w} \in \mathbb{R}^3$ dostaneme ako súčin

$$\mathrm{vol}_3(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = \mathrm{vol}_2(\boldsymbol{u}, \boldsymbol{v})\, \|\boldsymbol{w} - \boldsymbol{w}_T\|$$

plošného obsahu $\mathrm{vol}_2(\boldsymbol{u}, \boldsymbol{v})$ a dĺžky $\|\boldsymbol{w} - \boldsymbol{w}_T\|$ zložky vektora $\boldsymbol{w}$ kolmej na podpriestor $T = [\boldsymbol{u}, \boldsymbol{v}]$.

Podľa tejto schémy budeme pokračovať aj do vyšších dimenzií. Inak povedané, *k-rozmerný objem* rovnobežnostena vytvoreného k vektormi z vektorového priestoru so skalárnym súčinom V budeme definovať ako funkciu $\mathrm{vol}_k : V^k \to \mathbb{R}$ rekurziou cez k. Pre $k = 1$, $\boldsymbol{u} \in V$ kladieme

$$\mathrm{vol}_1(\boldsymbol{u}) = \|\boldsymbol{u}\|,$$

t. j. jednorozmerný objem je jednoducho dĺžka vektora. Ak $k > 1$ a $(k-1)$-rozmerný objem vol_{k-1} už máme definovaný, tak pre $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{k-1}, \boldsymbol{u}_k \in V$ položíme

$$\mathrm{vol}_k(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{k-1}, \boldsymbol{u}_k) = \mathrm{vol}_{k-1}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{k-1})\, \|\boldsymbol{u}_k - \mathrm{pr}_S(\boldsymbol{u}_k)\|\,,$$

kde $\mathrm{pr}_S(\boldsymbol{u}_k)$ je kolmý priemet vektora $\boldsymbol{u}_k$ do podpriestoru $S = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_{k-1}]$.

Dôkaz nasledujúceho jednoduchého tvrdenia prenechávame ako cvičenie čitateľovi.

15.1.1. Tvrdenie. *Nech V je vektorový priestor so skalárnym súčinom a $k \geq 1$. Potom pre ľubovoľné vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k \in V$ platí:*

(a) $\mathrm{vol}_k(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k) \geq 0$, *pričom rovnosť nastane práve vtedy, keď $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ sú lineárne závislé;*

(b) $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ *sú ortogonálne práve vtedy, keď* $\mathrm{vol}_k(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k) = \|\boldsymbol{u}_1\| \dots \|\boldsymbol{u}_k\|$;

(c) ak $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ sú lineárne nezávislé, tak

$$\mathrm{vol}_k(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k) = \mathrm{vol}_k(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k) = \|\boldsymbol{v}_1\| \dots \|\boldsymbol{v}_k\|\,,$$

kde vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ sú získané z vektorov $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ Gramovým-Schmidtovým ortogonalizačným procesom podľa vety 13.4.5.

Časť (c) nám dáva priamy návod na výpočet objemu $\mathrm{vol}_k(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$. Stačí utvoriť Gramovu maticu $\boldsymbol{G}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ a upraviť ju úpravami typu (1^+) na diagonálny tvar. Ak sa nám to podarí, tak na diagonále máme druhé mocniny noriem vektorov $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ – stačí teda zobrať druhú odmocninu ich súčinu. Ak sa nám to nepodarí, môže to byť len z toho dôvodu, že vektory $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ nie sú lineárne nezávislé – v takom prípade je objem ich rovnobežnostena 0. Objem $\mathrm{vol}_k(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ však možno vyjadriť aj bez Gramovho-Schmidtovho ortogonalizačného procesu len pomocou Gramovho determinantu.

15.1.2. Veta. *Nech V je vektorový priestor so skalárnym súčinom a $k \geq 1$. Potom pre ľubovoľné vektory $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k \in V$ platí*

$$\mathrm{vol}_k(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) = |\boldsymbol{G}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)|^{1/2}.$$

Dôkaz. Ak vektory $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ sú lineárne závislé, tak potrebný záver vyplýva z podmienky 15.1.1 (a) a dôsledku 13.2.2. Ak sú lineárne nezávislé, tak podľa predošlej úvahy a vety 12.2.3 spolu s poznámkou, ktorá ju predchádza, existuje horná trojuholníková matica $\boldsymbol{P} \in \mathbb{R}^{k\times k}$ s jednotkami na diagonále, taká, že

$$\boldsymbol{P}^{\mathsf{T}} \cdot \boldsymbol{G}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) \cdot \boldsymbol{P} = \mathrm{diag}\big(\|\boldsymbol{v}_1\|^2, \ldots, \|\boldsymbol{v}_k\|^2\big),$$

kde vektory $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ sú výsledkom Gramovej-Schmidtovej ortogonalizácie vektorov $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$. Keďže $|\boldsymbol{P}| = |\boldsymbol{P}^{\mathsf{T}}| = 1$, priamym výpočtom s použitím 15.1.1 (c). dostávame

$$\begin{aligned}\mathrm{vol}_k(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)^2 &= \|\boldsymbol{v}_1\|^2 \ldots \|\boldsymbol{v}_k\|^2 = \big|\mathrm{diag}\big(\|\boldsymbol{v}_1\|^2, \ldots, \|\boldsymbol{v}_k\|^2\big)\big| \\ &= \big|\boldsymbol{P}^{\mathsf{T}} \cdot \boldsymbol{G}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) \cdot \boldsymbol{P}\big| = |\boldsymbol{G}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)|.\end{aligned}$$

Zámena poradia i-teho a j-teho vektora sa na Gramovej matici $\boldsymbol{G}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ prejaví zámenou i-teho a j-teho stĺpca a zároveň i-teho a j-teho riadku – teda jej determinant, a preto ani objem $\mathrm{vol}_k(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$, sa tým nezmení. Doteraz uvažované objemy vol_k teda možno právom nazvať neorientovanými.

15.2 Orientácia

Skôr než sa začneme zaoberať orientovaným objemom v euklidovskom priestore, je potrebné najprv stručne pojednať o *orientácii* ako takej. Ukazuje sa, že tento pozoruhodný jav nezávisí na skalárnom súčine, ale možno sa s ním stretnúť v ľubovoľnom konečnorozmernom vektorovom priestore nad poľom $\mathbb{R}$.

Nech teda V je reálny vektorový priestor konečnej dimenzie $n \geq 1$. Hovoríme, že dve bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ priestoru V sú *súhlasne orientované*, ak matica prechodu $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ má kladný determinant.

Keďže $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\alpha}} = \boldsymbol{I}_n$, každá báza je súhlasne orientovaná sama so sebou, t. j. vzťah súhlasnej orientácie je *reflexívny.* Z rovnosti $\boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\alpha}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}^{-1}$ zasa vyplýva *symetria* tohto vzťahu. Konečne z rovnosti $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\gamma}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot \boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\gamma}}$ výplýva, že vzťah súhlasnej orientácie je *tranzitívny.* Podčiarknuté a zrátané, vzťah súhlasnej orientovanosti je *ekvivalenciou* na množine všetkých báz priestoru V. Keďže každá matica prechodu je regulárna, jej determinant môže byť len kladný alebo záporný. Tým sa nám množina všetkých báz priestoru V rozpadne na dve disjunktné triedy, z ktorých každá pozostáva so súhlasne orientovaných báz, kým dve bázy patriace do rôznych tried sú orientované nesúhlasne.

Orientácia konečnorozmerného reálneho vektorového priestoru V spočíva vo výbere jednej jeho bázy $\boldsymbol{\alpha}$, ktorú prehlásime za *kladne orientovanú*, rovnako ako všetky bázy orientované súhlasne s ňou – tieto tvoria jednu zo spomínaných tried. Druhá z týchto tried obsahuje bázy orientované nesúhlasne s $\boldsymbol{\alpha}$ – nazveme ich *záporne orientované* bázy. Vektorový priestor, v ktorom sme uskutočnili voľbu nejakej kladne orientovanej bázy nazveme *orientovaný.* Na konečnorozmernom reálnom vektorovom priestore tak možno zadať dve rôzne, navzájom opačné orientácie. Priestor $\mathbb{R}^n$ je prirodzené orientovať tak, aby kanonická báza $\boldsymbol{\varepsilon}^{(n)}$ mala kladnú orientáciu; tejto orientácii $\mathbb{R}^n$ hovoríme *kanonická.* V abstraktnom n-rozmernom priestore, kde nemáme žiadnu privilegovanú bázu, však aj táto pomoc pri voľbe orientácie odpadá – nanajvýš si môžeme hodiť mincou.

Orientovateľnosť konečnorozmerného reálneho vektorového priestoru V sa zakladá na tom, že každá nadrovina vo V má „dve strany“, t. j. delí V na dva polpriestory, pričom z jedného do druhého sa nemožno dostať spojitým pohybom bez toho, aby sme preťali deliacu nadrovinu. Navyše v euklidovskom priestore V nemožno útvar ležiaci v jednom z polpriestorov previesť na jeho „zrkadlový obraz“, t. j. na útvar s ním súmerne združený podľa nadroviny, nijakým zhodným zobrazením, ktoré možno realizovať spojitým pohybom vo V.

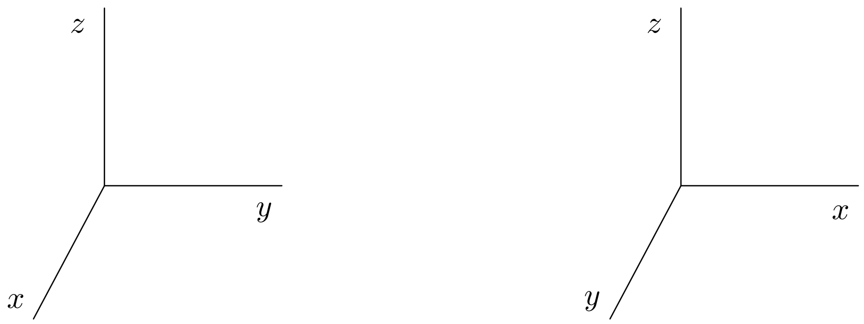
Intuitívne zodpovedá voľba orientácie priestoru V, t. j. pririeknutie jednej z dvoch možných orientácií nejakej jeho báze $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$, rozdeleniu V nadrovinou $[\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1}]$ na dva polpriestory a prehláseniu jedného z nich za kladný a druhého za záporný polpriestor. Podľa toho, do ktorého z nich smeruje vektor $\boldsymbol{u}_n$, bude i báza $\boldsymbol{\alpha}$ kladne alebo záporne orientovaná. Ak napr. prehlásime bázu $\boldsymbol{\alpha}$ za kladne orientovanú, bude báza $\boldsymbol{\alpha}' = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1}, -\boldsymbol{u}_n)$ orientovaná záporne.

Orientácia jednorozmerného priestoru pomocou nejakého vektora (kladne orientovanej bázy) $\boldsymbol{u} \neq \boldsymbol{0}$ teda zodpovedá vyznačeniu kladného smeru od počiatku $\mathbf{0}$, t. j. polpriamky $\{a\boldsymbol{u}; 0 < a \in \mathbb{R}\}$. Opačná polpriamka $\{a\boldsymbol{u}; 0 > a \in \mathbb{R}\}$ potom vyznačuje záporný smer.

V dvojrozmernom priestore sa na zadanie orientácie voľbou kladne orientovanej bázy $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ možno dívať aj ako na vyznačenie kladného zmyslu otáčania roviny okolo počiatku $\mathbf{0}$ od $\boldsymbol{u}_1$ k $\boldsymbol{u}_2$. Zmysel otáčania od $\boldsymbol{u}_2$ k $\boldsymbol{u}_1$ je potom záporný, čo sa intuitívne zhoduje s tým, že „opačná“ báza $(\boldsymbol{u}_2, \boldsymbol{u}_1)$ k báze $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ zadáva opačnú orientáciu dvojrozmerného priestoru.

Vo všeobecnom prípade sú bázy $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ a $(\boldsymbol{u}_{\sigma(1)}, \dots, \boldsymbol{u}_{\sigma(n)})$, kde $\sigma \in SS_n$, súhlasne orientované práve vtedy, keď σ je párna permutácia (dokážte).

V trojrozmernom priestore zadanie orientácie voľbou nejakej kladne orientovanej bázy zodpovedá napr. výberu pravej alebo ľavej strany. Pri tom môžeme využiť asymetriu našej fyziológie (napr. až na celkom ojedinelé výnimky majú ľudia srdce na ľavej strane, výrazná väčšina populácie sú praváci). Kladnú orientáciu možno fixovať napr. *pravidlom pravej ruky*: vystretý ukazovák, prostredník zohnutý kolmo k dlani a palec vztýčený kolmo na ich rovinu (v tomto poradí) tvoria kladne orientovanú „bázu“. Po stotožnení vektorového priestoru $\mathbb{R}^3$ s trojrozmerným fyzikálnym priestorom a smerového vektora $\boldsymbol{e}_1$ osi x s ukazovákom, smerového vektora $\boldsymbol{e}_2$ osi y s prostredníkom a smerového vektora $\boldsymbol{e}_3$ osi z s palcom pravej ruky tak dostávame tzv. *pravotočivú súradnú sústavu* v $\mathbb{R}^3$. Analogickým spôsobom dostaneme i *ľavotočivú súradnú sústavu*, ak zadáme kladnú orientáciu v $\mathbb{R}^3$ pomocou zrejmého *pravidla ľavej ruky*.



Obr. 15.2. Pravotočivá a ľavotočivá súradná sústava v $\mathbb{R}^3$

Všetky zákony klasickej fyziky sú invariantné voči zrkadlovej reflexii v priestore aj v čase, to znamená, že ich matematická formulácia sa nezmení, ak v nich vystupujúce veličiny vyjadrené v jednej báze vyjadríme vzhľadom na s ňou nesúhlasne orientovanú bázu. Výnimkou sú niektoré zákony štatistickej fyziky, predovšetkým *zákon rastu entropie*, ktorý nepripúšťa obrátenie smeru plynutia času. Otázka, či existujú fyzikálne zákony, ktoré nie sú invariantné voči zrkadlovej reflexii priestoru, bola na všeobecné prekvapenie zodpovedaná kladne v druhej polovici 50. rokov 20. storočia, keď sa experimentálne potvrdilo, že pri tzv. *slabých interakciách* sa nezachováva *parita*. Tento výsledok nám samozrejme nehovorí, ktorej z dvoch možných orientácií priestoru by sme mali dať prednosť, ale iba to, že matematický tvar istého fyzikálneho zákona sa môže zmeniť zmenou orientácie priestoru. Voľbou jednej z dvoch možných matematických foriem tohto zákona teda môžeme fixovať orientáciu „nášho“ trojrozmerného fyzikálneho priestoru aj menej antropomorfným spôsobom než len niektorým z pravidiel pravej alebo ľavej ruky. Neskôr sa zistilo, že pri slabých interakciách dochádza taktiež k (ešte podstatne slabšiemu) narušeniu invariancie voči zrkadlovej reflexii času, ktoré je ekvivalentné narušeniu tzv. *kombinovanej parity* (CP).

Podľa v súčasnosti prevládajúcich kozmologických predstáv práve narušeniu kombinovanej parity vďačíme za to, že vo veľmi ranom štádiu vývoja vesmíru vzniklo nepatrne viac *baryónov* (ťažkých častíc ako napr. protóny a neutróny) než ich zrkadlových dvojníkov, tzv. *antibaryónov*, takže postupne nedošlo k úplnej anihilácii všetkej hmoty s antihmotou (presnejšie baryónov s antibaryónmi) na žiarenie, čo teprv umožnilo vznik atómov, hviezd, planét, a napokon i nás samotných.

15.3 Orientovaný objem

V našich úvahách o orientovanom objeme sa v tomto paragrafe obmedzíme len na n-rozmerný orientovaný objem v n-rozmernom euklidovskom priestore.

Nech teda V je n-rozmerný euklidovský priestor. Vyberme v ňom pevnú *ortonormálnu* bázu $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$, ktorú prehlásime za kladne orientovanú. Pre ľubovoľnú usporiadanú n-ticu $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \in V^n$ definujeme n-rozmerný *orientovaný objem* rovnobežnostena vytvoreného vektormi $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ ako determinant

$$\mathbf{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = \begin{vmatrix} \langle \boldsymbol{x}_1, \boldsymbol{u}_1 \rangle & \dots & \langle \boldsymbol{x}_n, \boldsymbol{u}_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \boldsymbol{x}_1, \boldsymbol{u}_n \rangle & \dots & \langle \boldsymbol{x}_n, \boldsymbol{u}_n \rangle \end{vmatrix}$$

matice $\big(\langle \boldsymbol{x}_j, \boldsymbol{u}_i \rangle\big) = \big((\boldsymbol{x}_1)_{\boldsymbol{\alpha}}, \dots, (\boldsymbol{x}_n)_{\boldsymbol{\alpha}}\big) \in \mathbb{R}^{n\times n}$, ktorej stĺpce sú tvorené súradnicami vektorov $\boldsymbol{x}_j$ v báze $\boldsymbol{\alpha}$ (pozri tvrdenie 13.4.4 (a)).

Orientovaný objem sme teda definovali ako istý determinant, pričom za jeho jednotku sme zvolili orientovaný objem n-rozmernej kocky vytvorenej vektormi kladne orientovanej ortonormálnej bázy $\boldsymbol{\alpha}$. V euklidovskom priestore $\mathbb{R}^n$ so štandardným skalárnym súčinom a kanonickou orientáciou, t. j. pri voľbe $\boldsymbol{\alpha} = \boldsymbol{\varepsilon}^{(n)}$, pre ľubovoľnú maticu $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ platí

$$\det \boldsymbol{A} = \mathbf{vol}_n(\boldsymbol{s}_1(\boldsymbol{A}), \dots, \boldsymbol{s}_n(\boldsymbol{A})),$$

teda, presne v zhode s kapitolou 10, $\det \boldsymbol{A}$ je orientovaný n-rozmerný objem rovnobežnostena vytvoreného stĺpcami matice $\boldsymbol{A}$. Taktiež naopak, všetky podstatné vlastnosti orientovaného n-rozmerného objemu možno teraz odvodiť ako bezprostredné dôsledky príslušných vlastností determinantu, takže nemusíme čitateľa unavovať ich vymenúvaním. Zostáva sa presvedčiť, že

(1) orientovaný objem nezávisí od výberu konkrétnej kladne orientovanej ortonormálnej bázy $\boldsymbol{\alpha}$ (čo prenechávame ako cvičenie čitateľovi);

(2) medzi orientovaným a neorientovaným objemom naozaj je očakávaná súvislosť.

To sa udeje v nasledujúcej vete (v jej znení i v celom jej dôkaze zvislé zátvorky $|\ |$ označujú výlučne absolútnu hodnotu, kým determinant dôsledne značíme det).

15.3.1. Veta. *Nech V je orientovaný n-rozmerný euklidovský priestor. Potom pre ľubovoľné vektory $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n \in V$ platí*

$$\operatorname{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = \left|\mathbf{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)\right|;$$

pričom lineárne nezávislé vektory $\boldsymbol{x}_1, \dots .\boldsymbol{x}_n$ tvoria kladne orientovanú bázu vo V práve vtedy, keď $\mathbf{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) > 0$.

Dôkaz. Nech $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ je kladne orientovaná ortonormálna báza vo V. Označme

$$\boldsymbol{X} = \begin{pmatrix} \langle \boldsymbol{x}_1, \boldsymbol{u}_1 \rangle & \dots & \langle \boldsymbol{x}_n, \boldsymbol{u}_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \boldsymbol{x}_1, \boldsymbol{u}_n \rangle & \dots & \langle \boldsymbol{x}_n, \boldsymbol{u}_n \rangle \end{pmatrix}.$$

Podľa našej definície je $\mathbf{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = \det \boldsymbol{X}$. Z tvrdenia 13.4.4 (b) zase vyplýva $\boldsymbol{G}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = \boldsymbol{X}^\mathsf{T} \cdot \boldsymbol{X}$. Teda podľa vety 15.1.2 máme

$$\begin{aligned} \operatorname{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) &= \sqrt{\det \boldsymbol{G}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)} = \sqrt{\det(\boldsymbol{X}^\mathsf{T} \cdot \boldsymbol{X})} \\ &= \sqrt{\det \boldsymbol{X}^\mathsf{T} \det \boldsymbol{X}} = |\det \boldsymbol{X}| = \left|\mathbf{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)\right|. \end{aligned}$$

Keďže stĺpce matice $\boldsymbol{X}$ sú tvorené súradnicami vektorov $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ v báze $\boldsymbol{\alpha}$, ak sú tieto lineárne nezávislé, tak $\boldsymbol{X}$ je zároveň maticou prechodu z bázy $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ do bázy $\boldsymbol{\alpha}$ (pozri paragraf 7.5). Teda bázy $\boldsymbol{\alpha}$ a $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ sú orientované súhlasne práve vtedy, keď $\mathbf{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = \det \boldsymbol{X} > 0$.

15.3.2. Dôsledok. *Nech V je orientovaný n-rozmerný euklidovský priestor. Potom vektory $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n \in V$ sú lineárne nezávislé (t. j. tvoria bázu priestoru V) práve vtedy, keď $\mathbf{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \neq 0$.*

Orientovaný objem $\mathbf{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ sa zvykne nazývať aj *vonkajší súčin* vektorov $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ a v tejto súvislosti sa značí $\boldsymbol{(x_1 \dots x_n)}$. V kurze fyziky sa vonkajší súčin $\boldsymbol{(xyz)}$ v euklidovskom priestore $\mathbb{R}^3$ so štandardným skalárnym súčinom zavádza pomocou skalárneho a vektorového súčinu formulou

$$\boldsymbol{(xyz)} = \langle \boldsymbol{x} \times \boldsymbol{y}, \boldsymbol{z} \rangle = (\boldsymbol{x} \times \boldsymbol{y}) \cdot \boldsymbol{z}$$

a zvykne sa tiež nazývať *zmiešaný súčin* vektorov $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$. V nasledujúcom paragrafe si ukážeme, ako je tento súčin „zmiešaný“ zo skalárneho a vektorového súčinu vo všeobecnom n-rozmernom euklidovskom priestore. Postupovať však budeme opačným smerom – orientovaný objem (t. j. vonkajší súčin) spolu so skalárnym súčinom nám poslúžia ako východisko na definíciu vektorového súčinu.

15.4 Vektorový súčin

Predpokladajme i naďalej, že V je orientovaný euklidovský priestor dimenzie $n \geq 2$ a $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ je nejaká jeho kladne orientovaná ortonormálna báza.

Nech $0 \leq k \leq n$ a $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k \in V$ sú pevne zvolené vektory. Dosadením na prvých k miest orientovaného objemu (vonkajšieho súčinu) tieto vektory definujú vzťahom

$$F(\boldsymbol{y}_1, \dots, \boldsymbol{y}_{n-k}) = \mathbf{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_k, \boldsymbol{y}_1, \dots, \boldsymbol{y}_{n-k}) = \boldsymbol{(x_1 \dots x_k y_1 \dots y_{n-k})}$$

$(n-k)$-lineárne alternujúce zobrazenie $F : V^{n-k} \to \mathbb{R}$. V prípade $k = n$ nám, samozrejme, nezostáva miesto na dosadzovanie ypsilonov, takže F prirodzene stotožňujeme s hodnotou $\mathbf{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \in \mathbb{R}$; touto otázkou sme sa už zaoberali v predchádzajúcom paragrafe. V tomto paragrafe sa budeme zaoberať prípadom $k = n-1$.

Pevne zvolené vektory $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1}$ definujú vzťahom

$$\psi(\boldsymbol{y}) = \mathbf{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1}, \boldsymbol{y}) = \boldsymbol{(x_1 \dots x_{n-1} y)}$$

lineárny funkcionál $\psi : V \to \mathbb{R}$. Uvedomme si, že skalárny súčin je *regulárna* symetrická bilineárna forma na V. Preto podľa dôsledku 11.1.8 má každý lineárny funkcionál $\varphi : V \to \mathbb{R}$ tvar $\varphi(\boldsymbol{y}) = \langle \boldsymbol{v}, \boldsymbol{y} \rangle$ pre jednoznačne určený vektor $\boldsymbol{v} \in V$. Aplikované na náš konkrétny prípad to znamená, že existuje jediný vektor $\boldsymbol{v} \in V$ taký, že

$$\psi(\boldsymbol{y}) = \mathbf{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1}, \boldsymbol{y}) = \boldsymbol{(x_1 \dots x_{n-1} y)} = \langle \boldsymbol{v}, \boldsymbol{y} \rangle$$

pre všetky $\boldsymbol{y} \in V$. Tento jednoznačne určený vektor nazývame *vektorovým súčinom*, prípadne tiež *ortokomplementom* vektorov $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1}$, a značíme ho

$$\boldsymbol{v} = \boldsymbol{x}_1 \times \dots \times \boldsymbol{x}_{n-1}.$$

Z uvedenej defnície priamo vyplýva, že, rovnako ako orientovaný n-rozmerný objem, ani vektorový súčin nezávisí na výbere konkrétnej kladne orientovanej ortonormálnej bázy vo V.

V dimenzii $n = 2$ ide len o unárnu (jednomiestnu) operáciu $V \to V$ – z toho dôvodu nehovoríme o vektorovom súčine ale výlučne o *ortokomplemente* vektora $\boldsymbol{x}$, ktorý značíme $\boldsymbol{x}^\perp$ (keďže znak súčinu $\times$ nemáme medzi čo umiestniť). V dimenzii $n = 3$ ide o binárnu (dvojmiestnu) operáciu $V \times V \to V$ vektorového súčinu, známeho z kurzov fyziky. No v dimenzii $n \geq 4$ je už $n - 1 > 2$ a vektorový súčin je $(n-1)$-miestna operácia na V.

Vzťah medzi skalárnym (t. j. vnútorným), vektorovým a vonkajším súčinom zachytáva nasledujúca rovnosť, platná pre ľubovoľné $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}, \boldsymbol{x}_n \in V$,

$$\langle \boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1}, \boldsymbol{x}_n \rangle = \mathbf{vol}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}, \boldsymbol{x}_n) = \boldsymbol{(}\boldsymbol{x}_1 \ldots \boldsymbol{x}_{n-1} \boldsymbol{x}_n\boldsymbol{)}.$$

Čítaná zľava doprava je to už uvedená definícia vektorového súčinu $\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1}$ prostredníctvom orientovaného objemu a skalárneho súčinu. Môžeme ju však čítať aj sprava doľava – vtedy sa z nej stáva definícia vonkajšieho súčinu (orientovaného objemu) ako istej kompozície skalárneho a vektorového súčinu. Takto sa bežne postupuje v kurzoch fyziky, kde sa najprv zavedie skalárny a vektorový súčin v dimenzii $n = 3$ geometrickým spôsobom a zmiešaný (t. j. vonkajší) súčin príde na rad až po nich.

Pokúsme sa teraz vyjadriť súradnice vektorového súčinu $\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1}$ v báze $\boldsymbol{\alpha}$ pomocou súradníc jednotlivých vektorov $\boldsymbol{x}_j$.

Začneme najjednoduchším prípadom $n = 2$. Ortokomplement $\boldsymbol{x}^\perp$ vektora $\boldsymbol{x}$ je taký vektor, že

$$\langle \boldsymbol{x}^\perp, \boldsymbol{y} \rangle = \mathbf{vol}_2(\boldsymbol{x}, \boldsymbol{y}) = \begin{vmatrix} \langle \boldsymbol{x}, \boldsymbol{u}_1 \rangle & \langle \boldsymbol{y}, \boldsymbol{u}_1 \rangle \\ \langle \boldsymbol{x}, \boldsymbol{u}_2 \rangle & \langle \boldsymbol{y}, \boldsymbol{u}_2 \rangle \end{vmatrix}$$

platí pre každé $\boldsymbol{y} \in V$. Ak za $\boldsymbol{y}$ postupne dosadíme hodnoty $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, dostaneme

$$\langle \boldsymbol{x}^\perp, \boldsymbol{u}_1 \rangle = \mathbf{vol}_2(\boldsymbol{x}, \boldsymbol{u}_1) = \begin{vmatrix} \langle \boldsymbol{x}, \boldsymbol{u}_1 \rangle & 1 \\ \langle \boldsymbol{x}, \boldsymbol{u}_2 \rangle & 0 \end{vmatrix} = -\langle \boldsymbol{x}, \boldsymbol{u}_2 \rangle,$$
$$\langle \boldsymbol{x}^\perp, \boldsymbol{u}_2 \rangle = \mathbf{vol}_2(\boldsymbol{x}, \boldsymbol{u}_2) = \begin{vmatrix} \langle \boldsymbol{x}, \boldsymbol{u}_1 \rangle & 0 \\ \langle \boldsymbol{x}, \boldsymbol{u}_2 \rangle & 1 \end{vmatrix} = \langle \boldsymbol{x}, \boldsymbol{u}_1 \rangle.$$

Ak označíme $(\boldsymbol{x})_{\boldsymbol{\alpha}} = (x_1, x_2)^{\mathsf{T}}$, kde $x_1 = \langle \boldsymbol{x}, \boldsymbol{u}_1 \rangle$, $x_2 = \langle \boldsymbol{x}, \boldsymbol{u}_2 \rangle$ sú súradnice vektora $\boldsymbol{x}$ v báze $\boldsymbol{\alpha}$, tak pre súradnice vektora $\boldsymbol{x}^\perp$ dostaneme

$$(\boldsymbol{x}^\perp)_{\boldsymbol{\alpha}} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}.$$

Vektor $\boldsymbol{x}^{\perp}$ teda môžeme zapísať v tvare

$$\boldsymbol{x}^{\perp} = -x_2\boldsymbol{u}_1 + x_1\boldsymbol{u}_2 = \begin{vmatrix} x_1 & \boldsymbol{u}_1 \\ x_2 & \boldsymbol{u}_2 \end{vmatrix} = \det\big((\boldsymbol{x})_{\boldsymbol{\alpha}}, \boldsymbol{\alpha}^{\mathsf{T}}\big),$$

kde posledný determinant treba chápať ako úsporný zápis predošlej lineárnej kombinácie, ktorá je tak jeho čiste *formálnym* Laplaceovým rozvojom podľa druhého stĺpca.

Rovnako si budeme počínať pre ľubovoľné $n \geq 3$. Najprv položíme $x_{ij} = \langle \boldsymbol{x}_j, \boldsymbol{u}_i \rangle$ pre $1 \leq i \leq n$, $1 \leq j \leq n-1$, a označíme $\boldsymbol{X} = (x_{ij}) \in \mathbb{R}^{n\times(n-1)}$ maticu, ktorej stĺpce tvoria súradnice vektorov $\boldsymbol{x}_j$ v báze $\boldsymbol{\alpha}$. Vektorový súčin $\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1}$ je taký vektor, že

$$\langle \boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1}, \boldsymbol{y} \rangle = \mathbf{vol}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}, \boldsymbol{y}) = \det\big(\boldsymbol{X}, (\boldsymbol{y})_{\boldsymbol{\alpha}}\big)$$

platí pre každé $\boldsymbol{y} \in V$. Ak za $\boldsymbol{y}$ postupne dosadíme hodnoty $\boldsymbol{u}_i$, kde $1 \leq i \leq n$, pre súradnice vektora $\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1}$ v báze $\boldsymbol{\alpha}$ dostaneme

$$\begin{aligned} \langle \boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1}, \boldsymbol{u}_i \rangle &= \mathbf{vol}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}, \boldsymbol{u}_i) \\ &= \det(\boldsymbol{X}, \boldsymbol{e}_i) = (-1)^{n+i}|\boldsymbol{X}_i|, \end{aligned}$$

kde matica $\boldsymbol{X}_i \in \mathbb{R}^{(n-1)\times(n-1)}$ vznikne z matice $\boldsymbol{X}$ vynechaním i-teho riadku. Vektor $\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1}$ teda možno zapísať v tvare

$$\begin{aligned} \boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1} &= \sum_{i=1}^{n} (-1)^{n+i}|\boldsymbol{X}_i|\,\boldsymbol{u}_i \\ &= \begin{vmatrix} x_{11} & \ldots & x_{1\,n-1} & \boldsymbol{u}_1 \\ \vdots & \ddots & \vdots & \vdots \\ x_{n-1\,1} & \ldots & x_{n-1\,n-1} & \boldsymbol{u}_{n-1} \\ x_{n1} & \ldots & x_{n\,n-1} & \boldsymbol{u}_n \end{vmatrix} = \det\big(\boldsymbol{X}, \boldsymbol{\alpha}^{\mathsf{T}}\big). \end{aligned}$$

Pritom posledný determinant treba chápať najmä ako pomôcku na ľahké zapamätanie predošlej lineárnej kombinácie, ktorú z neho možno dostať *formálnym* Laplaceovým rozvojom podľa n-tého stĺpca.

V trojrozmernom euklidovskom priestore $\mathbb{R}^3$ so štandardným skalárnym súčinom sa vektory (kladne orientovanej) kanonickej bázy $\boldsymbol{\varepsilon}^{(3)}$ niekedy zvyknú značiť $\boldsymbol{e}_1 = \boldsymbol{i}$, $\boldsymbol{e}_2 = \boldsymbol{j}$, $\boldsymbol{e}_3 = \boldsymbol{k}$. Vektorový súčin vektorov $\boldsymbol{x} = (x_1, x_2, x_3)^{\mathsf{T}}$, $\boldsymbol{y} = (y_1, y_2, y_3)^{\mathsf{T}}$ v tomto prípade nadobúda známy tvar

$$\boldsymbol{x} \times \boldsymbol{y} = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix} = \begin{vmatrix} x_1 & y_1 & \boldsymbol{i} \\ x_2 & y_2 & \boldsymbol{j} \\ x_3 & y_3 & \boldsymbol{k} \end{vmatrix},$$

pričom v pravotočivom (ľavotočivom) súradnom systéme v $\mathbb{R}^3$ sú jeho smer a orientácia dané *pravidlom pravej (ľavej) ruky*: ak položíme dlaň príslušnej ruky v smere vektora $\boldsymbol{x}$ tak, že zohnuté prsty ukazujú v smere „kratšieho“ otočenia vektora $\boldsymbol{x}$ do vektora $\boldsymbol{y}$ okolo počiatku, tak vztýčený palec ukazuje v smere vektora $\boldsymbol{x} \times \boldsymbol{y}$.

Všetky základné vlastnosti operácie vektorového súčinu $V^{n-1} \to V$ možno teraz jednoducho odvodiť na základe jej definície, prípadne jej vzťahu k n-rozmernému orientovanému objemu a reprezentácie v tvare uvedeného formálneho determinantu.

15.4.1. Veta. *Nech V je orientovaný euklidovský priestor dimenzie $n \geq 2$ a $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1} \in V$. Potom*
(a) vektorový súčin je $(n-1)$-lineárne alternujúce zobrazenie $V^{n-1} \to V$;
(b) vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}$ sú lineárne závislé práve vtedy, keď $\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1} = \mathbf{0}$;
(c) ak vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}$ sú lineárne nezávislé, tak $\boldsymbol{v} = \boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1}$ je normálový vektor nadroviny $[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}]$ a vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}, \boldsymbol{v}$ tvoria kladne orientovanú bázu priestoru V;
(d) $\|\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1}\| = \mathrm{vol}_{n-1}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1})$.

Dôkaz. Nech $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ je nejaká kladne orientovaná ortonormálna báza priestoru V. Označme $\boldsymbol{X} = (x_{ij})_{n\times(n-1)}$, kde $x_{ij} = \langle \boldsymbol{x}_j, \boldsymbol{u}_i \rangle$. Potom $\boldsymbol{v} = \boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1} = \det(\boldsymbol{X}, \boldsymbol{\alpha}^\mathsf{T})$.

(a) je zrejmé z reprezentácie vektorového súčinu $\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1}$ v tvare uvedeného formálneho determinantu.

(b) Uvedomme si, že podľa dôsledku 11.1.8 je rovnosť $\boldsymbol{v} = \mathbf{0}$ ekvivalentná s podmienkou $(\forall \boldsymbol{y} \in V)(\langle \boldsymbol{v}, \boldsymbol{y} \rangle = 0)$. Podľa definície vektorového súčinu však platí

$$\langle \boldsymbol{v}, \boldsymbol{y} \rangle = \mathbf{vol}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}, \boldsymbol{y}).$$

Z dôsledku 15.3.2 potom vyplýva, že $\boldsymbol{v} = \mathbf{0}$ práve vtedy, keď pre každé $\boldsymbol{y} \in V$ sú vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}, \boldsymbol{y}$ lineárne závislé. Keďže $\dim V > n-1$, tento prípad zrejme nastane práve vtedy, keď už samotné vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}$ sú lineárne závislé.

(c) Podľa 15.3.2 pre ľubovoľné $1 \leq j \leq n-1$ platí

$$\langle \boldsymbol{v}, \boldsymbol{x}_j \rangle = \mathbf{vol}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_{n-1}, \boldsymbol{x}_j) = 0,$$

t. j. $\boldsymbol{v} \perp \boldsymbol{x}_j$. Navyše, ak $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}$ sú lineárne nezávislé, tak $\boldsymbol{v} \neq \mathbf{0}$, teda $\boldsymbol{v}$ je normála nadroviny $[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}]$.

Dokážeme, že báza $\boldsymbol{\beta} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}, \boldsymbol{v})$ priestoru V je súhlasne orientovaná s bázou $\boldsymbol{\alpha}$. Uvedomme si, že matica prechodu z bázy $\boldsymbol{\beta}$ do bázy $\boldsymbol{\alpha}$ má tvar $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \big(\boldsymbol{X}, (\boldsymbol{v})_{\boldsymbol{\alpha}}\big)$. Rozvojom jej determinantu podľa posledného stĺpca a s využitím našej znalosti súradníc

$$(\boldsymbol{v})_{\boldsymbol{\alpha}} = \big((-1)^{n+1}|\boldsymbol{X}_1|, \ \ldots, \ (-1)^{n+n}|\boldsymbol{X}_n| \big)^\mathsf{T},$$

kde $\boldsymbol{X}_i$ vznikne z matice $\boldsymbol{X}$ vynechaním i-teho riadku, dostávame

$$\left|\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}\right| = \sum_{i=1}^{n} (-1)^{n+i}|\boldsymbol{X}_i|\,(-1)^{n+i}|\boldsymbol{X}_i| = \sum_{i=1}^{n} |\boldsymbol{X}_i|^2 = \|\boldsymbol{v}\|^2 > 0,$$

lebo v dôsledku (b) je $\boldsymbol{v} \neq \boldsymbol{0}$. To však znamená, že bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ sú súhlasne orientované.

(d) Keďže $\boldsymbol{v} \in [\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1}]^{\perp}$, s využitím definície vektorového súčinu, vety 15.3.1, práve dokázanej druhej časti (c) a rekurzívnej definície k-rozmerného objemu môžeme písať

$$\begin{aligned}\|\boldsymbol{v}\|^2 &= \langle \boldsymbol{x}_1 \times \dots \times \boldsymbol{x}_{n-1}, \boldsymbol{v}\rangle = \mathbf{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1}, \boldsymbol{v})\\ &= \mathrm{vol}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1}, \boldsymbol{v}) = \mathrm{vol}_{n-1}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1})\, \|\boldsymbol{v}\|\,.\end{aligned}$$

Ak $\|\boldsymbol{v}\| = 0$, tak podľa (b) sú $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1}$ lineárne závislé, preto na základe tvrdenia 15.1.1 (a) tiež $\mathrm{vol}_{n-1}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1}) = 0$. V opačnom prípade požadovanú rovnosť dostaneme krátením oboch strán členom $\|\boldsymbol{v}\| \neq 0$.

Z časti (d) a vety 15.1.2 priamo dostávame nasledujúci dôsledok.

15.4.2. Dôsledok. *(a) V dvojrozmernom orientovanom euklidovskom priestore V pre každý vektor $\boldsymbol{x} \in V$ platí $\|\boldsymbol{x}^{\perp}\| = \|\boldsymbol{x}\|$.*

(b) V trojrozmernom orientovanom euklidovskom priestore V pre všetky nenulové vektory $\boldsymbol{x}, \boldsymbol{y} \in V$ platí $\|\boldsymbol{x} \times \boldsymbol{y}\| = \|\boldsymbol{x}\|\,\|\boldsymbol{y}\| \sin \sphericalangle(\boldsymbol{x}, \boldsymbol{y})$.

(c) Pre $n \geq 2$ v n-rozmernom orientovanom euklidovskom priestore V pre všetky $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1} \in V$ platí $\|\boldsymbol{x}_1 \times \dots \times \boldsymbol{x}_{n-1}\| = |\boldsymbol{G}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1})|^{1/2}$.

Cvičenia

15.1. V n-rozmernom euklidovskom priestore uvažujme n-rozmerné teleso, ktorého objem závisí od dĺžkového parametra R podľa vzťahu $V_n(R) = V_n(1) \cdot R^n$. Predstavme si toto teleso napr. ako krabicu so stenami z kartónu papiera alebo ovocie guľovitého tvaru so šupkou hrúbky ΔR.

(a) Predpokladajme, že hrúbka „šupky“ ΔR tvorí 1% parametra R. Dokážte, že pre dostatočne veľké n tvorí „šupka“ viac než 99% n-rozmerného objemu takého „ovocia“. Nájdite približnú hranicu, počnúc ktorým n to nastane.

(b) Za predpokladu, že šupka je voči celkovému rozmeru R veľmi tenká, určte približnú závislosť, ako sa mení pomer objemu šupky k celkovému objemu, t. j. výraz $\frac{\Delta V_n(R)}{V_n(R)} = \frac{V_n(R) - V_n(R-\Delta R)}{V_n(R)}$, ako funkcia pomeru $\frac{\Delta R}{R}$.

(c) Približne ako musí s rastúcou dimenziou n klesať pomer $\frac{\Delta R}{R}$, aby pomer $\frac{\Delta V_n(R)}{V_n(R)}$ mal konštantnú hodnotu C?

15.2. Nech $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ sú dve zhodne orientované ortonormálne bázy v n-rozmernom euklidovskom priestore V. Overte priamym výpočtom, že vyjadrenia orientovaného objemu $\operatorname{vol}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ resp. vektorového súčinu $\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1}$ ľubovoľných vektorov z V pomocou báz $\boldsymbol{\alpha}$ resp. $\boldsymbol{\beta}$ sa rovnajú.

15.3. Ako je to s orientáciou vektorového súčinu $\boldsymbol{x} \times \boldsymbol{y}$ v trojrozmernom euklidovskom priestore, keď vektory $\boldsymbol{x}$, $\boldsymbol{y}$ zvierajú uhol $180°$, teda obe otočenia $\boldsymbol{x}$ do $\boldsymbol{y}$ sú „rovnako dlhé“?

15.4. (a) Overte priamym výpočtom, že pre ľubovoľné nenulové vektory $\boldsymbol{x}$, $\boldsymbol{y}$ v trojrozmernom euklidovskom priestore platí $\|\boldsymbol{x} \times \boldsymbol{y}\| = \|\boldsymbol{x}\| \, \|\boldsymbol{y}\| \sin \sphericalangle(\boldsymbol{x}, \boldsymbol{y})$.

(b) Odvoďte z toho identitu $\langle \boldsymbol{x}, \boldsymbol{y} \rangle^2 + \|\boldsymbol{x} \times \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 \, \|\boldsymbol{y}\|^2$.

15.5. Nech V je trojrozmerný vektorový priestor. Dokážte, že pre ľubovoľné $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ platia rovnosti

(a) $\langle \boldsymbol{x} \times \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \times \boldsymbol{z} \rangle$;

(b) $\boldsymbol{x} \times (\boldsymbol{y} \times \boldsymbol{z}) = \langle \boldsymbol{x}, \boldsymbol{z} \rangle \boldsymbol{y} - \langle \boldsymbol{x}, \boldsymbol{y} \rangle \boldsymbol{z}$;

(c) $\boldsymbol{x} \times (\boldsymbol{y} \times \boldsymbol{z}) - (\boldsymbol{x} \times \boldsymbol{y}) \times \boldsymbol{z} = \boldsymbol{x} \langle \boldsymbol{y}, \boldsymbol{z} \rangle - \langle \boldsymbol{x}, \boldsymbol{y} \rangle \boldsymbol{z}$;

(d) $\boldsymbol{x} \times (\boldsymbol{y} \times \boldsymbol{z}) + \boldsymbol{y} \times (\boldsymbol{z} \times \boldsymbol{x}) + \boldsymbol{z} \times (\boldsymbol{x} \times \boldsymbol{y}) = \boldsymbol{0}$.

Rovnosť (d) sa nazýva *Jacobiho identita*.

15.6. Nech $\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_{n-1}$ sú afinne nezávislé body v n-rozmernom euklidovskom priestore. Potom rovnica nadroviny $\ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_{n-1})$ (vzhľadom na ľubovoľnú kladne orientovanú ortonormálnu bázu priestoru V) má tvar $\langle (\boldsymbol{p}_1 - \boldsymbol{p}_0) \times \ldots \times (\boldsymbol{p}_{n-1} - \boldsymbol{p}_0), \boldsymbol{x} - \boldsymbol{p}_0 \rangle = 0$. Dokážte.

15.7. Nech K je ľubovoľné pole. Potom $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^{\mathsf{T}} \cdot \boldsymbol{y}$ je regulárna symetrická bilineárna forma na vektorovom priestore K^n a kanonická báza $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ je „ortonormálna“ vzhľadom na tento „skalárny súčin“. Definujme vektorový súčin v K^n ako formálny determinant $\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1} = \det\bigl(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}, \boldsymbol{\varepsilon}^{\mathsf{T}}\bigr)$. Overte, že rovnica nadroviny $\ell(\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_{n-1}) \subseteq K^n$ má rovnaký tvar ako v predošlom cvičení.

16. Úvod do špeciálnej teórie relativity

V predchádzajúcich troch kapitolách sa nám podarilo zrekonštruovať značnú časť euklidovskej geometrie, zovšeobecnenej do ľubovoľnej konečnej dimenzie, z jedinej kladne definitnej symetrickej bilineárnej formy na reálnom vektorovom priestore. V tejto kapitole najprv stručne preskúmame geometriu konečnorozmerných reálnych vektorových priestorov vybavených indefinitnou regulárnou symetrickou bilineárnou formou. Potom si predvedieme, ako možno z jedinej takejto formy signatúry $(1, n, 0)$ odvodiť matematický aparát *špeciálnej teórie relativity*. Postupovať však budeme v opačnom smere, ako je zvykom vo fyzike. Nebudeme budovať matematický model analýzou fyzikálnej situácie, ale naopak, tento model, tzv. *Minkowského časopriestor*, nájdeme už hotový. Fyzika sa nám začne vynárať pri jeho matematickom štúdiu takpovediac samovoľne, keď pre niektoré javy a objekty, s ktorými sa v ňom stretneme, začneme používať fyzikálnu terminológiu. Pri tom, samozrejme, budeme dbať na to, aby takéto pomenúvanie bolo v zhode s našou fyzikálnou intuíciou. To nebude zďaleka také ľahké, ako by sa vari dalo čakať – čitateľovi sú iste aspoň zbežne známe niektoré „populárne“ dôsledky špeciálnej teórie relativity, ktoré našej každodennej fyzikálnej skúsenosti zdanlivo protirečia. Aj nimi sa tu budeme pomerne podrobne zaoberať.

Základom špeciálnej teórie relativity je dôsledne uplatnený *Galileov princíp relativity*, postulujúci ekvivalenciu popisu pohybu a mechanických dejov z pohľadu ktoréhokoľvek z navzájom rovnomerne priamočiaro sa pohybujúcich pozorovateľov. Einstein tento princíp rozšíril do postulátu ekvivalencie popisu prírody z hľadiska ktoréhokoľvek z takýchto pozorovateľov. Presnejšia formulácia *Einsteinovho princípu relativity* hovorí, že *všetky fyzikálne zákony majú rovnakú matematickú podobu nezávisle od inerciálnej sústavy, vzhľadom na ktorú ich formulujeme. Stálosť rýchlosti svetla* už možno považovať za dôsledok tohto princípu. Medzi fyzikálne zákony treba totiž zahrnúť aj rýchlosť, akou sa šíri svetelný signál vo vákuu. Tým sa z tejto hodnoty stáva fundamentálna konštanta, rovnaká pre všetky inerciálne sústavy.[1] Matematická podoba formúl teórie relativity si navyše vynucuje uznať *medzný charakter rýchlosti svetla*: relatívna rýchlosť pohybu hmotných objektov je vždy menšia než rýchlosť svetla.

Rozpisovať sa o epochálnom význame Einsteinovho objavu špeciálnej a potom všeobecnej teórie relativity by dnes už bolo nosením dreva do lesa. Patrí sa však poznamenať, že základy špeciálnej relativity možno do istej

[1] Niektorí fyzici považujú princíp stálosti rýchlosti svetla za nezávislý princíp.

miery nájsť už u Lorentza a jej značnú časť rozvinul prakticky súčasne s Einsteinom a nezávisle na ňom Poincaré. Výlučným Einsteinovým objavom je až *všeobecná teória relativity* – no pri jej formulácii už nevystačíme s matematickým aparátom lineárnej algebry. Vznik všeobecnej relativity však bol do značnej miery umožnený Minkowského formuláciou špeciálnej relativity, ktorá predstavuje jeden z prvých a rozhodujúcich momentov mimoriadne plodného a dodnes živého programu tzv. *geometrizácie fyziky.* Práve výklad špeciálnej teórie relativity v Minkowského geometrickom poňatí bude náplňou tejto kapitoly.

16.1 Pseudoeuklidovské priestory

Pseudoeuklidovským priestorom nazývame ľubovoľný konečnorozmerný vektorový priestor V nad poľom $\mathbb{R}$, vybavený regulárnou indefinitnou symetrickou bilineárnou formou. Túto formu nazývame *pseudoskalárny súčin* a jej hodnotu na vektoroch $\boldsymbol{x}, \boldsymbol{y} \in V$ značíme $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$, t. j. rovnako, ako sme značili skalárny súčin. *Signatúrou pseudoeuklidovského priestoru V* rozumieme signatúru príslušnej formy. Táto má tvar $(p, q, 0)$, kde $p, q \geq 1$ a $p + q = \dim V$, čo nám umožňuje vynechať z nej posledný člen 0 a hovoriť o nej len ako o signatúre (p, q); v takom prípade hovoríme tiež o *(p, q)-rozmernom pseudoeuklidovskom priestore.* Od tejto chvíle až do konca tohto paragrafu V označuje nejaký pevne zvolený pseudoeuklidovský priestor a $n = \dim V$.

Pseudoskalárny súčin vo V takisto spĺňa prvé tri podmienky z definície skalárneho súčinu zo začiatku paragrafu 13.1, ako aj ich o kúsok ďalej uvedené dva dôsledky. Podmienku kladnej definitnosti (z ktorej už vyplýva regularita) však treba nahradiť nasledujúcimi dvoma podmienkami

$$
\begin{aligned}
&(\exists\, \boldsymbol{x}, \boldsymbol{y})\big(\langle \boldsymbol{x}, \boldsymbol{x} \rangle < 0 < \langle \boldsymbol{y}, \boldsymbol{y} \rangle\big) \quad \text{(indefinitnosť)},\\
&(\forall\, \boldsymbol{y})\big(\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0\big) \;\Rightarrow\; \boldsymbol{x} = \boldsymbol{0} \quad \text{(regularita)},
\end{aligned}
$$

pričom ekvivalencia poslednej implikácie a regularity vyplýva z dôsledku 11.1.8.

Väčšinu pojmov, s ktorými sme sa zoznámili v euklidovských priestoroch, možno, niekedy s istými nevyhnutnými úpravami, zaviesť aj pre pseudoeuklidovské priestory. Taktiež celý rad výsledkov o euklidovských priestoroch si, opäť s istými modifikáciami, zachováva platnosť aj pre pseudoeuklidovské priestory. Keďže podrobné štúdium týchto priestorov nie je naším cieľom, nevydáme sa cestou systematickej revízie výsledkov troch predchádzajúcich kapitol. Obmedzíme sa len na niekoľko málo príkladov, ktoré nám budú užitočné v ďalších paragrafoch. To si však vyžiada zaviesť aj niekoľko pojmov a dokázať zopár výsledkov, ktoré nemajú priame analógie v euklidovských priestoroch.

Dvojmiestny vzťah ortogonality a ortokomplement množiny zavádzame rovnako ako v euklidovskom priestore

$$\boldsymbol{x} \perp \boldsymbol{y} \Leftrightarrow \langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0,$$
$$X^{\perp} = \{\boldsymbol{y} \in V;\ (\forall \boldsymbol{x} \in X)(\boldsymbol{x} \perp \boldsymbol{y})\},$$

pre $\boldsymbol{x}, \boldsymbol{y} \in V$, $X \subseteq V$.[2]

Gramovou maticou (usporiadanej k-tice) vektorov $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) \in V^n$ nazývame maticu

$$\boldsymbol{G}(\boldsymbol{\alpha}) = \boldsymbol{G}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) = \big(\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle\big)_{k \times k};$$

jej determinant $|\boldsymbol{G}(\boldsymbol{\alpha})|$ nazývame *Gramovým determinantom* vektorov $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$. Hovoríme, že *báza* $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ lineárneho podpriestoru $S \subseteq V$ je *ortonormálna*, ak pre všetky $i, j \leq k$ platí $\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = 0$, ak $i \neq j$, a $\langle \boldsymbol{u}_i, \boldsymbol{u}_i \rangle = \pm 1$, t. j. práve vtedy, keď jej Gramova matica $\boldsymbol{G}(\boldsymbol{\alpha})$ je diagonálna, len s prvkami ± 1 na diagonále.

Štandardný pseudoskalárny súčin signatúry (p, q) na (stĺpcovom) vektorovom priestore $\mathbb{R}^{p+q}$ je daný predpisom

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{p} x_i y_i - \sum_{i=p+1}^{p+q} x_i y_i = \boldsymbol{x}^{\mathsf{T}} \cdot \operatorname{diag}(\boldsymbol{I}_p, -\boldsymbol{I}_q) \cdot \boldsymbol{y}.$$

Ako vyplýva z výsledkov kapitol 11, 12, každý pseudoskalárny súčin tejto signatúry možno voľbou vhodnej ortonormálnej bázy, pri správnom poradí jej členov, upraviť na uvedený tvar. Pseudoeuklidovský priestor $\mathbb{R}^n$ so štandardným pseudoskalárnym súčinom signatúry (p, q) budeme značiť $\mathbb{R}^{(p,q)}$.

Lineárny podpriestor $S \subseteq V$ sa nazýva *kladne definitný*, *záporne definitný*, *indefinitný*, *regulárny*, resp. *singulárny*, ak bilineárna forma $\langle\ ,\ \rangle$ zúžená na S má príslušnú vlastnosť. Zrejme kladne alebo záporne definitný podpriestor je regulárny.

Podobne, nenulový *vektor* $\boldsymbol{u} \in V$ sa nazýva *kladne* resp. *záporne definitný*, ak $\langle \boldsymbol{u}, \boldsymbol{u} \rangle > 0$, resp. $\langle \boldsymbol{u}, \boldsymbol{u} \rangle < 0$, t. j. práve vtedy, keď ním generovaný lineárny podpriestor má príslušnú vlastnosť. (Rozmyslite si, prečo nemá zmysel hovoriť o indefinitných vektoroch.)

Nenulový *vektor* $\boldsymbol{u} \in V$ sa nazýva *izotropný*, ak $\langle \boldsymbol{u}, \boldsymbol{u} \rangle = 0$, t. j. ak $\boldsymbol{u} \perp \boldsymbol{u}$. V opačnom prípade hovoríme, že $\boldsymbol{u}$ je *anizotropný vektor*. Na rozdiel od euklidovských priestorov, v pseudoeuklidovských priestoroch existujú izotropné vektory (presvedčte sa o tom), ako aj netriviálne singulárne podpriestory (napr. každý podpriestor generovaný nenulovým izotropným vektorom je taký). Na druhej strane, každá ortonormálna báza lineárneho podpriestoru vo V nutne pozostáva len z anizotropných vektorov.

[2] Správne by sme mali samozrejme hovoriť o *pseudoortogonálnych vektoroch*, *pseudoortokomplemente množiny* a pod. Dáme však prednosť stručnosti na úkor presnosti.

16.1.1. Tvrdenie. *(a) Lineárny podpriestor $S \subseteq V$ je regulárny práve vtedy, keď má ortonormálnu bázu.*

(b) Ľubovoľnú ortonormálnu bázu lineárneho podpriestoru $S \subseteq V$ možno doplniť do ortonormálnej bázy celého priestoru V.

Dôkaz. (a) Nech S je regulárny a $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ je jeho ľubovoľná báza. Potom aj Gramova matica $\boldsymbol{G}(\boldsymbol{\alpha})$, ako matica bilineárnej formy $\langle\ ,\ \rangle$ zúženej na S vzhľadom na bázu $\boldsymbol{\alpha}$, je regulárna. Podľa vety 12.1.2 existuje regulárna matica $\boldsymbol{P} \in \mathbb{R}^{k\times k}$ taká, že $\boldsymbol{P}^{\mathsf{T}} \cdot \boldsymbol{G}(\boldsymbol{\alpha}) \cdot \boldsymbol{P}$ je diagonálna matica len s prvkami ± 1 na diagonále. Potom $\boldsymbol{\alpha} \cdot \boldsymbol{P}$ je ortonormálna báza podpriestoru S. Obrátená implikácia je triviálna.

(b) Nech $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k)$ je nejaká ortonormálna báza (regulárneho) podpriestoru S. Doplňme ju (hocakým spôsobom) do bázy $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_n)$ priestoru V. Potom aj Gramova matica $\boldsymbol{G}(\boldsymbol{\beta})$ je regulárna a jej ľavý horný roh rozmeru $k \times k$ je diagonálny len s ± 1 na diagonále, čiže túto jej časť už upravovať nemusíme. Preto dvojicami ERO a ESO možno celú maticu upraviť na diagonálnu maticu $\boldsymbol{Q}^{\mathsf{T}} \cdot \boldsymbol{G}(\boldsymbol{\beta}) \cdot \boldsymbol{Q}$ s ± 1 na diagonále tak, že ani jeden z prvých k riadkov resp. stĺpcov pôvodnej matice $\boldsymbol{G}(\boldsymbol{\beta})$ nezmení polohu, nebude vynásobený skalárom $\neq 0$, ani k nemu nepripočítame násobok iného riadku či stĺpca. To znamená, že regulárna matica $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ zodpovedajúca príslušným ESO má blokový tvar $\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{I}_k & \boldsymbol{Q}_1 \\ \boldsymbol{0} & \boldsymbol{Q}_2 \end{pmatrix}$. Preto bázy $\boldsymbol{\beta}$ a $\boldsymbol{\beta} \cdot \boldsymbol{Q}$ majú prvých k vektorov rovnakých, teda $\boldsymbol{\beta} \cdot \boldsymbol{Q}$ je hľadaná ortonormálna báza priestoru V.

16.1.2. Tvrdenie. *Nech $S \subseteq V$ je regulárny lineárny podpriestor. Potom aj $S^{\perp}$ je regulárny lineárny podpriestor a platí*

$$V = S \oplus S^{\perp}, \qquad S^{\perp\perp} = S.$$

Dôkaz. Podľa tvrdenia 16.1.1 má S nejakú ortonormálnu bázu $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$, ktorú možno doplniť do ortonormálnej bázy $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{u}_{k+1}, \dots, \boldsymbol{u}_n)$ celého priestoru V. Ľahko nahliadneme, že $S^{\perp} = [\boldsymbol{u}_{k+1}, \dots, \boldsymbol{u}_n]$. Z toho už priamo vyplýva regularita podpriestoru $S^{\perp}$ ako aj rovnosti $S \cap S^{\perp} = \{\boldsymbol{0}\}$, $S + S^{\perp} = V$ a $S^{\perp\perp} = S$.

16.1.3. Tvrdenie. *Ak $S \subseteq V$ je maximálny kladne definitný podpriestor, tak $S^{\perp}$ je maximálny záporne definitný podpriestor.*

Samozrejme tiež naopak, ak $S \subseteq V$ je maximálny záporne definitný podpriestor, tak $S^{\perp}$ je maximálny kladne definitný podpriestor.

Dôkaz. Maximalita kladne definitného podpriestoru S znamená, že lineárny podpriestor $S + [\boldsymbol{x}]$ nie je kladne definitný pre žiadny vektor $\boldsymbol{x} \in V \setminus S$.

Keďže S je regulárny, podľa predchádzajúceho tvrdenia je regulárny aj $S^\perp$. Nech $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$, $\boldsymbol{\alpha}' = (\boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_n)$ sú ľubovoľné ortonormálne bázy podpriestorov S resp. $S^\perp$. Potom $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ je zrejme ortonormálna báza celého V. Z kladnej definitnosti S vyplýva, že $\boldsymbol{G}(\boldsymbol{\alpha}) = \boldsymbol{I}_k$; z jeho maximality a Sylvestrovho zákona zotrvačnosti (veta 12.1.1) zas $\boldsymbol{G}(\boldsymbol{\alpha}') = -\boldsymbol{I}_{n-k}$. Preto $S^\perp$ je záporne definitný podpriestor, ktorý je v dôsledku vety 12.1.1 zrejme maximálny s touto vlastnosťou.

Podľa ostatných dvoch tvrdení je každý pseudoeuklidovský priestor priamym súčtom $V = S \oplus T$ maximálneho kladne definitného podpriestoru S a maximálneho záporne definitného podpriestoru T; tento rozklad však nie je zďaleka jednoznačný. Pseudoskalárny súčin na podpriestore S je priamo skalárnym súčinom, takže S je vlastne euklidovský priestor. Takisto T možno považovať za euklidovský priestor – stačí formálne zmeniť znamienko pseudoskalárneho súčinu a priradením $(\boldsymbol{x}, \boldsymbol{y}) \mapsto -\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ je už definovaný skalárny súčin na T. Štruktúra pseudoeuklidovského priestoru vzniká takpovediac prepojením dvoch euklidovských štruktúr opačných znamienok. V dôsledku toho sa môže Cauchyho-Schwartzova nerovnosť niekedy zmeniť na opačnú.

16.1.4. Tvrdenie. *Nech $\boldsymbol{u}, \boldsymbol{v} \in V$ sú anizotropné vektory. Potom platí*

(a) $\langle \boldsymbol{u}, \boldsymbol{v} \rangle^2 = \langle \boldsymbol{u}, \boldsymbol{u} \rangle \langle \boldsymbol{v}, \boldsymbol{v} \rangle$
práve vtedy, keď $\boldsymbol{u}$, $\boldsymbol{v}$ sú lineárne závislé alebo $[\boldsymbol{u}, \boldsymbol{v}]$ je dvojrozmerný singulárny podpriestor;

(b) $\langle \boldsymbol{u}, \boldsymbol{v} \rangle^2 < \langle \boldsymbol{u}, \boldsymbol{u} \rangle \langle \boldsymbol{v}, \boldsymbol{v} \rangle$
práve vtedy, keď $\boldsymbol{u}$, $\boldsymbol{v}$ sú lineárne nezávislé a podpriestor $[\boldsymbol{u}, \boldsymbol{v}]$ je kladne alebo záporne definitný;

(c) $\langle \boldsymbol{u}, \boldsymbol{v} \rangle^2 > \langle \boldsymbol{u}, \boldsymbol{u} \rangle \langle \boldsymbol{v}, \boldsymbol{v} \rangle$
práve vtedy, keď $[\boldsymbol{u}, \boldsymbol{v}]$ je indefinitný podpriestor.

Dodajme, že v prípade, keď niektorý z vektorov $\boldsymbol{u}$, $\boldsymbol{v}$ je izotropný, triviálne platí $\langle \boldsymbol{u}, \boldsymbol{v} \rangle^2 \geq 0 = \langle \boldsymbol{u}, \boldsymbol{u} \rangle \langle \boldsymbol{v}, \boldsymbol{v} \rangle$, pričom rovnosť nastane práve vtedy, keď $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$.

Dôkaz. Keďže $|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})| = \langle \boldsymbol{u}, \boldsymbol{u} \rangle \langle \boldsymbol{v}, \boldsymbol{v} \rangle - \langle \boldsymbol{u}, \boldsymbol{v} \rangle^2$, uvedená rovnosť z (a), resp. nerovnosti z (b), (c) sú postupne ekvivalentné s podmienkami $|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})| = 0$, $|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})| > 0$, resp. $|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})| < 0$.

Ak $\boldsymbol{u}$, $\boldsymbol{v}$ sú lineárne závislé, tak rovnosť $|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})| = 0$ možno jednoducho overiť priamym výpočtom. Ak sú nezávislé, tak $\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})$ je maticou pseudoskalárneho súčinu na podpriestore $[\boldsymbol{u}, \boldsymbol{v}]$ v báze $(\boldsymbol{u}, \boldsymbol{v})$. Z toho vyplýva:

(a) Podpriestor $[\boldsymbol{u}, \boldsymbol{v}]$ je singulárny práve vtedy, keď matica $\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})$ je singulárna, t. j. práve vtedy, keď $|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})| = 0$.

(b) Podľa Sylvestrovho kritéria (veta 12.2.4) je podpriestor $[\boldsymbol{u}, \boldsymbol{v}]$ kladne definitný práve vtedy, keď $\langle \boldsymbol{u}, \boldsymbol{u} \rangle > 0$ a $|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})| > 0$, a záporne definitný práve vtedy, keď $\langle \boldsymbol{u}, \boldsymbol{u} \rangle < 0$ a $|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})| > 0$. Keďže pre anizotropný vektor $\boldsymbol{u}$ iná možnosť nenastane, $|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})| > 0$ práve vtedy, keď $[\boldsymbol{u}, \boldsymbol{v}]$ je kladne alebo záporne definitný.

(c) Ako vidno z (b), $[\boldsymbol{u}, \boldsymbol{v}]$ je indefinitný práve vtedy, keď $|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})| < 0$.

16.2 Minkowského časopriestor

Vo fyzike, presnejšie v špeciálnej teórii relativity, sa pod Minkowského priestoročasom zvyčajne rozumie pseudoeuklidovský priestor $\mathbb{R}^4$ so štandardným pseudoskalárnym súčinom signatúry $(3, 1)$, prípadne $(1, 3)$. Z matematického hľadiska sú obe možnosti ekvivalentné – voľba jednej z nich je preto najmä vecou vkusu. My dáme prednosť druhej možnosti, t. j. pre $\boldsymbol{x} = (x_0, x_1, x_2, x_3)^{\mathsf{T}}$, $\boldsymbol{y} = (y_0, y_1, y_2, y_3)^{\mathsf{T}}$ položíme

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_0 y_0 - x_1 y_1 - x_2 y_2 - x_3 y_3 = \boldsymbol{x}^{\mathsf{T}} \cdot \operatorname{diag}(1, -\boldsymbol{I}_3) \cdot \boldsymbol{y}.$$

Pritom súradnicu x_0 interpretujeme ako čas a x_1, x_2, x_3 ako súradnice polohy v euklidovskom priestore. V súlade s tým budeme používať termím *časopriestor*. Tento pojem však ďalej rozšírime jednak na vyššie, jednak na nižšie dimenzie a na abstraktné pseudoeuklidovské priestory. *Minkowského časopriestorom* budeme teda nazývať ľubovoľný pseudoeuklidovský priestor V signatúry $(1, n)$, kde $n \geq 1$. $\mathbb{R}^{(1,n)}$ označuje Minkowského časopriestor $\mathbb{R}^{n+1}$ so štandardným pseudoskalárnym súčinom signatúry $(1, n)$. V našom výklade budú hrať dôležitú úlohu práve časopriestory „malej“ signatúry $(1, 1)$ a $(1, 2)$, ktoré ešte pripúšťajú názorné grafické znázornenie.

Na Minkowského časopriestor V budeme v prevažnej miere pozerať afinne, t. j. jeho prvky budeme častejšie považovať za body než za vektory – tentokrát ich však budeme nazývať *udalosťami* alebo tiež *svetobodmi*. Svetobody predstavujú idealizované okamžité bodové udalosti (ako napr. vyžiarenie fotónu atómom, či zrážku dvoch elementárnych častíc), pri ktorých abstrahujeme od toho, „čo sa stalo“, a zaznamenávame len ich čas a polohu.

Ak $\boldsymbol{x}, \boldsymbol{y} \in V$ sú dva svetobody, tak pseudoskalárny súčin $\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle$ nazývame štvorcom ich *časopriestorovej odľahlosti*. Ak $\boldsymbol{x} \neq \boldsymbol{y}$, tak podľa toho, či $\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle$ je väčšie, rovné alebo menšie ako 0 (t. j. vektor $\boldsymbol{x} - \boldsymbol{y}$ je kladne definitný, izotropný alebo záporne definitný), hovoríme, že udalosti $\boldsymbol{x}$, $\boldsymbol{y}$ sú *časovo*, *svetelne*, resp. *priestorovo odľahlé*. Miesto kladne definitný, záporne definitný, resp. izotropný vektor hovoríme tiež *časový*, *priestorový*, resp. *svetelný vektor*.[3] Množinu

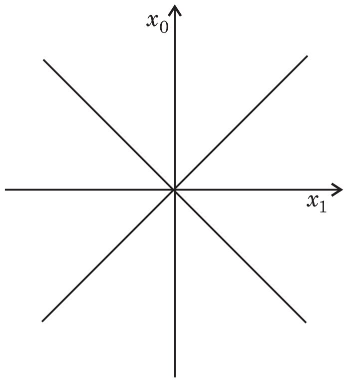
$$\mathrm{LC}(\boldsymbol{p}) = \{\boldsymbol{x} \in V;\ \langle \boldsymbol{x} - \boldsymbol{p}, \boldsymbol{x} - \boldsymbol{p} \rangle = 0\}$$

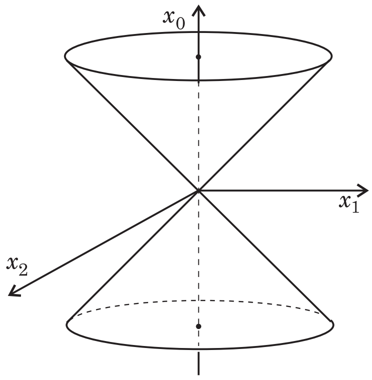
[3]Obvykle sa používajú možno výstižnejšie, no určite ťažkopádnejšie názvy *časupodobný*, *priestorupodobný* a *svetlupodobný vektor*.

všetkých udalostí, ktoré sú od daného svetobodu $\boldsymbol{p} \in V$ svetelne odľahlé, nazývame *svetelný kužeľ* (anglicky *light cone*) s počiatkom v $\boldsymbol{p}$. Tento názov je motivovaný tvarom svetelného kužeľa v Minkowského časopriestore $\mathbb{R}^{(1,2)}$, ktorý je znázornený na obrázku vpravo; vľavo vidíme svetelný kužeľ v Minkowského časopriestore $\mathbb{R}^{(1,1)}$, tvorený dvoma priamkami $x_1 = \pm x_0$.

V pozadí práve zavedeného názvoslovia stojí fyzikálna interpretácia Minkowského časopriestoru, ktorá bude v priebehu nášho výkladu vychádzať najavo čoraz zreteľnejšie. Zatiaľ si len všimnime, že vyslaniu svetelného signálu v istom okamihu z istého miesta možno priradiť istú udalosť v Minkowského časopriestore $\mathbb{R}^{(1,3)}$, ktorú si bez ujmy na všeobecnosti možno zvoliť za počiatok odpočtu času i súradnej sústavy v priestore. Tento signál sa šíri rovnakou rýchlosťou c všetkými smermi, takže v čase $t > 0$ bude vytvárať sférickú vlnoplochu s polomerom ct a rovnicou

$$x_1^2 + x_2^2 + x_3^2 = c^2t^2.$$





Obr. 16.1. Svetelný kužeľ v Minkowského časopriestoroch $\mathbb{R}^{(1,1)}$ a $\mathbb{R}^{(1,2)}$

Po voľbe rýchlosti svetla za jednotku rýchlosti ($c = 1$) a substitúcii $x_0 = ct = t$ vidíme, že všetky svetobody, do ktorých dospeje svetelný signál vyslaný v okamihu 0 z počiatku priestorovej súradnej sústavy, vytvárajú „hornú polovicu“ svetelného kužeľa, niekedy nazývanú tiež *svetelný kužeľ budúcnosti*,

$$\mathrm{LC}^+(\mathbf{0}) = \{\boldsymbol{x} \in \mathbb{R}^{(1,3)};\, x_0 \geq 0 \ \&\ \langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0\}.$$

Jeho „dolná polovica“, nazývaná aj *svetelný kužeľ minulosti*,

$$\mathrm{LC}^-(\mathbf{0}) = \{\boldsymbol{x} \in \mathbb{R}^{(1,3)};\, x_0 \leq 0 \ \&\ \langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0\}.$$

je tvorená svetobodmi, z ktorých svetelný signál dospel do počiatku priestorovej súradnej sústavy v okamihu 0. Hviezdy, ktoré vidíme na jasnej nočnej oblohe, sú rôzne vzdialené, preto svetlo z nich k nám letí rôzne dlho –

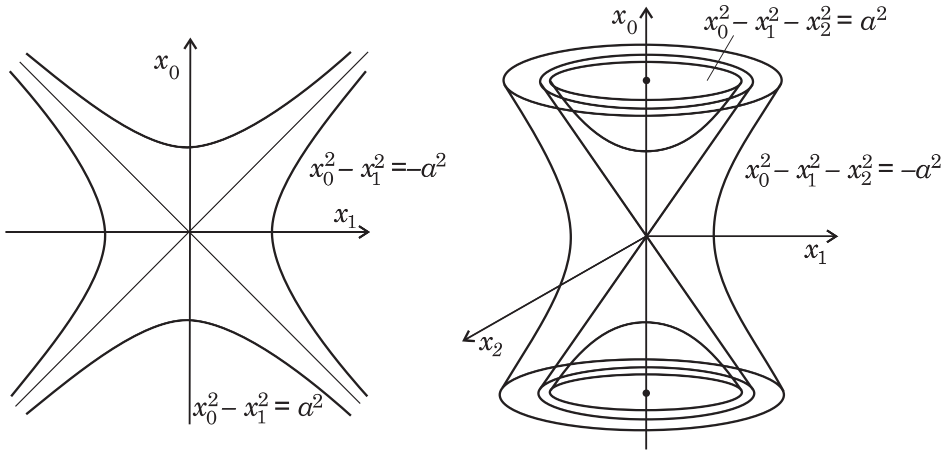
všetky takéto lúče však ležia na svetelnom kuželi minulosti $\mathrm{LC}^-(\mathbf{0})$. Svetobody, z ktorých bol svetelný signál vyslaný v čase $t < 0$ opäť vytvárajú sférickú vlnoplochu s polomerom $-ct$ a rovnakou rovnicou ako v predošlom prípade. Príkladom takejto vlnoplochy je belasá nebeská sféra, ktorej časť vidíme za jasného dňa nad hlavou.

Pri potlačení jednej priestorovej súradnice x_3 možno situáciu názorne ilustrovať v Minkowského časopriestore $\mathbb{R}^{(1,2)}$ (predchádzajúci obrázok vpravo). Miesto trojrozmerného priestoru si predstavme dvojrozmernú vodnú hladinu a miesto vyslania svetelného signálu hoďme kameň do vody. Vznikne vlnenie, ktorého čelo sa šíri po hladine v tvare kružnice a za čas $t > 0$ dospeje do vzdialenosti ct, kde c je rýchlosť jeho šírenia. „Hornú polovicu“ svetelného kužeľa si predstavme ako kruhové čelo vlny „unášané plynúcim časom“ – jeho stav v nejakom okamihu t je daný rezom kužeľa rovinou $x_0 = ct$.

Tento príklad navodzuje predstavu Minkovského časopriestoru signatúry $(1, n)$ ako n-rozmerného euklidovského priestoru „unášaného časom“ pozdĺž časovej osi. I keď táto predstava býva často užitočná, uvidíme, že v Minkowského časopriestore neexistuje privilegovaná časová os, ani kanonický, jednoznačný rozklad na časovú a priestorovú zložku, ako by sa nám mohlo zdať pri zbežnom pohľade na Minkowského časopriestor $\mathbb{R}^{(1,n)}$. Skutočnosť, že „časom unášaný fyzikálny priestor“ je euklidovský, čiže „plochý“, poukazuje na to, že špeciálna relativita skúma vlastne prázdny časopriestor, presnejšie, abstrahuje od gravitačného pôsosobenia v ňom rozloženej hmoty. Tieto otázky tematizuje až *všeobecná teória relativity*, ktorá gravitačné pôsobenie zachytáva opäť geometricky – ako spojite sa meniace zakrivenie časopriestoru.

Pre časový vektor $\boldsymbol{u} \in V$ možno definovať normu alebo dĺžku $\|\boldsymbol{u}\| = \sqrt{\langle \boldsymbol{u}, \boldsymbol{u} \rangle}$ rovnako ako v euklidovskom prípade. Pre priestorový vektor $\boldsymbol{v} \in V$ však kladieme $\|\boldsymbol{v}\| = \sqrt{-\langle \boldsymbol{v}, \boldsymbol{v} \rangle}$.

V euklidovskom priestore $\mathbb{R}^2$ vytvárajú vektory rovnakej dĺžky $r > 0$ (presnejšie ich konce) kružnicu s rovnicou $x_1^2 + x_2^2 = r^2$; v $\mathbb{R}^3$ je to sféra s rovnicou $x_1^2 + x_2^2 + x_3^2 = r^2$. Na rozdiel od toho v Minkowského časopriestore $\mathbb{R}^{(1,1)}$ vytvárajú časové vektory dĺžky $r > 0$ rovnoosú hyperbolu s rovnicou $x_0^2 - x_1^2 = r^2$; priestorové vektory dĺžky r zasa vytvárajú rovnoosú hyperbolu s rovnicou $x_0^2 - x_1^2 = -r^2$ (ďalší obrázok vľavo). V Minkowského časopriestore $\mathbb{R}^{(1,2)}$ vytvoria takéto časové vektory dvojdielny rotačný hyperboloid s rovnicou $x_0^2 - x_1^2 - x_2^2 = r^2$, ktorý leží „vovnútri“ svetelného kužeľa; zodpovedajúce priestorové vektory tvoria jednodielny rotačný hyperboloid s rovnicou $x_0^2 - x_1^2 - x_2^2 = -r^2$, ktorý obaľuje svetelný kužeľ „zvonka“ (obrázok vpravo). Do vyšších dimenzií, vrátane „nášho“ časopriestoru $\mathbb{R}^{(1,3)}$, bohužiaľ, už naša predstavivosť nesiaha. V istom zmysle si však nadplochy tvorené vektormi rovnakej dĺžky aj vo vyšších dimenziách zachovávajú niektoré podstatné vlastnosti, ktoré sme objavili v časopriestore signatúry $(1, 2)$: časové vektory dĺžky $r > 0$ stále vytvárajú dvojdielnu rotačnú hyperbolickú



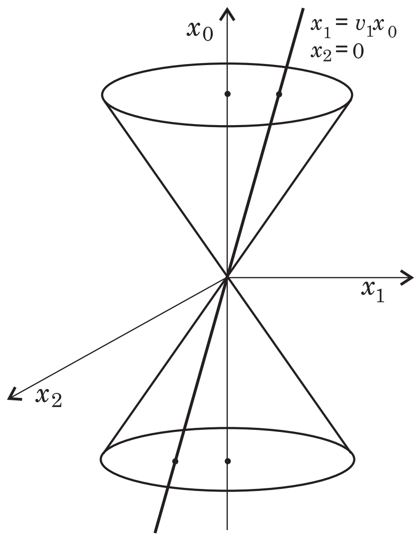
Obr. 16.2. Časové resp. priestorové vektory rovnakých dĺžok v $\mathbb{R}^{(1,1)}$ a $\mathbb{R}^{(1,2)}$

nadplochu umiestnenú „vnútri“ svetelného kužeľa, kým priestorové vektory dĺžky $r > 0$ tvoria jednodielnu rotačnú hyperbolickú nadplochu obaľujúcu svetelný kužeľ „zvonka“.

Varujeme však čitateľa, aby podobným obrázkom a z nich vychádzajúcim názorným geometrickým predstavám neprikladal väčšiu váhu, než im náleží – Minkowského časopriestory $\mathbb{R}^{(1,1)}$ a $\mathbb{R}^{(1,2)}$ sú na nich totiž zobrazené skreslene prostredníctvom euklidovskej geometrie. To vidno napr. už z toho, že časové vektory rovnakej dĺžky sú zobrazené ako vektory nerovnakej euklidovskej dĺžky. Len na okraj poznamenajme, že napr. na jednom diele hyperboloidu $x_0^2 - x_1^2 - x_2^2 = r^2$ v $\mathbb{R}^{(1,2)}$ sa realizuje dvojrozmerná *Bolyaiho-Lobačevského geometria*, zo zrejmých dôvodov nazývaná tiež *hyperbolickou*, ktorá je historicky prvým známym príkladom neeuklidovskej geometrie. Štúdium podobných, nesporne zaujímavých otázok však už nie je predmetom tohto kurzu.

16.3 Inerciálny pozorovateľ a jeho svetočiara

Inerciálneho pozorovateľa v Minkovského časopriestore V si predstavujeme ako ideálneho dispečera obrovského počtu rovnomerne priamočiaro sa pohybujúcich, no *navzájom nehybných*, čertovsky malých (presnejšie bodových) trpaslíkov, vyskytujúcich sa v každom svetobode a vybavených synchronizovanými hodinkami a jednotným metrom. Niekedy je však účelná predstava, že aj inerciálny pozorovateľ je len jedným z takýchto trpaslíkov, prípadne, že ktorýkoľvek z trpaslíkov môže prevziať úlohu dispečera. Matematicky však nebudeme zavádzať nijakých inerciálnych dispečerov ani trpaslíkov – úplne vystačíme s dráhami, ktoré opisujú vo V.



Obr. 16.3. Svetočiara inerciálneho pozorovateľa v $\mathbb{R}^{(1,2)}$

Na začiatok si uvedomme, akú dráhu opisuje v $\mathbb{R}^{(1,n)}$ nehybný bod. Keďže čas neustále plynie, i nehybný bod sa v $\mathbb{R}^{(1,n)}$ „pohybuje“ – a to po „zvislej“ priamke s parametrickými rovnicami $x_0 = t$, $x_1 = p_1$, …, $x_n = p_n$, kde $(p_1, \ldots, p_n)^\mathsf{T}$ sú jeho priestorové súradnice v okamihu $t = 0$. Jej smerový vektor je $\boldsymbol{e}_0 = (1, 0, \ldots, 0)^\mathsf{T}$. A akú dráhu v $\mathbb{R}^{(1,n)}$ opisuje bod pohybujúci sa rovnomerne priamočiaro rýchlosťou $\boldsymbol{v}$ so zložkami $v_1, \ldots, v_n$ v smere jednotlivých osí $x_1, \ldots, x_n$? Zrejme je to priamka s parametrickými rovnicami $x_0 = t$, $x_1 = p_1 + v_1 t$, …, $x_n = p_n + v_n t$, kde $(p_1, \ldots, p_n)^\mathsf{T}$ sú jeho priestorové súradnice v okamihu $t = 0$, a smerovým vektorom $(1, v_1, \ldots, v_n)^\mathsf{T}$. To si najlepšie znázorníme v $\mathbb{R}^{(1,2)}$, keď si zvolíme os x_1 v smere vektora rýchlosti $\boldsymbol{v}$, teda $\boldsymbol{v} = (v_1, 0)^\mathsf{T}$. Situácia pre $p_1 = p_2 = 0$ je znázornená na obrázku. Keďže rýchlosť pohybu hmotného bodu je menšia než rýchlosť svetla, t. j. $|v_1| < 1$, naša priamka leží „vovnútri“ svetelného kužeľa. Vo všeobecnom prípade máme $v_1^2 + \ldots + v_n^2 < 1$, čiže $(1, v_1, \ldots, v_n)^\mathsf{T}$ je časový vektor.

Teraz si uvedomme, že naše otázky boli chybne položené. Hovoriť o nehybnom alebo pohybujúcom sa pozorovateľovi ako takom nedáva rozumný fyzikálny zmysel. Neexistuje absolútny kľud ani absolútny pohyb, ale kľud i pohyb sú relatívne. Nejaký fyzikálny objekt môže byť v kľude alebo v pohybe len vzhľadom na nejaký iný objekt. Aspoň tak nás to učí klasická mechanika od čias Galileových. Aj tak však z našich chybne položených otázok možno vyťažiť netriviálny poznatok: *inerciálni pozorovatelia sa v Minkowského časopriestore pohybujú po priamkach s časovými smerovými vektormi.*

Svetočiarou inerciálneho pozorovateľa, alebo len *inerciálnou svetočiarou* nazývame ľubovoľnú *orientovanú* priamku (t. j. jednorozmerný afinný podpriestor) vo V tvaru

$$\mathrm{WL}(\boldsymbol{p},\boldsymbol{a}) = \boldsymbol{p} + [\boldsymbol{a}] = \{\boldsymbol{p} + t\boldsymbol{a};\ t \in \mathbb{R}\},$$

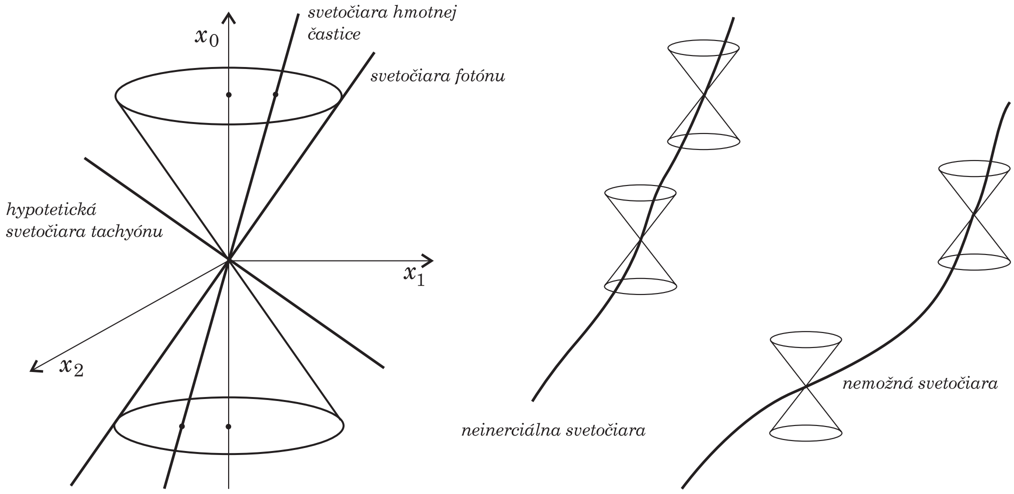
(z anglického *world line*), kde $\boldsymbol{p} \in V$ je svetobod a $\boldsymbol{a} \in V$ je časový vektor (čiže $\langle \boldsymbol{a},\boldsymbol{a}\rangle > 0$), s orientáciou zadanou vektorom $\boldsymbol{a}$. To znamená, že opačne orientovanú svetočiaru $\mathrm{WL}(\boldsymbol{p},-\boldsymbol{a})$ považujeme za rôznu od $\mathrm{WL}(\boldsymbol{p},\boldsymbol{a})$, hoci ako množiny bodov predstavujú tú istú priamku.

Tu treba upozorniť na ďalšie skreslenie, idúce na vrub zobrazenia v euklidovskej geometrii. Svetočiara $\mathrm{WL}(\boldsymbol{0},\boldsymbol{e}_0)$ v $\mathbb{R}^{(1,2)}$ je zobrazená ako os svetelného kužeľa $\mathrm{LC}(\boldsymbol{0})$, kým svetočiara $\mathrm{WL}(\boldsymbol{0},\boldsymbol{a})$ s iným časovým vektorom $\boldsymbol{a}$ vedie akoby bližšie popri jeho okraji. Z postulátu stálosti rýchlosti svetla však vyplýva, že svetočiary všetkých inerciálnych pozorovateľov „majú rovnako ďaleko k okraju svetelného kužeľa".

Formálne možno svetočiary $\mathrm{WL}(\boldsymbol{p},\boldsymbol{a})$ zaviesť rovnako pre ľubovoľné vektory $\boldsymbol{a} \neq \boldsymbol{0}$ vo V. Obmedzenie sa na časové smerové vektory je dôsledkom postulátu, podľa ktorého sa všetky hmotné objekty pohybujú rýchlosťou menšou než rýchlosť svetla. Svetočiary tvaru $\mathrm{WL}(\boldsymbol{p},\boldsymbol{a})$, kde $\boldsymbol{a}$ je svetelný vektor, t. j. $\langle \boldsymbol{a},\boldsymbol{a}\rangle = 0$, predstavujú pohyb svetelnou rýchlosťou – takto sa však môžu pohybovať len nehmotné častice (presnejšie, častice s nulovou kľudovou hmotnosťou), teda z dnes známych častíc iba fotóny (o neutrínach sú už dnes – zdá sa – všetci presvedčení, že sú hmotné). Svetočiary tvaru $\mathrm{WL}(\boldsymbol{p},\boldsymbol{a})$, kde $\boldsymbol{a}$ je priestorový vektor, t. j. $\langle \boldsymbol{a},\boldsymbol{a}\rangle < 0$, by zodpovedali pohybu nadsvetelnou rýchlosťou, teda – aspoň v rámci špeciálnej teórie relativity – nemajú fyzikálny význam. Hoci o časticiach pohybujúcich sa nadsvetelnými rýchlosťami, tzv. *tachyónoch*, sa v teoretickej fyzike stále špekuluje, všetky doterajšie pokusy objaviť ich skončili neúspešne.

Pohyb má podľa našich fyzikálnych predstáv vždy relatívny charakter. V abstraktnom Minkowského časopriestore, kde nemáme privilegovanú časovú os, sú všetky inerciálne svetočiary rovnocenné. Pokiaľ teda budeme hovoriť o pohybe inerciálneho pozorovateľa, vždy pôjde o jeho pohyb vzhľadom na iného inerciálneho pozorovateľa. Na druhej strane, niektoré *vlastnosti* pohybu sú absolútne. V Minkowského časopriestore V je to napr. vlastnosť „pohybovať sa rovnomerne priamočiaro" (t. j. opisovať inerciálnu svetočiaru vo V), a v dôsledku toho tiež vlastnosť „pohybovať sa premennou rýchlosťou". Svetočiara hmotného bodu v Minkowského časopriestore V pohybujúceho sa premennou rýchlosťou totiž nie je inerciálna. Zmena rýchlosti sa prejaví zakrivením príslušnej svetočiary. Nakoľko však okamžitá rýchlosť pohybu hmotného bodu je vždy menšia než rýchlosť svetla, dotykový vektor k takejto svetočiare v každom jej svetobode $\boldsymbol{p}$ je časový, t. j. leží „vovnútri" svetelného kužeľa $\mathrm{LC}(\boldsymbol{p})$.

K danej svetočiare $\mathrm{WL}(\boldsymbol{p},\boldsymbol{a})$ inerciálneho pozorovateľa v Minkowského časopriestore V existuje jednoznačne určený vektor $\boldsymbol{a}_0 = \langle \boldsymbol{a},\boldsymbol{a}\rangle^{-1/2}\boldsymbol{a}$ taký,



Obr. 16.4. Rôzne typy svetočiar v $\mathbb{R}^{(1,2)}$

že $\mathrm{WL}(\boldsymbol{p}, \boldsymbol{a}) = \mathrm{WL}(\boldsymbol{p}, \boldsymbol{a}_0)$ a $\langle \boldsymbol{a}_0, \boldsymbol{a}_0 \rangle = 1$, ktorý nazývame jej *časovým šípom* alebo *šípom času* (vo fyzike sa v Minkovského časopriestore $\mathbb{R}^{(1,3)}$ používa tiež názov *štvorrýchlosť*). To znamená, že parameter t vo vyjadrení svetobodov $\boldsymbol{x} = \boldsymbol{p} + t\boldsymbol{a}_0$ svetočiary $\mathrm{WL}(\boldsymbol{p}, \boldsymbol{a}_0)$ možno skutočne interpretovať ako vlastný čas príslušného inerciálneho pozorovateľa, počítaný od udalosti $\boldsymbol{p}$.

Matematicky tak možno inerciálneho pozorovateľa-dispečera stotožniť s časovým šípom $\boldsymbol{a}$ ľubovoľnej jeho svetočiary $\mathrm{WL}(\boldsymbol{p}, \boldsymbol{a})$ a jemu podriadených trpaslíkov s týmito svetočiarami. Ich zladenie, t. j. niečo ako synchronizácia, teda spočíva v totožnosti časových šípov všetkých takýchto svetočiar.

Svetočiara $\mathrm{WL}(\boldsymbol{p}, \boldsymbol{a})$ inerciálneho pozorovateľa tak predstavuje jeho vlastný tok času (a nie jeho pohyb – ten je možný len vzhľadom na iného inerciálneho pozorovateľa). Orientácia svetočiary zodpovedá orientácii času z minulosti do budúcnosti – prostriedkami špeciálnej teórie relativity ju však nemožno odlíšiť od orientácie z budúcnosti do minulosti, presnejšie, rozhodnúť, ktorá z nich je „tá pravá“. Možno však rozhodnúť, či sú dve inerciálne svetočiary orientované súhlasne alebo nesúhlasne, t. j. či ich vlastné časy plynú tým istým alebo opačným smerom.

Časové vektory $\boldsymbol{a}$, $\boldsymbol{b}$ sa nazývajú *súhlasne orientované*, ak $\langle \boldsymbol{a}, \boldsymbol{b} \rangle > 0$; ak $\langle \boldsymbol{a}, \boldsymbol{b} \rangle < 0$, hovoríme, že $\boldsymbol{a}$, $\boldsymbol{b}$ sú *nesúhlasne orientované*. (Samostatne si dokážte, že prípad $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = 0$ nemôže nastať.) Inerciálne svetočiary $\mathrm{WL}(\boldsymbol{p}, \boldsymbol{a})$, $\mathrm{WL}(\boldsymbol{q}, \boldsymbol{b})$ sú potom súhlasne resp. nesúhlasne orientované práve vtedy, keď sú súhlasne resp. nesúhlasne orientované ich časové vektory.

Podobne, hovoríme, že *svetelný vektor* $\boldsymbol{u} \neq \boldsymbol{0}$ je *súhlasne orientovaný s časovým vektorom* $\boldsymbol{a}$, ak $\langle \boldsymbol{a}, \boldsymbol{u} \rangle > 0$; ak $\langle \boldsymbol{a}, \boldsymbol{u} \rangle < 0$, hovoríme, že sú nesúhlasne orientované. (Prípad $\langle \boldsymbol{a}, \boldsymbol{u} \rangle = 0$ je opäť nemožný.)

Ukážeme si, že súhlasne orientované časové vektory $\boldsymbol{a}$, $\boldsymbol{b}$ určujú rovnakú orientáciu časových aj svetelných vektorov.

16.3.1. Tvrdenie. *Nech $\boldsymbol{a}, \boldsymbol{b}$ sú súhlasne orientované časové vektory v Minkowského časopriestore V. Potom pre ľubovoľný časový alebo svetelný vektor $\boldsymbol{u}$ platí $\langle \boldsymbol{a}, \boldsymbol{u} \rangle > 0$ práve vtedy, keď $\langle \boldsymbol{b}, \boldsymbol{u} \rangle > 0$,*

Dôkaz. Bez ujmy na všeobecnosti môžeme predpokladať, že $\boldsymbol{a}$, $\boldsymbol{b}$ sú súhlasne orientované časové šípy, t. j. $\langle \boldsymbol{a}, \boldsymbol{a} \rangle = \langle \boldsymbol{b}, \boldsymbol{b} \rangle = 1$ a $\langle \boldsymbol{a}, \boldsymbol{b} \rangle > 0$. Vzhľadom na symetriu úlohy stačí dokazovať implikáciu $\langle \boldsymbol{a}, \boldsymbol{u} \rangle > 0 \ \Rightarrow\ \langle \boldsymbol{b}, \boldsymbol{u} \rangle > 0$.

Priamym výpočtom sa možno presvedčiť, že vektory

$$\boldsymbol{b}' = \boldsymbol{b} - \langle \boldsymbol{a}, \boldsymbol{b} \rangle \boldsymbol{a}, \qquad\qquad \boldsymbol{u}' = \boldsymbol{u} - \langle \boldsymbol{a}, \boldsymbol{u} \rangle \boldsymbol{a}$$

patria do lineárneho podpriestoru $[\boldsymbol{a}]^\perp$, teda sú to priestorové vektory. Potom

$$0 > \langle \boldsymbol{b}', \boldsymbol{b}' \rangle = \langle \boldsymbol{b}, \boldsymbol{b} \rangle - 2\langle \boldsymbol{a}, \boldsymbol{b} \rangle^2 + \langle \boldsymbol{a}, \boldsymbol{b} \rangle^2 \langle \boldsymbol{a}, \boldsymbol{a} \rangle = 1 - \langle \boldsymbol{a}, \boldsymbol{b} \rangle^2,$$

a podobne, nakoľko $\langle \boldsymbol{u}, \boldsymbol{u} \rangle \geq 0$,

$$0 > \langle \boldsymbol{u}', \boldsymbol{u}' \rangle = \langle \boldsymbol{u}, \boldsymbol{u} \rangle - 2\langle \boldsymbol{a}, \boldsymbol{u} \rangle^2 + \langle \boldsymbol{a}, \boldsymbol{u} \rangle^2 \langle \boldsymbol{a}, \boldsymbol{a} \rangle = \langle \boldsymbol{u}, \boldsymbol{u} \rangle - \langle \boldsymbol{a}, \boldsymbol{u} \rangle^2 \geq -\langle \boldsymbol{a}, \boldsymbol{u} \rangle^2$$

Keďže $[\boldsymbol{b}', \boldsymbol{u}'] \subseteq [\boldsymbol{a}]^\perp$, je to záporne definitný podpriestor, a podľa bodov (a), (b) tvrdenia 16.1.4 platí

$$\langle \boldsymbol{b}', \boldsymbol{u}' \rangle^2 \leq \langle \boldsymbol{b}', \boldsymbol{b}' \rangle \langle \boldsymbol{u}', \boldsymbol{u}' \rangle.$$

Z toho dostávame odhad

$$|\langle \boldsymbol{b}', \boldsymbol{u}' \rangle| \leq \sqrt{|\langle \boldsymbol{b}', \boldsymbol{b}' \rangle \langle \boldsymbol{u}', \boldsymbol{u}' \rangle|} \leq \langle \boldsymbol{a}, \boldsymbol{u} \rangle \sqrt{\langle \boldsymbol{a}, \boldsymbol{b} \rangle^2 - 1},$$

z ktorého konečne vyplýva

$$\begin{aligned} \langle \boldsymbol{b}, \boldsymbol{u} \rangle &= \langle \boldsymbol{a}, \boldsymbol{b} \rangle \langle \boldsymbol{a}, \boldsymbol{u} \rangle + \langle \boldsymbol{b}', \boldsymbol{u}' \rangle \geq \langle \boldsymbol{a}, \boldsymbol{b} \rangle \langle \boldsymbol{a}, \boldsymbol{u} \rangle - |\langle \boldsymbol{b}', \boldsymbol{u}' \rangle| \\ &\geq \langle \boldsymbol{a}, \boldsymbol{u} \rangle \Big(\langle \boldsymbol{a}, \boldsymbol{b} \rangle - \sqrt{\langle \boldsymbol{a}, \boldsymbol{b} \rangle^2 - 1} \Big) > 0. \end{aligned}$$

Ak teraz definujeme *svetelný kužeľ budúcnosti*

$$\mathrm{LC}^{\boldsymbol{a}}(\boldsymbol{p}) = \{\boldsymbol{x} \in V;\ \langle \boldsymbol{a}, \boldsymbol{x} - \boldsymbol{p} \rangle > 0 \ \&\ \langle \boldsymbol{x} - \boldsymbol{p}, \boldsymbol{x} - \boldsymbol{p} \rangle = 0\}$$

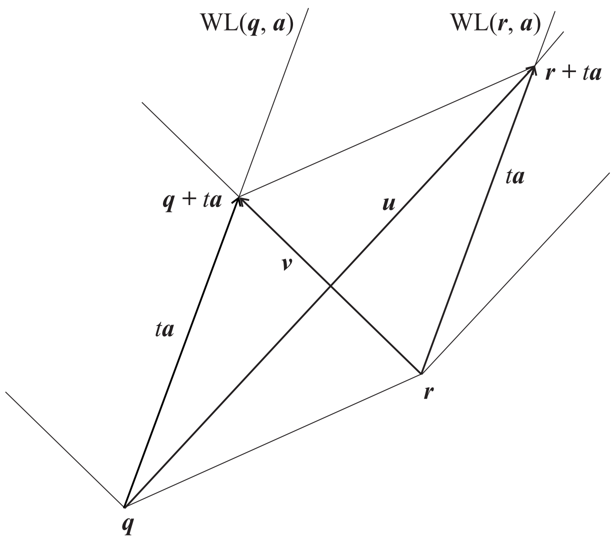
a *svetelný kužeľ minulosti*

$$\mathrm{LC}^{-\boldsymbol{a}}(\boldsymbol{p}) = \{\boldsymbol{x} \in V;\ \langle \boldsymbol{a}, \boldsymbol{x} - \boldsymbol{p} \rangle < 0 \ \&\ \langle \boldsymbol{x} - \boldsymbol{p}, \boldsymbol{x} - \boldsymbol{p} \rangle = 0\}$$

udalosti $\boldsymbol{p} \in V$ vzhľadom na časový vektor $\boldsymbol{a}$, tak z práve dokázaného tvrdenia vyplýva, že platí $\mathrm{LC}^{\boldsymbol{a}}(\boldsymbol{p}) = \mathrm{LC}^{\boldsymbol{b}}(\boldsymbol{p})$ a $\mathrm{LC}^{-\boldsymbol{a}}(\boldsymbol{p}) = \mathrm{LC}^{-\boldsymbol{b}}(\boldsymbol{p})$ pre ľubovoľné súhlasne orientované časové vektory $\boldsymbol{a}$, $\boldsymbol{b}$. Keďže $\langle \boldsymbol{a}, \boldsymbol{a} \rangle > 0$ a vektory $\boldsymbol{x}$, pre ktoré $\langle \boldsymbol{a}, \boldsymbol{x} \rangle = 0$, tvoria nadrovinu $[\boldsymbol{a}]^{\perp}$ oddeľujúcu obe „polovice“ svetelného kužeľa $\mathrm{LC}(\mathbf{0})$, súhlasná orientácia časových vektorov $\boldsymbol{a}$, $\boldsymbol{b}$ znamená, že ležia „vovnútri tej istej polovice“ svetelného kužeľa $\mathrm{LC}(\mathbf{0})$; nesúhlasne orientované časové vektory potom ležia „vovnútri opačných polovíc“ $\mathrm{LC}(\mathbf{0})$. Jedna z dvoch možných orientácií času v Minkowského časopriestore V je tak daná voľbou jediného časového vektora $\boldsymbol{a}$, ktorý prehlásime za kladne orientovaný. Pre $V = \mathbb{R}^{(1,n)}$ je týmto vektorom prirodzene časový šíp $\boldsymbol{e}_0$.

16.4 Relativita súčasnosti

Uvažujme inerciálneho pozorovateľa s časovým šípom $\boldsymbol{a}$ a pokúsme sa vyjasniť si, ktoré udalosti $\boldsymbol{q}$, $\boldsymbol{r}$ v Minkowského časopriestore V by mal tento pozorovateľ považovať za súčasné. Prirodzenou požiadavkou je, aby svetelný signál vyslaný zo svetobodu $\boldsymbol{q}$ a preťal svetočiaru $\mathrm{WL}(\boldsymbol{r}, \boldsymbol{a})$ vo svetobode zodpovedajúcom rovnakému času tejto svetočiary počítanému od udalosti $\boldsymbol{r}$, ako je čas počítaný od udalosti $\boldsymbol{q}$, v ktorom svetelný signál vyslaný zo svetobodu $\boldsymbol{r}$ pretne svetočiaru $\mathrm{WL}(\boldsymbol{q}, \boldsymbol{a})$ – pozri obrázok 16.5.



Obr. 16.5. Súčasné udalosti z hľadiska inerciálneho pozorovateľa s časovým šípom $\boldsymbol{a}$

Na základe toho *udalosti* $\boldsymbol{q}, \boldsymbol{r} \in V$ prehlásime za *súčasné z hľadiska inerciálneho pozorovateľa s časovým šípom* $\boldsymbol{a}$, ak $\boldsymbol{q} = \boldsymbol{r}$ alebo existujú svetelné vektory $\boldsymbol{u}$, $\boldsymbol{v}$, *súhlasne orientované* s časovým vektorom $\boldsymbol{a}$, a hodnota časového parametra $t \in \mathbb{R}$ také, že

$$(\boldsymbol{q}+\boldsymbol{u})-\boldsymbol{r}=(\boldsymbol{r}+\boldsymbol{v})-\boldsymbol{q}=t\boldsymbol{a},$$

alebo, čo je to isté,

$$\boldsymbol{u}-\boldsymbol{v}=2(\boldsymbol{r}-\boldsymbol{q}) \qquad \text{a} \qquad \boldsymbol{u}+\boldsymbol{v}=2t\boldsymbol{a}.$$

16.4.1. Tvrdenie. *Udalosti $\boldsymbol{q}$ a $\boldsymbol{r}$ v Minkowského časopriestore sú súčasné z hľadiska inerciálneho pozorovateľa s časovým šípom $\boldsymbol{a}$ práve vtedy, keď platí $\langle \boldsymbol{a}, \boldsymbol{r}-\boldsymbol{q}\rangle = 0$.*

Dôkaz. Keďže pre $\boldsymbol{q}=\boldsymbol{r}$ niet čo dokazovať, predpokladajme, že $\boldsymbol{q}\neq\boldsymbol{r}$.

Nech teda $\boldsymbol{q}$ a $\boldsymbol{r}$ sú súčasné udalosti vzhľadom na časový vektor $\boldsymbol{a}$ a $\boldsymbol{u}$, $\boldsymbol{v}$ a t sú svetelné vektory resp. časový parameter zaručené definíciou súčasnosti. Potom

$$4t\langle \boldsymbol{a}, \boldsymbol{r}-\boldsymbol{q}\rangle = \big\langle 2t\boldsymbol{a}, 2(\boldsymbol{r}-\boldsymbol{q})\big\rangle = \langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}-\boldsymbol{v}\rangle = \langle \boldsymbol{u}, \boldsymbol{u}\rangle - \langle \boldsymbol{v}, \boldsymbol{v}\rangle = 0.$$

Preto musí platiť $t=0$ alebo $\langle \boldsymbol{a}, \boldsymbol{r}-\boldsymbol{q}\rangle = 0$.

Keby nastala prvá možnosť $t=0$, tak by platilo $\boldsymbol{u}-\boldsymbol{v}\neq\boldsymbol{0}$ a $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{0}$, t. j. $\boldsymbol{u}$, $\boldsymbol{v}$ by boli nenulové, navzájom opačné svetelné vektory. Z nich však len jeden je súhlasne orientovaný s časovým vektorom $\boldsymbol{a}$. Preto musí nastať druhá možnosť $\langle \boldsymbol{a}, \boldsymbol{r}-\boldsymbol{q}\rangle = 0$.

Nech naopak platí $\langle \boldsymbol{a}, \boldsymbol{r}-\boldsymbol{q}\rangle = 0$. Položme

$$t=\sqrt{-\langle \boldsymbol{r}-\boldsymbol{q}, \boldsymbol{r}-\boldsymbol{q}\rangle} = \|\boldsymbol{r}-\boldsymbol{q}\|\,,$$
$$\boldsymbol{u}=\boldsymbol{r}-\boldsymbol{q}+t\boldsymbol{a}, \qquad \boldsymbol{v}=\boldsymbol{q}-\boldsymbol{r}+t\boldsymbol{a}.$$

Priamym výpočtom možno ľahko overiť, že $\boldsymbol{u}$, $\boldsymbol{v}$ a t spĺňajú všetky požadované podmienky.

Zrejme definíciu súčasnosti možno ekvivalentne formulovať pre ľubovoľný časový vektor $\boldsymbol{a}$, nielen časový šíp, a práve dokázané tvrdenie pritom zostane v platnosti.

Množina všetkých udalostí súčasných s udalosťou $\boldsymbol{q}$ z hľadiska inerciálneho pozorovateľa s časovým vektorom $\boldsymbol{a}$ tak tvorí *afinný podpriestor* $\boldsymbol{q}+[\boldsymbol{a}]^{\perp}$ v Minkowského časopriestore V, ktorý nazývame *okamžitým fyzikálnym priestorom* inerciálneho pozorovateľa nachádzajúceho sa v svetobode $\boldsymbol{q}$ svojej svetočiary $\mathrm{WL}(\boldsymbol{p}, \boldsymbol{a})$; každý jednotlivý z takýchto afinných podpriestorov nazývame *okamžitým fyzikálnym priestorom svetočiary* $\mathrm{WL}(\boldsymbol{p}, \boldsymbol{a})$.

Keďže $[\boldsymbol{a}]$ je zrejme maximálny kladne definitný lineárny podpriestor vo V, jeho ortokomplement $[\boldsymbol{a}]^{\perp}$ je podľa tvrdenia 16.1.3 záporne definitný a podľa tvrdenia 16.1.2 platí $V = [\boldsymbol{a}] \oplus [\boldsymbol{a}]^\perp$. Na $[\boldsymbol{a}]^\perp$ sa budeme dívať ako na euklidovský priestor vybavený skalárnym súčinom $-\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ a normou $\|\boldsymbol{x}\| = \sqrt{-\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$. Uvedomme si, že okamžité fyzikálne priestory $\boldsymbol{p} + [\boldsymbol{a}]^\perp$ nášho inerciálneho pozorovateľa sú vlastne tvorené tým istým euklidovským priestorom $[\boldsymbol{a}]^\perp$ „unášaným tokom jeho času“ pozdĺž jeho svetočiary a dohromady vytvárajú celý Minkowského časopriestor $V = \boldsymbol{p} + [\boldsymbol{a}] + [\boldsymbol{a}]^\perp$.

Všetky udalosti v okamžitom fyzikálnom priestore $\boldsymbol{p} + [\boldsymbol{a}]^\perp$ sa z hľadiska príslušného inerciálneho pozorovateľa odohrávajú súčasne. On sám však nemá ako odlíšiť svoj stav od kľudu, teda preňho je jeho okamžitý fyzikálny priestor stále ten istý a splýva so zameraním $[\boldsymbol{a}]^\perp$ jeho okamžitého fyzikálneho priestoru. Preto $[\boldsymbol{a}]^\perp$ predstavuje *subjektívny fyzikálny priestor* inerciálneho pozorovateľa so svetočiarou $\mathrm{WL}(\boldsymbol{p}, \boldsymbol{a})$. Pre inerciálneho pozorovateľa s iným časovým vektorom $\boldsymbol{b} \notin [\boldsymbol{a}]$ však platí $[\boldsymbol{a}]^\perp \neq [\boldsymbol{b}]^\perp$, čiže udalosti súčasné pre jedného z nich sa tak nemusia javiť druhému. Tento jav, označovaný ako *relativita súčasnosti*, je vari zo všetkých dôsledkov teórie relativity najťažšie uviesť do súladu s našou každodennou skúsenosťou a „zdravým rozumom“.

Pre časový vektor $\boldsymbol{a} \neq \boldsymbol{0}$ v $\mathbb{R}^{(1,1)}$ je $[\boldsymbol{a}]^\perp$ priamka súmerne združená s priamkou $[\boldsymbol{a}]$ podľa osi $x_0 = x_1$ (alebo, čo je to isté, podľa osi $x_0 = -x_1$). Teda okrem prípadu, keď $\boldsymbol{a}$ leží v smere osi x_0, priamky $[\boldsymbol{a}]$, $[\boldsymbol{a}]^\perp$ nie sú na seba euklidovsky kolmé. (Pozri obrázok 16.5.) Ako cvičenie si rozmyslite, ako je rovina $[\boldsymbol{a}]^\perp$ „súmerne združená“ s priamkou $[\boldsymbol{a}]$ podľa svetelného kužeľa v $\mathbb{R}^{(1,2)}$.

16.5 Inerciálne bázy a vzťažné sústavy

Nech $\boldsymbol{a}_0$ je ľubovoľný časový šíp a $(\boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ je nejaká ortonormálna báza podpriestoru $[\boldsymbol{a}_0]^\perp$. Potom $\boldsymbol{\alpha} = (\boldsymbol{a}_0, \boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ je zrejme ortonormálna báza pseudoeuklidovského priestoru V s Gramovou maticou $\boldsymbol{G}(\boldsymbol{\alpha}) = \mathrm{diag}(1, -\boldsymbol{I}_n)$, nazývanou tiež *Minkowského symbol* alebo *Minkowského metrický tenzor*. Takúto bázu nazývame *inerciálnou bázou* inerciálneho pozorovateľa s časovým šípom $\boldsymbol{a}_0$ a k nej prislúchajúcu sústavu súradníc nazývame *inerciálnou súradnou sústavou* prípadne *inerciálnou vzťažnou sústavou*. Všimnite si, že – na rozdiel od časového šípu $\boldsymbol{a}_0$ – priestorové vektory $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ uvedenej bázy nie sú určené jednoznačne, teda vo všeobecnosti existuje mnoho rôznych inerciálnych báz spojených s daným inerciálnym pozorovateľom.

Ak $\boldsymbol{\alpha}$ je inerciálna báza vo V, tak pre udalosti $\boldsymbol{x}, \boldsymbol{y} \in V$ so súradnicami $(\boldsymbol{x})_{\boldsymbol{\alpha}} = (x_0, x_1, \dots, x_n)^\mathsf{T}$, $(\boldsymbol{y})_{\boldsymbol{\alpha}} = (y_0, y_1, \dots, y_n)^\mathsf{T}$ platí

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_0 y_0 - x_1 y_1 - \ldots - x_n y_n,$$

čiže pseudoskalárny súčin vo V nadobúda tvar štandardného pseudoskalárneho súčinu v $\mathbb{R}^{(1,n)}$.

Fyzikálne si pod šípom času $\boldsymbol{a}_0$ inerciálnej bázy $(\boldsymbol{a}_0, \boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ treba predstaviť hodinky a pod priestorovými vektormi $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ sústavu n očíslovaných navzájom kolmých kovových tyčí jednotkovej dĺžky pevne zvarených v jednom spoločnom koncovom bode, ktoré slúžia na fixovanie jednotlivých súradných osí a meranie vzdialeností v ich smeroch (v „našom" časopriestore, samozrejme, $n = 3$). Zaviesť inerciálnu vzťažnú sústavu v Minkowského časopriestore znamená vybaviť každého jedného trpaslíka podriadeného príslušnému dispečerovi hodinkami a sústavou takýchto tyčí. Navyše všetky hodinky jednotlivých trpaslíkov musia byť navzájom synchronizované a sústavy ich tyčí paralelizované a zhodne orientované.

Nasledujúce zrejmé tvrdenie dodáva ďalšie oprávnenie spôsobu, akým sme definovali súčasnosť, okamžité fyzikálne priestory a subjektívny priestor inerciálneho pozorovateľa.

16.5.1. Tvrdenie. *Nech $\boldsymbol{\alpha} = (\boldsymbol{a}_0, \boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ je inerciálna báza pozorovateľa s časovým šípom $\boldsymbol{a}_0$. Potom udalosti $\boldsymbol{x}, \boldsymbol{y} \in V$ so súradnicami $(\boldsymbol{x})_{\boldsymbol{\alpha}} = (x_0, x_1, \dots, x_n)^{\mathsf{T}}$, $(\boldsymbol{y})_{\boldsymbol{\alpha}} = (y_0, y_1, \dots, y_n)^{\mathsf{T}}$ sú súčasné z hľadiska tohto pozorovateľa práve vtedy, keď $x_0 = y_0$, t. j. $\boldsymbol{x}$ a $\boldsymbol{y}$ sú súčasné udalosti vzhľadom na vzťažnú sústavu $\boldsymbol{\alpha}$.*

16.6 Paradox dvojčiat

Náš výklad začneme malým doplnkom k obrátenej Cauchyho-Schwartzovej nerovnosti z tvrdenia 16.1.4.

16.6.1. Tvrdenie. *Nech V je Minkovského časopriestor a aspoň jeden z vektorov $\boldsymbol{u}, \boldsymbol{v} \in V$ je časový. Potom*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle^2 \geq \langle \boldsymbol{u}, \boldsymbol{u} \rangle \langle \boldsymbol{v}, \boldsymbol{v} \rangle,$$

pričom rovnosť nastane práve vtedy, keď vektory $\boldsymbol{u}$, $\boldsymbol{v}$ sú lineárne závislé.

Dôkaz, Nech napr. $\boldsymbol{u} \neq \boldsymbol{0}$ je časový vektor. Potom $[\boldsymbol{u}]$ je maximálny kladne definitný lineárny podpriestor vo V, teda $[\boldsymbol{u}]^{\perp}$ je záporne definitný podpriestor podľa tvrdenia 16.1.3 a $V = [\boldsymbol{u}] \oplus [\boldsymbol{u}]^{\perp}$ podľa tvrdenia 16.1.2. Preto $\boldsymbol{v} = a\boldsymbol{u} + \boldsymbol{z}$ pre jednoznačne určený skalár a a vektor $\boldsymbol{z} \in [\boldsymbol{u}]^{\perp}$. Z úvah o ortogonalizácii, prípadne priamym výpočtom dostaneme

$$|\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{v})| = |\boldsymbol{G}(\boldsymbol{u}, \boldsymbol{z})| = \langle \boldsymbol{u}, \boldsymbol{u} \rangle \langle \boldsymbol{z}, \boldsymbol{z} \rangle.$$

Z toho vyplýva, že podpriestor $[\boldsymbol{u}, \boldsymbol{v}] = [\boldsymbol{u}, \boldsymbol{z}]$ je singulárny práve vtedy, keď $\boldsymbol{z} = \boldsymbol{0}$, t. j. práve vtedy, keď $\boldsymbol{u}$, $\boldsymbol{v}$ sú lineárne závislé. Preto ak $\boldsymbol{u}$, $\boldsymbol{v}$ sú lineárne nezávislé, tak podpriestor $[\boldsymbol{u}, \boldsymbol{v}]$ je indefinitný, a požadovaný záver vyplýva z tvrdenia 16.1.4 (c).

Dôsledkom práve dokázaného tvrdenia je nasledujúca „obrátená trojuholníková nerovnosť“.

16.6.2. Dôsledok. *Nech $\boldsymbol{u}$, $\boldsymbol{v}$ sú súhlasne orientované časové vektory v Minkovského časopriestore V. Potom aj $\boldsymbol{u}+\boldsymbol{v}$ je časový vektor a platí*

$$\|\boldsymbol{u}+\boldsymbol{v}\| \geq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|,$$

pričom rovnosť nastane práve vtedy, keď $\boldsymbol{u}$, $\boldsymbol{v}$ sú lineárne závislé.

Dôkaz. Keďže $\langle \boldsymbol{u}, \boldsymbol{v}\rangle > 0$, priamym výpočtom dostávame

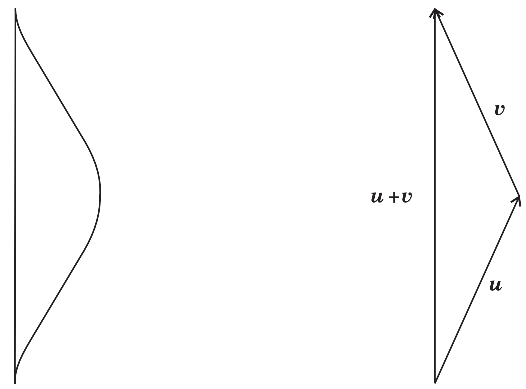
$$\begin{aligned}\langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}\rangle &= \langle \boldsymbol{u}, \boldsymbol{u}\rangle + 2\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \langle \boldsymbol{v}, \boldsymbol{v}\rangle \\ &\geq \|\boldsymbol{u}\|^2 + 2\,\|\boldsymbol{u}\|\,\|\boldsymbol{v}\| + \|\boldsymbol{v}\|^2 = \big(\|\boldsymbol{u}\| + \|\boldsymbol{v}\|\big)^2,\end{aligned}$$

pričom rovnosť zrejme nastane práve vtedy, keď $\langle \boldsymbol{u}, \boldsymbol{v}\rangle^2 = \langle \boldsymbol{u}, \boldsymbol{u}\rangle\langle \boldsymbol{v}, \boldsymbol{v}\rangle$.

Predstavme si teraz v Minkowského časopriestore dvoch trpaslíkov-dvojčatá. Nazvime ich trebárs Kýblik a Spachtoš. Spachtoš je inerciálny a celý čas nášho rozprávania prespí doma. Kýblik sa zatiaľ vyberie na vesmírny výlet raketou. Nejaký čas sa vzďaľuje rovnomerne zrýchleným pohybom, po dosiahnutí istej dosť veľkej rýchlosti sa dlho pohybuje rovnomerne priamočiaro, potom začne brzdiť a po dosiahnutí nulovej rýchlosti obráti svoju vesmírnu loď a začne sa vracať domov na Zem – najprv rovnomerne zrýchleným pohybom naberie istú veľkú rýchlosť, ktorou potom dlho letí rovnomerne priamočiaro, a keď sa priblíži k Zemi, začne brzdiť, až napokon pristane doma na priedomí, kde ho už čaká Spachtoš, ktorý sa práve zobudil a vyšiel von nadýchať sa čerstvého vzduchu. Ukážeme si, že Kýblik je po návrate mladší než Spachtoš, čiže z jeho hľadiska uplynul kratší čas než z hľadiska jeho spiaceho brata.

Príslušné úseky svetočiar oboch bratov sú znázornené na obrázku vľavo. Kým Spachtošov úsek je časťou inerciálnej svetočiary, Kýblikov úsek je neinerciálny – parabolicky zakrivené úseky zodpovedajú zrýchľovaniu resp. brzdeniu rakety, priame letu stálou rýchlosťou. Ak zrýchľovanie a brzdenie trvá v porovnaní s rovnomerným priamočiarym letom zanedbateľne krátko, príslušný úsek Kýblikovej svetočiary možno pre naše účely dostatočne presne aproximovať (fyzikálne neuskutočniteľnou) lomenou svetočiarou na obrázku vpravo. Jej úseky zodpovedajú vektorom $\boldsymbol{u}$, $\boldsymbol{v}$. Spachtošov úsek potom zodpovedá vektoru $\boldsymbol{u}+\boldsymbol{v}$. Zo Spachtošovho pohľadu uplynie čas $\|\boldsymbol{u}+\boldsymbol{v}\|$, kým z Kýblikovho $\|\boldsymbol{u}\| + \|\boldsymbol{v}\|$. Keďže, ako (u)vidíme, $\boldsymbol{u}$, $\boldsymbol{v}$ sú súhlasne orientované časové vektory, podľa práve dokázaného dôsledku platí $\|\boldsymbol{u}+\boldsymbol{v}\| \geq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$.

Celú situáciu možno popísať v $\mathbb{R}^{(1,1)}$. Ak si označíme v veľkosť rýchlosti Kýblikovej rakety v rovnomerných priamočiarych úsekoch, a t čas Spachtošovho spánku, máme $\boldsymbol{u} = (t/2, vt/2)$, $\boldsymbol{v} = (t/2, -vt/2)$, $\boldsymbol{u}+\boldsymbol{v} = (t, 0)$ a



Obr. 16.6. Svetočiaty spiaceho a cestujúceho dvojčaťa

$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = t^2(1 + v^2)/4 > 0$, čiže $\boldsymbol{u}$, $\boldsymbol{v}$ sú súhlasne orientované. Jednoduchým výpočtom možno dostať presnejší odhad

$$\|\boldsymbol{u}\| + \|\boldsymbol{v}\| = t\sqrt{1 - v^2} < t = \|\boldsymbol{u} + \boldsymbol{v}\|.$$

Pomer vlastných časov oboch bratov teda závisí na rýchlosti v.

Názvom *paradox dvojčiat* sa zvykne označovať zdanlivý rozpor, ktorý vzniká, ak sa na uvedenú situáciu pokúšame neuvážene aplikovať relativistický princíp ekvivalencie ľubovoľných inerciálnych sústav. Ak totiž zabudneme, že Kýblikova svetočiara je neinerciálna, a začneme celú situáciu posudzovať z jeho hľadiska, vyjde nám, že by nakoniec mal byť mladší vzhľadom na Kýblika sa pohybovavší Spachtoš. K tejto otázke sa ešte vrátime v paragrafe 16.8, venovanom dilatácii času.

16.7 Relatívna rýchlosť dvoch inerciálnych pozorovateľov

Uvažujme dvoch inerciálnych pozorovateľov so svetočiarami $\mathrm{WL}(\boldsymbol{p}, \boldsymbol{a})$, $\mathrm{WL}(\boldsymbol{q}, \boldsymbol{b})$ v Minkowského časopriestore V, a kvôli jednoduchosti predpokladajme, že vektory $\boldsymbol{a}$, $\boldsymbol{b}$ sú ich šípy času, t. j. $\langle \boldsymbol{a}, \boldsymbol{a} \rangle = \langle \boldsymbol{b}, \boldsymbol{b} \rangle = 1$.

Svetočiara $\mathrm{WL}(\boldsymbol{q}, \boldsymbol{b})$ druhého inerciálneho pozorovateľa pretína okamžitý fyzikálny priestor $\boldsymbol{p} + t\boldsymbol{a} + [\boldsymbol{a}]^\perp$ prvého pozorovateľa vo svetobode $\boldsymbol{q} + t'\boldsymbol{b}$, kde t' nájdeme z podmienky $(\boldsymbol{q} + t'\boldsymbol{b}) - (\boldsymbol{p} + t\boldsymbol{a}) \in [\boldsymbol{a}]^\perp$, čiže

$$0 = \langle \boldsymbol{a}, \boldsymbol{q} + t'\boldsymbol{b} - \boldsymbol{p} - t\boldsymbol{a} \rangle = \langle \boldsymbol{a}, \boldsymbol{b} \rangle t' + \langle \boldsymbol{a}, \boldsymbol{q} - \boldsymbol{p} \rangle - \langle \boldsymbol{a}, \boldsymbol{a} \rangle t = \langle \boldsymbol{a}, \boldsymbol{b} \rangle t' + \langle \boldsymbol{a}, \boldsymbol{q} - \boldsymbol{p} \rangle - t.$$

Z toho vyplýva

$$t' = \frac{t - \langle \boldsymbol{a}, \boldsymbol{q} - \boldsymbol{p} \rangle}{\langle \boldsymbol{a}, \boldsymbol{b} \rangle}.$$

Z pohľadu prvého inerciálneho pozorovateľa, t. j. v jeho subjektívnom fyzikálnom priestore $[\boldsymbol{a}]^\perp$, tomuto okamihu zodpovedá poloha

$$(\boldsymbol{q} + t'\boldsymbol{b}) - (\boldsymbol{p} + t\boldsymbol{a}) = (\boldsymbol{q} - \boldsymbol{p}) - \frac{\langle \boldsymbol{a}, \boldsymbol{q} - \boldsymbol{p}\rangle}{\langle \boldsymbol{a}, \boldsymbol{b}\rangle}\boldsymbol{b} + \frac{t}{\langle \boldsymbol{a}, \boldsymbol{b}\rangle}\boldsymbol{b} - t\boldsymbol{a}$$

druhého inerciálneho pozorovateľa.

Nejakému časovému intervalu $\Delta t = t_2 - t_1$ prvého inerciálneho pozorovateľa tak zodpovedá časový interval

$$\Delta t' = t_2' - t_1' = \frac{t_2 - t_1}{\langle \boldsymbol{a}, \boldsymbol{b}\rangle} = \frac{\Delta t}{\langle \boldsymbol{a}, \boldsymbol{b}\rangle}$$

druhého z nich. Za ten čas sa poloha druhého pozorovateľa v subjektívnom fyzikálnom priestore $[\boldsymbol{a}]^\perp$ prvého zmení o priestorový vektor $\Delta t\big(\langle \boldsymbol{a}, \boldsymbol{b}\rangle^{-1}\boldsymbol{b} - \boldsymbol{a}\big) \in [\boldsymbol{a}]^\perp$. Priestorový vektor

$$\boldsymbol{v} = \langle \boldsymbol{a}, \boldsymbol{b}\rangle^{-1}\boldsymbol{b} - \boldsymbol{a} \in [\boldsymbol{a}]^\perp$$

potom prirodzene predstavuje rýchlosť, akou sa druhý inerciálny pozorovateľ pohybuje v subjektívnom fyzikálnom priestore prvého. Keďže $\boldsymbol{v}$ nezávisí od času t, pohyb druhého inerciálneho pozorovateľa sa prvému skutočne javí ako rovnomerný priamočiary. Pre veľkosť tejto rýchlosti platí

$$v = \|\boldsymbol{v}\| = \sqrt{-\langle \boldsymbol{v}, \boldsymbol{v}\rangle} = \sqrt{-\langle \boldsymbol{a}, \boldsymbol{b}\rangle^{-2}\langle \boldsymbol{b}, \boldsymbol{b}\rangle + 2 - \langle \boldsymbol{a}, \boldsymbol{a}\rangle} = \sqrt{1 - \langle \boldsymbol{a}, \boldsymbol{b}\rangle^{-2}}.$$

Teda veľkosť v rýchlosti, ktorou sa z pohľadu prvého inerciálneho pozorovateľa pohybuje druhý z nich, možno vyjadriť pomocou pseudoskalárneho súčinu $\langle \boldsymbol{a}, \boldsymbol{b}\rangle$ ich časových šípov $\boldsymbol{a}$, $\boldsymbol{b}$. Taktiež naopak, pseudoskalárny súčin $\langle \boldsymbol{a}, \boldsymbol{b}\rangle$ možno až na znamienko vyjadriť pomocou veľkosti relatívnej rýchlosti v:

$$|\langle \boldsymbol{a}, \boldsymbol{b}\rangle| = \frac{1}{\sqrt{1 - v^2}},$$

v čom čitateľ asi spozná známy Lorentzov koeficient, hoci vo fyzike ho častejšie zapisujeme v tvare $\frac{1}{\sqrt{1-(v/c)^2}}$, kde c je rýchlosť svetla. Zrejme postulát medznej hodnoty rýchlosti svetla je ekvivalentný s požiadavkou reálnosti a konečnosti tohto výrazu. Ak navyše prijmeme prirodzený predpoklad, že $\boldsymbol{a}$, $\boldsymbol{b}$ sú súhlasne orientované, dostaneme

$$\langle \boldsymbol{a}, \boldsymbol{b}\rangle = \frac{1}{\sqrt{1 - v^2}}.$$

Ešte podotknime, že ak by sme si na začiatku nezjednodušili život podmienkou $\langle \boldsymbol{a}, \boldsymbol{a} \rangle = \langle \boldsymbol{b}, \boldsymbol{b} \rangle = 1$, čiže za $\boldsymbol{a}$, $\boldsymbol{b}$ by sme si vzali ľubovoľné súhlasne orientované časové vektory, poslednú rovnosť by sme dostali v tvare

$$\frac{\langle \boldsymbol{a}, \boldsymbol{b} \rangle}{\|\boldsymbol{a}\| \, \|\boldsymbol{b}\|} = \frac{1}{\sqrt{1-v^2}},$$

čo na základe analógie s euklidovskými priestormi navodzuje myšlienku, že Lorentzov koeficient predstavuje „kosínus" akéhosi „pseudouhla" vektorov $\boldsymbol{a}$, $\boldsymbol{b}$. Keďže je však uvedený výraz vždy ≥ 1, o obyčajný kosínus uhla ísť nemôže. Zvyšok paragrafu je venovaný upresneniu týchto úvah.

Z *Eulerových vzťahov*

$$\mathrm{e}^{\mathrm{i}\alpha} = \cos\alpha + \mathrm{i}\sin\alpha, \qquad \mathrm{e}^{-\mathrm{i}\alpha} = \cos\alpha - \mathrm{i}\sin\alpha,$$

ktoré tu nebudeme odvodzovať, vyplývajú nasledujúce vyjadrenia goniometrických funkcií pomocou exponenciály imaginárneho argumentu

$$\cos\alpha = \frac{\mathrm{e}^{\mathrm{i}\alpha} + \mathrm{e}^{-\mathrm{i}\alpha}}{2}, \qquad \sin\alpha = \frac{\mathrm{e}^{\mathrm{i}\alpha} - \mathrm{e}^{-\mathrm{i}\alpha}}{2\mathrm{i}}$$

pre ľubovoľné $\alpha \in \mathbb{R}$. Funkcie *hyperbolický kosínus* a *hyperbolický sínus* sú definované reálnou analógiou uvedených rovností

$$\cosh\theta = \frac{\mathrm{e}^{\theta} + \mathrm{e}^{-\theta}}{2}, \qquad \sinh\theta = \frac{\mathrm{e}^{\theta} - \mathrm{e}^{-\theta}}{2}$$

pre ľubovoľné $\theta \in \mathbb{R}$.

Všetko, čo potrebujeme v tejto chvíli vedieť, je jednotkový vzťah

$$\cosh^2\theta - \sinh^2\theta = 1$$

(overte si samostatne jednoduchým výpočtom), z ktorého vyplýva, že všetky dvojice $(\cosh\theta, \sinh\theta)$ ležia na jednej vetve rovnoosej hyperboly $x^2 - y^2 = 1$, $(x \geq 1)$, a že každý bod (x, y) tejto vetvy má uvedený tvar. Naozaj, stačí položiť $\theta = \ln(x + y)$. Druhá vetva $(x \leq -1)$ tejto hyperboly je tvorená dvojicami $(-\cosh\theta, \sinh\theta)$, pre $\theta \in \mathbb{R}$.

Pre súhlasne orientované časové vektory $\boldsymbol{a}$, $\boldsymbol{b}$ potom existuje jednoznačne určené reálne číslo θ, nazývané tiež *hyperbolický uhol* vektorov $\boldsymbol{a}$, $\boldsymbol{b}$, také, že

$$\cosh\theta = \frac{\langle \boldsymbol{a}, \boldsymbol{b} \rangle}{\|\boldsymbol{a}\| \, \|\boldsymbol{b}\|} = \frac{1}{\sqrt{1-v^2}}, \qquad \sinh\theta = \frac{\sqrt{-|\boldsymbol{G}(\boldsymbol{a}, \boldsymbol{b})|}}{\|\boldsymbol{a}\| \, \|\boldsymbol{b}\|} = \frac{v}{\sqrt{1-v^2}}.$$

Explicitné vyjadrenie pre θ je

$$\theta = \ln \frac{\langle \boldsymbol{a}, \boldsymbol{b} \rangle + \sqrt{-|\boldsymbol{G}(\boldsymbol{a}, \boldsymbol{b})|}}{\|\boldsymbol{a}\| \, \|\boldsymbol{b}\|} = \ln \sqrt{\frac{1+v}{1-v}}.$$

16.8 Relativistická dilatácia času

Ako sme odvodili v predošlom paragrafe, pre časový úsek $\varDelta t'$ druhého inerciálneho pozorovateľa, ktorý zodpovedá časovému úseku $\varDelta t$ s ním súhlasne orientovaného prvého pozorovateľa, platí

$$\varDelta t' = \frac{\varDelta t}{\langle \boldsymbol{a}, \boldsymbol{b} \rangle} = \varDelta t \sqrt{1 - v^2},$$

teda $\varDelta t' \leq \varDelta t$, pričom rovnosť nastane práve vtedy, keď $v = 0$, čo je ekvivalentné s rovnosťou $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = 1$, a na základe tvrdenia 16.1.4 s lineárnou závislosťou časových šípov $\boldsymbol{a}$, $\boldsymbol{b}$, čo v tomto prípade znamená $\boldsymbol{a} = \boldsymbol{b}$. Z hľadiska prvého pozorovateľa, ktorý sám seba považuje za nehybného, tak medzi dvoma okamihmi t_1 a t_2 uplynie dlhší čas, než z hľadiska druhého pozorovateľa medzi okamihmi t_1', t_2', v ktorých sa tento pozorovateľ nachádza v okamžitých fyzikálnych priestoroch $\boldsymbol{p} + t_1\boldsymbol{a} + [\boldsymbol{a}]^\perp$, resp. $\boldsymbol{p} + t_2\boldsymbol{a} + [\boldsymbol{a}]^\perp$ prvého. Tento efekt sa nazýva *relativistické spomalenie*, prípadne *relativistická dilatácia času.*

V uvedenej rovnosti sa však skrýva zdanlivý paradox, niekedy nazývaný *paradox času.* Z matematických dôvodov symetrie ako aj z fyzikálnych dôvodov rovnocennosti inerciálnych sústav by malo takisto platiť

$$\varDelta t = \frac{\varDelta t'}{\langle \boldsymbol{a}, \boldsymbol{b} \rangle} = \varDelta t' \sqrt{1 - v^2},$$

teda $\varDelta t \leq \varDelta t'$. Potom nevyhnutne $\varDelta t = \varDelta t'$, $v = 0$ a $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = 1$, t. j. $\boldsymbol{a} = \boldsymbol{b}$ pre ľubovoľné súhlasne orientované časové šípy $\boldsymbol{a}$, $\boldsymbol{b}$, čo je zrejmý nezmysel. Teda niekde v našich úvahách je asi chyba. Odhaliť ju nie je až také ťažké. Časový okamih prvého inerciálneho pozorovateľa, zodpovedajúci časovému okamihu t' druhého inerciálneho pozorovateľa z pohľadu druhého pozorovateľa, nie je pôvodný okamih t, ale okamih

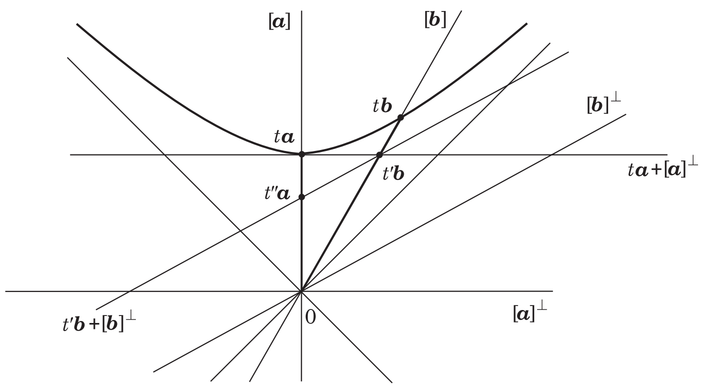
$$t'' = \frac{t' - \langle \boldsymbol{b}, \boldsymbol{p} - \boldsymbol{q} \rangle}{\langle \boldsymbol{a}, \boldsymbol{b} \rangle} = \frac{t - \langle \boldsymbol{a}, \boldsymbol{q} - \boldsymbol{p} \rangle + \langle \boldsymbol{a}, \boldsymbol{b} \rangle \langle \boldsymbol{b}, \boldsymbol{p} - \boldsymbol{q} \rangle}{\langle \boldsymbol{a}, \boldsymbol{b} \rangle^2}.$$

Teda časovému intervalu $\varDelta t' = t_2' - t_1'$ druhého inerciálneho pozorovateľa zodpovedá z jeho pohľadu časový interval

$$\varDelta t'' = \frac{\varDelta t'}{\langle \boldsymbol{a}, \boldsymbol{b} \rangle} = \frac{\varDelta t}{\langle \boldsymbol{a}, \boldsymbol{b} \rangle^2},$$

a nie $\varDelta t$, prvého pozorovateľa. Situácia pre špeciálny prípad $\boldsymbol{p} = \boldsymbol{q} = \boldsymbol{0}$, $t_1 = 0$, $t_2 = t = \varDelta t$ je znázornená na obrázku v $\mathbb{R}^{(1,1)}$.

Na obrázku nám asi udrie do očú, že dĺžka vektora $t'\boldsymbol{b}$ je *väčšia* ako dĺžka vektora $t\boldsymbol{a}$, aj keď $\|t'\boldsymbol{b}\| = t' = t\sqrt{1 - v^2} < t = \|t\boldsymbol{a}\|$. Nezaškodí preto znova



Obr. 16.7. Relativistická dilatácia času

pripomenúť, že vzdialenosti, ktoré nám vnucuje obrázok, sú euklidovské, netreba ich preto brať vážne – geometria Minkowského časopriestoru je totiž neeuklidovská. Z hľadiska tejto geometrie sú rovnako dlhé napr. vektory $t\boldsymbol{a}$ a $t\boldsymbol{b}$ (ich „konce“ ležia na tej istej hyperbole).

Experimentálny dôkaz dilatácie času poskytujú μ-*mezóny*, zvané tiež *mióny* (niečo ako „ťažké elektróny“) – elementárne častice s veľmi krátkou priemernou dobou života (asi $2,2 \cdot 10^{-6}$s). Mióny vznikajú vplyvom primárneho kozmického žiarenia v horných vrstvách atmosféry (t. j. vo výškach 10 a viac km) a prilietajú veľkými rýchlosťami (až 0, 998 rýchlosti svetla) na zemský povrch. Za čas $2,2 \cdot 10^{-6}$s by však ani rýchlosťou svetla $c \approx 300\,000\,\mathrm{km\,s^{-1}}$ nemali preletieť viac než 660 m. Z hľadiska pozemského pozorovateľa však času $\Delta t'$ v sústave letiaceho miónu zodpovedá čas $\Delta t = \Delta t'/\sqrt{1-v^2}$, kde v je jeho rýchlosť v pomere k rýchlosti svetla. Namerané hodnoty priemernej doby života miónov pri rôznych rýchlostiach (či už v kozmických lúčoch alebo v pozemských urýchľovačoch) sa so značnou presnosťou zhodujú s uvedeným vzťahom. Napríklad vlastnému času $2,2 \cdot 10^{-6}$s v sústave miónu letiaceho rýchlosťou $0,998\,c$, zodpovedá v sústave pozemského pozorovateľa čas asi $34,8 \cdot 10^{-6}$s, za ktorý mión preletí dráhu približne 10, 4 km.

Všimnite si, že vo formule pre dilatáciu času vystupuje rovnaký koeficient $\sqrt{1-v^2}$ ako v *približnom* kvantitatívnom odhade z paradoxu dvojčiat. Navyše dilatácia času je – popri neinerciálnosti Kýblikovej svetočiary – naozaj spoluzodpovedná za jeho „pomalšie starnutie“. To sú dôvody, pre ktoré sa oba tieto efekty často pletú. Ešte stále sa možno z času na čas stretnúť s naivnou kritikou špeciálnej teórie relativity, ktorá si berie na mušku práve paradox dvojčiat a používa pri tom už spomínaný argument, akým možno zdanlivo spochybniť dilatáciu času: „Keďže pohyb je relatívny, môžeme rovnako dobre z Kýblikovho hľadiska považovať Spachtoša za pohybujúceho sa a Kýblika za nehybného. Potom by mal viac zostarnúť Kýblik.“ Podobná symetria tu však nemá miesto. Aby sa mohli Kýblik a Spachtoš, opisujúci najprv tú istú svetočiaru, rozdeliť a potom opäť stretnúť, musí sa (aspoň) jeden z nich v istých úsekoch svojej svetočiary „zneinerciálniť“, t. j. pohybovať sa so zrýchlením. Rozdiel medzi inerciálnou Spachtošovou a neinerciálnou Kýblikovou svetočiarou má tak – v protiklade k našim inerciálnym pozorovateľom z tohto paragrafu – absolútny charakter a rozdiel veku sa pri stretnutí oboch bratov prejaví „hmatateľne“.

Tento efekt bol potvrdený aj experimentálne – neexperimentovalo sa však s dvojčatami ale s veľmi presnými hodinami, merajúcimi čas pomocou oscilácií v elektrónovom obale atómov cézia. Štvoro céziových hodín obletelo Zem v dvoch prúdových lietadlách – dvojo v smere od západu na východ, dvojo v smere od východu na západ – a po prílete ich porovnali s referenčnými hodinami, ktoré zostali „doma“ v Námornom observatóriu v USA. Namerané časové rozdiely sa veľmi dobre zhodovali s hodnotami predpovedanými teóriou. Poznamenajme však, že vzhľadom na účinky rotácie a gravitačného poľa Zeme je reálna situácia podstatne zložitejšia než náš umelý príklad s Kýblikom a Spachtošom, a jej matematický popis si vyžaduje i čo-to zo všeobecnej teórie relativity. V skutočnosti v porovnaní s pozemskými hodinami „omladli“ len hodiny, ktoré leteli smerom na východ – o $(59 \pm 10) \cdot 10^{-9}$s; naopak hodiny, ktoré leteli na západ, v porovnaní s pozemskými hodinami dokonca „zostarli“ o $(273 \pm 7) \cdot 10^{-9}$s. Teoretická predpoveď dávala hodnoty $(40 \pm 23) \cdot 10^{-9}$s, resp. $(275 \pm 21) \cdot 10^{-9}$s.

16.9 Lorentzova transformácia

Vráťme sa ešte raz k našim inerciálnym pozorovateľom v Minkowského časopriestore V so súhlasne orientovanými časovými šípmi $\boldsymbol{a}$, $\boldsymbol{b}$. Označme $\boldsymbol{a}_0 = \boldsymbol{a}$, $\boldsymbol{b}_0 = \boldsymbol{b}$ a predpokladajme, že $\boldsymbol{\alpha} = (\boldsymbol{a}_0, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$, $\boldsymbol{\beta} = (\boldsymbol{b}_0, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ sú ich inerciálne bázy. *Lorentzova transformácia* z inerciálnej bázy $\boldsymbol{\beta}$ do inerciálnej bázy $\boldsymbol{\alpha}$ je transformácia súradníc vzhľadom na tieto bázy, t. j. lineárne zobrazenie $\varphi\colon \mathbb{R}^{(1,n)} \to \mathbb{R}^{(1,n)}$ také, že

$$(\boldsymbol{u})_{\boldsymbol{\alpha}} = \varphi\big((\boldsymbol{u})_{\boldsymbol{\beta}}\big)$$

pre každé $\boldsymbol{u} \in V$, alebo, čo je to isté,

$$\boldsymbol{\alpha} \cdot \varphi(\boldsymbol{x}) = \boldsymbol{\beta} \cdot \boldsymbol{x}$$

pre každé $\boldsymbol{x} \in \mathbb{R}^{(1,n)}$ (stačí položiť $\boldsymbol{x} = (\boldsymbol{u})_{\boldsymbol{\beta}}$). Z uvedených vzťahov okamžite vidno, že Lorentzova transformácia φ je danými bázami $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ jednoznačne určená a jej matica vzhľadom na kanonickú ortonormálnu bázu $\boldsymbol{\varepsilon} = (\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ v $\mathbb{R}^{(1,n)}$ je zároveň maticou prechodu z bázy $\boldsymbol{\beta}$ do bázy $\boldsymbol{\alpha}$, t. j.

$$(\varphi)_{\boldsymbol{\varepsilon},\boldsymbol{\varepsilon}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}.$$

16.9.1. Tvrdenie. *Nech φ je Lorentzova transformácia z inerciálnej bázy $\boldsymbol{\beta}$ do inerciálnej bázy $\boldsymbol{\alpha}$ v Minkowského časopriestore V. Potom pre ľubovoľné vektory $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{(1,n)}$ platí*

$$\langle \varphi \boldsymbol{x}, \varphi \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle.$$

Inými slovami, Lorentzova transformácia zachováva štandardný pseudoskalárny súčin.

Dôkaz. Nech $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{(1,n)}$. Vzhľadom na inerciálnosť báz $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ platí

$$\langle \varphi \boldsymbol{x}, \varphi \boldsymbol{y} \rangle = \langle \boldsymbol{\alpha} \cdot \varphi \boldsymbol{x}, \boldsymbol{\alpha} \cdot \varphi \boldsymbol{y} \rangle = \langle \boldsymbol{\beta} \cdot \boldsymbol{x}, \boldsymbol{\beta} \cdot \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle,$$

pričom v dvoch krajných výrazoch ide o štandardný pseudoskalárny súčin v $\mathbb{R}^{(1,n)}$, kým v dvoch vnútorných výrazoch o pseudoskalárny súčin vo V.

Ak φ je Lorentzova transformácia z inerciálnej bázy $\boldsymbol{\beta}$ do inerciálnej bázy $\boldsymbol{\alpha}$, ktoré prislúchajú svetočiaram $\mathrm{WL}(\boldsymbol{b}, \boldsymbol{q})$ resp. $\mathrm{WL}(\boldsymbol{p}, \boldsymbol{a})$, pričom naši pozorovatelia si za počiatky odpočtu svojich súradníc zvolili udalosti $\boldsymbol{q}$ resp. $\boldsymbol{p}$, tak vzťah medzi súradnicami ľubovoľnej udalosti $\boldsymbol{z} \in V$ v takto zvolených inerciálnych súradných systémoch udáva afinná transformácia

$$(\boldsymbol{z} - \boldsymbol{p})_{\boldsymbol{\alpha}} = \varphi(\boldsymbol{z} - \boldsymbol{q})_{\boldsymbol{\beta}},$$

nazývaná tiež *Poincarého transformáciou.*

Preniknúť do štruktúry Lorentzových transformácií možno tak, že popíšeme štruktúru ich matíc. To je vo všeobecnom prípade pomerne náročná úloha. Ukážeme si však, že dané súhlasne orientované časové šípy $\boldsymbol{a} = \boldsymbol{a}_0$, $\boldsymbol{b} = \boldsymbol{b}_0$ možno vždy vhodne doplniť do inerciálnych báz $\boldsymbol{\alpha} = (\boldsymbol{a}_0, \boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$, $\boldsymbol{\beta} = (\boldsymbol{b}_0, \boldsymbol{b}_1, \dots, \boldsymbol{b}_n)$ tak, že vzhľadom na ne má matica Lorentzovej transformácie z $\boldsymbol{\beta}$ do $\boldsymbol{\alpha}$ obzvlášť jednoduchý a prehľadný tvar $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \mathrm{diag}(\boldsymbol{L}_v, \boldsymbol{I}_{n-1})$, kde $\boldsymbol{L}_v \in \mathbb{R}^{2\times 2}$ je matica Lorentzovej transformácie Minkowského časopriestoru $\mathbb{R}^{(1,1)}$, ktorej prvky možno vyjadriť výlučne pomocou veľkosti v relatívnej rýchlosti uvažovaných inerciálnych pozorovateľov. V takom prípade hovoríme o tzv. *špeciálnej Lorentzovej transformácii.* [4]

Ak $\boldsymbol{a} = \boldsymbol{b}$, nájdeme ľubovoľnú ortonormálnu bázu $\boldsymbol{\alpha} = \boldsymbol{\beta}$ časopriestoru V s prvým členom $\boldsymbol{a}_0 = \boldsymbol{b}_0$. Lorentzova transformácia z $\boldsymbol{\beta}$ do $\boldsymbol{\alpha}$ je potom identické zobrazenie s maticou $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \boldsymbol{I}_{n+1}$.

Ak $\boldsymbol{a} \neq \boldsymbol{b}$, tak ide o nezávislé vektory. V dôkaze tvrdenia 16.1.4 sme ukázali, že $[\boldsymbol{a}, \boldsymbol{b}]$ je indefinitný podpriestor vo V. Preto existuje $\boldsymbol{a}_1 \in [\boldsymbol{a}, \boldsymbol{b}]$ také, že $(\boldsymbol{a}_0, \boldsymbol{a}_1)$ je ortonormálna báza podpriestoru $[\boldsymbol{a}, \boldsymbol{b}]$. Potom nevyhnutne $\boldsymbol{a}_1 \in [\boldsymbol{a}]^{\perp}$ je priestorový vektor. Rovnako existuje $\boldsymbol{b}_1 \in [\boldsymbol{a}, \boldsymbol{b}] \cap [\boldsymbol{b}]^{\perp}$ také, že

[4] V anglojazyčnej literatúre sa požíva tiež názov *boost* (pozri paragraf 29.5).

$(\boldsymbol{b}_0, \boldsymbol{b}_1)$ je ortonormálna báza podpriestoru $[\boldsymbol{a}, \boldsymbol{b}]$. Ak si označíme $\boldsymbol{a}_2 = \boldsymbol{b}_2$, $\dots$, $\boldsymbol{a}_n = \boldsymbol{b}_n$ ľubovoľnú ortonormálnu bázu záporne definitného podpriestoru $[\boldsymbol{a}, \boldsymbol{b}]^\perp$, tak $\boldsymbol{\alpha} = (\boldsymbol{a}_0, \boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ a $\boldsymbol{\beta} = (\boldsymbol{b}_0, \boldsymbol{b}_1, \dots, \boldsymbol{b}_n)$ sú inerciálne bázy, ktoré majú posledných $n-1$ členov rovnakých. Ich matica prechodu má tvar $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \operatorname{diag}(\boldsymbol{P}_{\boldsymbol{a},\boldsymbol{b}}, \boldsymbol{I}_{n-1})$, kde $\boldsymbol{P}_{\boldsymbol{a},\boldsymbol{b}} \in \mathbb{R}^{2\times 2}$ označuje maticu prechodu z bázy $[\boldsymbol{b}_0, \boldsymbol{b}_1]$ do bázy $[\boldsymbol{a}_0, \boldsymbol{a}_1]$, ktorú vyjadríme explicitne. Na ten účel stačí poznať vektory $\boldsymbol{a}_1$, $\boldsymbol{b}_1$.

V paragrafe 16.7 sme odvodili tvar vektora $\boldsymbol{v}$ rýchlosti, ktorou sa v subjektívnom fyzikálnom priestore inerciálneho pozorovateľa s časovým šípom $\boldsymbol{a}$ pohybuje inerciálny pozorovateľ s časovým šípom $\boldsymbol{b}$:

$$\boldsymbol{v} = \langle \boldsymbol{a}, \boldsymbol{b} \rangle^{-1} \boldsymbol{b} - \boldsymbol{a}.$$

Keďže $\boldsymbol{v} \in [\boldsymbol{a}, \boldsymbol{b}] \cap [\boldsymbol{a}]^\perp$ a $\dim\big([\boldsymbol{a}, \boldsymbol{b}] \cap [\boldsymbol{a}]^\perp\big) = 1$, stačí položiť

$$\boldsymbol{a}_1 = v^{-1} \boldsymbol{v} = \frac{\langle \boldsymbol{a}, \boldsymbol{b} \rangle^{-1} \boldsymbol{b} - \boldsymbol{a}}{\sqrt{1 - \langle \boldsymbol{a}, \boldsymbol{b} \rangle^{-2}}} = \frac{\boldsymbol{b} - \langle \boldsymbol{a}, \boldsymbol{b} \rangle \boldsymbol{a}}{\sqrt{\langle \boldsymbol{a}, \boldsymbol{b} \rangle^2 - 1}},$$

kde $v = \|\boldsymbol{v}\| = \sqrt{1 - \langle \boldsymbol{a}, \boldsymbol{b} \rangle^{-2}}$. S využitím symetrie úlohy možno písať

$$\boldsymbol{b}_1 = \frac{-\langle \boldsymbol{a}, \boldsymbol{b} \rangle^{-1} \boldsymbol{a} + \boldsymbol{b}}{\sqrt{1 - \langle \boldsymbol{a}, \boldsymbol{b} \rangle^{-2}}} = \frac{-\boldsymbol{a} + \langle \boldsymbol{a}, \boldsymbol{b} \rangle \boldsymbol{b}}{\sqrt{\langle \boldsymbol{a}, \boldsymbol{b} \rangle^2 - 1}}.$$

Z rovnosti pre $\boldsymbol{a}_1$ si vyjadríme

$$\boldsymbol{b}_0 = \boldsymbol{b} = \langle \boldsymbol{a}, \boldsymbol{b} \rangle \boldsymbol{a}_0 + \sqrt{\langle \boldsymbol{a}, \boldsymbol{b} \rangle^2 - 1}\, \boldsymbol{a}_1,$$

čo po dosadení do rovnosti pre $\boldsymbol{b}_1$ a malých úpravách dáva

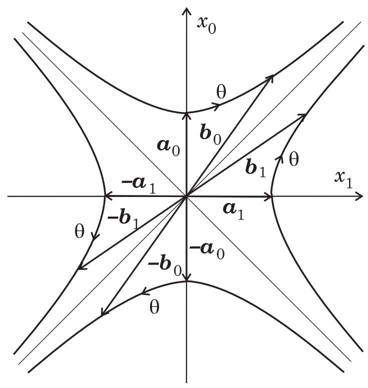
$$\boldsymbol{b}_1 = \sqrt{\langle \boldsymbol{a}, \boldsymbol{b} \rangle^2 - 1}\, \boldsymbol{a}_0 + \langle \boldsymbol{a}, \boldsymbol{b} \rangle \boldsymbol{a}_1.$$

To znamená, že

$$(\boldsymbol{b}_0, \boldsymbol{b}_1) = (\boldsymbol{a}_0, \boldsymbol{a}_1) \cdot \begin{pmatrix} \langle \boldsymbol{a}, \boldsymbol{b} \rangle & \sqrt{\langle \boldsymbol{a}, \boldsymbol{b} \rangle^2 - 1} \\ \sqrt{\langle \boldsymbol{a}, \boldsymbol{b} \rangle^2 - 1} & \langle \boldsymbol{a}, \boldsymbol{b} \rangle \end{pmatrix}.$$

Ak si ešte spomenieme na vzťah medzi pseudoskalárnym súčinom $\langle \boldsymbol{a}, \boldsymbol{b} \rangle$ a veľkosťou rýchlosti v, dostaneme dvojaké vyjadrenie matice prechodu $\boldsymbol{P}_{\boldsymbol{a},\boldsymbol{b}}$ alias Lorentzovej transformácie $\boldsymbol{L}_v$:

$$\boldsymbol{P}_{\boldsymbol{a},\boldsymbol{b}} = \begin{pmatrix} \langle \boldsymbol{a}, \boldsymbol{b} \rangle & \sqrt{\langle \boldsymbol{a}, \boldsymbol{b} \rangle^2 - 1} \\ \sqrt{\langle \boldsymbol{a}, \boldsymbol{b} \rangle^2 - 1} & \langle \boldsymbol{a}, \boldsymbol{b} \rangle \end{pmatrix} = \begin{pmatrix} \dfrac{1}{\sqrt{1 - v^2}} & \dfrac{v}{\sqrt{1 - v^2}} \\ \dfrac{v}{\sqrt{1 - v^2}} & \dfrac{1}{\sqrt{1 - v^2}} \end{pmatrix} = \boldsymbol{L}_v.$$



Obr. 16.8. Lorentzenova transformácia z inerciálnej bázy $(\boldsymbol{b}_0, \boldsymbol{b}_1)$

Po substitúcii $\theta = \ln\sqrt{(1+v)/(1-v)}$ dostávame ešte tretie vyjadrenie v tvare tzv. *hyperbolickej rotácie*

$$\boldsymbol{Rh}_\theta = \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix} = \boldsymbol{P}_{\boldsymbol{a},\boldsymbol{b}} = \boldsymbol{L}_v.$$

Všimnite si, že pre $\boldsymbol{a} = \boldsymbol{b}$, čiže $v = 0$, $\theta = 0$, dostávame $\boldsymbol{P}_{\boldsymbol{a},\boldsymbol{a}} = \boldsymbol{L}_0 = \boldsymbol{Rh}_0 = \boldsymbol{I}_2$, čo je v zhode so skôr prijatým riešením tohto špeciálneho prípadu. Všeobecný prípad je znázornený na obrázku.

Z fyzikálneho hľadiska je najdôležitejšia signatúra $(1,3)$, t. j. $n = 3$, kedy sa Lorentzova transformácia súradníc vzhľadom na inerciálne bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ obvykle uvádza v jednom z nasledujúcich tvarov:

$$\begin{aligned} x_0 = ct &= \frac{x_0' + (v/c)x_1}{\sqrt{1-(v/c)^2}} = \frac{ct' + (v/c)x_1'}{\sqrt{1-(v/c)^2}}, \\ x_1 &= \frac{(v/c)x_0' + x_1'}{\sqrt{1-(v/c)^2}} = \frac{vt' + x_1'}{\sqrt{1-(v/c)^2}}, \\ x_2 &= x_2', \\ x_3 &= x_3', \end{aligned}$$

kde c je rýchlosť svetla, v je relatívna rýchlosť pohybu pozorovateľov s inerciálnymi bázami $\boldsymbol{\alpha}$ a $\boldsymbol{\beta}$ v smere osi x_1 a $(x_0 = ct,\ x_1, x_2, x_3)^\mathsf{T}$, $(x_0' = ct',\ x_1', x_2', x_3')^\mathsf{T}$ sú časopriestorové súradnice ľubovoľnej udalosti vzhľadom na bázy $\boldsymbol{\alpha}$ resp. $\boldsymbol{\beta}$. Ešte raz však pripomíname, že takýto pomerne jednoduchý tvar má Lorentzova transformácia len pri „správnej“ voľbe priestorových vektorov oboch báz.

16.10 Relativistická kontrakcia dĺžky

Pomocou Lorentzovej transformácie možno jednoducho odvodiť ďalší zo známych relativistických efektov.

Nech $\mathrm{WL}(\boldsymbol{p},\boldsymbol{a})$, $\mathrm{WL}(\boldsymbol{q},\boldsymbol{b})$ sú súhlasne orientované svetočiary inerciálnych pozorovateľov a $\boldsymbol{\alpha} = (\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)$, resp. $\boldsymbol{\beta} = (\boldsymbol{b}_0, \boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n)$ sú s nimi spojené inerciálne bázy zvolené tak, ako v predchádzajúcom paragrafe, t. j. Lorentzova transformácia φ z $\boldsymbol{\beta}$ do $\boldsymbol{\alpha}$ má maticu tvaru $\operatorname{diag}(\boldsymbol{L}_v, \boldsymbol{I}_{n-1})$. Predpokladajme, že prvý pozorovateľ registruje nehybnú pevnú tyč dĺžky $l > 0$ v smere vektora $\boldsymbol{a}_1$. Matematicky ide o priestorový vektor $l\boldsymbol{a}_1$ prebiehajúci postupom času okamžité fyzikálne priestory $\boldsymbol{p} + t\boldsymbol{a} + [\boldsymbol{a}]^\perp$; jeho súradnice v okamihu t vzhľadom na bázu $\boldsymbol{\alpha}$ a počiatok $\boldsymbol{p}$ sú $(t, l, 0, \dots, 0)^\mathsf{T}$. Dĺžka tejto tyče sa druhému pozorovateľovi javí ako dĺžka

$$l' = \|\boldsymbol{u}\| = \sqrt{-\langle \boldsymbol{u}, \boldsymbol{u}\rangle},$$

vektora $\boldsymbol{u}$, ktorý leží v jeho subjektívnom fyzikálnom priestore $[\boldsymbol{b}]^\perp$ (čo značí, že pri jeho ľubovoľnom umiestnení sú oba jeho konce súčasné udalosti vzhľadom na bázu $\boldsymbol{\beta}$) a spĺňa podmienku

$$(t\boldsymbol{a}_0 + l\boldsymbol{a}_1)_{\boldsymbol{\alpha}} = (t, l, 0, \dots, 0) = \varphi\big((\boldsymbol{u})_{\boldsymbol{\beta}}\big)$$

pre nejaký okamih t vlastného času sústavy $\boldsymbol{\alpha}$. Keďže $\boldsymbol{u} \in [\boldsymbol{b}]^\perp$, jeho súradnice majú tvar $(\boldsymbol{u})_{\boldsymbol{\beta}} = (0, x_1', x_2', \dots, x_n')^\mathsf{T}$ a z podmienky pre ich Lorentzovu transformáciu okamžite vidíme, že $x_2' = \ldots = x_n' = 0$, ako aj

$$\begin{pmatrix} t \\ l \end{pmatrix} = \boldsymbol{L}_v \cdot \begin{pmatrix} 0 \\ x_1' \end{pmatrix} = \frac{x_1'}{\sqrt{1 - v^2}} \begin{pmatrix} v \\ 1 \end{pmatrix}.$$

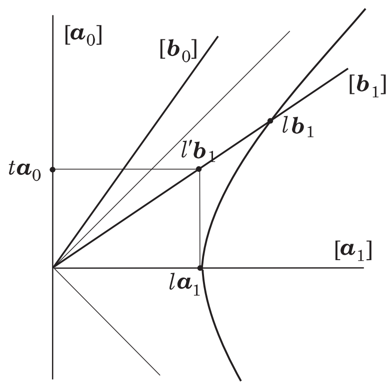
Potom zrejme $l' = x_1'$ a konečne dostávame ohlásený vzťah medzi dĺžkami tyče v oboch inerciálnych sústavách:

$$l = \frac{l'}{\sqrt{1 - v^2}} \geq l',$$

pričom rovnosť nastane práve vtedy, keď $v = 0$, teda dĺžka tyče sa javí najväčšia v tej inerciálnej sústave, vzhľadom na ktorú je tyč nehybná. Tento jav sa nazýva *relativistické skrátenie* alebo tiež *relativistická kontrakcia dĺžky* v smere pohybu. Nasledujúci obrázok, ktorý znázorňuje situáciu v $\mathbb{R}^{(1,1)}$, je plne analogický obrázku znázorňujúcemu relativistickú dilatáciu času. Opäť sa na ňom stretáme s disproporciou euklidovskej a Minkowského dĺžky vektorov $l'\boldsymbol{b}_1$ a $l\boldsymbol{a}_1$.

Ešte si všimnime, že onen okamih t vlastného času v sústave $\boldsymbol{\alpha}$ je jednoznačne určený:

$$t = l' \frac{v}{\sqrt{1 - v^2}} = lv.$$



Obr. 16.9. Relativistická kontrakcia dĺžky

Uvedená formulka je však lepšie čitateľná v obvyklom fyzikálnom tvare

$$ct = v\,\frac{l}{c}.$$

Inak povedané, za čas t preletí svetlo rovnakú vzdialenosť, akú urazíme rýchlosťou v za čas, ktorý potrebuje svetlo na prekonanie vzdialenosti l. Zrejme pre tyč „bežných rozmerov“ je čas t enormne malý.

Cvičenia

16.1. Nech $\boldsymbol{a}$, $\boldsymbol{b}$ sú nenulové vektory v Minkowského časopriestore V. Dokážte nasledujúce tvrdenia:

(a) Ak $\boldsymbol{a}$, $\boldsymbol{b}$ sú časové vektory tak $\langle \boldsymbol{a}, \boldsymbol{b}\rangle \neq 0$.

(b) Ak $\boldsymbol{a}$, $\boldsymbol{b}$ sú lineárne nezávislé svetelné vektory, tak $\langle \boldsymbol{a}, \boldsymbol{b}\rangle \neq 0$.

(c) Ak $\boldsymbol{a}$ je časový a $\boldsymbol{b}$ je svetelný vektor, tak $\langle \boldsymbol{a}, \boldsymbol{b}\rangle \neq 0$.

16.2. (a) Nenulové *svetelné vektory* $\boldsymbol{u}$, $\boldsymbol{v}$ v Minkowského časopriestore V nazveme *súhlasne orientovanými*, ak existuje časový vektor $\boldsymbol{a}$ taký, že $\boldsymbol{a}$, $\boldsymbol{u}$ aj $\boldsymbol{a}$, $\boldsymbol{v}$ sú súhlasne orientované. Dokážte, že nenulové svetelné vektory $\boldsymbol{u}$, $\boldsymbol{v}$ sú súhlasne orientované práve vtedy, keď $\langle \boldsymbol{u}, \boldsymbol{v}\rangle > 0$ alebo existuje *kladný* skalár $s \in \mathbb{R}$ taký, že $\boldsymbol{u} = s\boldsymbol{v}$.

(b) Dokážte, že Minkowského časopriestor $\mathbb{R}^{(1,n)}$ má bázu pozostávajúcu zo samých navzájom súhlasne orientovaných svetelných vektorov. Vypočítajte jej Gramovú maticu a vysvetlite, prečo žiadne dva vektory takejto bázy nemôžu byť ortogonálne.

(c) Dokážte, že v $\mathbb{R}^{(1,n)}$ existuje n-rozmerný singulárny podpriestor. (*Návod*: Uvažujte zároveň štandardný euklidovský skalárny súčin v $\mathbb{R}^{n+1}$ a vezmite euklidovský ortokomplement nejakého svetelného vektora.) Aká je signatúra pseudoskalárneho súčinu zúženého na tento podpriestor?

16.3. Doplňte chýbajúce výpočty v dôkazoch tvrdení 16.3.1 a 16.4.1.

16.4. Nech $\boldsymbol{a}$ je časový vektor v $\mathbb{R}^{(1,n)}$ a $S = [\boldsymbol{a}]^{\perp}$ je k nemu prislúchajúci subjektívny fyzikálny priestor. Označme $\boldsymbol{a}_S$ *euklidovský* kolmý priemet vektora $\boldsymbol{a}$ do podpriestoru S a M ľubovoľnú z dvojice priamok, ktoré tvoria prienik $[\boldsymbol{a}, \boldsymbol{a}_S] \cap \mathrm{LC}(\mathbf{0})$. Dokážte, že priamky $[\boldsymbol{a}]$, $[\boldsymbol{a}_S]$ sú (v euklidovskom zmysle) **s**úmerne združené podľa priamky M.

16.5. Načrtnite grafy funkcií $\beta(v) = \frac{1}{\sqrt{1-v^2}}$ a $1/\beta(v) = \sqrt{1-v^2}$ pre $-1 < v < 1$. Pri akej hodnote rýchlosti v dôjde k relativistickému skráteniu pohybujúcej sa tyče na polovicu jej dĺžky? Akému relativistickému skráteniu podlieha nadzvukové lietadlo pri rýchlosti $400\,\mathrm{m/s}$?

16.6. Overte rovnosť Lorentzovej transformácie $\boldsymbol{L}_v$ a hyperbolickej rotácie $\boldsymbol{Rh}_\theta$ pre $\theta = \ln\sqrt{\frac{1+v}{1-v}}$. Nájdite spätné vyjadrenie v ako vhodnej hyperbolickej funkcie argumentu θ.

16.7. Dokážte, že pre súčin Lorentzových transfomácií $\boldsymbol{L}_u$, $\boldsymbol{L}_v$ platí $\boldsymbol{L}_u \cdot \boldsymbol{L}_v = \boldsymbol{L}_w$, kde $w = \frac{u+v}{1+uv}$. Vysvetlite, prečo sa uvedená rovnosť nazýva *relativistickým pravidlom skladania rýchlostí.*

16.8. (a) Dokážte súčtové vzorce pre hyperbolické funkcie $\cosh x$ a $\sinh x$.

(b) Odvoďte súčtové pravidlo pre súčin (kompozíciu) hyberbolických rotácií $\boldsymbol{Rh}_\theta \cdot \boldsymbol{Rh}_\eta = \boldsymbol{Rh}_{\theta+\eta}$.

(c) Porovnajte toto pravidlo s relativistickým pravidlom skladania rýchlostí.

16.9. (a) Uvedomte si, že pri bežnej rotácii v $\mathbb{R}^2$ o uhol α rotuje jednotkový vektor po kružnici s rovnicou $x^2 + y^2 = 1$, pričom α je nielen dĺžka príslušného oblúka, ale aj polovica plošného obsahu výseku jednotkového kruhu opísaného týmto vektorom (tzv. *sprievodičom*).

(b) Overte, že pri hyberbolickej rotácii v $\mathbb{R}^{(1,1)}$ o hyperbolický uhol θ „rotuje" jednotkový časový vektor po rovnoosej hyperbole s rovnicou $x_0^2 - x_1^2 = 1$ a jednotkový priestorový vektor po rovnoosej hyperbole s rovnicou $x_0^2 - x_1^2 = -1$. Pritom pre $\theta > 0$ vektory „rotujú" smerom k asymptote s rovnicou $x_1 = x_0$ a pre $\theta < 0$ k asymptote s rovnicou $x_1 = -x_0$. Svetelné vektory si zachovávajú smer aj orientáciu. Nakreslite si obrázok (vzorom vám môže byť Obr. 16.8, na ktorom sú však vektory $\boldsymbol{a}_0$, $\boldsymbol{a}_1$ v špeciálnej polohe).

(c) Overte, že svetelný vektor $\boldsymbol{x}$ sa zobrazí do svojho násobku $\mathrm{e}^{\theta}\,\boldsymbol{x}$, ak leží na priamke $x_1 = x_0$, resp. $\mathrm{e}^{-\theta}\,\boldsymbol{x}$, ak leží na priamke $x_1 = -x_0$.

(d) Vyznačte v rovine plošný útvar opísaný ľubovoľným jednotkovým časovým alebo priestorovým vektorom pri hyperbolickej rotácii $\boldsymbol{Rh}_\theta$ a vypočítajte jeho plošný obsah pomocou vhodného určitého integrálu. Malo by vám vyjsť $\theta/2$. Vypočítajte aj dĺžku príslušného hyperbolického oblúka. Vyšla vám očakávaná hodnota θ? Porovnajte s (a).

16.10. Časové šípy $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ v Minkowského časopriestore V sú lineárne závislé práve vtedy, keď sú lineárne závislé vektory $\boldsymbol{u}$, $\boldsymbol{v}$ rýchlostí, ktorými sa pohybujú inerciálni pozorovatelia s časovými šípmi $\boldsymbol{a}$ resp. $\boldsymbol{b}$ vzhľadom na inerciálneho pozorovateľa s časovým šípom $\boldsymbol{c}$. Dokážte.

16.11. Vyložte relativistickú dilatáciu času a kontrakciu dĺžky ako odchýlku vhodných vektorov o „hyperbolický uhol".

17. Unitárne priestory

V paragrafe 12.1 sme videli, že kanonický diagonálny tvar ľubovoľnej symetrickej bilineárnej formy na konečnorozmernom vektorovom priestore nad poľom $\mathbb{C}$ všetkých komplexných čísel je jednoznačne určený jej hodnosťou. Zavádzať pre takéto formy niečo na spôsob signatúry a hovoriť o ich definitnosti vôbec nemá zmysel. Táto jednoduchosť v porovnaní s teóriou symetrických bilineárnych foriem na vektorových priestoroch nad poľom $\mathbb{R}$ má za následok, že symetrická regulárna bilineárna forma na komplexnom vektorovom priestore nevytvára geometrickú štruktúru analogickú reálnemu prípadu. Ukazuje sa však, že zdanlivo nepatrnou modifikáciou pojmu bilineárnej formy možno túto prekážku geometrizácie preklenúť.

Kapitolu začneme štúdiom tzv. *poldruhalineárnych foriem* na komplexných vektorových priestoroch. Pre takéto formy spĺňajúce istú mierne pozmenenú podmienku symetrie už možno prirodzene zaviesť pojmy signatúry a definitnosti a rozšíriť na ne platnosť tvrdení z kapitoly 12. Najdôležitejší bude pre nás opäť kladne definitný prípad, kedy hovoríme o *(komplexnom) skalárnom súčine*. Teória (konečnorozmerných) *unitárnych priestorov*, t. j. komplexných vektorových priestorov vybavených skalárnym súčinom, je natoľko priamočiarym zovšeobecnením teórie euklidovských priestorov, že väčšinu pojmov a výsledkov možno z jednej do druhej preniesť len s malými redakčnými úpravami. Preto miesto systematickej výstavby teórie unitárnych priestorov iba stručne naznačíme, ako to možno urobiť. Špeciálne zavedieme a stručne preskúmame tzv. *unitárne matice*, ktoré sú komplexnou analógiou ortogonálnych matíc.

Osobitný paragraf venujeme *diskrétnej Fourierovej transformácii*, ktorej význam prudko vzrástol po objave veľmi rýchlych algoritmov v 60. rokoch 20. storočia.

V záverečnej časti stručne načrtneme úlohu unitárnych priestorov v *kvantovej mechanike*. Tomu bude predchádzať krátke odbočenie do klasickej mechaniky a pojednanie o ťažkostiach, ktoré sa stavajú do cesty pokusom vytvoriť adekvátny fyzikálny obraz javov mikrosveta a matematický popis jeho zákonitostí.

17.1 Poldruhalineárne formy

Začneme banálnym pozorovaním: Absolútna hodnota $|x|$ reálneho čísla x, čiže jeho vzdialenosť od počiatku, je so súčinom xy, t. j. s najzákladnejšou bilineárnou formou, zviazaná vzťahom $|x|^2 = xx$. Taktiež absolútna hodnota

$|x|$ komplexného čísla x má geometrický význam jeho vzdialenosti od počiatku – uvedený vzťah však platí v modifikovanej podobe $|x|^2 = x\overline{x}$. Ak teda chceme budovať geometriu umožňujúcu vyjadriť vzdialenosti pomocou vhodných komplexných analógov reálnych bilineárnych a kvadratických foriem, súčin xy treba nahradiť výrazom $x\overline{y}$. Keďže $x\overline{y} = xy$ pre $x, y \in \mathbb{R}$, ide o prirodzené zovšeobecnenie reálneho prípadu.

Poldruhalineárnou formou na komplexnom vektorovom priestore V nazývame zobrazenie $F\colon V \times V \to \mathbb{C}$ také, že pre všetky $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$, $c \in \mathbb{C}$ platí

$$F(\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{z}) = F(\boldsymbol{x}, \boldsymbol{z}) + F(\boldsymbol{y}, \boldsymbol{z}), \qquad F(c\boldsymbol{x}, \boldsymbol{z}) = cF(\boldsymbol{x}, \boldsymbol{z}),$$
$$F(\boldsymbol{x}, \boldsymbol{y}+\boldsymbol{z}) = F(\boldsymbol{x}, \boldsymbol{y}) + F(\boldsymbol{x}, \boldsymbol{z}), \qquad F(\boldsymbol{x}, c\boldsymbol{y}) = \overline{c}F(\boldsymbol{x}, \boldsymbol{y}).$$

Hovoríme, že F je lineárne v prvej premennej a *semilineárne* v druhej premennej.

Ak V je konečnorozmerný a $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ je jeho báza, tak maticu

$$\boldsymbol{A} = [F]_{\boldsymbol{\alpha}} = \big(F(\boldsymbol{u}_j, \boldsymbol{u}_k)\big) \in \mathbb{C}^{n\times n},$$

rovnako ako v reálnom prípade, nazývame *maticou poldruhalineárnej formy* F v báze $\boldsymbol{\alpha}$. Potom pre ľubovoľné vektory $\boldsymbol{x}, \boldsymbol{y} \in V$ so súradnicami $(\boldsymbol{x})_{\boldsymbol{\alpha}} = (x_1, \dots, x_n)^{\mathsf{T}}$, $(\boldsymbol{y})_{\boldsymbol{\alpha}} = (y_1, \dots, y_n)^{\mathsf{T}}$ platí

$$F(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \overline{(\boldsymbol{y})_{\boldsymbol{\alpha}}} = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} x_j \overline{y}_k,$$

kde $\overline{(\boldsymbol{y})_{\boldsymbol{\alpha}}} = (\overline{y}_1, \dots, \overline{y}_k)^{\mathsf{T}}$. Pritom $\boldsymbol{A} = (a_{jk})_{n\times n} = [F]_{\boldsymbol{\alpha}}$ je jediná matica s touto vlastnosťou.

Ak $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ je druhá báza priestoru V, tak matice $\boldsymbol{A} = [F]_{\boldsymbol{\alpha}}$, $\boldsymbol{B} = [F]_{\boldsymbol{\beta}}$ sú zviazané vzťahom

$$\boldsymbol{B} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \overline{\boldsymbol{P}}_{\boldsymbol{\alpha},\boldsymbol{\beta}},$$

kde $\overline{\boldsymbol{P}}_{\alpha,\boldsymbol{\beta}} = (\overline{p}_{jk})$ je matica komplexne združená k matici prechodu $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = (p_{jk})$. Z toho vyplýva, že $\boldsymbol{A}$, $\boldsymbol{B} \in \mathbb{C}^{n\times n}$ sú maticami tej istej poldruhalineárnej formy vzhľadom na (možno) rôzne bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ práve vtedy, keď $\boldsymbol{B} = \boldsymbol{P}^{\mathsf{T}} \cdot \boldsymbol{A} \cdot \overline{\boldsymbol{P}}$ pre nejakú regulárnu maticu $\boldsymbol{P} \in \mathbb{C}^{n\times n}$, t. j. práve vtedy, keď existuje regulárna matica $\boldsymbol{Q} \in \mathbb{C}^{n\times n}$ taká, že

$$\boldsymbol{B} = \boldsymbol{Q}^* \cdot \boldsymbol{A} \cdot \boldsymbol{Q},$$

kde $\boldsymbol{Q}^* = \overline{\boldsymbol{Q}^{\mathsf{T}}} = \overline{\boldsymbol{Q}}^{\mathsf{T}}$ označuje maticu transponovanú a komplexne združenú k matici $\boldsymbol{Q}$ – hovoríme, že $\boldsymbol{Q}^*$ je *hermitovsky združená* alebo tiež *adjungovaná matica* k matici $\boldsymbol{Q}$. Naozaj, ak položíme $\boldsymbol{Q} = \overline{\boldsymbol{P}}$, tak $\boldsymbol{Q}^* = \boldsymbol{P}^{\mathsf{T}}$. V takom prípade hovoríme, že *matice* $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$ sú *hermitovsky kongruentné* a píšeme $\boldsymbol{A} \stackrel{*}{\equiv} \boldsymbol{B}$.

Keďže pole $\mathbb{C}$ všetkých komplexných čísel je rozšírením poľa $\mathbb{R}$ všetkých reálnych čísel, každý vektorový priestor V nad poľom $\mathbb{C}$ možno zároveň považovať za vektorový priestor nad poľom $\mathbb{R}$ (pozri príklad 1.6.1). Tento vektorový priestor budeme značiť $V_{\mathbb{R}}$ a nazývať *reálnym zúžením* alebo tiež *zreálnením* priestoru V.

Každé zobrazenie $F\colon V \times V \to \mathbb{C}$ určuje predpismi

$$F_0(\boldsymbol{x}, \boldsymbol{y}) = \operatorname{Re} F(\boldsymbol{x}, \boldsymbol{y}), \qquad F_1(\boldsymbol{x}, \boldsymbol{y}) = \operatorname{Im} F(\boldsymbol{x}, \boldsymbol{y}),$$

pre $\boldsymbol{x}, \boldsymbol{y} \in V$, dve zobrazenia $F_0 = \operatorname{Re} F$, $F_1 = \operatorname{Im} F\colon V \times V \to \mathbb{R}$; potom, samozrejme, $F(\boldsymbol{x}, \boldsymbol{y}) = F_0(\boldsymbol{x}, \boldsymbol{y}) + \mathrm{i}F_1(\boldsymbol{x}, \boldsymbol{y})$.

Dôkaz nasledujúceho tvrdenia prenechávame ako jednoduché cvičenie čitateľovi.

17.1.1. Tvrdenie. *Nech $F\colon V \times V \to \mathbb{C}$ je ľubovoľné zobrazenie. Potom F je poldruhalineárna forma práve vtedy, keď $F_0 = \operatorname{Re} F$, $F_1 = \operatorname{Im} F$ sú bilineárne formy na reálnom vektorovom priestore $V_{\mathbb{R}}$ a pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí*

$$F_0(\boldsymbol{x}, \boldsymbol{y}) = F_1(\mathrm{i}\boldsymbol{x}, \boldsymbol{y}), \qquad F_1(\boldsymbol{x}, \boldsymbol{y}) = F_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{y}).$$

To okrem iného znamená, že každá zo zložiek F_0, F_1 poldruhalineárnej formy F jednoznačne určuje druhú.

17.2 Hermitovské formy a hermitovské matice

Hovoríme, že *poldruhalineárna forma* F na komplexnom vektorovom priestore V je *hermitovská* alebo tiež *kososymetrická*, ak pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí

$$F(\boldsymbol{x}, \boldsymbol{y}) = \overline{F(\boldsymbol{y}, \boldsymbol{x})}.$$

Nasledujúce tvrdenie je bezprostredným dôsledkom našej definície.

17.2.1. Tvrdenie. *Poldruhalineárna forma $F\colon V \times V \to \mathbb{C}$ je hermitovská práve vtedy, keď jej zložky $F_0 = \operatorname{Re} F$, $F_1 = \operatorname{Im} F$ spĺňajú podmienky*

$$F_0(\boldsymbol{x}, \boldsymbol{y}) = F_0(\boldsymbol{y}, \boldsymbol{x}), \qquad F_1(\boldsymbol{x}, \boldsymbol{y}) = -F_1(\boldsymbol{y}, \boldsymbol{x})$$

pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$.

Podľa tvrdení 17.1.1 a 17.2.1 možno každú hermitovskú formu $F\colon V \times V \to \mathbb{C}$ rozložiť na súčet $F = F_0 + \mathrm{i}F_1$, pričom prvá z bilineárnych foriem F_0, $F_1\colon V_{\mathbb{R}} \times V_{\mathbb{R}} \to \mathbb{R}$ je symetrická, druhá antisymetrická a platí $F_1(\boldsymbol{x}, \boldsymbol{y}) =$

$F_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{y})$ alebo, čo vyjde vďaka (anti)symetrii narovnako, $F_0(\boldsymbol{x}, \boldsymbol{y}) = F_1(\mathrm{i}\boldsymbol{x}, \boldsymbol{y})$.

Štvorcová *matica* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ sa nazýva *hermitovská* alebo tiež *kososymetrická*, ak $\boldsymbol{A} = \boldsymbol{A}^*$, t. j. $a_{jk} = \overline{a}_{kj}$ pre všetky $j, k \le n$. Špeciálne, $a_{jj} = \overline{a}_{jj}$, čiže všetky diagonálne prvky hermitovskej matice sú reálne. Zrejme $\boldsymbol{A}$ je hermitovská práve vtedy, keď poldruhalineárna forma $F(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^\mathsf{T} \cdot \boldsymbol{A} \cdot \overline{\boldsymbol{y}}$ na (stĺpcovom) vektorovom priestore $\mathbb{C}^n$ je hermitovská.

Ak $F\colon V \times V \to \mathbb{C}$ je hermitovská forma, tak $F(\boldsymbol{x}, \boldsymbol{x}) = \overline{F(\boldsymbol{x}, \boldsymbol{x})}$ pre každé $\boldsymbol{x} \in V$, čiže všetky hodnoty $F(\boldsymbol{x}, \boldsymbol{x})$ sú reálne, teda má zmysel pýtať sa na ich znamienko. To nám v komplexnom prípade umožňuje zaviesť pre hermitovské formy a hermitovské matice všetky pojmy súvisiace s definitnosťou a so signatúrou rovnako, ako v reálnom prípade pre symetrické bilineárne formy a symetrické matice (pozri paragrafy 12.1 a 12.2). Navyše, v dôsledku antisymetrie formy $F_1 = \operatorname{Im} F$ platí

$$F(\boldsymbol{x}, \boldsymbol{x}) = F_0(\boldsymbol{x}, \boldsymbol{x}), \qquad F_1(\boldsymbol{x}, \boldsymbol{x}) = 0,$$

takže na $F(\boldsymbol{x}, \boldsymbol{x})$ sa možno dívať ako na kvadratickú formu na reálnom vektorovom priestore $V_\mathbb{R}$.

Signatúru hermitovskej matice $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, a tým aj hermitovskej formy $F\colon V^2 \to \mathbb{C}$ na konečnorozmernom vektorovom priestore V nad $\mathbb{C}$, možno zistiť jej úpravou na diagonálny tvar podľa schémy

$$\boldsymbol{A} \xrightarrow[\text{ERO}]{\overline{\text{ESO}}} \boldsymbol{B}$$

$$\boldsymbol{I}_n \xrightarrow{\text{ESO}} \boldsymbol{P}$$

Na matici $\boldsymbol{A}$ teda vždy vykonáme jednu ERO a jej zodpovedajúcu ESO s *komplexne združeným skalárom* (teda výmene dvoch riadkov zodpovedá výmena príslušných stĺpcov, avšak vynásobeniu j-teho riadku nenulovým skalárom $c \in \mathbb{C}$ zodpovedá vynásobenie j-teho stĺpca skalárom $\overline{c}$ a pripočítaniu c-násobku j-teho riadku ku k-temu riadku zodpovedá pripočítanie $\overline{c}$-násobku j-teho stĺpca ku k-temu stĺpcu). Na jednotkovej matici vykonáme príslušnú ESO s pôvodným skalárom. (Porovnaj s podobnou schémou z paragrafu 11.3). Potom $\boldsymbol{A} \overset{*}{\equiv} \boldsymbol{B} = \boldsymbol{P}^\mathsf{T} \cdot \boldsymbol{A} \cdot \overline{\boldsymbol{P}}$ a stĺpce matice $\boldsymbol{P}$ tvoria bázu $\boldsymbol{\beta}$ priestoru $\mathbb{C}^n$, vzhľadom na ktorú má hermitovská forma $\boldsymbol{x}^\mathsf{T} \cdot \boldsymbol{A} \cdot \overline{\boldsymbol{y}}$ diagonálnu (preto tiež nevyhnutne reálnu) maticu $\boldsymbol{B}$. Ak $\boldsymbol{A}$ bola maticou hermitovskej formy F na n-rozmernom vektorovom priestore V v báze $\boldsymbol{\alpha}$, tak diagonálna matica $\boldsymbol{B}$ je maticou formy F v báze $\boldsymbol{\beta} = \boldsymbol{\alpha} \cdot \boldsymbol{P}$, čiže $\boldsymbol{P} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ je maticou prechodu z bázy $\boldsymbol{\beta}$ do bázy $\boldsymbol{\alpha}$.

V podstate rovnako ako v paragrafe 11.3 sa dá dokázať, že uvedený postup vedie vždy k cieľu. Navyše možno dosiahnuť, aby matica $\boldsymbol{B}$ mala na diagonále len skaláry ± 1 a 0. Napriek istej nejednoznačnosti výsledných matíc $\boldsymbol{B}$ a $\boldsymbol{P}$, i v tomto prípade platí Sylvestrov zákon zotrvačnosti, teda signatúra hermitovskej formy či matice je dobre definovaná a sú splnené jednoduché analógie výsledkov pre reálne symetrické bilineárne formy a matice z paragrafu 12.1. Detaily prenechávame na samostatné premyslenie čitateľovi.

17.3 Komplexný skalárny súčin a unitárne priestory

Skalárnym alebo tiež *vnútorným súčinom* na komplexnom vektorovom priestore V nazývame ľubovoľnú kladne definitnú hermitovskú poldruhalineárnu formu na V; jej hodnotu na vektoroch $\boldsymbol{x}, \boldsymbol{y} \in V$ budeme značiť opäť $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$. Komplexný vektorový priestor V vybavený skalárnym súčinom nazývame *unitárny priestor*.

Nezávisle na znalosti uvedených pojmov možno skalárny súčin na V definovať ako binárnu operáciu $V \times V \to \mathbb{C}$, ktorá každej dvojici $(\boldsymbol{x}, \boldsymbol{y})$ vektorov z V priradí komplexné číslo $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ také, že pre všetky $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x}_1, \boldsymbol{x}_2 \in V$ a ľubovoľné $c \in \mathbb{C}$ platí:

$$
\begin{aligned}
\langle \boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y} \rangle &= \langle \boldsymbol{x}_1, \boldsymbol{y} \rangle + \langle \boldsymbol{x}_2, \boldsymbol{y} \rangle && \text{(aditivita)},\\
\langle c\boldsymbol{x}, \boldsymbol{y} \rangle &= c\langle \boldsymbol{x}, \boldsymbol{y} \rangle && \text{(homogenita)},\\
\langle \boldsymbol{x}, \boldsymbol{y} \rangle &= \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle} && \text{(kosá symetria)},\\
\boldsymbol{x} \neq \boldsymbol{0} \ &\Rightarrow\ \langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0 && \text{(kladná definitnosť)}.
\end{aligned}
$$

Spojenie aditivity a homogenity skalárneho súčinu dáva jeho linearitu ako funkcie prvej premennej (pri pevnej druhej premennej). Vďaka kosej symetrii z toho vyplýva semilinearita skalárneho súčinu ako funkcie druhej premennej (pri pevnej prvej premennej), t. j. rovnosti

$$
\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{y}_1 + \boldsymbol{y}_2 \rangle &= \langle \boldsymbol{x}, \boldsymbol{y}_2 \rangle + \langle \boldsymbol{x}, \boldsymbol{y}_2 \rangle,\\
\langle \boldsymbol{x}, c\boldsymbol{y} \rangle &= \overline{c}\langle \boldsymbol{x}, \boldsymbol{y} \rangle,
\end{aligned}
$$

pre všetky $\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2 \in V$ a $c \in \mathbb{C}$.[1] Rovnako z kosej symetrie vyplýva reálnosť výrazu $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \overline{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$ pre každé $\boldsymbol{x} \in V$, čo teprv dáva zmysel podmienke kladnej definitnosti; z poldruhalinearity potom vyplýva jej podrobnejší rozpis

$$
\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0 \ \&\ \big(\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}\big)
$$

[1] Poznamenávame, že vo fyzike sa v definícii komplexného skalárneho súčinu často žiada linearita v druhej a semilinerita v prvej zložke.

pre každé $\boldsymbol{x} \in V$.

Ako sme naznačili predchádzajúcimi dvoma odstavcami, ktoré verne sledujú formulácie prvých dvoch odstavcov z paragrafu 13.1, teória konečnorozmerných unitárnych priestorov je pomerne priamočiarym zovšeobecnením teórie euklidovských priestorov. Väčšinu pojmov, ktoré sme definovali pre euklidovské priestory, možno zaviesť aj pre (konečnorozmerné) unitárne priestory a väčšinu výsledkov o euklidovských priestoroch možno s malými modifikáciami dokázať aj pre (konečnorozmerné) unitárne priestory. S istou dávkou zjednodušenia možno povedať, že jediný formálny rozdiel spočíva v tom, že v komplexnom prípade si musíme dávať pozor na poradie činiteľov v skalárnom súčine, občas nad niektoré skaláry primaľovať pruh a rozlišovať medzi c^2 a $|c|^2$. Z toho dôvodu nebudeme čitateľa unavovať systematickým budovaním teórie unitárnych priestorov; miesto toho sa obmedzíme len na zopár základných pojmov a výsledkov, a kde to bude možné, odvoláme sa na zodpovedajúce analógie z euklidovských priestorov.

Keďže $0 \leq \langle \boldsymbol{x}, \boldsymbol{x} \rangle \in \mathbb{R}$, *dĺžku* alebo tiež *normu vektora* $\boldsymbol{x}$ v unitárnom priestore V možno definovať ako nezáporné reálne číslo

$$\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}.$$

Zrejme $\|\boldsymbol{x}\|$ je norma na vektorovom priestore $V_{\mathbb{R}}$ pochádzajúca od reálneho skalárneho súčinu $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_0 = \operatorname{Re}\langle \boldsymbol{x}, \boldsymbol{y} \rangle$, teda sú pre ňu splnené všetky tri definujúce podmienky (trojuholníková nerovnosť, pozitívna homogenita, oddeliteľnosť) z paragrafu 13.3. Navyše podmienka pozitívnej homogenity

$$\|c\boldsymbol{x}\| = |c|\, \|\boldsymbol{x}\|$$

platí pre všetky $\boldsymbol{x} \in V$, $c \in \mathbb{C}$ (a nielen pre $c \in \mathbb{R}$). Naozaj,

$$\|c\boldsymbol{x}\|^2 = \langle c\boldsymbol{x}, c\boldsymbol{x} \rangle = c\bar{c}\langle \boldsymbol{x}, \boldsymbol{x} \rangle = |c|^2\, \|\boldsymbol{x}\|^2 .$$

Pozornejší čitateľ možno v tejto chvíli pojal podozrenie, že poldruhalineárne formy, ako modifikácia bilineárnych foriem, boli vymyslené len nato, aby nám v práve vykonanom výpočte prešiel trik so skalárom c. Až na časticu „len“ mu toto podozrenie nehodláme vyvracať.

Z Cauchyho-Schwartzovej nerovnosti (veta 13.2.3) pre reálny skalárny súčin dostávame nerovnosť $|\operatorname{Re}\langle \boldsymbol{u}, \boldsymbol{v} \rangle| \leq \|\boldsymbol{u}\|\, \|\boldsymbol{v}\|$, z ktorej už vyplýva trojuholníková nerovnosť pre normu. S trochou úsilia však možno dokázať silnejší odhad.

17.3.1. Veta. *(Cauchyho-Schwartzova nerovnosť) Nech V je unitárny priestor. Potom pre ľubovoľné vektory $\boldsymbol{u}, \boldsymbol{v} \in V$ platí*

$$|\langle \boldsymbol{u}, \boldsymbol{v} \rangle| \leq \|\boldsymbol{u}\|\, \|\boldsymbol{v}\| ,$$

pričom rovnosť nastane práve vtedy, keď $\boldsymbol{u}$, $\boldsymbol{v}$ sú lineárne závislé.

Dôkaz. Označme si $c = \cos\alpha + \mathrm{i}\sin\alpha$, kde $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = |\langle \boldsymbol{u}, \boldsymbol{v}\rangle| (\cos\alpha + \mathrm{i}\sin\alpha)$ je vyjadrenie čísla $\langle \boldsymbol{u}, \boldsymbol{v}\rangle \in \mathbb{C}$ v goniometrickom tvare. Potom $|c| = 1$, $c^{-1} = \overline{c}$ a

$$|\langle \boldsymbol{u}, \boldsymbol{v}\rangle| = c^{-1}\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \overline{c}\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \langle \boldsymbol{u}, c\boldsymbol{v}\rangle = \operatorname{Re}\langle \boldsymbol{u}, c\boldsymbol{v}\rangle,$$

lebo ide o reálne číslo. Keďže $\operatorname{Re}\langle \boldsymbol{u}, \boldsymbol{v}\rangle$ je kladne definitná bilineárna forma na $V_{\mathbb{R}}$, podľa reálnej verzie Cauchyho-Schwartzovej nerovnosti platí

$$|\langle \boldsymbol{u}, \boldsymbol{v}\rangle| = \operatorname{Re}\langle \boldsymbol{u}, c\boldsymbol{v}\rangle \leq \|\boldsymbol{u}\|\,\|c\boldsymbol{v}\| = |c|\,\|\boldsymbol{u}\|\,\|\boldsymbol{v}\| = \|\boldsymbol{u}\|\,\|\boldsymbol{v}\|\,.$$

Rovnosť nastane práve vtedy, keď vektory $\boldsymbol{u}$, $c\boldsymbol{v}$ sú lineárne závislé nad $\mathbb{R}$. Z toho zrejme vyplýva lineárna závislosť vektorov $\boldsymbol{u}$, $\boldsymbol{v}$ nad $\mathbb{C}$. Naopak, ak $\boldsymbol{u}$, $\boldsymbol{v}$ sú lineárne závislé nad $\mathbb{C}$, tak rovnosť $|\langle \boldsymbol{u}, \boldsymbol{v}\rangle| = \|\boldsymbol{u}\|\,\|\boldsymbol{v}\|$ možno jednoducho overiť priamym výpočtom.

Pre nenulové vektory $\boldsymbol{u}, \boldsymbol{v} \in V$ predstavuje výraz $\langle \boldsymbol{u}, \boldsymbol{v}\rangle / \|\boldsymbol{u}\|\,\|\boldsymbol{v}\|$ komplexné číslo s absolútnou hodnotou ≤ 1. Analógia s reálnym prípadom nás zvádza pokúsiť sa na jeho základe nejako zadefinovať uhol vektorov $\boldsymbol{u}$, $\boldsymbol{v}$. Núkajú sa nám tri možnosti:

(1) Vychádzajúc z toho, že reálna časť $\operatorname{Re}\langle \boldsymbol{u}, \boldsymbol{v}\rangle$ skalárneho súčinu $\langle \boldsymbol{u}, \boldsymbol{v}\rangle$ je skalárnym súčinom na reálnom zúžení $V_{\mathbb{R}}$ unitárneho priestoru V, ktorý plne určuje jeho normu, môžeme sa sústrediť len na ňu a ignorovať imaginárnu časť. V takom prípade

$$\alpha = \arccos \frac{\operatorname{Re}\langle \boldsymbol{u}, \boldsymbol{v}\rangle}{\|\boldsymbol{u}\|\,\|\boldsymbol{v}\|}$$

predstavuje uhol, ktorý zvierajú vektory $\boldsymbol{u}$, $\boldsymbol{v}$ v priestore $V_{\mathbb{R}}$.

(2) Uhol vektorov $\boldsymbol{u}$, $\boldsymbol{v}$ môžeme definovať ako uhol priamok $[\boldsymbol{u}]$, $[\boldsymbol{v}]$, t. j. ako reálne číslo

$$\alpha = \arccos \frac{|\langle \boldsymbol{u}, \boldsymbol{v}\rangle|}{\|\boldsymbol{u}\|\,\|\boldsymbol{v}\|}.$$

V reálnom prípade by také niečo zodpovedalo nahradeniu odchýlky vektorov $\sphericalangle(\boldsymbol{u}, \boldsymbol{v})$ odchýlkou priamok $\sphericalangle([\boldsymbol{u}], [\boldsymbol{v}])$; to znamená, že medzi odchýlkami dvojíc $(\boldsymbol{u}, \boldsymbol{v})$ a $(\boldsymbol{u}, -\boldsymbol{v})$ by sme nerozlišovali a z pôvodných uhlov $\sphericalangle(\boldsymbol{u}, \boldsymbol{v})$, $\sphericalangle(\boldsymbol{u}, -\boldsymbol{v})$ by sme vybrali ten menší. (Nezabúdajme však, že priamky $[\boldsymbol{u}]$, $[\boldsymbol{v}]$ v komplexnom vektorovom priestore V sú z reálneho hľadiska dvojrozmerné, t. j. sú to vlastne roviny vo $V_{\mathbb{R}}$.)

(3) Prostriedkami teórie funkcií komplexnej premennej možno definičný obor reálnej funkcie arccos, t. j. interval $\langle -1, 1\rangle$, rozšíriť na jednotkový kruh $\{c \in \mathbb{C};\ |c| \leq 1\}$ (dokonca na celú komplexnú rovinu $\mathbb{C}$). Potom odchýlka $\sphericalangle(\boldsymbol{u}, \boldsymbol{v})$ by bola *komplexné číslo*

$$\alpha = \arccos \frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}{\|\boldsymbol{u}\|\,\|\boldsymbol{v}\|}$$

také, že
$$\cos\alpha = \frac{e^{i\alpha} + e^{-i\alpha}}{2} = \frac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}{\|\boldsymbol{u}\| \, \|\boldsymbol{v}\|}.$$
Tento spôsob by bol, samozrejme, matematicky najčistejší, hoci z geometrického hľadiska pre nás zatiaľ nie príliš názorný. Navyše by si vyžadoval istú znalosť teórie funkcií komplexnej premennej, ktorú u čitateľa nepredpokladáme.

My sa však nebudeme rozhodovať medzi uvedenými troma alternatívami, teda uhol dvoch nenulových vektorov v unitárnom priestore vôbec nebudeme skúmať ani definovať. Najdôležitejšou interpretácia výrazu $\langle \boldsymbol{u}, \boldsymbol{v} \rangle / \|\boldsymbol{u}\| \, \|\boldsymbol{v}\|$ v unitárnom priestore je totiž jeho interpretácia v kvantovej mechanike. Tou však nie je kosínus uhla, ale – akokoľek čudne to znie – „amplitúda pravdepodobnosti“. K tejto otázke sa vrátime až v záverečnom paragrafe 17.8.

I keď na uhol dvoch vektorov v unitárnom priestore sme rezignovali, nemienime rezignovať na vzťah *ortogonality* (kolmosti)
$$\boldsymbol{x} \perp \boldsymbol{y} \;\Leftrightarrow\; \langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$$
pre vektory $\boldsymbol{x}, \boldsymbol{y} \in V$ a operáciu *ortokomplementu*
$$X^{\perp} = \{\boldsymbol{y} \in V;\, (\forall \boldsymbol{x} \in X)(\boldsymbol{x} \perp \boldsymbol{y})\}$$
množiny $X \subseteq V$, ktoré majú rovnaké vlastnosti ako v reálnom prípade. Aj pojmy *ortogonálnej* či *ortonormálnej* množiny alebo usporiadanej k-tice vektorov, a najmä *ortogonálnej* a *ortonormálnej bázy* sú definované rovnako.

Gramovým-Schmidtovým ortogonalizačným procesom – či už ho vykonáme podľa vety 13.4.5 alebo diagonalizáciou Gramovej matice $\boldsymbol{G}(\boldsymbol{\alpha}) = \big(\langle \boldsymbol{u}_j, \boldsymbol{u}_k \rangle\big)$ nejakej bázy $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ priestoru V – možno dokázať, že konečnorozmerný unitárny priestor V má ortonormálnu bázu. Vzhľadom na takúto bázu $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ nadobúda skalárny súčin na V tvar tzv. *štandardného komplexného skalárneho súčinu* na $\mathbb{C}^n$, t. j. pre ľubovoľné vektory $\boldsymbol{x}, \boldsymbol{y} \in V$ so súradnicami $(\boldsymbol{x})_{\boldsymbol{\beta}} = (x_1, \ldots, x_n)^{\mathsf{T}}$, $(\boldsymbol{y})_{\boldsymbol{\beta}} = (y_1, \ldots, y_n)^{\mathsf{T}}$ platí tzv. *Parsevalova rovnosť*
$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = (\boldsymbol{x})_{\boldsymbol{\beta}}^{\mathsf{T}} \cdot \overline{(\boldsymbol{y})_{\boldsymbol{\beta}}} = \sum_{j=1}^{n} x_j \overline{y}_j = \sum_{j=1}^{n} \langle \boldsymbol{x}, \boldsymbol{v}_j \rangle \langle \boldsymbol{v}_j, \boldsymbol{y} \rangle,$$
keďže jednotlivé zložky súradníc vektorov $\boldsymbol{x}$, $\boldsymbol{y}$ sú
$$x_j = \langle \boldsymbol{x}, \boldsymbol{v}_j \rangle, \qquad y_j = \langle \boldsymbol{y}, \boldsymbol{v}_j \rangle.$$
Pre $\boldsymbol{x} = \boldsymbol{y}$ dostávame špeciálny prípad tejto rovnosti
$$\|\boldsymbol{x}\|^2 = (\boldsymbol{x})_{\boldsymbol{\beta}}^{\mathsf{T}} \cdot \overline{(\boldsymbol{x})_{\boldsymbol{\beta}}} = \sum_{j=1}^{n} |x_j|^2 = \sum_{j=1}^{n} |\langle \boldsymbol{x}, \boldsymbol{v}_j \rangle|^2$$
(porovnaj s tvrdením 13.4.4).

17.4 Unitárne matice

Unitárne matice sú komplexnou analógiou ortogonálnych matíc. *Matica* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ sa nazýva *unitárna,* ak platí

$$\boldsymbol{A}^* \cdot \boldsymbol{A} = \boldsymbol{I}_n,$$

t. j. $\boldsymbol{A}^{-1} = \boldsymbol{A}^*$. Prvá podmienka je zrejme ekvivalentná s rovnosťou $\boldsymbol{A}^{\mathsf{T}} \cdot \overline{\boldsymbol{A}} = \boldsymbol{I}_n$, ktorá opäť hovorí, že stĺpce matice $\boldsymbol{A}$ tvoria ortonormálnu bázu unitárneho priestoru $\mathbb{C}^n$ so štandardným skalárnym súčinom. Potom rovnako $\boldsymbol{A} \cdot \boldsymbol{A}^* = \boldsymbol{I}_n$, teda aj riadky matice $\boldsymbol{A}$ tvoria ortonormálnu bázu v $\mathbb{C}^n$.

Zrejme reálna štvorcová matica $\boldsymbol{A}$, uvažovaná zároveň ako matica $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, je unitárna práve vtedy, keď je ortogonálna. Podobne ako ortogonálne matice, aj unitárne matice možno charakterizovať jednak ako matice prechodu medzi ortonormálnymi bázami v konečnorozmerných unitárnych priestoroch, jednak ako matice, násobenie ktorými zachováva štandardný skalárny súčin resp. dĺžku v $\mathbb{C}^n$.

Dôkaz nasledujúcej vety možno dostať nepatrnou modifikáciou dôkazu vety 13.5.1.

17.4.1. Veta. *Nech V je n-rozmerný unitárny priestor, $\boldsymbol{\alpha}$ je ortonormálna a $\boldsymbol{\beta}$ je ľubovoľná báza priestoru V. Potom báza $\boldsymbol{\beta}$ je ortonormálna práve vtedy, keď matica prechodu $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ z bázy $\boldsymbol{\beta}$ do bázy $\boldsymbol{\alpha}$ je unitárna.*

17.4.2. Veta. *Nech $\mathbb{C}^n$ je stĺpcový unitárny priestor so štandardným skalárnym súčinom a $\boldsymbol{A} \in \mathbb{C}^n$. Potom nasledujúce podmienky sú ekvivalentné:*
(i) $\boldsymbol{A}$ je unitárna matica;
(ii) pre všetky $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ platí $\langle \boldsymbol{A}\cdot\boldsymbol{x}, \boldsymbol{A}\cdot\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$;
(iii) pre všetky $\boldsymbol{x} \in \mathbb{C}^n$ platí $\|\boldsymbol{A}\cdot\boldsymbol{x}\| = \|\boldsymbol{x}\|$.

Dôkaz. Implikácie (i) ⇒ (ii) a (ii) ⇒ (i) možno dokázať (takmer) rovnako ako v dôkaze vety 13.5.2 a implikácia (ii) ⇒ (iii) je opäť triviálna. Stačí teda dokázať (iii) ⇒ (ii).

Keďže $\|\boldsymbol{x}\|$ je zároveň norma na zreálnení unitárneho priestoru $\mathbb{C}^n$ pochádzajúca od reálneho skalárneho súčinu $\mathrm{Re}\langle \boldsymbol{x}, \boldsymbol{y}\rangle$, pre ľubovoľné $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ platí

$$\mathrm{Re}\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \frac{1}{2}\big(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2\big),$$

a podľa tvrdenia 17.2.1 tiež $\mathrm{Im}\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \mathrm{Re}\langle \boldsymbol{x}, \mathrm{i}\boldsymbol{y}\rangle$. Teda z podmienky (iii) vyplýva

$$\begin{aligned}\operatorname{Re}\langle \boldsymbol{A}\cdot\boldsymbol{x}, \boldsymbol{A}\cdot\boldsymbol{y}\rangle &= \frac{1}{2}\big(\|\boldsymbol{A}\cdot\boldsymbol{x}+\boldsymbol{A}\cdot\boldsymbol{y}\|^2 - \|\boldsymbol{A}\cdot\boldsymbol{x}\|^2 - \|\boldsymbol{A}\cdot\boldsymbol{y}\|^2\big)\\ &= \frac{1}{2}\big(\|\boldsymbol{A}\cdot(\boldsymbol{x}+\boldsymbol{y})\|^2 - \|\boldsymbol{A}\cdot\boldsymbol{x}\|^2 - \|\boldsymbol{A}\cdot\boldsymbol{y}\|^2\big)\\ &= \frac{1}{2}\big(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2\big) = \operatorname{Re}\langle \boldsymbol{x}, \boldsymbol{y}\rangle.\end{aligned}$$

V dôsledku toho

$$\begin{aligned}\operatorname{Im}\langle \boldsymbol{A}\cdot\boldsymbol{x}, \boldsymbol{A}\cdot\boldsymbol{y}\rangle &= \operatorname{Re}\langle \boldsymbol{A}\cdot\boldsymbol{x}, \mathrm{i}(\boldsymbol{A}\cdot\boldsymbol{y})\rangle = \operatorname{Re}\langle \boldsymbol{A}\cdot\boldsymbol{x}, \boldsymbol{A}\cdot(\mathrm{i}\boldsymbol{y})\rangle\\ &= \operatorname{Re}\langle \boldsymbol{x}, \mathrm{i}\boldsymbol{y}\rangle = \operatorname{Im}\langle \boldsymbol{x}, \boldsymbol{y}\rangle,\end{aligned}$$

teda $\langle \boldsymbol{A}\cdot\boldsymbol{y}, \boldsymbol{A}\cdot\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$.

Podrobnejší popis štruktúry unitárnych matíc opäť podáme len pre rády $n \leq 2$. Unitárne matice rozmeru 1×1 zrejme splývajú s komplexnými číslami s absolútnou hodnotou 1.

17.4.3. Veta. *Matica $\boldsymbol{A} \in \mathbb{C}^{2\times 2}$ je unitárna práve vtedy, keď má tvar*

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ -\overline{b}u & \overline{a}u \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{\mathrm{i}\alpha}\cos\vartheta & \mathrm{e}^{\mathrm{i}\beta}\sin\vartheta \\ -\,\mathrm{e}^{\mathrm{i}(\omega-\beta)}\sin\vartheta & \mathrm{e}^{\mathrm{i}(\omega-\alpha)}\cos\vartheta \end{pmatrix},$$

kde čísla $a, b, u \in \mathbb{C}$ vyhovujú podmienkam $|a|^2 + |b|^2 = |u| = 1$, resp. $\alpha, \beta, \vartheta, \omega \in \mathbb{R}$ sú ľubovoľné.

Dôkaz. Priamym výpočtom sa možno presvedčiť, že každá matica v ľubovoľnom z uvedených dvoch tvarov vyhovuje podmienke $\boldsymbol{A}^*\cdot\boldsymbol{A} = \boldsymbol{I}_2$, teda je unitárna.

Nech naopak $\boldsymbol{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ je unitárna. Potom matica $\boldsymbol{A}^*\cdot\boldsymbol{A} = \boldsymbol{I}_2$ má na mieste $(1,1)$ prvok $|a|^2+|b|^2 = 1$. Označme $u = \det\boldsymbol{A}$. Ľahko možno overiť, že $|u| = 1$, preto aj matice $\operatorname{diag}(1,\overline{u})$ a $\boldsymbol{B} = \operatorname{diag}(1,\overline{u})\cdot\boldsymbol{A} = \left(\begin{smallmatrix} a & b \\ c\overline{u} & d\overline{u} \end{smallmatrix}\right)$ sú unitárne (pozri cvičenie 17.15). Potom $\det\boldsymbol{B} = 1$, preto

$$\boldsymbol{B}^{-1} = \begin{pmatrix} d\overline{u} & -b \\ -c\overline{u} & a \end{pmatrix} \qquad \text{a} \qquad \boldsymbol{B}^* = \begin{pmatrix} \overline{a} & \overline{c}u \\ \overline{b} & \overline{d}u \end{pmatrix}.$$

Keďže $\boldsymbol{B}^{-1} = \boldsymbol{B}^*$, porovnaním príslušných zložiek oboch matíc dostávame $c = -\overline{b}u$, $d = \overline{a}u$.

Ak čísla a, b, u zapíšeme v goniometrickom tvare $a = |a|\,\mathrm{e}^{\mathrm{i}\alpha}$, $b = |b|\,\mathrm{e}^{\mathrm{i}\beta}$, $u = \mathrm{e}^{\mathrm{i}\omega}$, kde $\alpha, \beta, \omega \in \mathbb{R}$, a uvedomíme si, že z podmienky $|a|^2 + |b|^2 = 1$ vyplýva existencia $\vartheta \in \langle 0, \pi/2\rangle$ takého, že $|a| = \cos\vartheta$, $|b| = \sin\vartheta$, vidíme, že maticu $\boldsymbol{A}$ možno vyjadriť aj v druhom z uvedených tvarov.

17.5 Diskrétna Fourierova transformácia

Dohodnime sa, že zložky vektorov v priestore $\mathbb{C}^n$ ($n \geq 1$) budeme číslovať od 0 do $n-1$ miesto obvyklého spôsobu od 1 do n. Vektor $\boldsymbol{x} = (x_0, x_1, \ldots, x_{n-1})^\mathsf{T} \in \mathbb{C}^n$ si budeme predstavovať ako signál vysielaný v n diskrétnych časových okamihoch $t_0, t_1, \ldots, t_{n-1}$ s rovnakými rozostupmi $\Delta t = t_k - t_{k-1} > 0$ ($1 \leq k \leq n-1$). Takisto môže ísť o vzorku vybranú z nejakého signálu $X(t)$ vysielaného v spojitom časovom intervale $\langle a, b\rangle$. Ak je rozostup Δt dosť malý a n je dosť veľké tak, že interval $\langle t_0, t_{n-1}\rangle$ pokrýva relevantný úsek „vysielacieho času“ $\langle a, b\rangle$ (napr. ak sa signál $X(t)$ periodicky opakuje po úsekoch dĺžky $n\Delta t$), možno očakávať, že vzorka

$$\boldsymbol{x} = (x_0, x_1, \ldots, x_{n-1})^\mathsf{T} = \big(X(t_0), X(t_1), \ldots, X(t_{n-1})\big)^\mathsf{T}$$

je dostatočne reprezentatívna nato, aby sa z nej pôvodný signál $X(t)$ dal dostatočne verne zrekonštruovať. Na účely takejto rekonštrukcie, no taktiež pri rôznych iných typoch spracovania signálu sa požíva *Fourierova transformácia*, ktorá v ňom oddelí zložky rôznych frekvencií a určí ich amplitúdy a fázové posuny.

V tomto paragrafe sa zoznámime len s diskrétnou Fourierovou transformáciou, ktorá pracuje na konečnorozmernom unitárnom priestore $\mathbb{C}^n$ (so štandardným skalárnym súčinom). Vďaka objavu rýchlych algoritmov na jej výpočet v 60. rokoch 20. storočia sa však oblasť jej využitia značne rozšírila. Tzv. *rýchla Fourierova transformácia* sa dnes využíva napr. pri spracovaní obrazu či iných informácií uložených v počítačoch alebo na rýchle násobenie veľkých čísel.

Komplexnú jednotku

$$\omega = \omega_n = \mathrm{e}^{2\pi\mathrm{i}/n} = \cos\frac{2\pi}{n} + \mathrm{i}\sin\frac{2\pi}{n}$$

nazývame tiež n-tá primitívna odmocnina z jednotky. [2] Potom vektor

$$\boldsymbol{f}_k = \big(1, \omega^k, \ldots, \omega^{k(n-1)}\big)^\mathsf{T},$$

kde $k \in \mathbb{Z}_n = \{0, 1, \ldots, n-1\}$, predstavuje vybranú vzorku zo spojitého signálu

$$F_k(t) = \omega^{kt} = \mathrm{e}^{2\pi\mathrm{i}kt/n} = \cos\frac{2\pi kt}{n} + \mathrm{i}\sin\frac{2\pi kt}{n}$$

v okamihoch $t = 0, 1, \ldots, n-1$. Samotný signál $F_k(t)$ popisuje kosínusové a sínusové oscilácie s relatívnou frekvenciou k vzhľadom na základný oscilačný mód $F_1(t) = \omega^t$.

[2] Presnejšie sa *n-tou primitívnou odmocninou z jednotky* nazýva každé z čísel ω^k, kde $k \in \mathbb{Z}_n$ je nesúdeliteľné s n.

17.5.1. Tvrdenie. *Vektory $\boldsymbol{f}_k$ tvoria ortogonálnu bázu*

$$\boldsymbol{\phi} = (\boldsymbol{f}_0, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_{n-1})$$

unitárneho priestoru $\mathbb{C}^n$ a všetky majú rovnakú dĺžku $\|\boldsymbol{f}_k\| = \sqrt{n}$.

Dôkaz. Pre ľubovoľné $k, l \in \mathbb{Z}_n$ platí

$$\langle \boldsymbol{f}_k, \boldsymbol{f}_l \rangle = \sum_{t=0}^{n-1} \omega^{(k-l)t} = \begin{cases} \displaystyle\sum_{t=0}^{n-1} 1 = n, & \text{ak } k = l, \\ \dfrac{\omega^{(k-l)n} - 1}{\omega^{k-l} - 1} = 0, & \text{ak } k \neq l, \end{cases}$$

lebo $\omega^0 = \omega^n = 1$ a $\omega^{k-l} \neq 1$ pre $k \neq l$.

Popri kanonickej ortonormálnej báze $\boldsymbol{\varepsilon} = (\boldsymbol{e}_0, \ldots, \boldsymbol{e}_{n-1})$ tak v $\mathbb{C}^n$ máme ďalšiu významnú ortogonálnu bázu $\boldsymbol{\phi}$ a ortonormálnu bázu $\big(1/\sqrt{n}\big)\boldsymbol{\phi}$ tvorenú vektormi $\big(1/\sqrt{n}\big)\boldsymbol{f}_k$.

Diskrétnou Fourierovou transformáciou (DFT) na priestore $\mathbb{C}^n$ nazývame lineárne zobrazenie $F\colon \mathbb{C}^n \to \mathbb{C}^n$ dané predpisom

$$F(\boldsymbol{x}) = \widehat{\boldsymbol{x}} = \big(\langle \boldsymbol{x}, \boldsymbol{f}_0 \rangle, \ldots, \langle \boldsymbol{x}, \boldsymbol{f}_{n-1} \rangle\big)^{\mathsf{T}}$$

pre $\boldsymbol{x} = (x_0, \ldots, x_{n-1})^{\mathsf{T}} \in \mathbb{C}^n$. Rozpísané do zložiek to dáva

$$\widehat{x}_k = \langle \boldsymbol{x}, \boldsymbol{f}_k \rangle = \sum_{t=0}^{n-1} x_t\, \omega^{-kt}$$

pre $k \in \mathbb{Z}_n$.

Diskrétna Fourierova transformácia tak popisuje súradnice vektora $\boldsymbol{x} \in \mathbb{C}^n$ vzhľadom na bázu $\boldsymbol{\phi}$. Presnejšie, keďže báza $n^{-1/2}\boldsymbol{\phi}$ je ortonormálna, vďaka Parsevalovej rovnosti platí

$$\boldsymbol{x} = \sum_{k=0}^{n-1} \big\langle \boldsymbol{x}, n^{-1/2}\boldsymbol{f}_k \big\rangle\, n^{-1/2}\boldsymbol{f}_k = \frac{1}{n} \sum_{k=0}^{n-1} \widehat{x}_k \boldsymbol{f}_k = \frac{1}{n}\, \boldsymbol{\phi} \cdot \widehat{\boldsymbol{x}},$$

teda

$$(\boldsymbol{x})_{\boldsymbol{\phi}} = \frac{1}{n}\, \widehat{\boldsymbol{x}}.$$

To pre jednotlivé okamihy $t \in \mathbb{Z}_n$ dáva

$$x_t = \frac{1}{n} \sum_{k=0}^{n-1} \widehat{x}_k\, \omega^{kt} = \frac{1}{n} \sum_{k=0}^{n-1} \widehat{x}_k\, \mathrm{e}^{2\pi \mathrm{i} kt/n} = \frac{1}{n} \big\langle \widehat{\boldsymbol{x}}, \overline{\boldsymbol{f}}_t \big\rangle.$$

Ak si každé z komplexných čísel $\widehat{x}_k$ vyjadríme v goniometrickom tvare $\widehat{x}_k = A_k\,\mathrm{e}^{\mathrm{i}\alpha_k} = A_k(\cos\alpha_k + \mathrm{i}\sin\alpha_k)$, dostaneme vyjadrenie

$$x_t = \frac{1}{n}\sum_{k=0}^{n-1} A_k \exp\mathrm{i}\left(\frac{2\pi kt}{n} + \alpha_k\right),$$

v ktorom už zreteľne vidno amplitúdy A_k i fázové posuny α_k jednotlivých frekvenčných módov $\boldsymbol{f}_k$.

DFT sa často používa napr. na odfiltrovanie šumov z analyzovaného signálu alebo na kompresiu dát. Zložky vektora $\widehat{\boldsymbol{x}}$ zodpovedajúce vysokým frekvenciám k (považované za šum), prípadne veľmi malým amplitúdam A_k (považované za zanedbateľné) sa vymažú (t. j. nahradia nulami), čím miesto pôvodného vektora $\boldsymbol{x} = \frac{1}{n}\sum_{k\in\mathbb{Z}_n}\widehat{x}_k\boldsymbol{f}_k$ získame „prefiltrovaný“ resp. „stlačený“ vektor

$$\frac{1}{n}\sum_{k\in I}\widehat{x}_k\boldsymbol{f}_k,$$

kde I je vhodná podmnožina indexovej množiny $\mathbb{Z}_n$.

Vzhľadom na cyklický charakter DFT však vysoké frekvencie zodpovedajú hodnotám $k\in\mathbb{Z}_n$ blízkym k $n/2$. Naopak, keďže $-k \equiv_n n-k$, frekvencie blízke k „hornej hranici“ $n-1$ sú malé, rovnako ako frekvencie blízke k „dolnej hranici“ 0. To je jeden z dôvodov, prečo je často technicky výhodnejšie miesto nezáporných zvyškov $\mathbb{Z}_n = \{0,1,\dots,n-1\}$ pracovať s tzv. *absolútne najmenšími zvyškami* modulo n, t. j. s množinou všetkých celých čísel z intervalu $(-n/2, n/2\rangle$.

Zo spätného vyjadrenia

$$x_t = \frac{1}{n}\left\langle\widehat{\boldsymbol{x}},\overline{\boldsymbol{f}}_t\right\rangle = \frac{1}{n}\left\langle\widehat{\boldsymbol{x}},\boldsymbol{f}_{n-t}\right\rangle$$

navyše vyplýva, že aj inverzná lineárna transformácia $F^{-1}\colon\mathbb{C}^n\to\mathbb{C}^n$ k diskrétnej Fourierovej transformácii má tvar diskrétnej Fourierovej transformácie. Len sa od pôvodnej transformácie F líši multiplikatívnym faktorom $1/n$ a miesto primítívnej n-tej odmocniny z jednotky $\omega = \mathrm{e}^{2\pi\mathrm{i}/n}$ používa primitívnu odmocninu $\overline{\omega} = \omega^{-1} = \mathrm{e}^{-2\pi\mathrm{i}/n}$.

Pôvodnú DFT $\boldsymbol{x}\mapsto F(\boldsymbol{x}) = \widehat{\boldsymbol{x}}$ by sme mohli rovnako dobre definovať predpisom

$$\widehat{x}_k = d\,\langle\boldsymbol{x},\boldsymbol{f}_k\rangle = d\sum_{t=0}^{n-1} x_t\,\omega^{-kt},$$

kde d je ľubovoľné kladné reálne číslo (napr. veľmi dobre by sa hodilo $d=\Delta t$ alebo $d = 2\pi/n$). Inverzná DFT $\widehat{\boldsymbol{x}}\mapsto F^{-1}(\widehat{\boldsymbol{x}}) = \boldsymbol{x}$ by mala potom tvar

$$x_t = \frac{1}{nd}\left\langle\widehat{\boldsymbol{x}},\overline{\boldsymbol{f}}_t\right\rangle = \frac{1}{nd}\left\langle\widehat{\boldsymbol{x}},\boldsymbol{f}_{n-t}\right\rangle = \frac{1}{nd}\sum_{k=0}^{n-1}\widehat{x}_t\,\omega^{kt}.$$

Popri najbežnejšej voľbe $d = 1$, $1/nd = 1/n$ sa často používa aj duálna možnosť $d = 1/n$, $1/nd = 1$, prípadne symetrická možnosť $d = 1/nd = 1/\sqrt{n}$.

Jedna z najdôležitejších a často využívaných algebraických vlastností Fourierovej transformácie spočíva v tom, že DFT prevádza cyklickú konvolúciu vektorov v $\mathbb{C}^n$ na ich súčin po zložkách (a naopak).

Násobenie vektorov $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ (no rovnako v K^n nad ľubovoľným poľom K) po zložkách je práve to „prirodzené“ násobenie

$$\boldsymbol{xy} = (x_0, x_1, \dots, x_{n-1})^\mathsf{T} \, (y_0, y_1, \dots, y_{n-1})^\mathsf{T} = (x_0 y_0, \dots, x_{n-1} y_{n-1})^\mathsf{T},$$

ktorému sme sa doteraz priam úzkostlivo vyhýbali. Ľahko však nahliadneme, že je to komutatívna, asociatívna a bilineárna binárna operácia $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}^n$ s neutrálnym prvkom $(1, 1, \dots, 1)^\mathsf{T}$.

Cyklickou konvolúciou vektorov $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ nazývame vektor

$$\boldsymbol{z} = \boldsymbol{x} * \boldsymbol{y} = (z_0, z_1, \dots, z_{n-1})^\mathsf{T},$$

ktorého zložky pre $t \in \mathbb{Z}_n$ sú dané rovnosťou

$$z_t = \sum_{r+s=t} x_r y_s = \sum_{r=0}^{n-1} x_r y_{t-r} = \sum_{s=0}^{n-1} x_{t-s} y_s,$$

pričom sčítanie sumačných indexov r, s, t prebieha v $\mathbb{Z}_n$, t. j. cyklicky modulo n.

17.5.2. Veta. *Nech $F\colon \mathbb{C}^n \to \mathbb{C}^n$ je diskrétna Fourierova transformácia. Potom pre všetky vektory $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ platí*

$$F(\boldsymbol{x} * \boldsymbol{y}) = F(\boldsymbol{x}) \, F(\boldsymbol{y}) = \widehat{\boldsymbol{x}} \, \widehat{\boldsymbol{y}},$$

a tiež naopak

$$F(\boldsymbol{xy}) = \frac{1}{n} \big(F(\boldsymbol{x}) * F(\boldsymbol{y}) \big) = \frac{1}{n} \big(\widehat{\boldsymbol{x}} * \widehat{\boldsymbol{y}} \big).$$

Dôkaz. Označme $\boldsymbol{z} = \boldsymbol{x} * \boldsymbol{y}$. Potom pre každé $k \in \mathbb{Z}_n$ platí

$$\begin{aligned} \widehat{z}_k = \langle \boldsymbol{z}, \boldsymbol{f}_k \rangle &= \sum_{t=0}^{n-1} z_t \, \omega^{-kt} = \sum_{t=0}^{n-1} \sum_{r+s=t} x_r y_s \, \omega^{-kt} \\ &= \sum_{r=0}^{n-1} x_r \, \omega^{-kr} \sum_{s=0}^{n-1} y_s \, \omega^{-ks} = \langle \boldsymbol{x}, \boldsymbol{f}_k \rangle \, \langle \boldsymbol{y}, \boldsymbol{f}_k \rangle = \widehat{x}_k \widehat{y}_k. \end{aligned}$$

Dôkaz druhej rovnosti prenechávame ako cvičenie čitateľovi.

Cyklická konvolúcia silne pripomína súčin polynómov (pozri odstavec 1.6.3). Aj pre polynómy $f(x) = \sum_{i=0}^{p} a_i x^i$ $g(x) = \sum_{j=0}^{q} b_j x^j$ sú koeficienty ich súčinu $(fg)(x) = \sum_{k=0}^{p+q} c_k x^k$ dané konvolučnou formulou

$$c_k = \sum_{i+j=k} a_i b_j = \sum_{j=0}^{k} a_{k-j} b_j,$$

len s tým rozdielom, že súčty sumačných indexov nepočítame cyklicky modulo n ale obvyklým spôsobom. Ak si však zvolíme dosť veľké n (stačí, aby $p+q < n$) a oba polynómy reprezentujeme vektormi $\boldsymbol{a} = (a_0, a_1, \ldots, a_{n-1})^{\mathsf{T}}$, $\boldsymbol{b} = (b_0, b_1, \ldots, b_{n-1})^{\mathsf{T}}$ ich koeficientov (doplnenými o nulové členy), bude vektor $\boldsymbol{c} = (c_0, c_1, \ldots, c_{n-1})^{\mathsf{T}}$ koeficientov súčinu fg splývať s cyklickou konvolúciou $\boldsymbol{c} = \boldsymbol{a} * \boldsymbol{b}$. Používané algoritmy pre DFT sú také rýchle, že túto konvolúciu je výhodnejšie počítať nie priamo, ale podľa formuly

$$\boldsymbol{c} = \boldsymbol{a} * \boldsymbol{b} = F^{-1}\big(F(\boldsymbol{a})\, F(\boldsymbol{b})\big) = F^{-1}\big(\widehat{\boldsymbol{a}}\,\widehat{\boldsymbol{b}}\big).$$

V pozičnej číselnej sústave pri nejakom základe $m \geq 2$ predstavuje ciferný zápis $a_p \ldots a_1 a_0$, kde $a_i \in \mathbb{Z}_m$, celé číslo $a_0 + a_1 m + \ldots + a_p m^p$, čiže hodnotu polynómu $f(x)$ pre $x = m$. Rýchle algoritmy používané na násobenie veľmi veľkých celých čísel sa zakladajú na uvedenom spôsobe násobenia polynómov pomocou Fourierovej transformácie. Na záver treba ešte ošetriť prípadne vzniknuté „dvojciferné cifry“, teda to, čo v prípade základu $m = 10$ nazývame prechodom cez desiatku.

Popri jednorozmernej DFT, ktorá operuje na jednoparametrických dátových súboroch, sa často používa aj viacrozmerná DFT. My sa stručne zmienime len o DFT v dvoch rozmeroch.

Zvoľme kladné celé čísla m, n. Jednotlivé zložky matice $\boldsymbol{X} = (x_{st}) \in \mathbb{C}^{m\times n}$ môžu predstavovať napr. farbu (vlnovú dĺžku) alebo intenzitu vysvietenia bodu obrazovky s diskrétnymi súradnicami $(s, t) \in \mathbb{Z}_m \times \mathbb{Z}_n$.

Dvojrozmerná diskrétna Fourierova transformácia $F\colon \mathbb{C}^{m\times n} \to \mathbb{C}^{m\times n}$ priradí matici $\boldsymbol{X} = (x_{st}) \in \mathbb{C}^{m\times n}$ maticu $F(\boldsymbol{X}) = \widehat{\boldsymbol{X}} = (\widehat{x}_{kl}) \in \mathbb{C}^{m\times n}$ so zložkami

$$\widehat{x}_{kl} = \sum_{s=0}^{m-1}\sum_{t=0}^{n-1} x_{st}\, \omega_m^{-ks}\, \omega_n^{-lt} = \sum_{s=0}^{m-1}\sum_{t=0}^{n-1} x_{st} \exp\left[-2\pi \mathrm{i}\left(\frac{ks}{m} + \frac{lt}{n}\right)\right].$$

Inverzná dvojrozmerná DFT je potom daná formulou

$$x_{st} = \frac{1}{mn} \sum_{k=0}^{m-1}\sum_{l=0}^{n-1} \widehat{x}_{kl}\, \omega_m^{ks}\, \omega_n^{lt} = \frac{1}{mn} \sum_{k=0}^{m-1}\sum_{k=0}^{n-1} \widehat{x}_{kl} \exp\left[2\pi \mathrm{i}\left(\frac{ks}{m} + \frac{lt}{n}\right)\right].$$

Dvojrozmerná DFT sa hojne využíva pri digitálnom kódovaní a spracovaní obrazu napr. v počítačovej grafike. Podobne ako jednorozmerná DFT, aj ona umožňuje efektívne spracovanie a kompresiu dát či odfiltrovanie rôznych šumov napr. z digitálnej fotografie. Ak však chceme pomocou dvojrozmernej DFT dostatočne verne spracovať napr. signál na obrazovke s $\mu \times \nu$ pixlami, treba aby m a n boli vzhľadom na μ resp. ν dosť veľké (aspoň dvojnásobné, ale radšej $m \geq \mu^2$, $n \geq \nu^2$). Počiatok diskrétneho súradného systému pixlov si zvolíme v strede obrazovky, indexové množiny $\mathbb{Z}_m$, $\mathbb{Z}_n$ nahradíme množinami absolútne najmenších zvyškov modulo m resp. n a hodnoty signálu x_{st} „za okrajom obrazovky“ doplníme nulami.

17.6 Stavové priestory v klasickej mechanike*

Cieľom tohto paragrafu nie je systematický výklad klasickej mechaniky. Hodláme len zbežne zaviesť a motivovať pojem stavového priestoru a stručne naznačiť prednosti v ňom fungujúceho Hamiltonovho formalizmu. Hlavne nám však ide o vybudovanie aspoň akej-takej názornej predstavy, či aspoň analógie, o ktorú by sme sa mohli oprieť v poslednom paragrafe, kde sa chystáme zaviesť stavové priestory v kvantovej mechanike, čo je oblasť, kde naša intuícia z viacerých dôvodov zlyháva. Uvažujme hmotný bod pohybujúci sa vo fyzikálnom priestore, ktorý prostredníctvom voľby nejakej pravouhlej súradnicovej sústavy zvykneme stotožňovať s euklidovským priestorom $\mathbb{R}^3$. Jeho priestorové súradnice (zložky jeho rádius-vektora) v čase t označme $\boldsymbol{x}(t) = (x_1(t), x_2(t), x_3(t))$. Jeho okamžitá rýchlosť v čase t je potom daná ako derivácia

$$\boldsymbol{v}(t) = \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \left(\frac{\mathrm{d}x_1}{\mathrm{d}t}, \frac{\mathrm{d}x_2}{\mathrm{d}t}, \frac{\mathrm{d}x_3}{\mathrm{d}t}\right).$$

Ak navyše poznáme jeho hmotnosť m, tak jeho pohybový stav v čase t je jednoznačne určený dvoma vektormi: polohovým vektorom $\boldsymbol{x}(t)$ a vektorom hybnosti $\boldsymbol{p}(t) = m\boldsymbol{v}(t)$ (t. j. jednotlivé zložky v_j rýchlosti a p_j hybnosti v smere súradných osí sú zviazané vzťahmi $p_j = mv_j = m(\mathrm{d}x_j/\,\mathrm{d}t)$ pre $j = 1, 2, 3$). Inak povedané, okamžitý pohybový stav hmotnej častice je jednoznačne určený jediným v čase premenným vektorom (bodom) $(\boldsymbol{x}, \boldsymbol{p})$ v *stavovom priestore* $\mathbb{R}^6$. Tento priestor možno chápať ako priamy súčin $\mathbb{R}^3 \times \mathbb{R}^3$ dvoch exemplárov euklidovského priestoru $\mathbb{R}^3$ (so štandardným skalárnym súčinom), z ktorých prvý slúži na zaznamenávanie polohy a druhý hybnosti.

Kvôli jednoduchosti ďalej predpokladajme, že pohyb nášho hmotného bodu sa odohráva v tzv. *konzervatívnom* silovom poli, určenom potenciálnou energiou $U = U(\boldsymbol{x})$, ktorá je funkciou len jeho polohy $\boldsymbol{x}$ a nezávisí od času. Pomocou nej možno vyjadriť silu pôsobiacu na časticu v danom mieste ako

$$\boldsymbol{F} = -\operatorname{grad} U = -\left(\frac{\partial U}{\partial x_1}, \frac{\partial U}{\partial x_2}, \frac{\partial U}{\partial x_3}\right).$$

S využitím Newtonovej pohybovej rovnice

$$\boldsymbol{F} = \frac{\mathrm{d}(m\boldsymbol{v})}{\mathrm{d}t} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}$$

dostávame

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = -\operatorname{grad} U, \qquad \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{v} = \frac{1}{m}\boldsymbol{p}.$$

Celková energia hmotného bodu v danom mieste a čase je súčtom jeho kinetickej a potenciálnej energie

$$H = \frac{1}{2} m \|\boldsymbol{v}\|^2 + U(\boldsymbol{x}) = \frac{1}{2m} \|\boldsymbol{p}\|^2 + U(\boldsymbol{x}).$$

Výraz $H = H(\boldsymbol{x}, \boldsymbol{p})$ nazývame *Hamiltonovou funkciou* alebo *hamiltoniánom* príslušnej pohybujúcej sa sústavy. Jednoduchý výpočet dáva

$$\begin{aligned}
\frac{\partial H}{\partial x_j} &= \frac{\partial U}{\partial x_j} = -\frac{\mathrm{d}p_j}{\mathrm{d}t},\\
\frac{\partial H}{\partial p_j} &= \frac{1}{2m} \cdot \frac{\partial \|\boldsymbol{p}\|^2}{\partial p_j} = \frac{1}{2m} \cdot \frac{\partial (p_1^2 + p_2^2 + p_3^2)}{\partial p_j} = \frac{1}{m} p_j = \frac{\mathrm{d}x_j}{\mathrm{d}t},
\end{aligned}$$

pre $j = 1, 2, 3$. Ak si ešte zavedieme skrátené označenie

$$\left(\frac{\partial H}{\partial x_1}, \frac{\partial H}{\partial x_2}, \frac{\partial H}{\partial x_3}\right) = \frac{\partial H}{\partial \boldsymbol{x}}, \qquad \left(\frac{\partial H}{\partial p_1}, \frac{\partial H}{\partial p_2}, \frac{\partial H}{\partial p_3}\right) = \frac{\partial H}{\partial \boldsymbol{p}},$$

dostávame parciálne diferenciálne rovnice, popisujúce pohyb hmotného bodu, v tzv. *Hamiltonovom tvare*

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \frac{\partial H}{\partial \boldsymbol{p}}, \qquad \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = -\frac{\partial H}{\partial \boldsymbol{x}},$$

ktorý vyniká mimoriadnou eleganciou, symetriou a jednoduchosťou.

Podobne, okamžitý stav sústavy pohybujúcich sa n hmotných bodov je jednoznačne určený dvoma usporiadanými n-ticami: $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ ich polohových vektorov a $(\boldsymbol{p}_1, \dots, \boldsymbol{p}_n)$ ich hybností, ktoré možno výhodne reprezentovať ako blokovú maticu

$$(\boldsymbol{X}, \boldsymbol{P}) = \begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{p}_1 \\ \vdots & \vdots \\ \boldsymbol{x}_n & \boldsymbol{p}_n \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & x_{13} & p_{11} & p_{12} & p_{13} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & p_{n1} & p_{n2} & p_{n3} \end{pmatrix} \in \mathbb{R}^{n\times 6},$$

pričom jednotlivé zložky $x_{ij} = x_{ij}(t)$ matice $\boldsymbol{X} = \boldsymbol{X}(t)$, resp. $p_{ij} = p_{ij}(t)$ matice $\boldsymbol{P} = \boldsymbol{P}(t)$ predstavujú j-tu súradnicu polohy resp. hybnosti i-tej častice a sú zviazané vzťahom $p_{ij} = m_i(\mathrm{d}x_{ij}/\,\mathrm{d}t)$, kde m_i je jej hmotnosť. Teda okamžitý pohybový stav sústavy n hmotných bodov sme zachytili ako *jediný* od času závislý bod či vektor $(\boldsymbol{X}, \boldsymbol{P})$ v $6n$-rozmernom *stavovom priestore* $\mathbb{R}^{n\times 6}$, chápanom ako priamy súčin $2n$ exemplárov euklidovského priestoru $\mathbb{R}^3$, zodpovedajúcich zložkám matice $(\boldsymbol{x}_i, \boldsymbol{p}_i) \in (\mathbb{R}^3)^{n\times 2}$, nazývaným tiež *zovšeobecnené súradnice* a *zovšeobecnené hybnosti* sústavy.

Celkom analogicky ako v prípade jedinej častice, i pre n hmotných bodov v konzervatívnom silovom poli možno odvodiť, že ich pohyb sa riadi parciálnymi diferenciálnymi rovnicami

$$\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} = \frac{\partial H}{\partial \boldsymbol{P}}, \qquad \frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t} = -\frac{\partial H}{\partial \boldsymbol{X}}.$$

ktoré nazývame *Hamiltonovými rovnicami.* Uvedené výrazy označujú matice

$$\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} = \left(\frac{\mathrm{d}x_{ij}}{\mathrm{d}t}\right)_{n\times 3}, \qquad \frac{\partial H}{\partial \boldsymbol{X}} = \left(\frac{\partial H}{\partial x_{ij}}\right)_{n\times 3},$$

$$\frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t} = \left(\frac{\mathrm{d}p_{ij}}{\mathrm{d}t}\right)_{n\times 3}, \qquad \frac{\partial H}{\partial \boldsymbol{P}} = \left(\frac{\partial H}{\partial p_{ij}}\right)_{n\times 3},$$

a $H = H(\boldsymbol{X}, \boldsymbol{P})$ je opäť Hamiltonova funkcia, vyjadrujúca celkovú energiu sústavy. Jej explicitný tvar je

$$H = \sum_{i=1}^{n} \frac{1}{2m_i} \|\boldsymbol{p}_i\|^2 + \sum_{i=1}^{n} U_i(\boldsymbol{x}_i) + \sum_{i<j} U_{ij}(\boldsymbol{x}_i - \boldsymbol{x}_j),$$

kde prvý člen predstavuje súčet kinetických energií jednotlivých častíc, druhý súčet ich potenciálnych energií $U_i(\boldsymbol{x}_i)$ vo vonkajšom silovom poli a tretí súčet potenciálnych energií $U_{ij}(\boldsymbol{x}_i - \boldsymbol{x}_j)$ vzájomného pôsobenia i-tej a j-tej častice, ktoré (podľa tretieho Newtonovho zákona) závisia len od vektora $\boldsymbol{x}_i - \boldsymbol{x}_j$. Zvláštnu úlohu vo fyzike hrajú tzv. *centrálne sily*, pri ktorých energie $U_{ij} = U_{ij}\big(\|\boldsymbol{x}_i - \boldsymbol{x}_j\|\big)$ závisia len od dĺžok vektorov $\boldsymbol{x}_i - \boldsymbol{x}_j$.

Číslo $3n$ udáva *počet stupňov voľnosti* danej sústavy. Keby sme miesto hmotných bodov uvažovali reálne pevné telesá, ktoré sa okrem translačného pohybu môžu navyše otáčať okolo ľubovoľnej osi, pohybový stav každého z nich by bol okrem polohy a hybnosti charakterizovaný ešte troma uhlami a troma momentmi hybnosti vzhľadom na dané súradné osi. Celá sústava n telies by tak mala nie $3n$ ale $6n$ stupňov voľnosti a jej okamžitý stav by predstavoval nejaký bod v $12n$-rozmernom priestore (maticu rozmeru $n \times 12$).

Na konkrétnej dimenzii stavového priestoru však teraz nezáleží. Náš výklad klasickej mechaniky totiž v tejto chvíli ukončíme a sústredíme sa len na stručné zhodnotenie prínosu jej matematického modelovania v mnohorozmernom priestore. Predstavme si, že by sme pohyb sústavy n hmotných bodov či telies popisovali v trojrozmernom priestore. Prišlo by nám rozpletať motanicu n trajektórií a pre každú z nich zvlášť sledovať časový vývoj rýchlosti resp. hybnosti, nehovoriac už o uhloch a momentoch. Namiesto toho sme celý tento spletenec zakódovali do *trajektórie jediného bodu* v priestore dimenzie $2d$, kde d je počet stupňov voľnosti celej sústavy. Časový vývoj tejto trajektórie sa riadi Hamiltonovými rovnicami, to znamená, že na predpoveď budúceho alebo rekonštrukciu minulého stavu stačí poznať hamiltonián sústavy a jediný okamžitý stav – tzv. *počiatočnú podmienku*.

Teda opustením trojrozmerného fyzikálneho priestoru a využitím matematických priestorov vyšších dimenzií možno získať účinné prostriedky aj na popis dejov odohrávajúcich sa v pôvodnom trojrozmernom priestore. Rozbitie okov, ktorými nás pútajú tri rozmery fyzikálneho priestoru, a následný rozlet do vyšších dimenzií, vedený potrebami klasickej mechaniky, je dodnes plne nedoceneným titanským výkonom geniálneho intelektu Lagrangeovho. Bol to precedens, ktorý mal za následok, že sa matematika i fyzika v priestoroch ľubovoľných dimenzií postupne pevne a natrvalo usadili a zároveň tento ich „úlet“ plne legalizoval. Následné dotvorenie Lagrangeovej mechaniky Hamiltonom do uvedenej definitívnej podoby – akokoľvek matematicky elegantnej – už nenesie onú pečať oslobodzujúceho titanského činu.

17.7 Kvantová mechanika

Kvantová mechanika je teóriou mikrosveta – popisuje správanie atómov, elementárnych častíc (elektrónov, protónov, neutrónov atď.) a svetelných kvánt (fotónov), teda objektov, s ktorými ľudstvo nemalo a nemá nijakú priamu, bezprostrenú zmyslovú skúsenosť. V dôsledku toho si pre javy mikrosveta nemohlo vytvoriť náležité porozumenie ani jazykové pojmy, pomocou ktorých by sa o týchto javoch dalo primerane hovoriť. Zostávajú teda len analógie vypracované na základe klasických fyzikálnych predstáv a makroskopických prejavov atomárnych a subatomárnych objektov. Tieto prejavy sú však značne rôznorodé, ba priam protichodné, takže klasické fyzikálne teórie sú schopné adekvátne popísať vždy len istú ich stránku. Skĺbenie ich dielčich popisov do jedného logicky konzistentného celku sa v rámci klasickej fyziky ukazuje ako principiálne nemožné.

Ešte Newton si svetlo predstavoval ako prúd malých svetelných čiastočiek, tzv. *korpuskúl*. Avšak javy ako interferencia, ohyb a polarizácia svetla jasne svedčia o jeho vlnovej povahe. Aj lom svetla a spektrálny rozklad na rozhraní vzduch/voda resp. vzduch/sklo tak, ako ich pozorujeme, by sa dali v rámci korpuskulárnej teórie vysvetliť, len keby rýchlosť svetla vo vzduchu bola menšia než v druhom prostredí a zároveň rýchlosť červeného svetla vo vode resp. skle bola najnižšia a rýchlosť fialového svetla najvyššia spomedzi rýchlostí v rámci viditeľného spektra. Experimenty však ukázali, že je to naopak. Preto svetlo dnes chápeme prevažne ako elektromagnetické vlnenie o vlnovej dĺžke z istého rozsahu. Ale napr. tzv. *fotoelektrický jav* (t. j. uvoľnenie elektrónov z atómov pôsobením svetla) sa v rámci vlnovej teórie svetla nepodarilo vysvetliť. Podarilo sa to až Einsteinovi pomocou Planckovej kvantovej hypotézy. Za tým účelom však bolo potrebné považovať svetlo za prúd diskrétnych energetických kvánt (s energiou $E = h\nu$, kde $h \approx 6,626 \cdot 10^{-34}\,\mathrm{J\,s}$ je *Planckova konštanta* a ν je frekvencia príslušného svetelného žiarenia). Tieto svetelné kvantá dostali názov *fotóny*. Svetlo má teda – aspoň v rámci súčasných fyzikálnych predstáv – zároveň vlnovú i korpuskulárnu povahu. No čestnejšie by asi bolo povedať, že pravdepodobne nie je ani jedným ani druhým, len sa prejavuje raz ako vlnenie, raz ako prúd fotónov.

Naopak, elektróny si predstavujeme ako malé, hmotné, elektricky nabité guľôčky, obiehajúce okolo atómového jadra, podobne ako planéty okolo Slnka, alebo sa len tak voľne „poflakujúce“ v kovoch. Ale v mnohých prípadoch sa elektróny správajú ako vlnenie: napr. homogénny zväzok elektrónov sa pri prechode kryštalickou mriežkou, rovnako ako svetelný lúč, ohýba a vytvára difrakčný obrazec. Napokon i samotný planetárny model atómu je v rozpore s predstavou elektrónu ako nabitej častice. Elektrón pohybujúci sa po uzavretej krivke (teda premennou rýchlosťou) v elektrickom poli jadra by totiž musel vyžarovaním strácať energiu a rýchlo spadnúť na jadro. Väčšina atómov je však značne stabilných a vlnová teória elektrónu vie túto ich stabilitu vysvetliť. Podobne je tomu aj s inými elementárnymi časticami – tak ako svetlo sa raz prejavujú ako vlnenie, inokedy ako častice. Najskôr nebudú ani jedno ani druhé, presnejšie otázka, „zmenšením akého makroskopického javu sú naozaj“, zrejme vôbec nemá zmysel. Avšak práve *vlnovo-korpuskulárny dualizmus* kvantovej mechaniky umožňuje opísať a vysvetliť ich prejavy podstatne úplnejším spôsobom, než je v silách ktoréhokoľvek jedného z oboch hľadísk.

Fyzikálne experimenty majú v zásade štatistický charakter. Aj pri zabezpečení prakticky identických podmienok sa pri opakovaní pokusu objavuje istý rozptyl nameraných hodnôt pozorovaných veličín. V klasickej fyzike zvykneme malé odchýlky od priemeru pripisovať na vrub nepresnosti meracích prístrojov alebo si ich vykladáme ako dôsledky istých (malých) zmien v podmienkach pokusu, ku ktorým dochádza pri jeho opakovaní a ktorým nevieme zabrániť. Stále si zachovávame presvedčenie, že s „absolútne presnými“ prístrojmi a pri „absolútne identických“ podmienkach by rozptyl nameraných hodnôt bol nulový. Veľké odchýlky nás vedú k hypotézam o *skrytých parametroch*, t. j. neznámych fyzikálnych veličinách ktoré, spolu s veličinami fixovanými podmienkami pokusu, majú tiež vplyv na jeho výsledok. Neraz vedú takéto pokusy k objaveniu dovtedy skrytých parametrov a ich zahrnutiu do novoformulovaných fyzikálnych zákonov. Krátko povedané, ani značný rozptyl nameraných hodnôt pri klasických experimentoch nijako nenarúša našu dôveru v kauzalitu a matematický determinizmus zákonitostí fyzikálneho sveta.

Na atomárnej a subatomárnej úrovni však táto naša viera už ďalej neobstojí. Sám výsledok kvantovomechanického experimentu je vlastne prírodou uskutočneným štatistickým spracovaním (spriemerovaním) veľkého množstva individuálnych pokusov s identickými časticami. Napríklad difrakčný obrazec vytvorený homogénnym zväzkom elektrónov prechádzajúcich kryštálom nám vlastne premennou intenzitou zobrazuje čosi ako rozdelenie pravdepodobnosti dopadu jednotlivého elektrónu na rôzne miesta detekčnej plochy. Podobne, intenzity spektrálnych čiar nejakého atómu predstavujú, veľmi zjednodušene povedané, rozdelenie pravdepodobnosti pobytu jeho elektrónov na rozličných orbitách, t. j. energetických hladinách.

Doteraz všetky pokusy vysvetliť chovanie elementárnych častíc v rámci deterministickej paradigmy klasickej fyziky zlyhali. Niečo ako „priamočiary analóg klasickej teórie“ vedie k strate lokálneho charakteru, teda – obrazne povedané – si vyžaduje pripustiť okamžité (t. j. šíriace sa nekonečnou rýchlosťou) vzájomné pôsobenie častíc na diaľku. Zavedenie skrytých parametrov, ktoré malo lokálny charakter teórie zachrániť, sa ukázalo neudržateľné, keď sa na jednej strane deduktívne odvodilo, že z nich – nezávisle od konkrétneho tvaru teórie – vyplývajú tzv. *Bellove nerovnosti*, a na druhej strane sa experimentálne zistilo, že pravdepodobnosti atómových procesov týmto nerovnostiam nevyhovujú.

Navyše pokusy o spresnenie podmienok experimentu alebo výsledkov merania narazili na principiálne neprekonateľnú bariéru. Každé meranie totiž ovplyvňuje skúmaný jav. Tieto vplyvy možno na makroskopickej úrovni zanedbať, no na atomárnej a subatomárnej úrovni už hrajú významnú úlohu. Napríklad lúč svetla prakticky nijako neovplyvní pohyb letiaceho kameňa, no z hľadiska elektrónu, ktorého polohu zisťujeme pomocou svetelného žiarenia, je energia naň dopadajúcich fotónov nezanedbateľná. Čím presnejšie chceme fixovať polohu elektrónu, tým kratšiu vlnovú dĺžku λ, teda tým väčšiu frekvenciu $\nu = c/\lambda$, musí mať použité svetlo. Tým väčšia je aj energia $E = h\nu$ jednotlivého fotónu a jeho vplyv na „osvetlený“ elektrón. K týmto otázkam sa ešte vrátime, keď sa budeme zaoberať vzťahom meraných veličín a istých lineárnych operátorov v kvantovej mechanike.

Kvantová mechanika teda nepopisuje správanie elementárnych častíc „ako takých“, ale výsledky ich meraní, t. j. ich interakcií s vhodnými makroskopickými objektmi – meracími prístrojmi. Výsledkami takýchto meraní (okrem výnimočných prípadov) nie sú konkrétne hodnoty meraných veličín ale pravdepodobnostné rozdelenia, určujúce s akou pravdepodobnosťou nadobúda meraná veličina tú-ktorú hodnotu. Priraďovať objektom mikrosveta jednoznačne určené „objektívne“ hodnoty makroskopických veličín (ako poloha, hybnosť, rýchlosť, energia a pod.), ktoré by im prislúchali nezávisle od spôsobu merania, nemá podľa všeobecne prevládajúceho poňatia kvantovej mechaniky vôbec žiadny zmysel. V zhode s tým nie je úlohou kvantovej mechaniky predpovedať určité hodnoty a časový vývoj meraných veličín, ale predpovedať ich pravdepodobnostné rozdelenia a časový vývoj týchto rozdelení. Ten je podriadený *Schrödingerovej rovnici* (t. j. istej parciálnej diferenciálnej rovnici, ktorá je komplexným analógom vlnovej rovnice) pre tzv. *vlnovú funkciu* (t. j. určitý druh v čase premennej *amplitúdy pravdepodobnosti*).

Základné postuláty kvantovej mechaniky nemajú charakter v názorných pojmoch sformulovaných a experimentálne overených axióm. Práve naopak, niektoré z týchto princípov zo zásadných dôvodov ani experinentálne overiť nemožno. Kvantovej mechanike vlastne poriadne nerozumieme a diskusie o tej „pravej“ interpretácii jej matematického formalizmu sprevádzajú túto disciplínu od čias jej vzniku až podnes. Čitateľ by preto nemal k štúdiu tohto a nasledujúceho paragrafu pristupovať s príliš optimistickým očakávaním, že po ich prečítaní nebodaj kvantovej mechanike porozumie. Náš cieľ je ďaleko skromnejší. Autor tejto učebnice si bude považovať za úspech, ak sa mu podarí čitateľa presvedčiť, že komplexný skalárny súčin a unitárne priestory sú naozaj na niečo dobré, a nie sú to len výplody bezuzdnej fantázie zovšeobecňovaniachtivých matematikov.

Kvantová mechanika totiž – a to je na nej asi to najzáhadnejšie – skvele funguje. Jej makroskopicky overiteľné predpovede sa plnia s udivujúcou presnosťou, siahajúcou zrejme ďalej než presnosť našich meraní. Práve na tomto fakte je založená naša dôvera v jej matematický formalizmus. Aplikácie kvantovej mechaniky sa dnes už stali neoddiskutovateľnou realitou nášho sveta. Pritom nemáme na mysli len tie najznámejšie príklady vojenského použitia, ako sú nukleárne zbrane, ani ekologicky kontroverzné využitie v jadrovej energetike, ale tiež aplikácie ďaleko subtílnejšie, napr. lasery, masery, polovodiče, supravodivosť, diagnostické a terapeutické vyžitie rádioizotopov v medicíne, či známu uhlíkovú metódu určovania veku skamenelín a archeologických nálezov. V súčasnosti značnú pozornosť vzbudzuje zatiaľ prevažne v teoretickej rovine sa rozvíjajúca oblasť informatiky známa pod názvom *quantum computing*. Už teraz však možno povedať, že zostrojenie fungujúceho kvantového počítača by asi od základu zmenilo všeobecne uznávané názory na možnosti počítačov a, okrem iného, by si zrejme vynútilo fundamentálne prepracovanie používaných systémov kódovania, dekódovania a ochrany informácií.

17.8 Stavovové priestory v kvantovej mechanike*

Matematický model kvantovej mechaniky sa zakladá na niekoľkých východzích princípoch, ktoré nemožno odvodiť z jednoduchších predpokladov, a je preto nutné ich postulovať. V tomto paragrafe sa v krátkosti zoznámime len s dvoma z nich, ktoré budeme stručne nazývať *princípom superpozí-*

cie a *princípom amplitúd pravdepodobnosti*. Stretnutie s ďalšími princípmi kvantovej mechaniky odložíme až na dobu, keď podrobnejšie preskúmame štruktúru a vlastnosti tzv. *samoadjungovaných lineárnych operátorov* v unitárnych priestoroch, ktorými hodláme reprezentovať merateľné fyzikálne veličiny kvantovomechanických sústav.

V analógii s klasickou mechanikou je každému kvantovomechanickému systému (ako napr. elektrón, atóm vodíka a pod.) priradený istý *stavový priestor* V, ktorého prvky nazývame *stavovými vektormi*. V tomto prípade však ide o unitárny priestor, v typickom prípade nekonečnorozmerný. Konečnorozmerné unitárne priestory sa vyskytujú len ako stavové priestory pevne lokalizovaných častíc – takým je napr. dvojrozmerný priestor spinových stavov elektrónu. V nekonečnorozmernom prípade navyše pristupuje požiadavka *úplnosti* priestoru V, ktorá je v konečnorozmernom prípade splnená automaticky. Dodajme, že úplný unitárny priestor sa nazýva (komplexný) *Hilbertov priestor*. Popri nich sa študujú aj reálne Hilbertove priestory t. j. úplné reálne vektorové priestory so skalárnym súčinom.[3]

Stavy kvantovomechanického systému však nezodpovedajú priamo vektorom jeho stavového priestoru V, ale jeho jednorozmerným podpriestorom, t. j. priamkam alebo *lúčom* $[\boldsymbol{u}]$, kde $\mathbf{0} \neq \boldsymbol{u} \in V$. Lineárne závislé nenulové vektory $\boldsymbol{u}, \boldsymbol{v} \in V$ tak reprezentujú ten istý stav. Uvedené stavy sa presnejšie nazývajú *čistými stavmi*; okrem nich sa vyskytujú ešte tzv. *zmiešané stavy*, ktorými sa tu ale nebudeme zaoberať. Každý (čistý) stav kvantovomechanického systému tak možno reprezentovať nejakým vektorom $\boldsymbol{u}$ jeho stavovového priestoru jednotkovej dĺžky $\|\boldsymbol{u}\| = 1$. Táto reprezentácia je jednoznačná až na tzv. *fázový činiteľ*, t. j. číslo tvaru $\mathrm{e}^{\mathrm{i}\vartheta} = \cos\vartheta + \mathrm{i}\sin\vartheta$, kde $0 \leq \vartheta < 2\pi$. Trochu nepresne nazývame (čistými) stavmi aj jednotkové vektory $\boldsymbol{u} \in V$.[4]

Súčet $\boldsymbol{u} + \boldsymbol{v}$ nazývame *superpozíciou stavov* (predstavovaných vektormi) $\boldsymbol{u}, \boldsymbol{v} \in V$. Trochu všeobecnejšie nazývame superpozíciou stavových vektorov $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ s koeficientmi $a_1, \ldots, a_n$ ich lineárnu kombináciu $a_1\boldsymbol{u}_1 + \ldots + a_n\boldsymbol{u}_n$. Podobne ako stavové vektory $\boldsymbol{u}_j$, ani koeficienty a_j nie sú určené jednoznačne a nemajú bezprostredný fyzikálny význam. Značne reduntantná reprezentácia stavov kvantovomechanického systému Hilbertovým priestorom V je vedená väčšmi potrebou zabezpečiť hladké fungovanie matematického for-

[3] Topologická podmienka *úplnosti* (reálneho či komplexného) vektorového prietoru so skalárnym súčinom znamená, že ľubovoľná postupnosť $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ vo V spĺňajúca *Cauchyho-Bolzanovu podmienku*, t. j. $(\forall\, \varepsilon > 0)(\exists\, m \in \mathbb{N})(\forall\, n \geq m)(\|\boldsymbol{x}_n - \boldsymbol{x}_m\| < \varepsilon)$, už konverguje k nejakému vektoru $\boldsymbol{x} \in V$, t. j. $\lim_{n\to\infty} \|\boldsymbol{x}_n - \boldsymbol{x}\| = 0$.

[4] Treba si však uvedomiť, že priamka alebo lúč $[\boldsymbol{u}] = \mathbb{C}\,\boldsymbol{u}$ vo vektorovom priestore nad poľom $\mathbb{C}$ je vlastne dvojrozmerný podpriestor čiže rovina, ak ju chápeme ako vektorový priestor nad poľom $\mathbb{R}$. Na druhej strane, niektorí fyzici zastávajú stanovisko, že stotožňovať vektory $\boldsymbol{u}$ a $\mathrm{e}^{\mathrm{i}\vartheta}\,\boldsymbol{u}$ líšiace sa fázovým činiteľom $\mathrm{e}^{\mathrm{i}\vartheta} \neq 1$ je scestné, lebo práve v závislosti na ňom môžu interferovať rôznymi spôsobmi. Pre nich sú teda čistými stavmi priamo jednotkové vektory $\boldsymbol{u} \in V$.

malizmu, než snahou o „bezprostredné zobrazenie skutočnosti".

Predstavme si kvôli jednoduchosti hypotetickú elementárnu časticu, ktorá môže zaujímať n rôznych polôh v priestore, prípadne n energetických hladín, ktoré očíslujeme od 1 po n. Zdanlivo, t. j. z klasického hľadiska, by jej stavový priestor mal byť totožný s množinou $\{1, 2, \ldots, n\}$. No kvantová mechanika vo všeobecnosti neumožňuje určiť, v ktorom zo stavov $1, \ldots, n$ sa častica nachádza. Môže nám poskytnúť len pravdepodobnosti $p_1, \ldots, p_n$ výskytu v jednotlivých stavoch. Takéto pravdepodobnostné rozdelenie $\boldsymbol{p} = (p_1, \ldots, p_n)$, teda nový kandidát na to, čo by sme mohli nazvať stavom systému, leží vo vektorovom priestore $\mathbb{R}^n$. Takže sa zdá, že pri pravdepodobnostnej formulácii kvantovej mechaniky by sme mali vystačiť s reálnymi vektorovými priestormi (a s reálnym skalárnym súčinom). Komplexné unitárne priestory, v našom prípade to bude $\mathbb{C}^n$, sú potrebné na zachytenie vlnového charakteru kvantovomechanických javov. Pri zobrazení v $\mathbb{R}^n$ (či v inom reálnom vektorovom priestore), by boli všetky vlnenia nevyhnutne vo fáze, takže typické vlnové javy, ako napr. interferenciu, by nebolo možné zachytiť. Normovaný stavový vektor $\boldsymbol{u} = (u_1, \ldots, u_n) \in \mathbb{C}^n$ teda v sebe nesie informáciu nielen o pravdepodobnostnom rozdelení $\boldsymbol{p} = \big(|u_1|^2, \ldots, |u_n|^2\big)$ ale aj o fázach „pravdepodobnostnej vlny", presnejšie *amplitúdy pravdepodobnosti*,

$$\boldsymbol{u} = \big(|u_1|\, \mathrm{e}^{\mathrm{i}\vartheta_1}, \ldots, |u_n|\, \mathrm{e}^{\mathrm{i}\vartheta_n}\big).$$

V typickom prípade sa stavový priestor kvantovomechanickej sústavy realizuje ako nekonečnorozmerný Hilbertov priestor istých komplexných funkcií definovaných na klasickom „fyzikálnom" priestore sústavy. Napríklad stavový priestor V pre realistickejšiu časticu – elektrón v trojrozmernom priestore $\mathbb{R}^3$ – pozostáva z určitých „dostatočne dobre" integrovateľných komplexných funkcií $\mathbb{R}^3 \to \mathbb{C}$ so skalárnym súčinom[5]

$$\langle f, g\rangle = \iiint_{\mathbb{R}^3} f(x, y, z)\, \overline{g}(x, y, z)\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z.$$

Istotne netreba zvlášť podotýkať, že ide o nekonečnorozmerný priestor. Pravdepodobnosť, že sa elektrón s normovanou stavovou funciou f nachádza v nejakej (*merateľnej*) množine $A \subseteq \mathbb{R}^3$, t. j. v istej časti priestoru, je vyjadrená trojným integrálom

$$\iiint_A |f(x, y, z)|^2\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z.$$

[5]Komplexnú funkciu reálnej premmennej $h\colon M \to \mathbb{C}$, kde $M \subseteq \mathbb{R}$ je nejaký (či už ohraničený alebo neohraničený) interval, možno písať v tvare $h = h_0 + \mathrm{i}h_1$, kde $h_0(x) = \operatorname{Re} h(x)$ a $h_1(x) = \operatorname{Im} h(x)$, čiže $h_0, h_1\colon M \to \mathbb{R}$. Potom h je integrovateľná (v akomkoľvek bližšie špecifikovanom zmysle) práve vtedy, keď h_0 aj h_1 sú integrovateľné, pričom $\int_M h(x)\, \mathrm{d}x = \int_M h_0(x)\, \mathrm{d}x + \mathrm{i} \int_M h_1(x)\, \mathrm{d}x$. Podobne pre funkcie $h\colon M \to \mathbb{C}$, kde $M \subseteq \mathbb{R}^n$.

Samotná funkcia f taká, že $\|f\|^2 = \langle f, f\rangle = 1$, predstavuje svojimi jednotlivými hodnotami $f(x,y,z)$ *amplitúdu hustoty pravdepodobnosti*, že sa daná častica nachádza v tom-ktorom bode $(x,y,z) \in \mathbb{R}^3$. [6]

Ak sa kvantovomechanický systém, reprezentovaný stavovým priestorom V, nachádza v istom okamihu v stave $[\boldsymbol{u}]$, tak pravdepodobnosť, že ho „vzápätí“ nájdeme v stave $[\boldsymbol{v}]$, sa rovná číslu

$$\frac{|\langle \boldsymbol{u}, \boldsymbol{v}\rangle|^2}{\|\boldsymbol{u}\|^2\,\|\boldsymbol{v}\|^2} = \frac{|\langle \boldsymbol{u}, \boldsymbol{v}\rangle|^2}{\langle \boldsymbol{u}, \boldsymbol{u}\rangle\langle \boldsymbol{v}, \boldsymbol{v}\rangle}.$$

Všimnite si, že uvedený výraz naozaj závisí len na stavoch $[\boldsymbol{u}]$, $[\boldsymbol{v}]$ a nie na samotných vektoroch $\boldsymbol{u}$, $\boldsymbol{v}$. Keďže ide o nezáporné reálne číslo, ktoré je podľa Cauchyho-Schwartzovej nerovnosti ≤ 1, môžeme ho oprávnene interpretovať ako pravdepodobnosť. Pre normované vektory $\boldsymbol{u}$, $\boldsymbol{v}$ je táto pravdepodobnosť rovná priamo $|\langle \boldsymbol{u}, \boldsymbol{v}\rangle|^2$, čo je vlastne druhá mocnina dĺžky kolmého priemetu $\mathrm{pr}_{[\boldsymbol{v}]}(\boldsymbol{u}) = \langle \boldsymbol{u}, \boldsymbol{v}\rangle\boldsymbol{v}$ vektora $\boldsymbol{u}$ do podpriestoru $[\boldsymbol{v}]$. Samotné komplexné číslo $\langle \boldsymbol{u}, \boldsymbol{v}\rangle$, nazývané *amplitúdou pravdepodobnosti* prechodu zo stavu $\boldsymbol{u}$ do stavu $\boldsymbol{v}$, nemá ani v tomto prípade bezprostredný fyzikálny význam. Umožňuje však elegantné výpočty.

Ak je stavový priestor V konečnorozmerný a $\boldsymbol{\gamma} = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ je jeho ortonormálna báza, tak ľubovoľný vektor $\boldsymbol{u} = a_1\boldsymbol{w}_1 + \ldots + a_n\boldsymbol{w}_n \in V$ je normovaný práve vtedy, keď pre jeho súradnice $a_j = \langle \boldsymbol{u}, \boldsymbol{w}_j\rangle$ platí

$$|a_1|^2 + \ldots + |a_n|^2 = 1.$$

Z postulátov kvantovej mechaniky vyplýva, že číslo $|a_j|^2$ predstavuje pravdepodobnosť, že systém v stave $\boldsymbol{u}$ prejde do niektorého z bázických stavov $\boldsymbol{w}_j$. Uvedená rovnosť tak hovorí, že všetky tieto pravdepodobnosti dávajú v súčte jednotku. [7] Ak je aj vektor $\boldsymbol{v} = b_1\boldsymbol{w}_1 + \ldots + b_n\boldsymbol{w}_n$ normovaný, tak pre amplitúdu pravdepodobnosti prechodu zo stavu $\boldsymbol{u}$ do stavu $\boldsymbol{v}$ dostávame

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \sum_{j=1}^{n} \langle \boldsymbol{u}, \boldsymbol{w}_j\rangle\langle \boldsymbol{w}_j, \boldsymbol{v}\rangle = \sum_{j=1}^{n} a_j\bar{b}_j.$$

Inak povedané, amplitúda pravdepodobnosti prechodu $\boldsymbol{u} \to \boldsymbol{v}$ je súčtom súčinov amplitúd pravdepodobnosti prechodov $\boldsymbol{u} \to \boldsymbol{w}_j \to \boldsymbol{v}$ cez jednotlivé bázické stavy. Posledná formula umožňuje ďalekosiahle zovšeobecnenie.

[6] Priamo o amplitúde pravdepodobnosti nemôžeme hovoriť, lebo pravdepodobnosť, že sa častica nachádza v bode (x,y,z) nie je $|f(x,y,z)|^2$ ale 0.

[7] K takýmto prechodom, pri ktorých si systém „musí vybrať“ jeden zo stavov istej ortonormálnej bázy, dochádza napr. pri meraní hodnôt fyzikálnych veličín kvantovo-mechanického systému.

Ľubovoľnú konečnú postupnosť $(\boldsymbol{u}, \boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k, \boldsymbol{v})$ normovaných vektorov v stavovom priestore V nazývame *trajektóriou* dĺžky $k+1$ z $\boldsymbol{u}$ do $\boldsymbol{v}$. Táto postupnosť zrejme predstavuje obdobu klasickej trajektórie. Jej amplitúdou pravdepodobnosti rozumieme súčin $\langle \boldsymbol{u}, \boldsymbol{x}_1 \rangle \langle \boldsymbol{x}_1, \boldsymbol{x}_2 \rangle \dots \langle \boldsymbol{x}_{k-1}, \boldsymbol{x}_k \rangle \langle \boldsymbol{x}_k, \boldsymbol{v} \rangle$, chápaný ako „komplexná pravdepodobnosť" prechodu systému zo stavu $\boldsymbol{u}$ do stavu $\boldsymbol{v}$ po naznačenej trajektórii. Matematickou indukciou podľa k možno dokázať, že pre ľubovoľné stavy $\boldsymbol{u}, \boldsymbol{v}$ a ortonormálne bázy $\boldsymbol{\gamma}_1 = (\boldsymbol{w}_{11}, \dots, \boldsymbol{w}_{1n})$,..., $\boldsymbol{\gamma}_k = (\boldsymbol{w}_{k1}, \dots, \boldsymbol{w}_{kn})$ platí

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{j_1, \dots, j_k = 1}^{n} \langle \boldsymbol{u}, \boldsymbol{w}_{1j_1} \rangle \langle \boldsymbol{w}_{1j_1}, \boldsymbol{w}_{2j_2} \rangle \dots \langle \boldsymbol{w}_{k-1 j_{k-1}}, \boldsymbol{w}_{kj_k} \rangle \langle \boldsymbol{w}_{kj_k}, \boldsymbol{v} \rangle.$$

Inak povedané, amplitúdu pravdepodobnosti prechodu $\boldsymbol{u} \to \boldsymbol{v}$ možno vypočítať ako súčet amplitúd pravdepodobnosti cez všetky trajektórie $(\boldsymbol{u}, \boldsymbol{w}_{1j_1}, \dots, \boldsymbol{w}_{kj_k}, \boldsymbol{v})$ pevnej dĺžky $k+1$.

Na značne sofistikovanom nekonečnorozmernom zovšeobecnení poslednej formuly je založené vyjadrenie amplitúd pravdepodobnosti pomocou tzv. *Feynmanovho integrálu* cez trajektórie v stavovom priestore. Ako zaujímavosť poznamenajme, že hoci fyzici pri svojich výpočtoch Feynmanov integrál hojne a úspešne používajú, matematikom sa dodnes nepodarilo preň vypracovať plne uspokojivú teóriu a zaradiť ho do teoreticko-množinového rámca súčasnej matematiky.

Cvičenia

17.1. Nech V je konečnorozmerný vektorový priestor nad poľom $\mathbb{C}$ s bázou tvorenou vektormi $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$. Potom vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n, \mathrm{i}\boldsymbol{v}_1, \dots, \mathrm{i}\boldsymbol{v}_n$ tvoria bázu jeho reálneho zúženia $V_{\mathbb{R}}$. Dokážte. Odvoďte z toho vzťah $\dim V_{\mathbb{R}} = 2 \dim V$.

17.2. Pre ľubovoľné vektory $\boldsymbol{x}$, $\boldsymbol{y}$ v unitárnom priestore V platí $\langle \boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y} \rangle = \|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2 - 2\mathrm{i}\,\mathrm{Im}\langle \boldsymbol{x}, \boldsymbol{y} \rangle$. Dokážte.

17.3. Nech U, V sú vektorové priestory nad poľom $\mathbb{C}$.

(a) Definujte jadro a obraz semilineárneho zobrazenia $\varphi\colon V \to U$ rovnako ako pre lineárne zobrazenia a dokážte, že $\mathrm{Ker}\,\varphi$ je lineárny podpriestor vo V a $\mathrm{Im}\,\varphi$ je lineárny podpriestor v U.

(b) Za predpokladu, že V je konečnorozmerný, dokážte, že veta o dimenzii jadra a obrazu platí aj pre semilineárne zobrazenia $\varphi\colon V \to U$.

(c) Za predpokladu, že oba priestory U, V sú konečnorozmerné, definujte maticu semilineárneho zobrazenia $\varphi\colon V \to U$ vzhľadom na bázy $\boldsymbol{\beta}$ vo V a $\boldsymbol{\alpha}$ v U rovnako ako pre lineárne zobrazenia. Dokážte, že pre ľubovoľný vektor $\boldsymbol{x} \in V$ platí $(\varphi\boldsymbol{x})_{\boldsymbol{\alpha}} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot \overline{(\boldsymbol{x})}_{\boldsymbol{\beta}}$.

(d) Ako sa mení matica semilineárneho zobrazenia $\varphi\colon V \to U$ pri zmene báz v priestoroch U, V?

17.4. Nech U, V sú konečnorozmerné vektorové priestory nad poľom $\mathbb{C}$ a $F\colon U \times V \to \mathbb{C}$ je poldruhalineárna forma.

(a) Definujte maticu formy F vzhľadom na bázy $\boldsymbol{\alpha}$ v U a $\boldsymbol{\beta}$ vo V rovnakým spôsobom ako pre bilineárne formy. Dokážte, že pre ľubovoľné vektory $\boldsymbol{x} \in U$, $\boldsymbol{y} \in V$ platí $F(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x})^{\mathsf{T}} \cdot [F]_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot \overline{(\boldsymbol{y})_{\boldsymbol{\beta}}}$.

(b) Ako sa mení matica formy F pri zmene báz v priestoroch U, V?

(c) Označme $\overline{V}^*$ množinu všetkých semilineárnych zobrazení $V \to \mathbb{C}$. Dokážte, že $\overline{V}^*$ je lineárny podpriestor vektorového priestoru $\mathbb{C}^V$ a platí $\dim \overline{V}^* = \dim V$.

(d) Dokážte, že predpisom $F^{\times}(\boldsymbol{x})(\boldsymbol{y}) = F(\boldsymbol{x}, \boldsymbol{y})$ pre $\boldsymbol{x} \in U$, $\boldsymbol{y} \in V$ je definované *lineárne* zobrazenie $F^{\times}: U \to \overline{V}^*$.

(e) Definujte hodnosť formy F rovnako ako pre bilineárne formy. Ako súvisí hodnosť $h(F)$ s vlastnosťami zobrazenia $F^{\times}: U \to \overline{V}^*$? Porovnajte s vetou 11.1.7.

17.5. (a) Nech $F\colon V \times V \to \mathbb{C}$ je poldruhalineárma forma a $F_0 = \operatorname{Re} F$, $F_1 = \operatorname{Im} F$. Dokážte, že $F_0, F_1\colon V_{\mathbb{R}} \times V_{\mathbb{R}} \to \mathbb{R}$ sú bilineárne formy.

(b) Dokážte tvrdenia 17.1.1 a 17.2.1.

17.6. Nech $F\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ je poldruhalineárna forma na (stĺpcovom) vektorovom priestore $\mathbb{C}^n$ a $\boldsymbol{A} = [F]_{\boldsymbol{\varepsilon}}$ je jej matica vzhľadom na kanonickú bázu $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$. Označme $F_0 = \operatorname{Re} F$, $F_1 = \operatorname{Im} F$, $A_0 = \operatorname{Re} A$, $A_1 = \operatorname{Im} A$. Ako vyzerajú matice bilineárnych foriem $F_0, F_1\colon \mathbb{C}^n_{\mathbb{R}} \times \mathbb{C}^n_{\mathbb{R}} \to \mathbb{R}$ v báze $\boldsymbol{\varepsilon}^+ = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n, \mathrm{i}\boldsymbol{e}_1, \ldots, \mathrm{i}\boldsymbol{e}_n)$? Čo sa dá povedať navyše, ak F je hermitovská?

17.7. Sformulujte pravidlá pre diagonalizáciu hermitovských matíc $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ a dokážte analógy príslušných tvrdení z paragrafov 11.3 a 12.1, 12.2.

17.8. Ak $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ sú ortonormálne bázy konečnorozmerného unitárneho priestoru V, tak pre ľubovoľné vektory $\boldsymbol{x}, \boldsymbol{y} \in V$ platí $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = (\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot \overline{\boldsymbol{P}}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot \overline{(\boldsymbol{y})_{\boldsymbol{\beta}}}$. Dokážte.

17.9. Dokážte analóg tvrdenia 13.2.1 pre Gramove matice v unitárnych priestoroch.

17.10. (a) Popíšte lineárne zobrazenie ortogonálnej projekcie do konečnorozmerného lineárneho podpriestoru S v unitárnom priestore V a odvoďte jeho základné vlastnosti.

(b) Zovšeobecnite výsledky o vzdialenosti vektora od lineárneho podpriestoru resp. bodu od afinného podpriestoru v konečnorozmernom unitárnom priestore.

(c) Napíšte maticu ortogonálnej projekcie do lineárneho podpriestoru $S = [s_1(\boldsymbol{A}), \ldots, s_k(\boldsymbol{A})]$ generovaného stĺpcami matice $\boldsymbol{A} \in \mathbb{C}^{n\times k}$ v unitárnom priestore $\mathbb{C}^n$ so štandardným skalárnym súčinom vzhľadom na kanonickú ortonormálnu bázu $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$.

17.11. Upravte Parsevalove rovnosti tak, aby platili aj pre ortogonálne bázy v unitárnych priestoroch.

17.12. Integrál komplexnej funkcie reálnej premennej $f(x) = f_0(x) + \mathrm{i} f_1(x)$, kde $f_0, f_1\colon \langle a, b\rangle \to \mathbb{R}$ definujeme rovnosťou $\int_a^b f(x)\,\mathrm{d}x = \int_a^b f_0(x)\,\mathrm{d}x + \mathrm{i}\int_a^b f_1(x)\,\mathrm{d}x$.

(a) Dokážte, že formulou $\langle f, g\rangle = \int_a^b f(x)\,\overline{g}(x)\,\mathrm{d}x$ je definovaný komplexný skalárny súčin na vektorovom priestore všetkých spojitých funkcií $f\colon \langle a, b\rangle \to \mathbb{C}$.

(b) Dokážte, že funkcie $f_n\colon \langle 0, 2\pi\rangle \to \mathbb{C}$, kde $f_n(x) = \mathrm{e}^{\mathrm{i}nx}$ pre $n \in \mathbb{Z}$, $x \in \langle 0, 2\pi\rangle$, sú navzájom ortogonálne. Porovnajte s cvičeniami 13.7 a 13.8.

17.13. (a) Overte, že oba tvary matice $\boldsymbol{A} \in \mathbb{C}^{2\times 2}$ z vety 17.4.3 naozaj predstavujú unitárne matice.

(b) Pre akú voľbu parametrov α, β, ϑ, ω v druhom z týchto vyjadrení dostaneme maticu rotácie $\boldsymbol{R}_\gamma$ resp. maticu osovej súmernosti $\boldsymbol{S}_\gamma$, kde $\gamma \in \mathbb{R}$?

17.14. Nech $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ je štvorcová komplexná matica. Dokážte nasledujúce tvrdenia:
(a) Ak $\boldsymbol{A}$ je hermitovská, tak $\det \boldsymbol{A}$ je reálne číslo.

(b) Ak $\boldsymbol{A}$ je unitárna, tak $\det \boldsymbol{A}$ je komplexná jednotka.

(c) Ak $u_1, \dots, u_n \in \mathbb{C}$, tak matica $\operatorname{diag}(u_1, \dots, u_n)$ je hermitovská práve vtedy, keď $u_1, \dots, u_n \in \mathbb{R}$.

(d) Ak $u_1, \dots, u_n \in \mathbb{C}$, tak matica $\operatorname{diag}(u_1, \dots, u_n)$ je unitárna práve vtedy, keď $|u_1| = \ldots = |u_n| = 1$.

17.15. Ak $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$ sú unitárne matice, tak aj matice $\boldsymbol{A}^{-1}$ a $\boldsymbol{A}\cdot\boldsymbol{B}$ sú unitárne. Dokážte. Platí niečo podobné aj pre hermitovské matice?

17.16. Zovšeobecnite tvrdenia o *Iwasawovom rozklade* a *QR-rozklade* z cvičení 13.11 resp. 13.13 na komplexné matice a dokážte ich:

(a) Každú regulárnu maticu $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ možno jednoznačne rozložiť na súčin $\boldsymbol{A} = \boldsymbol{P}\cdot\boldsymbol{D}\cdot\boldsymbol{T}$, kde $\boldsymbol{P} \in \mathbb{C}^{n\times n}$ je unitárna matica, $\boldsymbol{D} \in \mathbb{R}^{n\times n}$ je diagonálna matica s kladnými reálnymi číslami na diagonále a $\boldsymbol{T} \in \mathbb{C}^{n\times n}$ je horná trojuholníková matica s jednotkami na diagonále.

(b) Každú maticu $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ kde $m \geq n$ možno rozložiť na súčin $\boldsymbol{A} = \boldsymbol{Q}\cdot\boldsymbol{R}$, kde $\boldsymbol{Q} \in \mathbb{C}^{m\times m}$ je unitárna matica a $\boldsymbol{R} \in \mathbb{C}^{m\times n}$ je horná trojuholníková matica s nulovými poslednými $m - n$ riadkami. Ak $\boldsymbol{A}$ má (maximálnu) hodnosť n, tak $\boldsymbol{Q}$ a $\boldsymbol{R}$ sú určené jednoznačne za dodatočnej podmienky, že diagonálne prvky r_{ii}, $i \leq n$, matice $\boldsymbol{R}$ sú kladné reálne čísla.

17.17. Nájdite Iwasawove a QR-rozklady nasledujúcich matíc:

(a) $\begin{pmatrix} 1+2\mathrm{i} & -1 \\ 1+3\mathrm{i} & \mathrm{i} \end{pmatrix}$ (b) $\begin{pmatrix} \mathrm{i} & -1+\mathrm{i} & 2+\mathrm{i} \\ 1 & 3 & 2\mathrm{i} \\ 1 & 0 & -\mathrm{i} \end{pmatrix}$

17.18. (a) S využitím cvičenia 11 napíšte Parsevalove rovnosti vzhľadom na ortogonálnu bázu $\boldsymbol{\phi} = (\boldsymbol{f}_0, \dots, \boldsymbol{f}_{n-1})$ z tvrdenia 17.5.1.

(b) Overte, že pre ľubovoľné vektory $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ platí $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \frac{1}{n}\langle \widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}\rangle$ a $\|\boldsymbol{x}\|^2 = \frac{1}{n}\|\widehat{\boldsymbol{x}}\|^2$.

(c) Odvoďte z toho, že „znormovaná“ matica diskrétnej Fourierovej transformácie $\boldsymbol{F} = \frac{1}{\sqrt{n}}\big(\omega^{-jk}\big) \in \mathbb{C}^{n\times n}$, kde $\omega = \mathrm{e}^{2\pi\mathrm{i}/n}$, je unitárna. Ako vyzerá k nej inverzná matica?

17.19. (a) Operácia $*$ cyklickej konvolúcie na priestore $\mathbb{C}^n$ je asociatívna, komutatívna a bilineárna. Dokážte priamo aj pomocou vzťahu s násobením po zložkách sprostredkovaného diskrétnou Fourierovou transformáciou.

(b) Má operácia $*$ neutrálny prvok? Ak áno, ako vyzerá?

17.20. Dokážte druhú rovnosť z vety 17.5.2 – jednak priamo, jednak ako dôsledok prvej rovnosti a vlastností DFT.

17.21. Nájdite vhodnú ortogonálnu bázu priestoru všetkých komplexných matíc $\mathbb{C}^{m\times n}$ so štandardným skalárnym súčinom $\langle \boldsymbol{X}, \boldsymbol{Y}\rangle = \sum_{j,k} x_{jk}\overline{y}_{jk}$ pozostávajúcu z matíc $\boldsymbol{\Omega}_{rs} \in \mathbb{C}^{m\times n}$ ($r \in \mathbb{Z}_m$, $s \in \mathbb{Z}_n$) tak, aby pre každú maticu $\boldsymbol{X} \in \mathbb{C}^{m\times n}$ platilo $\widehat{x}_{rs} = \langle \boldsymbol{X}, \boldsymbol{\Omega}_{rs}\rangle$. Zovšeobecnite na viacrozmerné matice.

17.22. Vyjadrite formulu „Feynmanovho integrálu“
$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{j_1,\dots,j_k=1}^{n} \langle \boldsymbol{u}, \boldsymbol{w}_{1j_1} \rangle \langle \boldsymbol{w}_{1j_1}, \boldsymbol{w}_{2j_2} \rangle \dots \langle \boldsymbol{w}_{k-1j_{k-1}}, \boldsymbol{w}_{kj_k} \rangle \langle \boldsymbol{w}_{kj_k}, \boldsymbol{v} \rangle$
zo záveru paragrafu 17.8 pomocou matíc prechodu medzi bázami $\boldsymbol{\gamma}_j$.

Časť III

Lineárne operátory

18. Vlastné hodnoty a vlastné vektory

Touto kapitolou začíname najdôležitejšiu partiu nášho kurzu lineárnej algebry a geometrie. Jej ústredné pojmy ako *vlastná hodnota*, *vlastný vektor* a *spektrum lineárneho operátora* hrajú kľúčovú úlohu nielen v samotnej lineárnej algebre, ale aj v jej aplikáciách – či už v iných oblastiach matematiky, vo fyzike, i v ďalších disciplínach.

Po šiestich kapitolách, ktoré sa odohrávali nad poľom $\mathbb{R}$ prípadne $\mathbb{C}$, sa opäť vraciame k vektorovým priestorom nad ľubovoľným poľom K.

18.1 Matica lineárneho operátora a podobnosť matíc

Pripomeňme, že *lineárnym operátorom* na vektorovom priestore V alebo tiež *lineárnou transformáciou* priestoru V nazývame ľubovoľné lineárne zobrazenie $\varphi\colon V \to V$. Ak V je konečnorozmerný, tak lineárny operátor $\varphi\colon V \to V$ je injektívny práve vtedy, keď je surjektívny, čo je ekvivalentné s rovnosťou $h(\varphi) = \dim V$ (pozri dôsledok 6.2.4).

Maticou lineárnej transformácie $\varphi\colon V \to V$ vzhľadom na bázu $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ nazývame maticu

$$(\varphi)_{\boldsymbol{\alpha}} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\alpha}} = \big((\varphi \boldsymbol{u}_1)_{\boldsymbol{\alpha}}, \dots, (\varphi \boldsymbol{u}_n)_{\boldsymbol{\alpha}}\big) \in K^{n\times n},$$

tvorenú súradnicami obrazov $\varphi(\boldsymbol{u}_j)$ vektorov $\boldsymbol{u}_j$ bázy $\boldsymbol{\alpha}$ vzhľadom na *tú istú* bázu $\boldsymbol{\alpha}$.

Jedným z vedúcich zámerov tejto i nasledujúcich dvoch kapitol bude dosiahnuť vhodnou voľbou bázy $\boldsymbol{\alpha}$ čo najjednoduchší tvar matice $\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha}}$ lineárneho operátora φ. Poznamenajme, že pokiaľ by sme netrvali na prirodzenej požiadavke vyjadrovať súradnice vzorov aj obrazov vektorov $\boldsymbol{x} \in V$ vzhľadom na *tú istú bázu* priestoru V, šlo by o špeciálny prípad vety 7.5.4: vždy by sme mohli (navyše jednoduchým spôsobom) zvoliť bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ priestoru V tak, aby matica φ vzhľadom na ne mala blokový tvar

$$(\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \begin{pmatrix} \boldsymbol{I}_h & \boldsymbol{0}_{h,n-h} \\ \boldsymbol{0}_{n-h,h} & \boldsymbol{0}_{n-h,n-h} \end{pmatrix},$$

kde $h = h(\varphi)$. Niečo jednoduchšie si ťažko možno predstaviť – my by sme už tradične boli spokojní s diagonálnou maticou $\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha}}$. Zdanlivo nevinná požiadavka $\boldsymbol{\alpha} = \boldsymbol{\beta}$ však značne zužuje možnosti našej voľby, čo – ako uvidíme – dramaticky komplikuje situáciu.

Analogickú úlohu sme už riešili v kapitole 11 pre symetrické bilineárne formy; vtedy sa nám však podarilo ukázať, že (až na prípad polí charakteristiky 2) možno voľbou vhodnej bázy vždy dosiahnuť diagonálny tvar matice príslušnej formy. Poznamenajme už vopred, že pre lineárne operátory sa nám niečo podobné nepodarí. Jednako preskúmame štruktúru lineárnych operátorov na konečnorozmerných vektorových priestoroch do takej miery, že dokážeme charakterizovať operátory diagonalizovateľné vo vhodnej báze, ako aj identifikovať (ba čiastočne tiež odstrániť – ale to až v nasledujúcej kapitole) prekážky diagonalizovateľnosti u tých ostatných.

Na začiatok si uvedomme vzťah medzi maticami lineárneho operátora vzhľadom na rôzne bázy. Ako zvláštny prípad vety 7.6.1 dostávame:

18.1.1. Veta. *Nech $\varphi\colon V \to V$ je lineárna transformácia konečnorozmerného vektorového priestoru V a $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ sú jeho dve bázy. Potom*

$$(\varphi)_{\boldsymbol{\beta}} = \boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\alpha}} \cdot (\varphi)_{\boldsymbol{\alpha}} \cdot \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}.$$

Štvorcové matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{n\times n}$ sa nazývajú *podobné*, označenie $\boldsymbol{A} \approx \boldsymbol{B}$, ak existuje regulárna matica $\boldsymbol{P} \in K^{n\times n}$ taká, že platí

$$\boldsymbol{B} = \boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P}.$$

Zrejme podobné matice majú rovnakú hodnosť. Čitateľ si iste sám bez ťažkostí overí, že pre ľubovoľné matice $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C} \in K^{n\times n}$ platí

$$\begin{gathered}
\boldsymbol{A} \approx \boldsymbol{A},\\
\boldsymbol{A} \approx \boldsymbol{B} \;\Rightarrow\; \boldsymbol{B} \approx \boldsymbol{A},\\
\boldsymbol{A} \approx \boldsymbol{B} \;\&\; \boldsymbol{B} \approx \boldsymbol{C} \;\Rightarrow\; \boldsymbol{A} \approx \boldsymbol{C}.
\end{gathered}$$

To znamená, že vzťah podobnosti je *reflexívny*, *symetrický* a *tranzitívny*, čiže je to *ekvivalencia* na množine $K^{n\times n}$. Ekvivalencia podobnosti nám asi pripomína inú ekvivalenciu na množine $K^{n\times n}$: totiž kongruenciu matíc $\boldsymbol{A} \equiv \boldsymbol{B}$, s ktorou ju však neslobodno zamieňať (pozri paragraf 11.3).

Keďže $\boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\alpha}} = \boldsymbol{P}^{-1}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ a každá regulárna matica je maticou prechodu medzi vhodnou dvojicou báz, nasledujúca veta je bezprostredným dôsledkom vety 18.1.1.

18.1.2. Veta. *Nech V je n-rozmerný vektorový priestor nad poľom K. Potom pre ľubovoľné matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{n\times n}$ nasledujúce podmienky sú ekvivalentné:*

(i) $\boldsymbol{A}$, $\boldsymbol{B}$ sú maticami tej istej lineárnej transformácie $\varphi\colon V \to V$ vzhľadom na nejaké dve (možno no nie nutne rôzne) bázy priestoru V;

(ii) $\boldsymbol{A} \approx \boldsymbol{B}$.

Stopu matice $\boldsymbol{A} \in K^{n\times n}$, označenie $\operatorname{tr} \boldsymbol{A}$ (z anglického *trace*), definujeme ako súčet jej diagonálnych prvkov, t. j.

$$\operatorname{tr} \boldsymbol{A} = a_{11} + \ldots + a_{nn} = \sum_{i=1}^{n} a_{ii}.$$

18.1.3. Tvrdenie. *Nech* $\boldsymbol{A} \in \boldsymbol{K}^{m\times n}$, $\boldsymbol{B} \in K^{n\times m}$. *Potom*

$$\operatorname{tr}(\boldsymbol{A} \cdot \boldsymbol{B}) = \operatorname{tr}(\boldsymbol{B} \cdot \boldsymbol{A}).$$

Dôkaz. Označme $\boldsymbol{A}\cdot\boldsymbol{B} = (c_{ij})_{m\times m}$, $\boldsymbol{B}\cdot\boldsymbol{A} = (d_{jk})_{n\times n}$. Jednoduchým výpočtom dostávame

$$\operatorname{tr}(\boldsymbol{A} \cdot \boldsymbol{B}) = \sum_{i=1}^{m} c_{ii} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{m} b_{ji}a_{ij} = \sum_{j=1}^{n} d_{jj} = \operatorname{tr}(\boldsymbol{B} \cdot \boldsymbol{A}).$$

Vety 10.3.2 a 10.3.3 o determinante súčinu matíc a determinante inverznej matice, resp. tvrdenie 18.1.3 majú nasledujúci bezprostredný

18.1.4. Dôsledok. *Podobné matice majú rovnaký determinant aj stopu.*

Hovoríme, že determinant a stopa sú *invariantmi podobnosti matíc.* Ak teda matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{n\times n}$ majú rôzne determinanty alebo rôzne stopy (pričom najmä túto druhú podmienku možno veľmi ľahko nahliadnuť), tak nemôžu byť podobné. Na druhej strane však ani rovnosť determinantu a stopy ešte nezaručuje ich podobnosť.

18.2 Vlastné hodnoty a vlastné vektory

Lineárny operátor $\varphi\colon V \to V$ na konečnorozmernom vektorovom priestore V sa nazýva *diagonalizovateľný*, ak existuje nejaká báza priestoru V, vzhľadom na ktorú má φ diagonálnu maticu.

Nech teda $\varphi\colon V \to V$ je diagonalizovateľný lineárny operátor a $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ je taká báza priestoru V, že matica $\boldsymbol{B} = (\varphi)_{\boldsymbol{\beta}}$ je diagonálna so skalármi $\lambda_1, \ldots, \lambda_n \in K$ na diagonále. Potom pre bázické vektory $\boldsymbol{v}_i$ platí

$$\varphi(\boldsymbol{v}_i) = \lambda_i \boldsymbol{v}_i.$$

Ukazuje sa, že tento vzťah medzi lineárnym operátorom φ skalárom λ_i a vektorom $\boldsymbol{v}_i$ má kľúčový význam.

Vopred zdôrazňujeme, že nasledujúce dve definície sa vzťahujú rovnako na konečno- i nekonečnorozmerné vektorové priestory.

Hovoríme, že skalár $\lambda \in K$ je *vlastná* alebo tiež *charakteristická hodnota* lineárneho operátora $\varphi\colon V \to V$, ak existuje vektor $\mathbf{0} \neq \boldsymbol{v} \in V$, pre ktorý platí $\varphi(\boldsymbol{v}) = \lambda\boldsymbol{v}$. V prípade vektorových priestorov nad číselnými poľami, ako napr. $\mathbb{R}$ alebo $\mathbb{C}$, zvykneme hovoriť o *vlastnom čísle* lineárneho operátora.

Hovoríme, že $\mathbf{0} \neq \boldsymbol{v} \in V$ je *vlastný* alebo tiež *charakteristický vektor* lineárneho operátora $\varphi\colon V \to V$, ak existuje skalár $\lambda \in K$, pre ktorý platí $\varphi(\boldsymbol{v}) = \lambda\boldsymbol{v}$. Ak V je vektorový priestor funkcií, zvykneme hovoriť o *vlastnej funkcii* lineárneho operátora.

Obe uvedené definície hovoria vlastne o tom istom. Ak $\lambda \in K$ je vlastná hodnota operátora φ, tak každý nenulový vektor $\boldsymbol{v} \in V$ taký, že $\varphi(\boldsymbol{v}) = \lambda\boldsymbol{v}$, je vlastný vektor operátora φ. Naopak, ak $\boldsymbol{v} \in V$ je vlastný vektor, tak skalár λ, pre ktorý platí $\varphi(\boldsymbol{v}) = \lambda\boldsymbol{v}$, je vlastná hodnota. Hovoríme, že $\boldsymbol{v}$ je *vlastný vektor prislúchajúci k vlastnej hodnote* λ, resp. že λ je *vlastná hodnota prislúchajúca k vlastnému vektoru* $\boldsymbol{v}$. Ešte si všimnite, že vlastná hodnota prislúchajúca k danému vlastnému vektoru je určená jednoznačne; na druhej strane, ako uvidíme, k danej vlastnej hodnote môže prislúchať viacero, dokonca lineárne nezávislých vektorov.

Vlastnou (charakteristickou) hodnotou (vlastným číslom), resp. vlastným (charakteristickým) vektorom štvorcovej matice $\boldsymbol{A} \in K^{n\times n}$ nazývame vlastnú hodnotu, resp. vlastný vektor lineárneho operátora $K^n \to K^n$ daného predpisom $\boldsymbol{x} \mapsto \boldsymbol{A}\cdot\boldsymbol{x}$. Vlastná hodnota $\lambda \in K$ a k nej prislúchajúci vlastný vektor $\mathbf{0} \neq \boldsymbol{v} \in K^n$ matice $\boldsymbol{A}$ sú tak zviazané vzťahom $\boldsymbol{A}\cdot\boldsymbol{v} = \lambda\boldsymbol{v}$.

Z vety 18.1.2 vyplýva, že vlastné hodnoty podobných matíc sú vlastnými hodnotami toho istého lineárneho operátora, preto

18.2.1. Tvrdenie. *Podobné matice majú rovnaké vlastné hodnoty.*

Jednorozmerný podpriestor $[\boldsymbol{v}]$ generovaný vlastným vektorom $\boldsymbol{v}$ lineárneho operátora je špeciálnym prípadom tzv. invariantného podpriestoru. Hovoríme, že lineárny podpriestor S vektorového priestoru V je *invariantným podpriestorom* lineárneho operátora $\varphi\colon V \to V$, ak platí $\varphi(S) \subseteq S$, t. j. $\varphi(\boldsymbol{x}) \in S$ pre každé $\boldsymbol{x} \in S$. Ak lineárny operátor φ je fixovaný kontextom, hovoríme jednoducho o invariantnom podpriestore.

Triviálny podpriestor $\{\mathbf{0}\}$ a nevlastný podpriestor V sú vždy invariantné. Zrejme jednorozmerný podpriestor $[\boldsymbol{v}]$ je invariantný práve vtedy, keď $\boldsymbol{v}$ je vlastný vektor príslušného operátora. Jednorozmerné podpriestory generované vlastnými vektormi lineárneho operátora sú teda príkladmi netriviálnych, a ak $\dim V > 1$, tak i vlastných invariantných podpriestorov.

Ak S je invariantný podpriestor lineárnej transformácie $\varphi\colon V \to V$, tak zúženie φ na S je opäť lineárnou transformáciou $\varphi \restriction S\colon S \to S$ na vektorovom priestore S. Ak $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_n)$ je báza priestoru V taká, že jej prvých k vektorov tvorí bázu invariantného podpriestoru S, tak matica φ v tejto báze má blokový tvar

$$\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha}} = \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{M} \\ \boldsymbol{0}_{n-k,k} & \boldsymbol{A}_2 \end{pmatrix},$$

kde $\boldsymbol{A}_1 \in K^{k\times k}$ je matica lineárnej transformácie $\varphi \restriction S\colon S \to S$ v báze $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ a $\boldsymbol{M} \in K^{k\times(n-k)}$, $\boldsymbol{A}_2 \in K^{(n-k)\times(n-k)}$.

Ak $V = S \oplus T$ je dokonca priamym súčtom invariantných podpriestorov S, T, tak V má bázu $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{u}_{k+1}, \dots, \boldsymbol{u}_n)$, ktorej prvých k vektorov tvorí bázu S a zvyšných $n-k$ vektorov tvorí bázu T. Vzhľadom na takúto bázu má matica φ blokovo diagonálny tvar

$$\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha}} = \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{0}_{k,n-k} \\ \boldsymbol{0}_{n-k,k} & \boldsymbol{A}_2 \end{pmatrix} = \operatorname{diag}(\boldsymbol{A}_1, \boldsymbol{A}_2),$$

kde $\boldsymbol{A}_1 \in K^{k\times k}$ je matica lineárnej transformácie $\varphi \restriction S\colon S \to S$ v báze $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ a $\boldsymbol{A}_2 \in K^{(n-k)\times(n-k)}$ je matica lineárnej transformácie $\varphi \restriction T\colon T \to T$ v báze $(\boldsymbol{u}_{k+1}, \dots, \boldsymbol{u}_n)$. Toto pozorovanie možno zrejmým spôsobom zovšeobecniť na priamy súčet ľubovoľného konečného počtu invariantných podpriestorov. (Detaily prenechávame na samostatné premyslenie čitateľovi.)

Z vykonaných úvah priamo vyplýva nasledujúca charakterizácia diagonalizovateľných lineárnych operátorov.

18.2.2. Veta. *Nech φ je lineárny operátor na konečnorozmernom vektorovom priestore V. Potom nasledujúce podmienky sú ekvivalentné:*
(i) φ je diagonalizovateľný;
(ii) existuje báza priestoru V pozostávajúca z vlastných vektorov operátora φ;
(ii) V je priamym súčtom jednorozmerných invariantných podpriestorov lineárneho operátora φ.

Samozrejme, matica operátora φ v báze vlastných vektorov $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ má tvar $(\varphi)_{\boldsymbol{\beta}} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, kde λ_i je vlastné hodnota prislúchajúca k vlastnému vektoru $\boldsymbol{v}_i$.

Podčiarkujeme, že nasledujúce tvrdenie platí aj bez predpokladu konečnorozmernosti priestoru V.

18.2.3. Tvrdenie. *Nech $\lambda_1, \dots, \lambda_k$ sú navzájom rôzne vlastné hodnoty lineárneho operátora $\varphi\colon V \to V$. Potom k nim prislúchajúce vlastné vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ sú lineárne nezávislé.*

Dôkaz. Predpokladajme, že $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ sú lineárne závislé. Potom existuje $j \le k$ také, že vektor $\boldsymbol{v}_j$ je lineárnou kombináciou predchádzajúcich; zvoľme najmenšie také j. Keďže $\boldsymbol{v}_1 \neq \mathbf{0}$, $j \geq 2$ a žiaden z vektorov $\boldsymbol{v}_1, \dots, \boldsymbol{v}_{j-1}$ nie je lineárnou kombináciou predchádzajúcich, sú to lineárne nezávislé vektory. Pre nejaké skaláry $c_1, \dots, c_{j-1}$ platí $\boldsymbol{v}_j = c_1\boldsymbol{v}_1 + \ldots + c_{j-1}\boldsymbol{v}_{j-1}$. Nakoľko $\boldsymbol{v}_j \neq \mathbf{0}$, aspoň jeden z týchto skalárov je $\neq 0$. Vektor $\varphi(\boldsymbol{v}_j)$ si vyjadríme dvoma spôsobmi:

$$\varphi(\boldsymbol{v}_j) = c_1\varphi(\boldsymbol{v}_1) + \ldots + c_{j-1}\varphi(\boldsymbol{v}_{j-1}) = c_1\lambda_1\boldsymbol{v}_1 + \ldots + c_{j-1}\lambda_{j-1}\boldsymbol{v}_{j-1},$$

$$\varphi(\boldsymbol{v}_j) = \lambda_j\boldsymbol{v}_j = \lambda_j(c_1\boldsymbol{v}_1 + \ldots + c_{j-1}\boldsymbol{v}_{j-1}) = c_1\lambda_j\boldsymbol{v}_1 + \cdots + c_{j-1}\lambda_j\boldsymbol{v}_{j-1}.$$

V dôsledku toho

$$c_1(\lambda_1 - \lambda_j)\boldsymbol{v}_1 + \ldots + c_{j-1}(\lambda_{j-1} - \lambda_j)\boldsymbol{v}_{j-1} = \mathbf{0},$$

a keďže $\lambda_i \neq \lambda_j$ pre všetky $i \leq j-1$, aspoň jeden z koeficientov $c_i(\lambda_i - \lambda_j)$ je rôzny od nuly. To je však spor s nezávisloťou vektorov $\boldsymbol{v}_1, \dots, \boldsymbol{v}_{j-1}$.

Práve dokázané tvrdenie spolu s vetou 18.2.2 majú za bezprostredný dôsledok prvú časť nasledujúceho tvrdenia.

18.2.4. Tvrdenie. *Nech φ je lineárny operátor na n-rozmernom vektorovom priestore V. Ak φ má n navzájom rôznych vlastných hodnôt $\lambda_1, \dots, \lambda_n$, tak je φ je diagonalizovateľný v báze im prislúchajúcich vlastných vektorov. Navyše každý vlastný vektor $\boldsymbol{v}_i$ prislúchajúci k vlastnej hodnote λ_i je určený jednoznačne až na skalárny násobok.*

Dôkaz. Zostáva overiť záverečnú podmienku jednoznačnosti. Nech teda $j \leq n$ a $\boldsymbol{w}$ je tiež vlastný vektor prislúchajúci k vlastnej hodnote λ_j. Keďže $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ tvoria bázu V, $\boldsymbol{w} = \sum_{i=1}^n c_i\boldsymbol{v}_i$ pre nejaké koeficienty $c_i \in K$. Dokážeme, že $\boldsymbol{w} = c_j\boldsymbol{v}_j$. V opačnom prípade by aj $\boldsymbol{w} - c_j\boldsymbol{v}_j \neq \mathbf{0}$ bol vlastným vektorom operátora φ prislúchajúcim k λ_j. Podľa predchádzajúceho tvrdenia sú vektory $\boldsymbol{w} - c_j\boldsymbol{v}_j$, a $\boldsymbol{v}_i$, $i \neq j$, lineárne nezávislé. To je však spor so skutočnosťou, že

$$\boldsymbol{w} - c_j\boldsymbol{v}_j = \sum_{i \neq j} c_i\boldsymbol{v}_i$$

je lineárnou kombináciou ostatných vektorov.

18.3 Charakteristický polynóm

V tomto paragrafe si predvedieme, ako možno k danej štvorcovej matici $\boldsymbol{A} \in K^{n\times n}$ nájsť jej vlastné hodnoty a k nim prislúchajúce vlastné vektory. Reprezentácia lineárneho operátora na konečnorozmernom vektorovom priestore pomocou jeho matice v nejakej (dokonca ľubovoľnej) báze nám potom umožní vyriešiť analogickú úlohu aj preň.

Maticu $\boldsymbol{A}-x\boldsymbol{I}$ nazývame *charakteristickou maticou* matice $\boldsymbol{A} \in K^{n\times n}$; jej *charakteristickým polynómom* nazývame determinant charakteristickej matice, t. j. polynóm

$$\mathrm{ch}_{\boldsymbol{A}}(x) = \det(\boldsymbol{A} - x\boldsymbol{I}_n) = \begin{vmatrix} a_{11}-x & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-x & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-x \end{vmatrix}$$

v premennej x s koeficientmi z poľa K, t. j. $\mathrm{ch}_{\boldsymbol{A}}(x) \in K[x]$. Charakteristický polynóm je zrejme polynóm stupňa n s koeficientom $(-1)^n$ pri najvyššej mocnine x^n. *Charakteristickou rovnicou* matice $\boldsymbol{A}$ nazývame rovnicu $\mathrm{ch}_{\boldsymbol{A}}(x) = 0$, t. j.

$$\det(\boldsymbol{A} - x\boldsymbol{I}_n) = 0.$$

Niektorí autori definujú charakteristickú maticu ako $x\boldsymbol{I}-\boldsymbol{A}$ a charakteristický polynóm ako $\det(x\boldsymbol{I}-\boldsymbol{A})$, čiže ako $(-1)^n$ krát „náš“ charakteristický polynóm – zrejme ide o nepodstatný rozdiel.

Význam práve definovaných pojmov je daný nasledujúcou vetou.

18.3.1. Veta. *Nech $\boldsymbol{A} \in K^{n\times n}$. Potom skalár $\lambda \in K$ je vlastnou hodotou matice $\boldsymbol{A}$ práve vtedy, keď*

$$\det(\boldsymbol{A} - \lambda\boldsymbol{I}_n) = 0,$$

t. j. práve vtedy, keď λ vyhovuje charakteristickej rovnici matice $\boldsymbol{A}$.

Dôkaz. Skalár $\lambda \in K$ je vlastnou hodnotou matice $\boldsymbol{A}$ práve vtedy, keď $\boldsymbol{A}\cdot\boldsymbol{v} = \lambda\boldsymbol{v}$ pre nejaký vektor $\boldsymbol{0} \neq \boldsymbol{v} \in K^n$, čiže práve vtedy, keď homogénna sústava lineárnych rovníc $(\boldsymbol{A} - \lambda\boldsymbol{I})\cdot\boldsymbol{v} = \boldsymbol{0}$ má aspoň jedno *nenulové* riešenie $\boldsymbol{v} \in K^n$. To nastane práve vtedy, keď matica $\boldsymbol{A}-\lambda\boldsymbol{I}$ je singulárna, t. j. $\det(\boldsymbol{A}-\lambda\boldsymbol{I}) = 0$.

Zopakujme si ešte raz, čo sme sa naučili v tomto dôkaze a nie je zahrnuté v znení vety: vlastné vektory štvorcovej matice $\boldsymbol{A}$ prislúchajúce k jej vlastnej hodnote λ sú práve všetky nenulové riešenia homogénnej sústavy s maticou $\boldsymbol{A} - \lambda\boldsymbol{I}$; pritom práve singularita uvedenej matice zaručuje ich existenciu.

18.3.2. Veta. *Nech $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$. Ak $\boldsymbol{A} \approx \boldsymbol{B}$, tak $\mathrm{ch}_{\boldsymbol{A}} = \mathrm{ch}_{\boldsymbol{B}}$; inými slovami, podobné matice majú rovnaký charakteristický polynóm.*

Dôkaz. Nech $\boldsymbol{A}, \boldsymbol{B}$ sú podobné a $\boldsymbol{P}$ je regulárna matica taká, že $\boldsymbol{B} = \boldsymbol{P}^{-1}\cdot\boldsymbol{A}\cdot\boldsymbol{P}$. Keďže aj $x\boldsymbol{I} = \boldsymbol{P}^{-1}\cdot x\boldsymbol{I}\cdot\boldsymbol{P}$, s použitím rovností pre determinant súčinu matíc a determinant inverznej matice (vety 10.3.2 a 10.3.3) dostávame

$$\begin{aligned}\mathrm{ch}_{\boldsymbol{B}}(x) &= \det(\boldsymbol{B} - x\boldsymbol{I}) = \det\big(\boldsymbol{P}^{-1}\cdot\boldsymbol{A}\cdot\boldsymbol{P} - \boldsymbol{P}^{-1}\cdot x\boldsymbol{I}\cdot\boldsymbol{P}\big)\\ &= \det\big(\boldsymbol{P}^{-1}\cdot(\boldsymbol{A} - x\boldsymbol{I})\cdot\boldsymbol{P}\big) = \det\boldsymbol{P}^{-1}\det(\boldsymbol{A} - x\boldsymbol{I})\det\boldsymbol{P}\\ &= \det(\boldsymbol{A} - x\boldsymbol{I}) = \mathrm{ch}_{\boldsymbol{A}}(x).\end{aligned}$$

To znamená, že i charakteristický polynóm je *invariantnom* podobnosti matíc. Táto jeho vlastnosť nám umožňuje korektne zadefinovať aj *charakteristický polynóm* $\mathrm{ch}_\varphi(x)$ *lineárnej transformácie* φ konečnorozmerného vektorového priestoru V ako charakteristický polynóm matice tejto transformácie vzhľadom na ľubovoľnú bázu priestoru V. Vlastné hodnoty takejto lineárnej transformácie sú potom totožné s vlastnými hodnotami jej matice.

Tvrdenie 18.2.1 teraz priamo vyplýva z vety 18.3.2. Keby sme boli schopní nahliadnuť, že koeficienty pri mocninách x^0 resp. x^{n-1} v charakteristickom polynóme $\mathrm{ch}_{\boldsymbol{A}}(x)$ matice $\boldsymbol{A} \in K^{n\times n}$ sú $\det\boldsymbol{A}$ resp. $(-1)^{n-1}\operatorname{tr}\boldsymbol{A}$, čiže

$$\mathrm{ch}_{\boldsymbol{A}}(x) = \det\boldsymbol{A} - \ldots + (-1)^{n-1}(a_{11} + \ldots + a_{nn})x^{n-1} + (-1)^n x^n$$

(čo nie je až také ťažké), mohli by sme okamžite dostať aj dôsledok 18.1.4 z práve dokázanej vety. Ani čitateľ, ktorý to nahliadnuť nedokáže, si však nemusí zúfať. Tieto výsledky nám totiž onedlho spadnú do lona samy, ako vedľajšie plody nášho štúdia.

18.4 Príklady

Pokúsme sa teraz na niekoľkých veľmi jednoduchých príkladoch (všetky sa týkajú matíc najnižšieho netriviálneho rozmeru 2×2) ilustrovať metódu výpočtu vlastných hodnôt λ matice $\boldsymbol{A}$ riešením jej charakteristickej rovnice $\mathrm{ch}_{\boldsymbol{A}}(x) = 0$ a následný výpočet vlastných vektorov riešením homogénnych sústav so singulárnymi maticami $\boldsymbol{A} - \lambda\boldsymbol{I}$. Zároveň sa pri tom zoznámime s rôznymi možnosťami, ktoré môžu nastať, a pripravíme si tak pôdu pre ďalšie úvahy.

18.4.1. Príklad. Súmernosť roviny podľa osi prechádzajúcej počiatkom a zvierajúcej s osou x uhol α je lineárny operátor $\boldsymbol{S}_\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$, ktorý má vzhľadom na kanonickú bázu $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2)$ maticu

$$\boldsymbol{S}_\alpha = \begin{pmatrix}\cos 2\alpha & \sin 2\alpha\\ \sin 2\alpha & -\cos 2\alpha\end{pmatrix}$$

(pozri príklad 6.4.4). Charakteristický polynóm

$$\begin{aligned}\det(\boldsymbol{S}_\alpha - x\boldsymbol{I}_2) &= \begin{vmatrix}\cos 2\alpha - x & \sin 2\alpha\\ \sin 2\alpha & -\cos 2\alpha - x\end{vmatrix}\\ &= x^2 - \cos^2 2\alpha - \sin^2 2\alpha = x^2 - 1\end{aligned}$$

má dva korene $x_{1,2} = \pm 1$. K nim prislúchajúce vlastné vektory nájdeme riešením homogénnych sústav s maticami

$$\boldsymbol{S}_\alpha - \boldsymbol{I} = \begin{pmatrix} \cos 2\alpha - 1 & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha - 1 \end{pmatrix} \sim \begin{pmatrix} \sin\alpha & -\cos\alpha \\ 0 & 0 \end{pmatrix},$$

resp.

$$\boldsymbol{S}_\alpha + \boldsymbol{I} = \begin{pmatrix} \cos 2\alpha + 1 & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha + 1 \end{pmatrix} \sim \begin{pmatrix} \cos\alpha & \sin\alpha \\ 0 & 0 \end{pmatrix}.$$

Oba podpriestory riešení sú jednorozmerné, generované vektormi $(\cos\alpha, \sin\alpha)^\mathsf{T}$ resp. $(-\sin\alpha, \cos\alpha)^\mathsf{T}$. To znamená, že operátor operátor $\boldsymbol{S}_\alpha$ má vzhľadom na bázu tvorenú stĺpcami matice

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

diagonálnu maticu $\operatorname{diag}(1, -1)$. Ešte si všimnite, že $(\cos\alpha, \sin\alpha)^\mathsf{T}$ je smerový vektor našej osi súmernosti a $(-\sin\alpha, \cos\alpha)^\mathsf{T}$ je smerový vektor kolmice na ňu v počiatku. Uvedomte si, že tento výsledok sa presne zhoduje s geometrickým názorom.

18.4.2. Príklad. Otočenie roviny okolo počiatku o uhol α je lineárny operátor $\boldsymbol{R}_\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$, ktorý má v kanonickej báze $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2)$ maticu

$$\boldsymbol{R}_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

(pozri príklad 6.4.3). Charakteristický polynóm

$$\begin{aligned} \det(\boldsymbol{R}_\alpha - x\boldsymbol{I}_2) &= \begin{vmatrix} \cos\alpha - x & -\sin\alpha \\ \sin\alpha & \cos\alpha - x \end{vmatrix} \\ &= x^2 - 2x\cos\alpha + \cos^2\alpha + \sin^2\alpha = x^2 - 2x\cos\alpha + 1 \end{aligned}$$

má diskriminant $D = 4\cos^2\alpha - 4 = -4\sin^2\alpha$. Okrem prípadov, keď $\sin\alpha = 0$, t. j. $\boldsymbol{R}_\alpha = \pm\boldsymbol{I}_2$ (identické zobrazenie resp. stredová súmernosť podľa počiatku, ktorými sa budeme zaoberať v rámci ďalšieho príkladu), je $D < 0$, teda charakteristický polynóm nemá reálne korene. Preto ani $\boldsymbol{R}_\alpha$ nemá reálne vlastné hodnoty a nie je podobná so žiadnou diagonálnou maticou nad $\mathbb{R}$.

Na druhej strane, keďže $\mathbb{R} \subseteq \mathbb{C}$, na $\boldsymbol{R}_\alpha$ sa môžeme dívať ako na komplexnú maticu z $\mathbb{C}^{2\times 2}$; ako taká určuje vzhľadom na kanonickú bázu $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2)$ v $\mathbb{C}^2$ lineárny operátor $\mathbb{C}^2 \to \mathbb{C}^2$. V poli $\mathbb{C}$ jej charakteristický polynóm už má dva korene $x_{1,2} = \cos\alpha \pm \mathrm{i}\sin\alpha = \mathrm{e}^{\pm\mathrm{i}\alpha}$, ktorým zodpovedajúce vlastné vektory dostaneme riešením homogénnych sústav s maticami

$$\boldsymbol{R}_\alpha - \mathrm{e}^{\mathrm{i}\alpha}\boldsymbol{I} = \begin{pmatrix} -\mathrm{i}\sin\alpha & -\sin\alpha \\ \sin\alpha & -\mathrm{i}\sin\alpha \end{pmatrix} \sim \begin{pmatrix} 1 & -\mathrm{i} \\ 0 & 0 \end{pmatrix},$$

resp.
$$\boldsymbol{R}_\alpha - \mathrm{e}^{-\mathrm{i}\alpha}\boldsymbol{I} = \begin{pmatrix} \mathrm{i}\sin\alpha & -\sin\alpha \\ \sin\alpha & \mathrm{i}\sin\alpha \end{pmatrix} \sim \begin{pmatrix} 1 & \mathrm{i} \\ 0 & 0 \end{pmatrix}.$$
Oba podpriestory riešení sú jednorozmerné, generované vektormi $(1, -\mathrm{i})^\mathsf{T}$ resp. $(1, \mathrm{i})^\mathsf{T}$. To znamená, že operátor $\mathbb{C}^2 \to \mathbb{C}^2$ daný predpisom $\boldsymbol{x} \mapsto \boldsymbol{R}_\alpha \cdot \boldsymbol{x}$ má vzhľadom na bázu tvorenú stĺpcami matice
$$\begin{pmatrix} 1 & 1 \\ -\mathrm{i} & \mathrm{i} \end{pmatrix}$$
diagonálnu maticu $\operatorname{diag}(\mathrm{e}^{\mathrm{i}\alpha}, \mathrm{e}^{-\mathrm{i}\alpha})$.

18.4.3. Príklad. Rovnoľahlosť v rovine so stredom v počiatku a koeficientom podobnosti $c \in \mathbb{R}$ je lineárny operátor $\mathbb{R}^2 \to \mathbb{R}^2$, ktorý má v kanonickej báze $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2)$ diagonálnu maticu $c\boldsymbol{I}_2$ (pozri príklad 6.4.5). Jej charakteristický polynóm
$$\det(c\boldsymbol{I}_2 - x\boldsymbol{I}_2) = (c - x)^2$$
má jeden dvojnásobný reálny koreň $x_{1,2} = c$. Podpriestor riešení homogénnej sústavy s maticou $c\boldsymbol{I}_2 - c\boldsymbol{I}_2 = \mathbf{0}_{2,2}$ je samozrejme celé $\mathbb{R}^2$. To znamená, že naša rovnoľahlosť má v *ľubovoľnej* báze priestoru $\mathbb{R}^2$ diagonálnu maticu $c\boldsymbol{I}_2$. Väčšinou, pokiaľ z nejakých dôvodov nedáme prednosť inej voľbe, si v takom prípade zvykneme vybrať kanonickú bázu $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2)$.

18.4.4. Príklad. Skosenie roviny v smere osi x s parametrom $a \in \mathbb{R}$ je lineárny operátor $\mathbb{R}^2 \to \mathbb{R}^2$ s maticou
$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$
vzhľadom na kanonickú bázu $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2)$ (pozri príklad 6.4.6). Keďže pre $a = 0$ ide o identické zobrazenie, ktoré má v ľubovoľnej báze maticu $\boldsymbol{I}_2$ (čo je špeciálny prípad predošlého príkladu), budeme ďalej predpokladať, že $a \neq 0$. Charakteristický polynóm
$$\begin{vmatrix} 1 - x & a \\ 0 & 1 - x \end{vmatrix} = (1 - x)^2$$
má jeden dvojnásobný reálny koreň $x_{1,2} = 1$. K nemu prislúchajúce vlastné vektory nájdeme riešením homogénnej sústavy s maticou
$$\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$
Podpriestor riešení je jednorozmerný, generovaný vektorom $\boldsymbol{e}_1 = (1, 0)^\mathsf{T}$, preto skosenie v smere osi x s nenulovým parametrom a nie je diagonalizovateľný lineárny operátor.

18.4.5. Príklad. Hyperbolická rotácia Minkowského „časopriamky“ $\mathbb{R}^{(1,1)}$ o hyperbolický uhol $\theta \in \mathbb{R}$ je lineárny operátor $\boldsymbol{Rh}_\theta \colon \mathbb{R}^2 \to \mathbb{R}^2$, ktorý má v kanonickej báze $\boldsymbol{\varepsilon} = (\boldsymbol{e}_0, \boldsymbol{e}_1)$ maticu

$$\boldsymbol{Rh}_\theta = \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix}$$

(pozri paragraf 16.7). Charakteristický polynóm

$$\begin{aligned} \det(\boldsymbol{Rh}_\theta - x\boldsymbol{I}_2) &= \begin{vmatrix} \cosh\theta - x & \sinh\theta \\ \sinh\theta & \cosh\theta - x \end{vmatrix} \\ &= x^2 - 2x\cosh\theta + \cosh^2\theta - \sinh^2\theta = x^2 - 2x\cosh\theta + 1 \end{aligned}$$

má diskriminant $D = 4\cosh^2\theta - 4 = 4\sinh^2\theta \geq 0$ a dva reálne korene

$$x_{1,2} = \cosh\theta \pm \sinh\theta = \mathrm{e}^{\pm\theta}\,.$$

Pre $\theta = 0$ je $\boldsymbol{Rh}_\theta = \boldsymbol{I}_2$, čo je už známy prípad. Pre $\theta \neq 0$ dostávame dve rôzne vlastné čísla. Príslušné vlastné vektory nájdeme riešením homogénnych sústav s maticami

$$\boldsymbol{Rh}_\theta - \mathrm{e}^\theta\boldsymbol{I}_2 = \begin{pmatrix} \cosh\theta - \mathrm{e}^\theta & \sinh\theta \\ \sinh\theta & \cosh\theta - \mathrm{e}^\theta \end{pmatrix} = \begin{pmatrix} -\sinh\theta & \sinh\theta \\ \sinh\theta & -\sinh\theta \end{pmatrix} \sim \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix},$$

resp.

$$\boldsymbol{Rh}_\theta - \mathrm{e}^{-\theta}\boldsymbol{I}_2 = \begin{pmatrix} \cosh\theta - \mathrm{e}^{-\theta} & \sinh\theta \\ \sinh\theta & \cosh\theta - \mathrm{e}^{-\theta} \end{pmatrix} = \begin{pmatrix} \sinh\theta & \sinh\theta \\ \sinh\theta & \sinh\theta \end{pmatrix} \sim \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Oba podpriestory riešení sú jednorozmerné, generované vlastnými vektormi $(1,1)^\mathsf{T}$, resp. $(1,-1)^\mathsf{T}$. Všimnite si, že ide o svetelné vektory. V nimi tvorenej báze, danej stĺpcami matice

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

má hyperbolická rotácia $\boldsymbol{Rh}_\theta$ diagonálnu maticu $\operatorname{diag}(\mathrm{e}^\theta, \mathrm{e}^{-\theta})$. To mimochodom platí aj pre $\theta = 0$.

18.5 Lineárne operátory na nekonečnorozmerných priestoroch

V tomto paragrafe (ako napokon ani v celom kurze) nie je našim cieľom systematické štúdium lineárnych operátorov na nekonečnorozmerných priestoroch. Obmedzíme sa len na dva poučné príklady, na ktorých sa výrazne prejavia rozdiely medzi konečno- a nekonečnorozmerným prípadom.

18.5.1. Príklad. Symbolom $\mathcal{C}^{(\infty)}(\mathbb{R})$ sa zvykne označovať množina všetkých funkcií $f\colon \mathbb{R} \to \mathbb{R}$, ktoré majú na celom $\mathbb{R}$ spojité derivácie všetkých rádov. Zrejme $\mathcal{C}^{(\infty)}(\mathbb{R})$ je lineárny podpriestor reálneho vektorového priestoru $\mathcal{C}(\mathbb{R})$ všetkých spojitých funkcií $\mathbb{R} \to \mathbb{R}$ s operáciami definovanými po zložkách (pozri príklady 4.1.3 a 6.1.8). Potom pre každú funkciu $f \in \mathcal{C}^{(\infty)}(\mathbb{R})$ aj jej derivácia $D(f) = f'$ patrí do $\mathcal{C}^{(\infty)}(\mathbb{R})$, teda $D\colon \mathcal{C}^{(\infty)}(\mathbb{R}) \to \mathcal{C}^{(\infty)}(\mathbb{R})$ je lineárny operátor. Podmienka $D(f) = \lambda f$ pre jeho vlastnú hodnotu a príslušnú vlastnú funkciu nie je nič iného než diferenciálna rovnica

$$f'(x) = \lambda f(x),$$

ktorá má pre každé λ riešenie

$$f(x) = f(0)\,\mathrm{e}^{\lambda x}\,.$$

To však v reči tejto kapitoly znamená, že *každé* reálne číslo λ je vlastnou hodnotou operátora D a prislúcha mu jednorozmerný vlastný podpriestor generovaný funkciou $\mathrm{e}^{\lambda x}$.

18.5.2. Príklad. Určitý integrál chápaný ako funkcia hornej medze, ktorý spojitej funkcii $f\colon \langle a, b\rangle \to \mathbb{R}$ priradí predpisom $F(x) = \int_a^x f(t)\,\mathrm{d}t$ jej primitívnu funkciu $I(f) = F$, definuje lineárny operátor $I\colon \mathcal{C}\langle a, b\rangle \to \mathcal{C}\langle a, b\rangle$ na vektorovom priestore $\mathcal{C}\langle a, b\rangle$ všetkých spojitých reálnych funkcií na intervale $\langle a, b\rangle$ (pozri príklad 6.1.9). Podmienka $I(f) = \lambda f$ pre jeho vlastnú hodnotu a príslušnú vlastnú funkciu má tvar integrálnej rovnice

$$\int_a^x f(t)\,\mathrm{d}t = \lambda f(x).$$

Ak $\lambda = 0$, tak derivovaním oboch strán podľa x zistíme, že jediná spojitá funkcia f, ktorá ju spĺňa, je identicky rovná nule. Teda 0 nie je vlastné číslo operátora I.

Nech teda $\lambda \neq 0$. Keďže funkcia na pravej strane je diferencovateľná, musí byť diferencovateľná aj f, a po derivovaní oboch strán podľa x dostávame diferenciálnu rovnicu $f(x) = \lambda f'(x)$, čiže

$$f'(x) = \frac{1}{\lambda} f(x),$$

ktorá, podobne ako v predchádzajúcom príklade, má riešenie

$$f(x) = f(a)\,\mathrm{e}^{\frac{x-a}{\lambda}}\,.$$

Dosadením $x = a$ do pôvodného vzťahu dostávame

$$\lambda f(a) = \int_a^a f(t)\,\mathrm{d}t = 0,$$

teda $f(a) = 0$, čo má opäť za následok $f(x) = 0$ pre každé x. Teda ani žiadne reálne $\lambda \neq 0$ nie je vlastným číslom operátora I.

Poučení príkladom 18.4.2 by sme sa mohli pokúšať nájsť nejaké komplexné vlastné čísla operátora I. Ale už zbežný pohľad na práve vykonané úvahy nám ukáže, že reálnosť skalára λ v nich nehrala podstatnú úlohu. Teda z rovnakých dôvodov I nemá ani komplexné vlastné čísla.

Na druhej strane, lineárny operátor $I_a : \mathcal{C}\langle a, b\rangle \to \mathcal{C}\langle a, b\rangle$ daný predpisom

$$I_a(f)(x) = f(a) + \int_a^x f(t)\,\mathrm{d}t$$

pre $f \in \mathcal{C}\langle a, b\rangle$, $x \in \langle a, b\rangle$ má jediné vlastné číslo $\lambda = 1$. Všetky riešenia príslušnej integrálnej rovnice $I_a(f) = f$ majú tvar

$$f(x) = f(a)\,\mathrm{e}^{x-a}\,.$$

To znamená, že tvoria jednorozmerný vlastný podpriestor generovaný vlastnou funkciou e^{x-a}. Presvedčte sa o tom.

Cvičenia

18.1. Nájdite vlastné čísla a k nim príslúchajúce vlastné vektory nasledujúcich matíc z $\mathbb{R}^{2\times 2}$:

(a) $\boldsymbol{A} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, (b) $\boldsymbol{B} = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}$, (c) $\boldsymbol{C} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$,

(d) $\boldsymbol{D} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$, (e) $\boldsymbol{E} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$, (f) $\boldsymbol{B} = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$,

V prípade, že vám nevyjdú reálne vlastné čísla, riešte úlohu nad poľom $\mathbb{C}$.

18.2. Nájdite vlastné čísla a k nim príslúchajúce vlastné vektory nasledujúcich matíc z $\mathbb{R}^{3\times 3}$:

(a) $\boldsymbol{A} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, (b) $\boldsymbol{B} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}$,

(c) $\boldsymbol{C} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 3 \\ 1 & 3 & 1 \end{pmatrix}$, (d) $\boldsymbol{D} = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}$.

18.3. Nech $\boldsymbol{A} \in K^{n\times n}$, $\boldsymbol{B} \in K^{m\times m}$ sú štvorcové matice. Dokážte, že nasledujúce dvojice matíc sú podobné a zakaždým nájdite regulárnu maticu, ktorá zaručí ich podobnosť: (a) $\operatorname{diag}(\boldsymbol{A}, \boldsymbol{B}) \approx \operatorname{diag}(\boldsymbol{B}, \boldsymbol{A})$; (b) $\boldsymbol{A} \approx \boldsymbol{A}^{\mathsf{T}}$.

18.4. Dokážte dôsledok 18.1.4 o invariantnosti determinantu a stopy vzhľadom na ekvivalenciu podobnosti matíc.

18.5. (a) Dokážte, že rovnosťou $\langle \boldsymbol{A}, \boldsymbol{B}\rangle = \operatorname{tr}\bigl(\boldsymbol{A} \cdot \boldsymbol{B}^{\mathsf{T}}\bigr)$ je definovaný skalárny súčin na reálnom vektorovom priestore $\mathbb{R}^{m\times n}$.

(b) Dokážte, že rovnosťou $\langle \boldsymbol{A}, \boldsymbol{B} \rangle = \operatorname{tr}(\boldsymbol{A} \cdot \boldsymbol{B}^*)$ je definovaný skalárny súčin na komplexnom vektorovom priestore $\mathbb{C}^{m\times n}$.

18.6. (a) Nájdite príklady matíc rozmeru 2×2, resp. 3×3, ktoré maju rovnaký determinant aj stopu, ale nie sú podobné.

(b) Nájdite príklady matíc rozmeru 2×2, resp. 3×3, ktoré maju rovnaký charakteristický polynóm, ale nie sú podobné. Čím sa líšia obe úlohy?

(c) Dokážte že dve matice rozmeru 2×2, ktoré majú rovnaký determinant aj stopu, majú rovnaký charakteristický polynóm.

(d) Nájdite príklad dvoch matíc rozmeru 3×3, ktoré majú rovnaký determinat aj stopu, ale rôzne charakteristické polynómy.

18.7. Nech $\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ je reálna matica. Dokážte postupne nasledujúce tvrdenia:

(a) $\boldsymbol{A}$ má dve rôzne reálne vlastné čísla práve vtedy, keď $(a-d)^2 > -4bc$.

(b) $\boldsymbol{A}$ má dvojnásobné vlastné číslo práve vtedy, keď $(a-d)^2 = -4bc$.

(c) Ak $\boldsymbol{A}$ je symetrická, tak obe jej vlastné čísla sú reálne.

(d) Ak $\boldsymbol{A}$ je symetrická, tak $\boldsymbol{A}$ má dvojnásobné vlastné číslo práve vtedy, keď $a = d$ a $b = c = 0$.

18.8. (a) Ak aspoň jedna z matíc $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$ je regulárna, tak matice $\boldsymbol{A} \cdot \boldsymbol{B}$, $\boldsymbol{B} \cdot \boldsymbol{A}$ sú podobné. Dokážte.

(b) Nájdite príklad matíc $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{2\times 2}$ takých, že matice $\boldsymbol{A}\cdot\boldsymbol{B}$, $\boldsymbol{B}\cdot\boldsymbol{A}$ nie sú podobné.

(c) Ak $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$, tak matice $\boldsymbol{A}\cdot\boldsymbol{B}$, $\boldsymbol{B}\cdot\boldsymbol{A}$ majú rovnaké vlastné hodnoty. Dokážte.

(d) Ak $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times m}$, tak matice $\boldsymbol{A}\cdot\boldsymbol{B}$, $\boldsymbol{B}\cdot\boldsymbol{A}$ majú rovnaké *nenulové* vlastné hodnoty. Dokážte.
(e) Nájdite príklad matíc $\boldsymbol{A} \in \mathbb{R}^{3\times 2}$, $\boldsymbol{B} \in \mathbb{R}^{2\times 3}$ takých, že 0 je vlastnou hodnotou matice $\boldsymbol{A}\cdot\boldsymbol{B}$, no nie je vlastnou hodnotou matice $\boldsymbol{B}\cdot\boldsymbol{A}$. Mohlo by to byť aj naopak?

18.9. (a) Nájdite vlastné čísla a vlastné vektory lineárnych operátorov $D, \Delta\colon \mathbb{R}^{(n)}[x] \to \mathbb{R}^{(n)}[x]$ z cvičení 6.9 resp. 6.10.

(b) Zovšeobecnite úlohu (a) na lineárne operátory $D, \Delta\colon K[x] \to K[x]$ dané predpismi $D(f)(x) = f'(x)$ resp. $\Delta(f)(x) = f(x+1) - f(x)$ pre polynómy $f(x)$ akéhokoľvek stupňa nad ľubovoľným poľom K.

18.10. (a) Ak $\boldsymbol{v} \in K^n$ je vlastný vektor matice $\boldsymbol{A} \in K^{n\times n}$ prislúchajúci k jej vlastnej hodnote $\lambda \in K$ a $k \in \mathbb{N}$, tak $\boldsymbol{v}$ je aj vlastným vektorom matice $\boldsymbol{A}^k$ prislúchajúcim k jej vlastnej hodnote λ^k. Dokážte.

(b) Nech $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$. Musí mať každé vlastné číslo matice $\boldsymbol{A}^2$ tvar $\mu = \lambda^2$, kde $\lambda \in \mathbb{C}$ je vlastné číslo matice $\boldsymbol{A}$? Zmení sa niečo, ak predpokladáme, že $\boldsymbol{A} \in \mathbb{C}^{2\times 2}$?

(c) Za predpokladu, že $\boldsymbol{A}$ je regulárna, sformulujte a dokážte tvrdenie analogické (a) aj pre záporné exponenty $k \in \mathbb{Z}$.

19. Spektrum lineárneho operátora

V tejto kapitole budeme pokračovať v štúdiu štruktúry lineárnych operátorov na konečnorozmerných vektorových priestoroch. Zavedieme dôležitý, ešte v predchádzajúcej kapitole avizovaný pojem spektra lineárneho operátora, ako i pojmy algebraickej a geometrickej násobnosti vlastnej hodnoty, ktoré nám umožnia klasifikovať prípadné prekážky jeho diagonalizovateľnosti.

Ďalej si ujasníme, aký vplyv má riešiteľnosť polynomických rovníc v základnom poli na spektrum lineárneho operátora. Vhodným rozšírením tohto poľa možno dosiahnuť, aby každý lineárny operátor na n-rozmernom priestore mal n vlastných hodnôt, ak každú z nich počítame toľkokrát, aká je jej algebraická násobnosť.

19.1 Spektrum lineárneho operátora a matice

Aby sme sa mohli pohnúť ďalej, je potrebné si pripomenúť zopár základných poznatkov o polynómoch. Aspoň v prípade polí $\mathbb{Q}$ a $\mathbb{R}$, a možno aj $\mathbb{C}$, by malo ísť o záležitosti známe zo stredoškolskej matematiky; prechod k ľubovoľnému poľu však nepredstavuje žiadnu ťažkosť.

Hovoríme, že polynóm $f(x) \in K[x]$ *delí* polynóm $g(x) \in K[x]$, ak existuje polynóm $p(x) \in K[x]$ taký, že $g(x) = f(x)\,p(x)$. Zrejme, ak $f(x)$ delí $g(x)$, tak stupeň $f(x)$ je menší alebo rovný stupňu $g(x)$.

Nech $f(x) \in K[x]$ je polynóm stupňa $n \geq 1$. Skalár $\lambda \in K$ je *koreňom polynómu* $f(x)$ (t. j. $f(\lambda) = 0$) práve vtedy, keď polynóm $x - \lambda$ delí polynóm $f(x)$. Hovoríme, že skalár $\lambda \in K$ je *m-násobný koreň polynómu* $f(x) \in K[x]$, ak $(x - \lambda)^m$ je najvyššia mocnina polynómu $x - \lambda$, ktorá ešte delí $f(x)$. Miesto 1-násobný hovoríme *jednoduchý koreň*. Z úvahy o stupňoch vyplýva, že polynóm $f(x)$ stupňa $n \geq 1$ má nanajvýš n koreňov, ak každý z nich počítame aj s jeho násobnosťou. Presnejšie, ak $\lambda_1, \ldots, \lambda_k$ sú všetky navzájom rôzne korene polynómu $f(x)$ stupňa n a $m_1, \ldots, m_k$ sú ich násobnosti, tak $m_1 + \ldots + m_k \leq n$.

Ďalej budene striedavo hovoriť raz o maticiach a inokedy o lineárnych operátoroch, podľa toho, čo bude pre nás v danej chvíli výhodnejšie. Na čitateľa nechávame, aby si podľa potreby sám urobil preklad z maticovej reči do operátorovej alebo naopak.

Skalár $\lambda \in K$ sa nazýva *m-násobná vlastná hodnota* matice $\boldsymbol{A} \in K^{n\times n}$, ak λ je m-násobným koreňom jej charakteristického polynómu $\mathrm{ch}_{\boldsymbol{A}}(x)$; hovoríme tiež, že *algebraická násobnosť vlastnej hodnoty* λ matice $\boldsymbol{A}$ je m. Miesto 1-násobná hovoríme *jednoduchá vlastná hodnota*. Podobne definujeme aj pojem

m-násobnej vlastnej hodnoty a algebraickej násobnosti vlastnej hodnoty pre lineárne operátory na konečnorozmerných priestoroch.

Z našich úvah o koreňoch polynómov vyplýva, že ak matica $\boldsymbol{A} \in K^{n\times n}$ má k navzájom rôznych vlastných hodnôt $\lambda_1, \dots, \lambda_k \in K$ s algebraickými násobnosťami $m_1, \dots, m_k$, tak $m_1 + \ldots + m_k \leq n$. Inak povedané, matica rádu n má nanajvýš n vlastných hodnôt, ak každú z nich počítame s jej násobnosťou.

Spektrom lineárneho operátora φ na konečnorozmernom vektorovom priestore nazývame množinu všetkých jeho vlastných hodnôt a označujeme ju $\operatorname{Spec}\varphi$. Podotýkame, že takto definujeme spektrum len v konečnorozmernom prípade; definícia spektra lineárneho operátora na nekonečnorozmernom vektorovom priestore je podstane zložitejšia záležitosť, ktorá presahuje rámec lineárnej algebry. Rovnako definujeme aj *spektrum matice* $\boldsymbol{A} \in K^{n\times n}$, ktoré značíme $\operatorname{Spec}\boldsymbol{A}$.

Algebraickou váhou spektra $\operatorname{Spec}\varphi$ nazývame súčet algebraických násobností všetkých vlastných hodnôt $\lambda \in \operatorname{Spec}\varphi$. Hovoríme, že lineárny operátor φ má *jednoduché spektrum*, ak sa jeho algebraická váha rovná počtu jeho prvkov, t. j. práve vtedy, keď všetky vlastné hodnoty operátora φ sú jednoduché.

Popri algebraickej násobnosti zavedieme aj tzv. geometrickú násobnosť. Spomeňme si, že množina všetkých lineárnych operátorov na vektorovom priestore V sama tvorí vektorový priestor $\mathcal{L}(V,V)$ nad poľom K, ktorého prvkom je aj identický operátor $\mathrm{id}_V \colon V \to V$ (pozri paragraf 6.5). Preto pre $\varphi \in \mathcal{L}(V,V)$ a $\lambda \in K$ aj zobrazenie $\varphi - \lambda\,\mathrm{id}_V$, dané predpisom $\boldsymbol{x} \mapsto \varphi(\boldsymbol{x}) - \lambda\boldsymbol{x}$, je lineárny operátor na V; stručne ho budeme značiť $\varphi - \lambda$. Zrejme $\boldsymbol{0} \neq \boldsymbol{v} \in V$ je vlastný vektor operátora φ prislúchajúci k jeho vlastnej hodnote λ práve vtedy, keď $\boldsymbol{v} \in \operatorname{Ker}(\varphi - \lambda)$.

Lineárny podpriestor $\operatorname{Ker}(\varphi - \lambda) \subseteq V$ nazývame *vlastný podpriestor lineárneho operátora* $\varphi \colon V \to V$ prislúchajúci k jeho vlastnej hodnote $\lambda \in K$. Zrejme pre všetky $\boldsymbol{v} \in \operatorname{Ker}(\varphi - \lambda)$ platí $\varphi(\boldsymbol{v}) = \lambda\boldsymbol{v}$, teda $\operatorname{Ker}(\varphi - \lambda)$ je invariantný podpriestor operátora φ.

Geometrickou násobnosťou vlastnej hodnoty λ lineárneho operátora φ nazývame dimenziu $\dim\operatorname{Ker}(\varphi - \lambda)$ jeho vlastného podpriestoru prislúchajúceho k λ.

Geometrická násobnosť vlastnej hodnoty λ matice $\boldsymbol{A} \in K^{n\times n}$ sa zrejme rovná číslu $\dim\mathcal{R}(\boldsymbol{A} - \lambda\boldsymbol{I})$. Pripomíname, že $\mathcal{R}(\boldsymbol{A} - \lambda\boldsymbol{I})$ označuje podpriestor riešení homogénnej sústavy s maticou $\boldsymbol{A} - \lambda\boldsymbol{I}$, takže platí

$$1 \leq \dim\mathcal{R}(\boldsymbol{A} - \lambda\boldsymbol{I}) = n - h(\boldsymbol{A} - \lambda\boldsymbol{I}) \leq n$$

(pozri paragraf 9.1).

Geometrickou váhou spektra lineárneho operátora φ na konečnorozmernom vektorovom priestore nazývame súčet geometrických násobností všetkých jeho vlastných hodnôt.

Pojmy ako vlastný podpriestor, algebraická či geometrická násobnosť (ak za ňu pripustíme aj nulu) možno formálne rovnako zaviesť pre lineárny operátor φ a ľubovoľný skalár $\lambda \in K$, nielen pre jeho vlastné hodnoty. No $\lambda \in K$ je vlastná hodnota práve vtedy, keď $\operatorname{Ker}(\varphi - \lambda) \neq \{\mathbf{0}\}$, čo je ekvivalentné s nenulovosťou tak geometrickej ako aj algebraickej násobnosti λ vzhľadom na φ. Teda 0-násobné, t. j. uvnevlastné hodnoty lineárneho operátora neprispievajú k algebraickej ani ku geometrickej váhe jeho spektra.

Pokúsme sa teraz nejako usúvzťažniť množstvo nových pojmov, ktoré sme práve definovali. Začneme jedným pomocným tvrdením.

19.1.1. Lema. *Nech φ je lineárny operátor na konečnorozmernom vektorovom priestore V a $S \subseteq V$ je jeho invariantný podpriestor. Označme $\varphi_1 = \varphi \restriction S$ zúženie operátora φ na podpriestor S. Potom charakteristický polynóm $\mathrm{ch}_{\varphi_1}(x)$ delí charakteristický polynóm $\mathrm{ch}_\varphi(x)$. Ak navyše $T \subseteq V$ je invariantný podpriestor taký, že $V = S \oplus T$ a $\varphi_2 = \varphi \restriction T$, tak $\mathrm{ch}_\varphi(x) = \mathrm{ch}_{\varphi_1}(x)\,\mathrm{ch}_{\varphi_2}(x)$.*

Dôkaz. Nech $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_n)$ je báza priestoru V, pričom jej prvých k členov tvorí bázu podpriestoru S. Ako sme si ujasnili v paragrafe 18.2, vzhľadom na takúto bázu má φ maticu v blokovom tvare

$$\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha}} = \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{M} \\ \mathbf{0}_{n-k,k} & \boldsymbol{A}_2 \end{pmatrix},$$

kde $\boldsymbol{A}_1 \in K^{k\times k}$ je matica operátora $\varphi_1\colon S \to S$ v báze $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ a $\boldsymbol{M} \in K^{k\times(n-k)}$, $\boldsymbol{A}_2 \in K^{(n-k)\times(n-k)}$. Pomocou tvrdenia 10.2.2 dostávame

$$\begin{aligned} \mathrm{ch}_\varphi(x) = \mathrm{ch}_{\boldsymbol{A}}(x) = |\boldsymbol{A} - x\boldsymbol{I}_n| &= \begin{vmatrix} \boldsymbol{A}_1 - x\boldsymbol{I}_k & \boldsymbol{M} \\ \mathbf{0}_{n-k,k} & \boldsymbol{A}_2 - x\boldsymbol{I}_{n-k} \end{vmatrix} \\ &= |\boldsymbol{A}_1 - x\boldsymbol{I}_k|\,|\boldsymbol{A}_2 - x\boldsymbol{I}_{n-k}| \\ &= \mathrm{ch}_{\boldsymbol{A}_1}(x)\,\mathrm{ch}_{\boldsymbol{A}_2}(x) = \mathrm{ch}_{\varphi_1}(x)\,\mathrm{ch}_{\boldsymbol{A}_2}(x). \end{aligned}$$

Druhá časť lemy je už triviálnym dôsledkom našich úvah.

19.1.2. Tvrdenie. *Nech λ je vlastná hodnota lineárneho operátora φ na konečnorozmernom vektorovom priestore V. Potom jej geometrická násobnosť je menšia alebo rovná jej algebraickej násobnosti.*

Dôkaz. Nech $S = \operatorname{Ker}(\varphi - \lambda)$ je vlastný podpriestor operátora φ prislúchajúci k λ a $\varphi_1 = \varphi \restriction S$. Potom $k = \dim S$ je geometrická násobnosť λ vzhľadom na φ; algebraickú násobnosť λ vzhľadom na φ označíme m. Nech $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$ je ľubovoľná báza S. Keďže každé $\boldsymbol{v}_i$ je vlastný vektor operátora φ prislúchajúci k λ, zrejme $(\varphi_1)_{\boldsymbol{\beta}} = \lambda\boldsymbol{I}_k$. Preto $\mathrm{ch}_{\varphi_1}(x) = (\lambda - x)^k$.

Podľa predchádzajúcej lemy tento polynóm delí $\mathrm{ch}_\varphi(x)$. Nakoľko $(x-\lambda)^m$ je najvyššia mocnina $x-\lambda$, ktorá delí $\mathrm{ch}_\varphi(x)$, platí $k \le m$.

Ak algebraická násobnosť skalára λ vzhľadom na operátor φ je ≥ 1, t. j. ak λ je vlastná hodnota, tak aj geometrická násobnosť λ vzhľadom na φ je aspoň 1 (teda nie je 0). Ako ukazuje nasledujúci veľmi dôležitý príklad, až na toto minimálne obmedzenie môže byť rozdiel medzi algebraickou a geometrickou násobnosťou vlastnej hodnoty lineárneho operátora ľubovoľne veľký.

19.1.3. Príklad. Označme $\boldsymbol{J}_n \in K^{n\times n}$ štvorcovú maticu rádu n, ktorej prvky na miestach $(i, i+1)$ sú rovné 1 pre $1 \le i \le n-1$ a všetky ostatné prvky sú rovné 0. Zrejme $h(\boldsymbol{J}_n) = n-1$. Ďalej položme

$$\boldsymbol{J}_n(\lambda) = \lambda \boldsymbol{I}_n + \boldsymbol{J}_n = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & \dots & 0 & \lambda \end{pmatrix},$$

pre $\lambda \in K$. Teda $\boldsymbol{J}_n(\lambda)$ je tvorená diagonálou z n lámbd, vedľajšou diagonálou vpravo od nej z $n-1$ jednotiek a zvyšok sú nuly. Tak napríklad

$$\boldsymbol{J}_1(\lambda) = (\lambda), \qquad \boldsymbol{J}_2(\lambda) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \qquad \boldsymbol{J}_3(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \qquad \text{atď.}$$

Každá matica tvaru $\boldsymbol{J}_n(\lambda)$ sa nazýva *Jordanova bunka*, prípadne *Jordanov blok* rádu n prislúchajúci skaláru λ. Zrejme aj $\boldsymbol{J}_n = \boldsymbol{J}_n(0)$ je Jordanova bunka.

Charakteristický polynóm Jordanovej bunky $\boldsymbol{J}_n(\lambda)$ je

$$\det(\boldsymbol{J}_n(\lambda) - x\boldsymbol{I}_n) = \det \boldsymbol{J}_n(\lambda - x) = (\lambda - x)^n.$$

Táto matica má jedinú vlastnú hodnotu $x = \lambda$ s algebraickou násobnosťou n. Na druhej strane, podpriestor riešení homogénnej sústavy s maticou $\boldsymbol{J}_n(\lambda) - \lambda \boldsymbol{I}_n = \boldsymbol{J}_n$ je jednorozmerný, generovaný vlastným vektorom $\boldsymbol{e}_1 = (1, 0, \dots, 0)^{\mathsf{T}}$. Teda geometrická násobnosť vlastnej hodnoty λ matice $\boldsymbol{J}_n(\lambda)$ je stále len 1, bez ohľadu na to, aké veľké je n. Preto $\boldsymbol{J}_n(\lambda)$ pre $n \ge 2$ nie je podobná so žiadnou diagonálnou maticou.

Ako sme už spomínali, podobné matice sú maticami toho istého lineárneho operátora. Tvrdenie 18.2.1 tak možno zosilniť do nasledujúcej podoby.

19.1.4. Tvrdenie. *Podobné matice majú rovnaké spektrum, vrátane algebraickej i geometrickej násobnosti každej vlastnej hodnoty.*

Pomocou pojmov spektra a algebraickej a geometrickej násobnosti teraz môžeme podstatne spresniť charakterizáciu diagonalizovateľných operátorov z vety 18.2.2.

19.1.5. Veta. *Nech φ je lineárny operátor na konečnorozmernom vektorovom priestore V dimenzie n. Potom nasledujúce podmienky sú ekvivalentné:*
(i) φ je diagonalizovateľný;
(ii) geometrická váha spektra $\operatorname{Spec}\varphi$ *sa rovná n;*
(iii) algebraická váha spektra $\operatorname{Spec}\varphi$ *sa rovná n a algebraická násobnosť každej vlastnej hodnoty sa rovná jej geometrickej násobnosti;*
(iv) algebraická i geometrická váha spektra $\operatorname{Spec}\varphi$ *sa rovná n.*

Dôkaz. Z vety 18.2.2 vyplývajú implikácie (iv) $\Rightarrow$ (i) $\Rightarrow$ (ii); (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) sú zasa dôsledkom tvrdenia 19.1.2.

Práve vyslovená veta nám umožňuje rozdeliť prekážky, ktoré bránia diagonalizácii nejakého lineárneho operátora φ na n-rozmernom vektorovom priestore V, do dvoch kategórií:
(1) algebraická váha spektra $\operatorname{Spec}\varphi$ je menšia než $n = \dim V$, inak povedané, φ má „málo“ vlastných hodnôt v poli K, i keď každú z nich počítame aj s jej algebraickou násobnosťou;
(2) geometrická váha spektra $\operatorname{Spec}\varphi$ je menšia než jeho algebraická váha, čiže geometrická násobnosť niektorých vlastných hodnôt nedosahuje ich algebraickú násobnosť.

Uvedené dva typy prekážok sa môžu u lineárnych operátorov vyskytovať každá zvlášť i obe súčasne. V druhej polovici tejto kapitoly si ukážeme, ako možno prekážky prvého typu prekonať prechodom do „bohatšieho“ poľa (niečo také sme už naznačili v príklade 18.4.2). Prekážky druhého typu majú zásadnejší charakter a definitívne vylučujú diagonalizáciu. Ako však uvidíme v nasledujúcej kapitole, i v tomto prípade sa možno do značnej miery priblížiť diagonálnemu tvaru pomocou blokovo diagonálnych matíc zložených zo Jordanových buniek.

19.2 Schurova veta o triangularizácii

Zatiaľ aspoň dokážeme, že v maticiach s „dostatočne mnoho“ vlastnými hodnotami možno vynulovať všetky prvky pod diagonálou. Z toho priamo vyplýva, že každý lineárny operátor s plnou algebraickou váhou spektra má vo vhodnej báze hornú trojuholníkovú maticu (pozri koniec paragraf 10.2), ktorej diagonálu tvorí spektrum operátora, vrátane algebraickej násobnosti každej vlastnej hodnoty.

19.2.1. Veta. (Schurova veta o triangularizácii) *Nech matica $\boldsymbol{A} \in K^{n\times n}$ má spektrum algebraickej váhy n. Potom $\boldsymbol{A}$ je podobná s hornou trojuholníkovou maticou $\boldsymbol{C} = \boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P}$, ktorej diagonálu tvoria všetky vlastné hodnoty matice $\boldsymbol{A}$, pričom každá sa tu vyskytuje toľkokrát, aká je jej algebraická násobnosť. Navyše možno zabezpečiť, aby matica prechodu $\boldsymbol{P}$ bola v prípade poľa $K = \mathbb{R}$ ortogonálna a v prípade poľa $K = \mathbb{C}$ unitárna.*

Dôkaz. Najprv sa sústredíme iba na prvú časť vety. Dokážeme len existenciu hornej trojuholníkovej matice $\boldsymbol{C} \approx \boldsymbol{A}$. Zvyšok vety je už nevyhnutným dôsledkom jej trojuholníkového tvaru. Budeme postupovať indukciou podľa n.

Pre $n = 1$ má už samotná matica $\boldsymbol{A} = (a)$ žiadaný tvar. Nech teda $n \geq 2$ a predpokladajme, že každá matica z $K^{(n-1)\times(n-1)}$ s plnou algebraickou váhou spektra je podobná s nejakou hornou trojuholníkovou maticou. Nech λ je niektorá z vlastných hodnôt matice $\boldsymbol{A}$. Keďže k nej príslušný vlastný vektor generuje $\boldsymbol{A}$-invariantný podpriestor, $\boldsymbol{A}$ je podobná s blokovou maticou

$$\boldsymbol{B} = \begin{pmatrix} \lambda & \boldsymbol{z} \\ \boldsymbol{0} & \boldsymbol{B}_1 \end{pmatrix},$$

kde $\boldsymbol{0} \in K^{(n-1)\times 1}$, $\boldsymbol{z} \in K^{1\times(n-1)}$, $\boldsymbol{B}_1 \in K^{(n-1)\times(n-1)}$. Preto $\boldsymbol{B} = \boldsymbol{R}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{R}$ pre nejakú regulárnu maticu $\boldsymbol{R} \in K^{n\times n}$. Keďže $\boldsymbol{A}$ aj $\boldsymbol{B}$ majú spektrum algebraickej váhy n, algebraická váha spektra $\boldsymbol{B}_1$ je $n-1$. Podľa indukčného predpokladu $\boldsymbol{B}_1$ je podobná s hornou trojuholníkovou maticou $\boldsymbol{C}_1$, teda $\boldsymbol{C}_1 = \boldsymbol{Q}^{-1} \cdot \boldsymbol{B}_1 \cdot \boldsymbol{Q}$ pre vhodnú regulárnu maticu $\boldsymbol{Q} \in K^{(n-1)\times(n-1)}$. Potom aj $\operatorname{diag}(1, \boldsymbol{Q}) \in K^{n\times n}$ je regulárna. Pomocou pravidla pre násobenie blokových matíc (pozri paragraf 2.2.3) dostávame

$$\begin{aligned}
\boldsymbol{A} &\approx (\boldsymbol{R} \cdot \operatorname{diag}(1, \boldsymbol{Q}))^{-1} \cdot \boldsymbol{A} \cdot (\boldsymbol{R} \cdot \operatorname{diag}(1, \boldsymbol{Q})) \\
&= \operatorname{diag}(1, \boldsymbol{Q})^{-1} \cdot \boldsymbol{R}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{R} \cdot \operatorname{diag}(1, \boldsymbol{Q}) \\
&= \operatorname{diag}\big(1, \boldsymbol{Q}^{-1}\big) \cdot \begin{pmatrix} \lambda & \boldsymbol{z} \\ \boldsymbol{0} & \boldsymbol{B}_1 \end{pmatrix} \cdot \operatorname{diag}(1, \boldsymbol{Q}) \\
&= \begin{pmatrix} \lambda & \boldsymbol{z} \cdot \boldsymbol{Q} \\ \boldsymbol{0} & \boldsymbol{Q}^{-1} \cdot \boldsymbol{B}_1 \cdot \boldsymbol{Q} \end{pmatrix} = \begin{pmatrix} \lambda & \boldsymbol{z} \cdot \boldsymbol{Q} \\ \boldsymbol{0} & \boldsymbol{C}_1 \end{pmatrix},
\end{aligned}$$

čo je opäť horná trojuholníková matica.

Podľa prvej časti vety existuje regulárna matica $\boldsymbol{P}$ taká, že matica $\boldsymbol{C} = \boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P}$ je horná trojuholníková. V prípade poľa $\mathbb{R}$ resp. $\mathbb{C}$ označme $\boldsymbol{S}$ maticu, ktorá vznikne ortogonalizáciou stĺpcov matice $\boldsymbol{P}$ podľa Gramovho-Schmidtovho procesu (vzhľadom na štandardný skalárny súčin v $\mathbb{R}^n$ resp.

v $\mathbb{C}^n$) a ich následným znormovaním. Potom $\boldsymbol{S}$ je ortogonálna resp. unitárna matica, a vzhľadom na to, že k-ty vektor vznikajúci pri Gramovom-Schmidtovom procese je vždy lineárnou kombináciou prvých $k-1$ pôvodných vektorov, $\boldsymbol{S} = \boldsymbol{P} \cdot \boldsymbol{T}$ pre jednoznačne určenú regulárnu hornú trojuholníkovú maticu $\boldsymbol{T}$. Potom

$$\boldsymbol{T}^{-1} \cdot \boldsymbol{C} \cdot \boldsymbol{T} = \boldsymbol{T}^{-1} \cdot \boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P} \cdot \boldsymbol{T} = \boldsymbol{S}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{S}$$

je horná trojuholníková matica, podobná s maticou $\boldsymbol{A}$ prostredníctvom ortogonálnej resp. unitárnej matice $\boldsymbol{S}$.

Preformulovaním práve dokázanej vety do reči lineárnych operátorov okamžite dostávame nasledujúci dôsledok.

19.2.2. Dôsledok. *Nech φ je lineárny operátor na n-rozmernom vektorovom priestore V nad poľom K so spektrom s algebraickou váhou n. Potom existuje báza $\boldsymbol{\beta}$ priestoru V, vzhľadom na ktorú má φ hornú trojuholníkovú maticu s diagonálou tvorenou jeho vlastnými hodnotami. V prípade poľa $K = \mathbb{R}$ a euklidovského priestoru V resp. poľa $K = \mathbb{C}$ a unitárneho priestoru V možno navyše zabezpečiť, aby báza $\boldsymbol{\beta}$ bola ortonormálna.*

Charakteristický polynóm hornej trojuholníkovej matice $\boldsymbol{C} \in K^{n\times n}$ s diagonálou $(\lambda_1, \dots, \lambda_n)$ je

$$\begin{aligned}\mathrm{ch}_{\boldsymbol{C}}(x) &= (\lambda_1 - x) \dots (\lambda_n - x)\\ &= \lambda_1 \dots \lambda_n - \dots + (-1)^{n-1}(\lambda_1 + \dots + \lambda_n)x^{n-1} + (-1)^n x^n,\end{aligned}$$

a jeho koeficienty pri mocninách x^0 a x^{n-1} možno tentokrát (na rozdiel od všeobecnej matice) nahliadnuť naozaj bez ťažkostí. Taktiež platí

$$\det \boldsymbol{C} = \lambda_1 \dots \lambda_n, \qquad \mathrm{tr}\, \boldsymbol{C} = \lambda_1 + \dots + \lambda_n.$$

Kedže podobné matice majú rovnaký determinant, stopu aj charakteristický polynóm, z práve dokázanej vety 19.2.1 taktiež vyplýva

19.2.3. Tvrdenie. *Nech matica $\boldsymbol{A} \in K^{n\times n}$ má n (nie nutne rôznych) vlastných hodnôt $\lambda_1, \dots, \lambda_n$. Potom*

$$\det \boldsymbol{A} = \lambda_1 \dots \lambda_n, \qquad \mathrm{tr}\, \boldsymbol{A} = \lambda_1 + \dots + \lambda_n$$

a v charakteristickom polynóme $\mathrm{ch}_{\boldsymbol{A}}(x)$ sú koeficienty pri mocninách x^0 a x^{n-1} rovné $\det \boldsymbol{A}$ resp. $(-1)^{n-1}\, \mathrm{tr}\, \boldsymbol{A}$.

19.2.4. Príklad. Uvažujme reálnu maticu

$$\boldsymbol{A} = \begin{pmatrix} 12 & 0 & -28 \\ 11 & 1 & -29 \\ 7 & 0 & -16 \end{pmatrix}.$$

Jej charakteristický polynóm

$$\mathrm{ch}_{\boldsymbol{A}}(x) = \begin{vmatrix} 12-x & 0 & -28 \\ 11 & 1-x & -29 \\ 7 & 0 & -16-x \end{vmatrix} = 4 - 3x^2 - x^3 = (1-x)(2+x)^2$$

má jednoduchý koreň $x_1 = 1$ a dvojnásobný koreň $x_{2,3} = -2$. Homogénna sústava s maticou

$$\boldsymbol{A} - \boldsymbol{I}_3 = \begin{pmatrix} 11 & 0 & -28 \\ 11 & 0 & -29 \\ 7 & 0 & -17 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

má jednorozmerný podpriestor riešení generovaný vektorom $(0,1,0)^\mathsf{T}$. Homogénna sústava s maticou

$$\boldsymbol{A} + 2\boldsymbol{I}_3 = \begin{pmatrix} 14 & 0 & -28 \\ 11 & 3 & -29 \\ 7 & 0 & -14 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -7/3 \\ 0 & 0 & 0 \end{pmatrix}$$

má opäť len jednorozmerný podpriestor riešení generovaný vektorom $(6,7,3)^\mathsf{T}$.

Matica $\boldsymbol{A}$ teda má jednoduché vlastné číslo 1, ku ktorému prislúcha vlastný vektor $(0,1,0)^\mathsf{T}$ a algebraicky dvojnásobné vlastné číslo -2 s geometrickou násobnosťou 1, ku ktorému prislúcha vlastný vektor $(6,7,3)^\mathsf{T}$. Preto $\boldsymbol{A}$ nie je podobná so žiadnou diagonálnou maticou. Ukážeme, ako k nej možno nájsť podobnú hornú trojuholníkovú maticu.

Vlastný vektor $(0,1,0)^\mathsf{T}$ prislúchajúci k vlastnému číslu 1 doplníme do bázy $\mathbb{R}^3$ tvorenej stĺpcami matice

$$\boldsymbol{P} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Potom

$$\boldsymbol{A} \approx \boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P} = \begin{pmatrix} 1 & 11 & -29 \\ 0 & 12 & -28 \\ 0 & 7 & -16 \end{pmatrix} = \boldsymbol{B}.$$

O matici $\boldsymbol{B}_1 \in \mathbb{R}^{2\times 2}$ tvoriacej pravý dolný matice $\boldsymbol{B}$ už vieme, že má jediné algebraicky dvojnásobné vlastné číslo -2. K nemu príslušný vlastný vektor nájdeme riešením homogénnej sústavy s maticou

$$\boldsymbol{B}_1 + 2\boldsymbol{I}_2 = \begin{pmatrix} 14 & -28 \\ 7 & -14 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix}.$$

Podpriestor riešení je generovaný vektorom $(2,1)^{\mathsf{T}}$. Doplnením tohto vektora do bázy priestoru $\mathbb{R}^2$ dostaneme (napr.) maticu

$$\boldsymbol{Q} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}.$$

Potom

$$\begin{pmatrix} 12 & -28 \\ 7 & -16 \end{pmatrix} \approx \boldsymbol{Q}^{-1} \cdot \begin{pmatrix} 12 & -28 \\ 7 & -16 \end{pmatrix} \cdot \boldsymbol{Q} = \begin{pmatrix} -2 & 7 \\ 0 & -2 \end{pmatrix}.$$

Všimnite si, že vlastná hodnota -2 sa nám zákonite objavila aj v pravom dolnom rohu „sama od seba“.

Stĺpce matice

$$\boldsymbol{R} = \boldsymbol{P} \cdot \operatorname{diag}(1, \boldsymbol{Q}) = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

tvoria bázu, vzhľadom na ktorú má lineárny operátor $\boldsymbol{x} \mapsto \boldsymbol{A} \cdot \boldsymbol{x}$ na $\mathbb{R}^3$ hornú trojuholníkovú maticu. Kedže

$$(11, -29) \cdot \boldsymbol{Q} = (-7, 11)$$

($\boldsymbol{z} = (11, -29)$ je prvý riadok matice $\boldsymbol{B}$ bez prvého člena), bez ďalšieho násobenia vidíme, že táto matica má tvar

$$\boldsymbol{R}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{R} = \begin{pmatrix} 1 & -7 & 11 \\ 0 & -2 & 7 \\ 0 & 0 & -2 \end{pmatrix} \approx \boldsymbol{A}.$$

Ešte poznamenajme, že pre maticu $\boldsymbol{A}$ rádu $n > 3$ by sme museli procedúru vyčleňovania vlastných čísel a príslušných vlastných podpriestorov opakovať viackrát, pokým by sme sa nedopracovali k hornej trojuholníkovej matici rádu 2 (ak by sme náhodou nemali šťastie a nedospeli už skôr k hornej trojuholníkovej matici väčšieho rádu). Až potom by sme mohli začať z dielčích matíc prechodu spätne skladať výslednú maticu prechodu a pomocou nej získať hornú trojuholníkovú maticu podobnú s $\boldsymbol{A}$.

19.3 Rozšírenia polí, algebraicky uzavreté polia

Pri ďalšom štúdiu spektra lineárnych operátorov sa nezaobídeme bez niektorých hlbších vedomostí o štruktúre polí a polynómov nad nimi. Ak by sme trvali na úplnosti a sebestačnosti nášho výkladu, museli by sme v tejto chvíli značne odbočiť od našej hlavnej témy a začať rozvíjať inú časť algebry. Miesto toho iba sformulujeme niekoľko základných výsledkov, o ktoré sa budeme ďalej opierať. Ich dôkazy vynecháme, takže ich vlastne len predložíme čitateľovi na uverenie. Ich formulácia si však vyžaduje zaviesť niekoľko nových pojmov, pri ktorých nám prídu vhod i niektoré poznatky z lineárnej algebry.

Pripomíname (pozri paragraf 1.2), že pole K sa nazýva *podpoľom* poľa L, prípadne pole L sa nazýva *rozšírením poľa* K, ak $K \subseteq L$, obe polia majú tú istú nulu 0 a jednotku 1 a pre všetky $a, b \in K$ sa ich súčet $a + b$ a súčin ab v K rovná ich súčtu $a + b$ resp. súčinu ab v L. Z toho už vyplýva, že pre $0 \neq a \in K$ sa inverzný prvok a^{-1} v K zhoduje s inverzným prvkom a^{-1} v L.

Ak pole L je rozšírením poľa K, tak L možno považovať za vektorový priestor nad poľom K (pozri príklad 1.6.1). Hovoríme, že pole L je *konečným rozšírením poľa* K, ak L je konečnorozmerný vektorový priestor nad K; jeho dimenziu nazývame *stupňom rozšírenia* L nad K a značíme ju

$$[L : K] = \dim_K L.$$

Teda konečné rozšírenie je také, ktoré má konečný stupeň. *Prvok* $c \in L$ sa nazýva *algebraický* nad K, ak c je koreňom nejakého polynómu $f(x) \in K[x]$ stupňa ≥ 1. *Rozšírenie* L poľa K sa nazýva *algebraické*, ak každý prvok $c \in L$ je algebraický nad K.

19.3.1. Tvrdenie. *Každé konečné rozšírenie poľa K je algebraické nad K.*

Dôkaz uvádzame ako jednoduchú ukážku použitia metód lineárnej algebry v teórii polí. Nech L je konečné rozšírenie poľa K stupňa $[L : K] = n$. Zvoľme ľubovoľné $c \in L$. Potom $(n + 1)$-tica $(1, c, c^2, \ldots, c^n)$ prvkov z L nemôže byť lineárne nezávislá nad K, preto existujú $a_0, a_1, \ldots, a_n \in K$, nie všetky rovné 0, také, že

$$a_0 + a_1 c + \ldots + a_n c^n = 0.$$

Ale to znamená, že c je koreňom polynómu $f(x) = a_0 + a_1 x + \ldots + a_n x^n \in K[x]$.

Poznamenajme, že obrátená implikácia neplatí, t. j. algebraické rozšírenie daného poľa môže mať nekonečný stupeň. Z dôkazu navyše vyplýva, že každý prvok konečného rozšírenia L poľa K stupňa $[L : K] = n$ je koreňom polynómu $f(x) \in K[x]$ stupňa $\leq n$.

Pre naše potreby má kľúčový význam nasledujúca skutočnosť.

19.3.2. Tvrdenie. *Nech $f(x) = a_0 + a_1x + \ldots + a_nx^n \in K[x]$, $a_n \neq 0$, je polynóm stupňa $n \geq 1$ nad poľom K. Potom existuje konečné (teda nutne algebraické) rozšírenie L poľa K také, že $f(x)$ má v $L[x]$ rozklad na lineárne faktory*

$$f(x) = a_n(x - \lambda_1)\ldots(x - \lambda_n).$$

Inak povedané, $f(x)$ má práve n koreňov $\lambda_1, \ldots, \lambda_n \in L$, ak každý z nich počítame aj s jeho násobnosťou.

Rozšírenie L možno navyše zvoliť tak, že každý jeho prvok má tvar $g(\lambda_1, \ldots, \lambda_n)$ pre nejaký polynóm $g(x_1, \ldots, x_n)$ v n premenných s koeficientmi z K – pole L, ktoré je týmito podmienkami určené jednoznačne až na izomorfizmus polí, nazývame *rozkladové pole polynómu* $f(x)$. Pre stupeň rozkladového poľa platí $[L : K] \leq n!$. (Prenechávame čitateľovi, aby si sám sformuloval pojem izomorfizmu polí.)

Pole L sa nazýva *algebraicky uzavreté*, alebo tiež *algebraicky úplné*, ak každý polynóm $f(x) \in L[x]$ stupňa aspoň 1 má v L aspoň jeden koreň. Pomerne jednoducho možno nahliadnuť (prípadne dokázať matematickou indukciou podľa stupňa n), že pole L je algebraicky úplné práve vtedy, keď každý polynóm $f(x) \in L[x]$ stupňa aspoň 1 má v $L[x]$ rozklad na lineárne faktory (skúste sami).

Algebraické rozšírenie poľa K, ktoré je navyše samo algebraicky uzavreté, sa nazýva *algebraický uzáver poľa* K. Nie celkom elementárnymi prostriedkami možno dokázať nasledujúcu vetu.

19.3.3. Veta. *Ku každému poľu K existuje jeho algebraický uzáver.*

Podobne ako rozkladové pole polynómu, aj algebraický uzáver daného poľa je určený jednoznačne až na izomorfizmus polí.

Na druhej strane, algebraický uzáver daného poľa nemusí byť jeho konečným rozšírením. V jednom dôležitom prípade však tomu tak je. Pole $\mathbb{C}$ všetkých komplexných čísel má ako rozšírenie poľa $\mathbb{R}$ všetkých reálnych čísel stupeň $[\mathbb{C} : \mathbb{R}] = 2$. Aby sme sa teda utvrdili v tom, že je algebraickým uzáverom poľa $\mathbb{R}$, stačí sa odvolať na nasledujúci klasický a hlboký výsledok, známy ako *základná veta algebry.*

19.3.4. Veta. *Pole $\mathbb{C}$ všetkých komplexných čísel je algebraicky úplné.*

Zhrňme si teraz stručne niektoré dôsledky práve vykonaných úvah pre spektrá matíc. Predovšetkým, spektrum matice $\boldsymbol{A} \in K^{n\times n}$ závisí na príslušnom poli. Presnejšie, takúto maticu možno považovať aj za maticu $\boldsymbol{A} \in L^{n\times n}$, kde L je ľubovoľné rozšírenie poľa K. Ak označíme $\operatorname{Spec}_K \boldsymbol{A}$, $\operatorname{Spec}_L \boldsymbol{A}$ príslušné spektrá (t. j. množiny vlastných hodnôt matice $\boldsymbol{A}$ v poli K resp. L), tak zrejme $\operatorname{Spec}_K \boldsymbol{A} \subseteq \operatorname{Spec}_L \boldsymbol{A}$. Na druhej strane, každé $\lambda \in \operatorname{Spec}_K \boldsymbol{A}$ má zrejme rovnakú algebraickú i geometrickú násobnosť, ktorá nezávisí na príslušnom poli.

Z tvrdenia 19.3.2 priamo vyplýva

19.3.5. Tvrdenie. *Ku každej matici $\boldsymbol{A} \in K^{n\times n}$ existuje konečné rozšírenie L poľa K také, že charakteristický polynóm matice $\boldsymbol{A}$ možno v $L[x]$ rozložiť na súčin lineárnych faktorov*

$$\mathrm{ch}_{\boldsymbol{A}}(x) = (\lambda_1 - x)\dots(\lambda_n - x).$$

Inak povedané, $\boldsymbol{A}$ má n vlastných hodnôt $\lambda_1, \dots, \lambda_n \in L$, vrátane ich algebraickej násobnosti. Potom pre každé rozšírenie M poľa L už platí $\mathrm{Spec}_M \boldsymbol{A} = \mathrm{Spec}_L \boldsymbol{A}$.

Aplikáciou tvrdenia 19.3.5 na Schurovu vetu 19.2.1 okamžite dostávame

19.3.6. Tvrdenie. *Ku každej matici $\boldsymbol{A} \in K^{n\times n}$ existuje konečné rozšírenie L poľa K také, že $\boldsymbol{A}$ je nad L podobná s hornou trojuholníkovou maticou $\boldsymbol{B} \in L^{n\times n}$, ktorej diagonálu tvoria vlastné hodnoty matice $\boldsymbol{A}$ v L, vrátane ich algebraickej násobnosti.*

Napokon si ešte uvedomme, že tvrdenia 19.3.5 a 19.3.6 sú automaticky (t. j. pre $L = K$) splnené pre štvorcové matice nad algebraicky uzavretým poľom K. Špeciálne to platí pre štvorcové matice nad poľom $\mathbb{C}$.

19.4 Komplexifikácia

Pozrime sa teraz na výsledky sformulovné v závere predchádzajúceho paragrafu z hľadiska vedúceho zámeru tejto kapitoly, ktorým je diagonalizácia matice lineárneho operátora $\varphi\colon V \to V$ na konečnorozmernom vektorovom priestore V. Zdá sa, že sme na tejto ceste naozaj kus pokročili a skutočne ukázali, ako možno prekonať prípadnú prekážku spočívajúcu v nedostatočnej algebraickej váhe spektra. Pokiaľ by sme sa zaujímali len o samotné matice, bolo by tomu tak. Nezabúdajme však, že v prvom rade nám ide o lineárne operátory a matice používame len ako viac-menej pomocné objekty slúžiace na ich popis.

Ak L je rozšírenie poľa K, tak každá matica $\boldsymbol{A} \in K^{n\times n} \subseteq L^{n\times n}$ prirodzene určuje lineárny operátor operátor $\varphi\colon K^n \to K^n$ ako aj lineárny operátor $\psi\colon L^n \to L^n$. Hoci φ aj ψ sú oba definované tým istým predpisom $\boldsymbol{x} \mapsto \boldsymbol{A}\cdot\boldsymbol{x}$, (okrem triviálneho prípadu, keď $L = K$) sú to *rôzne* operátory, lebo operujú na rôznych vektorových priestoroch, navyše nad rôznymi poľami. Napriek tomu akosi podvedome cítime, že tým „správnym“ analógom vektorového priestoru K^n nad poľom K je práve vektorový priestor L^n nad poľom L.

Potom operátoru $\varphi\colon K^n \to K^n$, kde $\varphi(\boldsymbol{x}) = \boldsymbol{A}\cdot\boldsymbol{x}$ pre $\boldsymbol{x} \in K^n$, „prirodzene“ zodpovedá operátor $\psi\colon L^n \to L^n$, kde $\psi(\boldsymbol{x}) = \boldsymbol{A}\cdot\boldsymbol{x}$ pre $\boldsymbol{x} \in L^n$. No v situácii, keď $\varphi\colon V \to V$ je lineárny operátor na „abstraktnom“ vektorovom priestore V nad poľom K, otázka, akým vektorovým priestorom W nad poľom L máme nahradiť V a aký lineárny operátor $\psi\colon W \to W$ by mal zodpovedať operátoru φ, už nie je zďaleka taká jasná.

Príslušnú konštrukciu možno elegantne popísať v jazyku tenzorových súčinov (pozri paragraf 32.3). Aj elementárnymi prostriedkami, ktoré máme v tejto chvíli k dispozícii však dokážeme zvládnuť pre nás najdôležitejší špeciálny prípad – totiž prechod od $\mathbb{R}$ k $\mathbb{C}$. Budeme si počínať v podstate rovnako ako pri dôverne známej konštrukcii poľa $\mathbb{C}$ z poľa $\mathbb{R}$.

Komplexifikáciou vektorového priestoru V nad poľom $\mathbb{R}$ nazveme vektorový priestor $V^{\mathbb{C}}$ nad poľom $\mathbb{C}$, ktorého prvkami sú všetky formálne lineárne kombinácie tvaru $\boldsymbol{z} = \boldsymbol{x} + \mathrm{i}\boldsymbol{y}$, kde $\boldsymbol{x}, \boldsymbol{y} \in V$ a $\mathrm{i} \in \mathbb{C}$ je imaginárna jednotka. Pre $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v} \in V$ kladieme

$$\boldsymbol{x} + \mathrm{i}\boldsymbol{y} = \boldsymbol{u} + \mathrm{i}\boldsymbol{v} \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{u} \ \&\ \boldsymbol{y} = \boldsymbol{v}.$$

Píšeme tiež $\boldsymbol{x} = \operatorname{Re}\boldsymbol{z}$, $\boldsymbol{y} = \operatorname{Im}\boldsymbol{z}$ a vektory $\boldsymbol{x}, \boldsymbol{y} \in V$ nazývame *reálnou* resp. *imaginárnou časťou vektora* $\boldsymbol{z} \in V^{\mathbb{C}}$. Sčítanie takýchto vektorov definujeme „po zložkách“, čiže

$$(\boldsymbol{x} + \mathrm{i}\boldsymbol{y}) + (\boldsymbol{u} + \mathrm{i}\boldsymbol{v}) = (\boldsymbol{x} + \boldsymbol{u}) + \mathrm{i}(\boldsymbol{y} + \boldsymbol{v}).$$

Konečne násobenie skalárom $a + \mathrm{i}b \in \mathbb{C}$, kde $a, b \in \mathbb{R}$, definujeme vzťahom

$$(a + \mathrm{i}b)(\boldsymbol{x} + \mathrm{i}\boldsymbol{y}) = (a\boldsymbol{x} - b\boldsymbol{y}) + \mathrm{i}(a\boldsymbol{y} + b\boldsymbol{x}).$$

Čitateľ si určite aj sám ľahko overí, že komplexifikácia $V^{\mathbb{C}}$ reálneho vektorového priestoru V s takto definovanou rovnosťou a operáciami naozaj tvorí vektorový priestor nad poľom $\mathbb{C}$. Ešte dodajme, že – rovnako ako v samotnom poli $\mathbb{C}$ – *komplexne združeným vektorom* k vektoru $\boldsymbol{z} = \boldsymbol{x} + \mathrm{i}\boldsymbol{y} \in V^{\mathbb{C}}$ nazývame vektor $\overline{\boldsymbol{z}} = \boldsymbol{x} - \mathrm{i}\boldsymbol{y}$.
Poznámka Komplexifikácia $V^{\mathbb{C}}$ reálneho vektorového priestoru V sa tiež zvykne značiť $\mathbb{C} \otimes_{\mathbb{R}} V$, prípadne len $\mathbb{C} \otimes V$, t. j. ako tenzorový súčin $\mathbb{C}$ a V nad poľom $\mathbb{R}$. V tejto chvíli však ide naozaj len o „nevinné“ označenie a práve popísaná konštrukcia ani jej ďalšie využitie si nijakú znalosť tenzorových súčinov nevyžadujú.

19.4.1. Tvrdenie. *Nech V je reálny vektorový priestor. Potom*

$$\dim_{\mathbb{C}} V^{\mathbb{C}} = \dim_{\mathbb{R}} V.$$

Dôkaz. Nech $\dim_{\mathbb{R}} V = n$ a $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ je nejaká báza priestoru V nad $\mathbb{R}$. Ponechávame na čitateľa, aby si sám overil, že tie isté vektory $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ tvoria aj bázu $V^{\mathbb{C}}$ nad $\mathbb{C}$.[1] Taktiež ľahko nahliadneme, že $\dim_{\mathbb{R}} V = \infty$ má za následok $\dim_{\mathbb{C}} V^{\mathbb{C}} = \infty$.

Komplexifikácia vektorového priestoru nad $\mathbb{R}$ je v istom zmysle opačnou konštrukciou k zreálneniu vektorového priestoru nad $\mathbb{C}$ (pozri paragraf 17.1). Vzťahu oboch konštrukcií sa budeme podrobnejšie venovať v cvičeniach.

Komplexifikáciou lineárneho zobrazenia $\varphi\colon V \to U$ medzi reálnymi vektorovými priestormi V, U nazývame lineárne zobrazenie $\varphi^{\mathbb{C}}\colon V^{\mathbb{C}} \to U^{\mathbb{C}}$ dané predpisom

$$\varphi^{\mathbb{C}}(\boldsymbol{x} + \mathrm{i}\boldsymbol{y}) = \varphi(\boldsymbol{x}) + \mathrm{i}\varphi(\boldsymbol{y}),$$

pre $\boldsymbol{x}, \boldsymbol{y} \in V$. Špeciálne $\varphi^{\mathbb{C}}(\boldsymbol{x}) = \varphi(\boldsymbol{x})$ pre $\boldsymbol{x} \in V$. Opäť sa ľahko overí, že $\varphi^{\mathbb{C}}$ je naozaj lineárne zobrazenie medzi vektorovými priestormi nad $\mathbb{C}$. Taktiež nasledujúce tvrdenie je zrejmé z práve vyslovenej definície a predchádzajúceho tvrdenia.

19.4.2. Tvrdenie. *Nech U, V sú reálne vektorové priestory s bázami $\boldsymbol{\alpha}$ resp. $\boldsymbol{\beta}$ a $\varphi\colon V \to U$ je lineárne zobrazenie. Potom*

$$(\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \left(\varphi^{\mathbb{C}}\right)_{\boldsymbol{\alpha},\boldsymbol{\beta}}.$$

Inak povedané φ a jeho komplexifikácia $\varphi^{\mathbb{C}}\colon V^{\mathbb{C}} \to U^{\mathbb{C}}$ majú tú istú maticu vzhľadom na bázy $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$.

Ak ľubovoľný vektor $\boldsymbol{z} = (x_1 + \mathrm{i}y_1, \dots, x_n + \mathrm{i}y_n)^{\mathsf{T}} \in \mathbb{C}^n$ rozložíme na reálnu a imaginárnu časť

$$\boldsymbol{z} = \boldsymbol{x} + \mathrm{i}\boldsymbol{y} = \operatorname{Re}\boldsymbol{z} + \mathrm{i}\operatorname{Im}\boldsymbol{z},$$

kde $\boldsymbol{x} = \operatorname{Re}\boldsymbol{z} = (x_1, \dots, x_n)^{\mathsf{T}}$, $\boldsymbol{y} = \operatorname{Im}\boldsymbol{z} = (y_1, \dots, y_n)^{\mathsf{T}}$ sú reálne vektory tvorené reálnymi resp. imaginárnymi časťami kanonických súradníc vektora $\boldsymbol{z}$, prirodzene tým stotožníme komplexifikáciu $(\mathbb{R}^n)^{\mathbb{C}}$ priestoru $\mathbb{R}^n$ s priestorom $\mathbb{C}^n$. Za takýchto okolností je komplexifikáciou lineárneho zobrazenia $\varphi\colon \mathbb{R}^n \to \mathbb{R}^m$ s maticou $\boldsymbol{A}$ (vzhľadom na kanonické bázy) zobrazenie $\varphi^{\mathbb{C}}\colon \mathbb{C}^n \to \mathbb{C}^m$, ktoré má (v kanonických bázach) rovnakú maticu $\boldsymbol{A}$. Vidíme teda, že naša „abstraktná“ konštrukcia komplexifikácie sa plne zhoduje s „konkrétnymi komplexifikáciami“ $\mathbb{C}^n$ priestorov $\mathbb{R}^n$ a lineárnych zobrazení medzi takýmito priestormi.

Z hľadiska vedúceho zámeru tejto kapitoly spočíva hlavný význam komplexifikácie v nasledujúcom tvrdení.

[1] Striktne podľa definície by sme mali hovoriť o vektoroch $\boldsymbol{u}_1 + \mathrm{i}\boldsymbol{0}, \dots, \boldsymbol{u}_n + \mathrm{i}\boldsymbol{0}$. Prirodzene však píšeme $\boldsymbol{x} + \mathrm{i}\boldsymbol{0} = \boldsymbol{x}$, $\boldsymbol{0} + \mathrm{i}\boldsymbol{y} = \mathrm{i}\boldsymbol{y}$, a V stotožňujeme s množinou $\{\boldsymbol{x} + \mathrm{i}\boldsymbol{0};\, \boldsymbol{x} \in V\} \subseteq V^{\mathbb{C}}$.

19.4.3. Veta. *Nech V je vektorový priestor konečnej dimenzie n nad poľom $\mathbb{R}$ a $\varphi\colon V \to V$ je lineárny operátor. Potom lineárny operátor $\varphi^{\mathbb{C}}\colon V^{\mathbb{C}} \to V^{\mathbb{C}}$ má, vrátane algebraickej násobnosti, n vlastných čísel $\lambda_1, \dots, \lambda_n \in \mathbb{C}$.*

Samozrejme, reálne spomedzi vlastných čísel $\lambda_1, \dots, \lambda_n$ operátora $\varphi^{\mathbb{C}}$ sú priamo vlastnými číslami pôvodného operátora φ. S istou dávkou zjednodušenia však nazývame aj hodnoty λ_j s nenulovou imaginárnou časťou *komplexnými vlastnými číslami* (reálneho) operátora φ.

Inak povedané, prekážku diagonalizácie matice lineárneho operátora na konečnorozmernom vektorovom priestore nad $\mathbb{R}$, spočívajúcu v nedostatočnej algebraickej váhe jeho spektra, vieme prekonať komplexifikáciou príslušného priestoru a operátora. Poznamenajme, že nahradením poľa $\mathbb{C}$ rozkladovým poľom príslušného charakteristického polynómu a istým zovšeobecnením metódy komplexifikácie možno rovnaký cieľ dosiahnuť aj pre lineárne operátory na konečnorozmerných vektorových priestoroch nad ľubovoľným poľom. Tieto otázky však už presahujú rámec nášho kurzu.

19.5 Geometrický význam komplexných vlastných čísel

Geometrický význam reálnych vlastných čísel lineárnej transformácie φ vektorového priestoru V nad poľom $\mathbb{R}$ je zrejmý: ak $\lambda \in \mathbb{R}$ je vlastné číslo, tak príslušný vlastný vektor $\boldsymbol{v} \in V$ (i každý vektor vlastnej priamky $[\boldsymbol{v}]$) sa zobrazením φ zobrazí na svoj λ-násobok. Pri našej „podvedomej" geometrickej interpretácii priestoru V (keď doňho mimovoľne vnášame euklidovskú štruktúru), sa dĺžka vektora $\boldsymbol{v}$ zmení na jej $|\lambda|$-násobok; ak $\lambda > 0$, tak orientácia vektora $\boldsymbol{v}$ sa zachová, ak $\lambda < 0$, zmení sa na opačnú. Tentokrát sa však nemusíme odvolávať na nijaké „podvedomé" geometrické súvislosti. Z podmienky homogenity (pozri paragraf 13.3) totiž vyplýva, že rovnosť $\|\varphi(\boldsymbol{v})\| = |\lambda|\,\|\boldsymbol{v}\|$ platí pre *ľubovoľnú* normu na V.

V príklade 18.4.2 sme však videli, že aj lineárna transformácia reálneho vektorového priestoru môže mať komplexné vlastné čísla. Cieľom tohto paragrafu je ukázať, že uvedený príklad je svojim spôsobom typický, presnejšie, za výskytom komplexných vlastných čísel reálnej lineárnej transformácie sa vždy skrýva nejaké otočenie (pri spomínanej mimovoľnej euklidovskej interpretácii). Tým by sa mal z úlohy komplexných vlastných čísel definitívne vytratiť akýkoľvek mystický ráz.

19.5.1. Tvrdenie. *Nech $\lambda \in \mathbb{C}$ je vlastné číslo matice $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ a $\boldsymbol{w}$ je k nemu prislúchajúci vlastný vektor. Potom aj $\overline{\lambda}$ je vlastné číslo matice $\boldsymbol{A}$ a prislúcha k nemu vlastný vektor $\overline{\boldsymbol{w}}$. Navyše vlastné čísla λ a $\overline{\lambda}$ majú rovnakú algebraickú i geometrickú násobnosť.*

Dôkaz. Najprv dokážeme, že komplexné číslo λ je m-násobným koreňom polynómu $f(x) = a_0 + a_1x + \ldots + a_nx^n \in \mathbb{R}[x]$ práve vtedy, keď $\overline{\lambda}$ je jeho m-násobným koreňom. Stačí ukázať, že z deliteľnosti $f(x)$ mocninou $(x - \lambda)^m$ vyplýva jeho deliteľnosť mocninou $(x - \overline{\lambda})^m$. Nech teda $f(x) = (x - \lambda)^m p(x)$ pre nejaký polynóm $p(x) = c_0 + c_1x + \ldots + c_kx^k \in \mathbb{C}[x]$. Potom zrejme

$$(x - \overline{\lambda})^m(\overline{c}_0 + \overline{c}_1x + \ldots + \overline{c}_kx^k) = \overline{a}_0 + \overline{a}_1x + \ldots + \overline{a}_nx^n = f(x),$$

lebo pre koeficienty polynómu $f(x) \in \mathbb{R}[x]$ platí $\overline{a}_j = a_j$. Ak teda λ je algebraicky m-násobné vlastné číslo matice $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, tak aj $\overline{\lambda}$ je jej vlastné číslo s tou istou algebraickou násobnosťou.

Nech $\boldsymbol{w} \in \mathbb{C}^n$ je vlastný vektor prislúchajúci k λ, čiže $\boldsymbol{A} \cdot \boldsymbol{w} = \lambda\boldsymbol{w}$. Potom tiež $\overline{\boldsymbol{A}} \cdot \overline{\boldsymbol{w}} = \overline{\lambda}\overline{\boldsymbol{w}}$. Keďže $\overline{\boldsymbol{A}} = \boldsymbol{A}$ je reálna, vyplýva z toho $\boldsymbol{A} \cdot \overline{\boldsymbol{w}} = \overline{\lambda}\overline{\boldsymbol{w}}$, t. j. $\overline{\boldsymbol{w}}$ je vlastný vektor matice $\boldsymbol{A}$ prislúchajúci k $\overline{\lambda}$. Teda zobrazenie $\boldsymbol{w} \mapsto \overline{\boldsymbol{w}}$ je semilineárny izomorfizmus (pozri cvičenia ku kapitole 17) vlastných podpriestorov $\mathcal{R}(\boldsymbol{A} - \lambda\boldsymbol{I}) \to \mathcal{R}(\boldsymbol{A} - \overline{\lambda}\boldsymbol{I})$, preto tieto priestory majú rovnakú dimenziu a vlastné čísla λ, $\overline{\lambda}$ majú rovnakú geometrickú násobnosť.

Teda komplexné vlastné čísla s nenulovou imaginárnou časťou reálnych matíc resp. lineárnych transformácií reálnych vektorových priestoroch sa vždy vyskytujú v komplexne združených pároch λ, $\overline{\lambda}$, rovnako ako k nim príslušné vlastné vektory. Ukážeme si, že za takýmto výskytom sa vždy skrýva dvojrozmerný invariantný podpriestor, na ktorom príslušná matica operuje ako rotácia o uhol α zložená s rovnoľahlosťou s koeficientom $r > 0$, kde $\lambda = r(\cos\alpha + \mathrm{i}\sin\alpha)$ je goniometrický tvar čísla λ.

Nech teda $\varphi\colon V \to V$ je lineárny operátor na reálnom vektorovom priestore V, $\lambda = a + \mathrm{i}b$, $\overline{\lambda} = a - \mathrm{i}b$, kde $a, b \in \mathbb{R}$, $b \neq 0$, sú komplexne združené vlastné čísla jeho komplexifikácie $\varphi^{\mathbb{C}}\colon V^{\mathbb{C}} \to V^{\mathbb{C}}$ a $\boldsymbol{w} \in V^{\mathbb{C}}$ resp. $\overline{\boldsymbol{w}} \in V^{\mathbb{C}}$ sú k nim prislúchajúce vlastné vektory. Potom zúženie operátora $\varphi^{\mathbb{C}}$ na jeho dvojrozmerný invariantný podpriestor $[\boldsymbol{w}, \overline{\boldsymbol{w}}] \subseteq V^{\mathbb{C}}$ má v báze $(\boldsymbol{w}, \overline{\boldsymbol{w}})$ maticu $\operatorname{diag}(\lambda, \overline{\lambda})$. Označme

$$\boldsymbol{u} = \operatorname{Re}\boldsymbol{w} = \operatorname{Re}\overline{\boldsymbol{w}} = \frac{1}{2}(\boldsymbol{w} + \overline{\boldsymbol{w}}), \qquad \boldsymbol{v} = -\operatorname{Im}\boldsymbol{w} = \operatorname{Im}\overline{\boldsymbol{w}} = \frac{1}{2\mathrm{i}}(\overline{\boldsymbol{w}} - \boldsymbol{w})$$

vektory z pôvodného reálneho priestoru V. Potom $\boldsymbol{w} = \boldsymbol{u} - \mathrm{i}\boldsymbol{v}$, $\overline{\boldsymbol{w}} = \boldsymbol{u} + \mathrm{i}\boldsymbol{v}$. Ľahko nahliadneme, že $\boldsymbol{u}$, $\boldsymbol{v}$ sú lineárne nezávislé (nad $\mathbb{R}$), teda tvoria bázu dvojrozmerného podpriestoru $[\boldsymbol{u}, \boldsymbol{v}] \subseteq V$. Počítajme

$$\varphi(\boldsymbol{u}) = \varphi^{\mathbb{C}}(\boldsymbol{u}) = \frac{1}{2}\big(\varphi^{\mathbb{C}}(\boldsymbol{w}) + \varphi^{\mathbb{C}}(\overline{\boldsymbol{w}})\big) = \frac{1}{2}(\lambda\boldsymbol{w} + \overline{\lambda}\overline{\boldsymbol{w}}) = \operatorname{Re}(\lambda\boldsymbol{w}) = a\boldsymbol{u} + b\boldsymbol{v},$$
$$\varphi(\boldsymbol{v}) = \varphi^{\mathbb{C}}(\boldsymbol{v}) = \frac{1}{2\mathrm{i}}\big(\varphi^{\mathbb{C}}(\overline{\boldsymbol{w}}) - \varphi^{\mathbb{C}}(\boldsymbol{w})\big) = \frac{1}{2\mathrm{i}}(\overline{\lambda}\overline{\boldsymbol{w}} - \lambda\boldsymbol{w}) = \operatorname{Im}(\overline{\lambda}\overline{\boldsymbol{w}}) = -b\boldsymbol{u} + a\boldsymbol{v}.$$

Vidíme, že $[\boldsymbol{u},\boldsymbol{v}] \subseteq V$ je dvojrozmerný invariantný podpriestor operátora φ a zúženie φ na $[\boldsymbol{u},\boldsymbol{v}]$ má v báze $(\boldsymbol{u},\boldsymbol{v})$ reálnu maticu, ktorú možno pomocou goniometrického vyjadrenia vlastných čísel $\lambda = r(\cos\alpha + \mathrm{i}\sin\alpha) = r\,\mathrm{e}^{\mathrm{i}\alpha}$, $\overline{\lambda} = r(\cos\alpha - \mathrm{i}\sin\alpha) = r\,\mathrm{e}^{-\mathrm{i}\alpha}$ (t. j. $a = r\cos\alpha$, $b = r\sin\alpha$) zapísať v tvare

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} r\cos\alpha & -r\sin\alpha \\ r\sin\alpha & r\cos\alpha \end{pmatrix} = r\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = r\boldsymbol{R}_\alpha.$$

Tým sme vlastne dokázali nasledujúcu vetu.

19.5.2. Veta. *Nech φ je lineárna transformácia reálneho vektorového priestoru V, $\lambda = a + \mathrm{i}b = r(\cos\alpha + \mathrm{i}\sin\alpha)$ je jej komplexná vlastná hodnota s nenulovou imaginárnou časťou, ku ktorej vo $V^{\mathbb{C}}$ prislúcha vlastný vektor $\boldsymbol{u} - \mathrm{i}\boldsymbol{v}$, kde $\boldsymbol{u},\boldsymbol{v} \in V$. Potom $(\boldsymbol{u},\boldsymbol{v})$ je bázou invariantného podpriestoru $[\boldsymbol{u},\boldsymbol{v}] \subseteq V$ a zúženie φ na $[\boldsymbol{u},\boldsymbol{v}]$ má v tejto báze reálnu maticu $\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right) = r\boldsymbol{R}_\alpha$.*

V úvode ku kapitole 13 sme zdôrazňovali, že hovoriť o kvantitatívnych geometrických veličinách ako dĺžka či uhol v abstraktných vektorových priestoroch, hoc aj nad poľom $\mathbb{R}$, nedáva dobrý zmysel. Až prítomnosť skalárneho súčinu v takomto priestore nám to umožňuje. Teda, prísne vzaté, zatiaľ stále nemáme právo tvrdiť, že φ na invariantnom podpriestore $[\boldsymbol{u},\boldsymbol{v}]$ pôsobí ako kompozícia otočenia o uhol α a rovnoľahlosti s koeficientom r a stredom v počiatku. Na $[\boldsymbol{u},\boldsymbol{v}]$ si však možno (navyše jednoznačne) zvoliť skalárny súčin tak, aby vektory $\boldsymbol{u}$, $\boldsymbol{v}$ vzhľadom naň tvorili ortonormálnu bázu. Potom lineárna transformácia $\varphi \restriction [\boldsymbol{u},\boldsymbol{v}]$ euklidovského priestoru $[\boldsymbol{u},\boldsymbol{v}]$ naozaj zodpovedá kompozícii rotácie $\boldsymbol{R}_\alpha$ a rovnoľahlosti $r\boldsymbol{I}_2$. (Pozri cvičenia 19.11 a 19.12.)

Ak si uvedomíme, že každý polynóm $f(x) \in \mathbb{R}[x]$ nepárneho stupňa má aspoň jeden reálny koreň, môžeme zaznamenať ešte jeden dôsledok vety 19.5.2.

19.5.3. Dôsledok. *Nech φ je lineárny operátor na konečnorozmernom vektorovom priestore V nad poľom $\mathbb{R}$. Potom φ má aspoň jeden invariantný podpriestor dimenzie 1 alebo 2. Ak $\dim V$ je nepárne číslo, tak φ má aspoň jedno reálne vlastné číslo a k nemu prislúchajúci vlastný vektor generuje invariantný podpriestor dimenzie 1.*

Cvičenia

19.1. Nech $f(x), g(x)$ sú polynómy nad poľom K, pričom $g(x)$ má stupeň aspoň 1. Polynóm $r(x)$ sa nazýva *zvyškom po delení* polynómu $f(x)$ polynómom, ak $g(x)$ delí polynóm $f(x) - r(x)$ a $r(x)$ je nižšieho stupňa ako $g(x)$. Nech ďalej $\lambda \in K$. Dokážte nasledujúce tvrdenia:

(a) Zvyšok po delení polynómu $f(x)$ lineárnym polynómom $x - \lambda$ je $f(\lambda)$.

(b) λ je koreňom polynómu $f(x)$, t. j. $f(\lambda) = 0$, práve vtedy, keď $x - \lambda$ delí $f(x)$.

19.2. Nech $\boldsymbol{A} \in K^{m\times m}$, $\boldsymbol{B} \in K^{n\times n}$ sú štvorcové matice. Označme $\boldsymbol{C} = \operatorname{diag}(\boldsymbol{A}, \boldsymbol{B})$. Dokážte nasledujúce tvrdenia:

(a) Ak $\boldsymbol{u} \in K^m$ resp. $\boldsymbol{v} \in K^n$ je vlastný vektor matice $\boldsymbol{A}$ resp. $\boldsymbol{B}$ prislúchajúci k jej vlastnej hodnote λ, tak $\binom{\boldsymbol{u}}{\boldsymbol{0}_n}$ resp. $\binom{\boldsymbol{0}_m}{\boldsymbol{v}}$ je vlastný vektor matice $\boldsymbol{C}$ prislúchajúci k tej istej jej vlastnej hodnote.

(b) $\operatorname{Spec} \boldsymbol{C} = \operatorname{Spec} \boldsymbol{A} \cup \operatorname{Spec} \boldsymbol{B}$.

(c) Algebraická aj geometrická násobnosť skalára λ ako vlastnej hodnoty matice $\boldsymbol{C}$ je súčtom príslušných násobností λ ako vlastných hodnôt matíc $\boldsymbol{A}$, $\boldsymbol{B}$.

Modfikujte (ak treba) tvrdenia (a) – (c) tak, aby zostali v platnosti aj pre maticu tvaru $\boldsymbol{C}' = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{M} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix}$, kde $\boldsymbol{M} \in K^{m\times n}$ je nenulová matica.

19.3. Nech $\boldsymbol{A} = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 0 & 1 \\ -b^2 & 2b \end{pmatrix}$, kde $a, b \in \mathbb{R}$.

(a) Nájdite spektrum a vlastné vektory matice $\boldsymbol{C} = \operatorname{diag}(\boldsymbol{A}, \boldsymbol{B})$. Je matica $\boldsymbol{C}$ podobná s nejakou diagonálnou maticou?

(b) Určte algebraickú a geometrickú násobnosť všetkých vlastných hodnôt matice $\boldsymbol{C}$. Urobte diskusiu vzhľadom na a, b.

(c) (Zároveň nápoveda k (b).) Dokážte, že pre $a = b$ sú matice $\boldsymbol{A}$, $\boldsymbol{B}$ podobné a nájdite aspoň jednu maticu $\boldsymbol{P} \in \mathbb{R}^{2\times 2}$, pre ktorú platí $\boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P} = \boldsymbol{B}$. Dokážete nájsť všetky také matice $\boldsymbol{P}$?

19.4. Dopracujte príklad 19.2.4 tak, aby ste našli hornú trojuholníkovú maticu podobnú s maticou $\boldsymbol{A}$ prostredníctvom ortogonálnej matice prechodu.

19.5. Nech φ je lineárny operátor na konečnorozmernom vektorovom priestore V so spektrom plnej algebraickej váhy $n = \dim V$. Zo Schurovej vety 19.2.1 odvoďte existenciu postupnosti invariantných podpriestorov $\{\boldsymbol{0}\} = V_0 \subseteq V_1 \subseteq \ldots \subseteq V_{n-1} \subseteq V_n = V$, v ktorej $\dim V_k = k$ pre $0 \le k \le n$.

19.6. Ako určíte spektrum (hornej či dolnej) trojuholníkovej matice vrátane algebraickej násobnosti každej jej vlastnej hodnoty bez počítania jej charakteristického polynómu? Dajú sa podobne priamočiaro určiť aj geometrické násobnosti jej vlastných hodnôt?

19.7. (a) Zadefinujte pojem homomorfizmu polí ako zobrazenia, ktoré zachováva sčítanie a násobenie.

(b) Nech $h\colon K \to L$ je homomorfizmus polí. Potom $h(1) = 1$ a $h(a^{-1}) = h(a)^{-1}$ pre každé $a \in K$. Dokážte.

(c) Zadefinujte pojem izomorfizmu polí a dokážte, že vzťah izomorfnosti je ekvivalenciou na obore všetkých polí.

(d) Dokážte, že každý homomorfizmus polí $h\colon K \to L$ je injektívne zobrazenie, teda je to izomorfizmus poľa K na podpole poľa L.

(e) Dokážte, že každé pole charakteristiky ∞ obsahuje podpole izomorfné s poľom $\mathbb{Q}$ všetkých racionálnych čísel.

(f) Dokážte, že každé pole prvočíselnej charakteristiky p obsahuje podpole izomorfné s poľom $\mathbb{Z}_p$. Odvoďte z toho, že počet prvkov každého konečného poľa L je mocninou prvočísla, presnejšie, $\# L = p^k$, kde $p = \operatorname{char} L$ a $k = [L : \mathbb{Z}_p]$

19.8. Nech polia L_1, L_2 sú algebraicky uzavreté algebraické rozšírenia poľa K. Potom $L_1 \cong L_2$. Dokážte.

19.9. Nech p je prvočíslo. Dokážte, že žiadne konečné rozšírenie poľa $\mathbb{Z}_p$ nie je algebraicky uzavreté. (*Návod*: Nech L je q-prvkové rozšírenie poľa $\mathbb{Z}_p$. Porovnajte počet všetkých normovaných kvadratických polynómov $x^2 + ax + b$ nad L s počtom všetkých kvadratických polynómov $(x - \alpha)(x - \beta)$ rozložiteľných nad L na lineárne faktory.)

19.10. (a) Nech V je vektorový priestor nad poľom $\mathbb{R}$ s bázou tvorenou vektormi $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$. Potom vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ tvoria zároveň bázu jeho komplexifikácie $V^{\mathbb{C}}$. Dokážte. Odvoďte z toho vzťah $\dim_{\mathbb{R}} V = \dim_{\mathbb{C}} V^{\mathbb{C}}$. Porovnajte s cvičením 17.1.

(b) Zovšeobecnite tvrdenie z cvičenia 17.1 do nasledujúcej podoby: Nech L je konečné rozšírenie poľa K stupňa $k = [L : K]$ a W je konečnorozmerný vektorový priestor nad poľom L s bázou tvorenou vektormi $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n \in W$. Nech ďalej skaláry $u_1 = 1, u_2 \dots, u_k \in L$ tvoria bázu L ako vektorového priestoru nad poľom K. Potom systém vektorov $u_i \boldsymbol{w}_j \in W$, kde $i \le k$, $j \le n$, tvorí bázu W chápaného ako vektorový priestor nad poľom K. Odvoďte z toho vzťah $\dim_K W = kn = [L : K] \dim_L W$.

19.11. Nech V je reálny vektorový priestor so skalárnym súčinom. Pre $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v} \in V$ položme $\langle \boldsymbol{x} + \mathrm{i}\boldsymbol{y}, \boldsymbol{u} + \mathrm{i}\boldsymbol{v} \rangle = (\langle \boldsymbol{x}, \boldsymbol{u} \rangle + \langle \boldsymbol{y}, \boldsymbol{v} \rangle) + \mathrm{i}(\langle \boldsymbol{y}, \boldsymbol{u} \rangle - \langle \boldsymbol{x}, \boldsymbol{v} \rangle)$. Dokážte nasledujúce tvrdenia:

(a) Uvedenou formulou je definovaný komplexný skalárny súčin na komplexifikácii $V^{\mathbb{C}}$ priestoru V.

(b) Uvedená formula predstavuje jedinú možnosť, ako definovať na $V^{\mathbb{C}}$ komplexný skalárny súčin, ktorý na V splýva s pôvodným skalárnym súčinom.

(c) Pre každý vektor $\boldsymbol{w} = \boldsymbol{u} + \mathrm{i}\boldsymbol{v} \in V^{\mathbb{C}}$, kde $\boldsymbol{u}, \boldsymbol{v} \in V$, platí $\|\boldsymbol{w}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 = \|\overline{\boldsymbol{w}}\|^2$.

(d) Vektory $\boldsymbol{w}_1, \overline{\boldsymbol{w}}_1, \dots, \boldsymbol{w}_k, \overline{\boldsymbol{w}}_k$, kde $\boldsymbol{w}_j = \boldsymbol{u}_j + \mathrm{i}\boldsymbol{v}_j$ pre $\boldsymbol{u}_j, \boldsymbol{v}_j \in V$, sú ortonormálne v priestore $V^{\mathbb{C}}$ práve vtedy, keď vektory $\sqrt{2}\,\boldsymbol{u}_1, \sqrt{2}\,\boldsymbol{v}_1, \dots, \sqrt{2}\,\boldsymbol{u}_k, \sqrt{2}\,\boldsymbol{v}_k$ sú ortonormálne vo V.

19.12. Nech V je vektorový priestor nad poľom $\mathbb{R}$ a $\varphi\colon V \to V$ je lineárna transformácia s komplexnou vlastnou hodnotou $\lambda = a + \mathrm{i}b$, kde $a, b \in \mathbb{R}$, $b \neq 0$, ktorej prislúcha vlastný vektor $\boldsymbol{w} \in V^{\mathbb{C}}$. Dokážte nasledujúce tvrdenia:

(a) Vektory $\boldsymbol{u} = \mathrm{Re}\,\boldsymbol{w}$, $\boldsymbol{v} = \mathrm{Im}\,\boldsymbol{w}$ sú lineárne nezávislé v pôvodnom priestore V.

(b) Nech navyše na priestore V je definovaný skalárny súčin, vzhľadom na ktorý sú vektory $\boldsymbol{u}, \boldsymbol{v} \in V$ ortonormálne. Rozšírme tento skalárny súčin do komplexného skalárneho súčinu na $V^{\mathbb{C}}$ ako v cvičení 19.11. Potom vektory $\boldsymbol{w}, \overline{\boldsymbol{w}} \in V^{\mathbb{C}}$ sú ortogonálne s rovnakou normou $\sqrt{2}$.

(c) Akú maticu má operátor φ zúžený na svoj invariantný podpriestor $[\boldsymbol{u}, \boldsymbol{v}]$ v každej z báz $(\boldsymbol{u}, \boldsymbol{v})$, $(\boldsymbol{u}, -\boldsymbol{v})$, $(-\boldsymbol{u}, \boldsymbol{v})$, $(-\boldsymbol{u}, -\boldsymbol{v})$, $(\boldsymbol{v}, \boldsymbol{u})$, $(-\boldsymbol{v}, \boldsymbol{u})$, $(\boldsymbol{v}, -\boldsymbol{u})$, $(-\boldsymbol{v}, -\boldsymbol{u})$? Zapíšte tieto matice s využitím goniometrického tvaru vlastného čísla λ. Sú všetky rôzne?

19.13. Nájdite vlastné čísla a k nim prislúchajúce vlastné vektory nasledujúcich reálnych matíc. V prípade komplexných vlastných čísel nájdite dvojrozmerné invariantné podpriestory zodpovedajúce komplexne združeným dvojiciam vlastných hodnôt a ich bázy zodpovedajúce k nim prislúchajúcim komplexne združeným dvojiciam vlastných vektorov:

(a) $\boldsymbol{A} = \begin{pmatrix} -1 & 2 \\ -2 & 1 \end{pmatrix}$; (b) $\boldsymbol{B} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$;

(c) $\boldsymbol{C} = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix}$; (d) $\boldsymbol{D} = \begin{pmatrix} 3/5 & 0 & 4/5 \\ 0 & 1 & 0 \\ -4/5 & 0 & 3/5 \end{pmatrix}$;

(e) $\boldsymbol{E} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$; (f) $\boldsymbol{F} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$;

19.14. Označme $\boldsymbol{C} = \boldsymbol{C}(a_0, a_1, \dots, a_{n-1}) = \begin{pmatrix} a_0 & a_1 & \dots & a_{n-1} \\ a_{n-1} & a_0 & \dots & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \dots & a_0 \end{pmatrix}$ maticu, ktorá vznikne postupnými cyklickými permutáciami usporiadanej n-tice komplexných čísel $\boldsymbol{a} = (a_0, a_1, \dots, a_{n-1})$ – nazývame ju *cirkulantnou maticou* prvkov $a_0, a_1, \dots, a_{n-1}$. Dokážte nasledujúce tvrdenia:

(a) Spektrum cirkulantnej matice $\boldsymbol{C}$ je tvorené komplexnými číslami $g(1)$, $g(\omega)$,..., $g(\omega^{n-1})$, kde $\omega = \mathrm{e}^{2\pi \mathrm{i}/n} = \cos(2\pi/n) + \mathrm{i}\sin(2\pi/n)$ je primitívna n-tá odmocnina z jednotky a g označuje polynóm $g(x) = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1}$, pričom vlastný vektor prislúchajúci k vlastnej hodnote $\lambda_k = g(\omega^k)$ má tvar $\boldsymbol{f}_k = \left(1, \omega^k, \dots, \omega^{(n-1)k}\right)^{\mathsf{T}}$ pre $0 \le k \le n-1$.

(b) Potom prvky spektra $g(\omega^k) = \langle \boldsymbol{a}, \overline{\boldsymbol{f}}_k \rangle = \langle \boldsymbol{a}, \boldsymbol{f}_{n-k} \rangle = \widehat{a}_{n-k}$ sú vlastne (v cyklicky opačnom slede napísané) Fourierove koeficienty vektora $\boldsymbol{a} \in \mathbb{C}^n$ a $\boldsymbol{C} = \boldsymbol{F}^* \cdot \operatorname{diag}\left(\widehat{a}_0, \widehat{a}_{n-1}, \dots, \widehat{a}_1\right) \cdot \boldsymbol{F}$, čiže $\boldsymbol{C}$ je podobná s uvedenou diagonálnou maticou Fourierových koeficientov prostredníctvom znormovanej (teda unitárnej) matice $\boldsymbol{F} = n^{-1/2}\left(\omega^{-jk}\right)$ diskrétnej Fourierovej transformácie $F\colon \mathbb{C}^n \to \mathbb{C}^n$ (pozri paragraf 17.5 a cvičenie 17.18).

(c) Pre determinant cirkulantnej matice, nazývaný tiež *cirkulant* prvkov $a_0, a_1, \dots, a_{n-1}$, platí $\det \boldsymbol{C} = g(1)g(\omega)\dots g(\omega^{n-1}) = \widehat{a}_0\widehat{a}_1 \dots \widehat{a}_{n-1}$.

19.15. Vysvetlite nasledujúci postup riešenia sústavy lineárnych rovníc $\boldsymbol{C} \cdot \boldsymbol{x} = \boldsymbol{b}$, kde $\boldsymbol{C}(a_0, a_1, \dots, a_{n-1})$ je cirkulatná matica a $\boldsymbol{x} = (x_0, x_1, \dots, x_{n-1})^{\mathsf{T}}$, $\boldsymbol{b} = (b_0, b_1, \dots, b_{n-1})^{\mathsf{T}}$.

(a) Uvedenú sústavu možno pomocou konvolúcie zapísať v tvare $\boldsymbol{a}^{\varrho} * \boldsymbol{x} = \boldsymbol{b}$, kde $\boldsymbol{a}^{\varrho} = (a_0, a_{n-1}, \dots, a_1)^{\mathsf{T}}$ je prvý stĺpec matice $\boldsymbol{C}$.

(b) Aplikujúc DFT na obe strany tejto rovnice dostaneme $F(\boldsymbol{a}^{\varrho})F(\boldsymbol{x}) = F(\boldsymbol{b})$, alebo inak zapísané $\widehat{\boldsymbol{a}^{\rho}}\widehat{\boldsymbol{x}} = \widehat{\boldsymbol{b}}$, kde $\boldsymbol{xy}$ označuje prirodzený súčin po zložkách v $\mathbb{C}^n$.

(c) Ak sú teda všetky zložky vektora $F(\boldsymbol{a}^{\varrho}) = \widehat{\boldsymbol{a}^{\varrho}}$ nenulové, môžeme riešenie sústavy vyjadriť pomocou inverznej Fourierovej transformácie $\boldsymbol{x} = F^{-1}\left(F(\boldsymbol{b})/F(\boldsymbol{a}^{\varrho})\right) = F^{-1}\left(\widehat{\boldsymbol{b}}/\widehat{\boldsymbol{a}^{\rho}}\right)$, kde $\boldsymbol{x}/\boldsymbol{y}$ označuje prirodzené delenie po zložkách v $\mathbb{C}^n$.

Aj keď uvedená formula tomu možno nenasvedčuje, ak použijeme rýchlu Forurierovu transformáciu, bude takýto výpočet podstatne rýchlejší než Gaussova-Jordanova eliminácia.

20. Jordanov kanonický tvar

V tejto kapitole si ukážeme, že i nediagonalizovateľné lineárne operátory či matice možno voľbou vhodnej bázy upraviť na tzv. *Jordanov kanonický tvar*, ktorý je – aspoň na pohľad – veľmi blízky diagonálnemu. Dôkaz tohto výsledku je však podstatne náročnejší než všetky dôkazy, s ktorými sme sa doteraz v tomto kurze stretli. Preto najprv iba sformulujeme príslušné vety a predvedieme, ako sa úprava na Jordanov kanonický tvar v niektorých jednoduchých prípadoch robí. S takýmito vedomosťami vystačíme vo väčšine učebnicových príkladov. Jednako pre náročnejšieho čitateľa uvádzame úplný dôkaz, ktorý nám zaberie celé dva paragrafy 20.4 a 20.5. Na jeho základe potom popíšeme ďalšiu metódu úpravy matice na Jordanov kanonický tvar. S niektorými aplikáciami výsledkov o Jordanovom kanonickom tvare sa oboznámime až v dvoch nasledujúcich kapitolách.

20.1 Jordanov kanonický tvar matice

Hovoríme, že matica $\boldsymbol{A} \in K^{n\times n}$ je v *Jordanovom kanonickom tvare*, skrátene JKT, ak má blokovo diagonálny tvar

$$\boldsymbol{A} = \operatorname{diag}\big(\boldsymbol{J}_{n_1}(\lambda_1), \dots, \boldsymbol{J}_{n_k}(\lambda_k)\big),$$

kde $\boldsymbol{J}_{n_i}(\lambda_i)$ sú Jordanove bunky rozmerov $n_i \times n_i$ prislúchajúce skalárom $\lambda_i \in K$ (pozri príklad 19.1.3). Zrejme v takom prípade je $n_1 + \ldots + n_k = n$ a $\boldsymbol{A}$ má charakteristický polynóm

$$\det(\boldsymbol{A} - x\boldsymbol{I}) = (\lambda_1 - x)^{n_1} \dots (\lambda_k - x)^{n_k}.$$

Vidíme, že skalár $\lambda \in K$ je vlastnou hodnotou matice $\boldsymbol{A}$ práve vtedy, keď sa nachádza v zozname $\lambda_1, \dots, \lambda_k$. Keďže skaláry $\lambda_1, \dots, \lambda_k$ nemusia byť nevyhnutne rôzne, algebraická násobnosť λ vzhľadom na $\boldsymbol{A}$ je súčet veľkostí blokov s hodnotou λ na diagonále, čiže $\sum_{\lambda_i=\lambda} n_i$. Ako vyplýva z príkladu 19.1.3, každému bloku $\boldsymbol{J}_{n_i}(\lambda_i)$, bez ohľadu na veľkosť n_i, zodpovedá len jednorozmerný vlastný podpriestor – preto geometrická násobnosť λ vzhľadon na $\boldsymbol{A}$ je rovná počtu takýchto blokov, t. j. počtu prvkov množiny $\{i \leq k;\ \lambda_i = \lambda\}$.

Jordanovým kanonickým tvarom matice $\boldsymbol{A} \in K^{n\times n}$ nazývame ľubovoľnú maticu $\boldsymbol{J} \in L^{n\times n}$ v JKT, kde pole L je nejaké rozšírenie poľa K, podobnú (nad poľom L) s maticou $\boldsymbol{A}$. *Upraviť maticu* $\boldsymbol{A}$ *na Jordanov kanonický tvar* znamená nájsť s ňou podobnú maticu $\boldsymbol{J} \in L^{n\times n}$ v Jordanovom kanonickom tvare a regulárnu maticu $\boldsymbol{P} \in L^{n\times n}$ takú, že $\boldsymbol{J} = \boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P}$. Potom lineárny operátor $\boldsymbol{x} \mapsto \boldsymbol{A} \cdot \boldsymbol{x}$ na (stĺpcovom) vektorovom priestore L^n má v báze tvorenej stĺpcami matice $\boldsymbol{P}$ maticu $\boldsymbol{J}$ v JKT.

Kľúčové výsledky tejto kapitoly možno zhrnúť do nasledujúcich dvoch viet.

20.1.1. Veta. *Nech $\varphi\colon V \to V$ je lineárny operátor na vektorovom priestore V konečnej dimenzie n nad poľom K. Ak φ má nad K spektrum algebraickej váhy n, tak existuje taká báza $\boldsymbol{\beta}$ priestoru V, vzhľadom na ktorú má φ maticu $(\varphi)_{\boldsymbol{\beta}}$ v Jordanovom kanonickom tvare. Pritom Jordanov kanonický tvar matice zobrazenia φ je určený jednoznačne až na poradie Jordanových blokov.*

20.1.2. Veta. *Nech matica $\boldsymbol{A} \in K^{n\times n}$ má nad poľom K spektrum algebraickej váhy n. Potom $\boldsymbol{A}$ je podobná s maticou $\boldsymbol{J} \in K^{n\times n}$ v Jordanovom kanonickom tvare. Pritom matica $\boldsymbol{J}$ je určená jednoznačne až na poradie Jordanových blokov.*

20.1.3. Dôsledok. *Nech matice $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$ majú v poli K spektrum algebraickej váhy n. Potom $\boldsymbol{A} \approx \boldsymbol{B}$ práve vtedy, keď $\boldsymbol{A}$ a $\boldsymbol{B}$ majú rovnaký Jordanov kanonický tvar.*

Všimnite si, že predpoklad o plnej algebraickej váhe spektra je splnený práve vtedy, keď K obsahuje rozkladové pole charakteristického polynómu lineárneho operátora φ, prípadne matice $\boldsymbol{A}$. Táto podmienka je automaticky splnená nad algebraicky uzavretým poľom K, špeciálne nad poľom komplexných čísel $\mathbb{C}$. Aj v prípade, že tento predpoklad nie je splnený, možno pole K vnoriť do jeho konečného rozšírenia L, nad ktorým už splnený bude. Podrobnosti, ako to možno urobiť, sme si aspoň v prípade rozšírenia $\mathbb{R} \subseteq \mathbb{C}$ objasnili v paragrafe 19.4 venovanom komplexifikácii.

Obe uvedené vety sú zrejme ekvivalentné, preto stačí dokázať len jednu z nich, prípadne dokazovať obe naraz a striedavo používať maticovú či operátorovú formuláciu, podľa toho, ktorá nám práve väčšmi vyhovuje. Dôkazu, ktorý je pomerne náročný, venujeme paragrafy 20.4 a 20.5. Považujeme však za potrebné predoslať mu niekoľko príkladov, na ktorých budeme ilustrovať *metódu* úpravy matice na JKT.

Celá metóda je založená na pozorovaní, čo robí matica $\boldsymbol{J}_n = \boldsymbol{J}_n(\lambda) - \lambda \boldsymbol{I}_n$ s kanonickou bázou $\boldsymbol{\varepsilon}^{(n)} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ v K^n. Odpoveď na túto otázku možno stručne a prehľadne vyjadriť schémou

$$\boldsymbol{J}_n\colon \boldsymbol{e}_n \mapsto \boldsymbol{e}_{n-1} \mapsto \dots \mapsto \boldsymbol{e}_2 \mapsto \boldsymbol{e}_1 \mapsto \boldsymbol{0},$$

t. j. $\boldsymbol{J}_n \cdot \boldsymbol{e}_1 = \boldsymbol{0}$ (inak povedané, $\boldsymbol{e}_1$ je vlastný vektor matice $\boldsymbol{J}_n(\lambda)$), a $\boldsymbol{J}_n \cdot \boldsymbol{e}_i = \boldsymbol{e}_{i-1}$ pre $1 < i \leq n$.

Ak má teda lineárny operátor $\varphi\colon V \to V$ mať v nejakej báze $\boldsymbol{\beta}$ maticu v JKT

$$(\varphi)_{\boldsymbol{\beta}} = \boldsymbol{J} = \operatorname{diag}\bigl(\boldsymbol{J}_{n_1}(\lambda_1), \dots, \boldsymbol{J}_{n_k}(\lambda_k)\bigr),$$

musí sa táto báza dať rozložiť na k reťazcov

$$\begin{aligned}
\boldsymbol{\beta}_1 &= (\boldsymbol{v}_{11}, \dots, \boldsymbol{v}_{1n_1}),\\
\boldsymbol{\beta}_2 &= (\boldsymbol{v}_{21}, \dots, \boldsymbol{v}_{2n_2}),\\
&\dots\dots\dots\dots\dots\\
\boldsymbol{\beta}_k &= (\boldsymbol{v}_{k1}, \dots, \boldsymbol{v}_{kn_k}),
\end{aligned}$$

zodpovedajúcich jednotlivým Jordanovým bunkám $\boldsymbol{J}_{n_1}(\lambda_1), \boldsymbol{J}_{n_2}(\lambda_2), \dots, \boldsymbol{J}_{n_k}(\lambda_k)$. Na každom reťazci $\boldsymbol{\beta}_i$ pôsobí lineárny operátor $\varphi - \lambda_i \operatorname{id}_V = \varphi - \lambda_i$ rovnako ako matica $\boldsymbol{J}_{n_i}$ na kanonickej báze $\boldsymbol{\varepsilon}^{(n_i)} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_{n_i})$, t. j. podľa schémy

$$\varphi - \lambda_i\colon \boldsymbol{v}_{in_i} \mapsto \boldsymbol{v}_{i\,n_i-1} \mapsto \dots \mapsto \boldsymbol{v}_{i2} \mapsto \boldsymbol{v}_{i1} \mapsto \mathbf{0}.$$

Potom prvý vektor $\boldsymbol{v}_{i1}$ každého reťazca $\boldsymbol{\beta}_i$ (v našej schéme prvý nenulový vektor sprava) je vlastným vektorom operátora φ prislúchajúcim k vlastnej hodnote λ_i. Celý reťazec je zasa tvorený postupnými obrazmi posledného vektora $\boldsymbol{v}_{in_i}$ (v našej schéme prvého vektora zľava) v zobrazení $\varphi - \lambda_i$, t. j.

$$\boldsymbol{\beta}_i = \bigl((\varphi - \lambda_i)^{n_i-1}(\boldsymbol{v}_{in_i}), (\varphi - \lambda_i)^{n_i-2}(\boldsymbol{v}_{in_i}), \dots, (\varphi - \lambda_i)(\boldsymbol{v}_{in_i}), \boldsymbol{v}_{in_i}\bigr).$$

Bázu zloženú z takýchto reťazcov nazývame *Jordanovou bázou* príslušného lineárneho operátora alebo matice.

Úprava matice (lineárneho operátora) na JKT teda spočíva hlavne v nájdení Jordanovej bázy, t. j. bázy zloženej z reťazcov vektorov prislúchajúcich jednotlivým vlastným číslam a vlastným vektorom tejto matice. Také niečo však predpokladá znalosť spektra, preto jeho určenie (prípadne doplnenie do plnej algebraickej váhy vo vhodnom rozšírení základného poľa) musí konštrukcii takejto bázy nevyhnutne predchádzať.

Vopred však poznamenávame, že priamočiara metóda budovania bázických reťazcov od vlastných vektorov „sprava doľava“, dobre funguje len pre matice malých rozmerov (maximálne tak 6×6), prípadne pre matice, v ktorých JKT sa vyskytujú len malé Jordanove bloky (maximálne do rozmeru 3×3). Pre matice väčších rozmerov, kde sa kombinatorika možných rozmerov blokov stáva pestrejšou, je už neúnosne zložitá. Neskôr, na základe dôkazu viet 20.1.1 a 20.1.2, sa zoznámime aj s metódou postupujúcou pri vytváraní bázických reťazcov „zľava doprava“. V nasledujúcej kapitole si ešte stručne priblížime elegantnú metódu založenú na úprave charakteristickej matice $\boldsymbol{A} - x\boldsymbol{I}$ pomocou ERO a ESO, využívajúcu niektoré hlbšie poznatky o maticiach, ktorých prvky sú polynómy nad daným poľom.

Samostatnú poznámku si vyžaduje otázka určenia spektra analyzovanej matice alebo operátora. Naše príklady budú vopred umelo pripravené tak, aby sme príslušnú charakteristickú rovnicu vedeli ľahko vyriešiť (väčšinou budú dokonca všetky jej korene malé celé čísla). V „reálnom živote“ to však tak byť nemusí. Vo všeobecnosti neexistujú vzorce, ktoré by vyjadrovali korene polynómu $f(x) \in \mathbb{C}[x]$ stupňa ≥ 5 ako funkcie jeho koeficientov zostavené pomocou operácií sčítania, odčítania, násobenia, delenia a mocnín s racionálnym exponentom. [1] Ani pre polynómy stupňa 3 a 4, pre ktoré takéto vzorce existujú, však nie sú pre svoju ťažkopádnosť prakticky upotrebiteľné. Takže použiteľné explicitné vzorce máme k dispozícii len na riešenie rovníc stupňa 1 a 2.

Tým vzrastá význam približných numerických metód výpočtu vlastných čísel a vektorov a JKT. Tieto otázky však už nie sú predmetom nášho kurzu. Momentálne nie je naším cieľom výpočtovo zvládnuť uvedenú problematiku v celej všeobecnosti, ale porozumieť jej základným súvislostiam. Keďže práve výsledky a metódy lineárnej algebry hrajú v modernej numerickej matematike významnú úlohu, je takéto prvotné porozumenie jedným z nevyhnutných predpokladov zvládnutia pokročilejších numerických metód.

Jeden numerický aspekt výpočtu JKT však nemožno v tejto súvislosti nespomenúť. Numerické metódy väčšinou dávajú len približné výsledky s istou vopred zadanou presnosťou. Navyše často pracujú so vstupnými údajmi získanými meraním, teda už od začiatku zaťaženými určitými chybami. Na druhej strane, „takmer všetky“ štvorcové matice nad $\mathbb{C}$ sú podobné s diagonálnymi. [2] Pri hocako malej náhodnej zmene prvkov sa takáto matica s pravdepodobnosťou hraničiacou s istotou stane podobnou diagonálnej. Teda matice s nediagonálnym JKT sú vlastne atypickými výnimkami s nekonečne málo pravdepodobným výskytom. To jednak činí numerický výpočet JKT značne chúlostivou záležitosťou, jednak navodzuje otázku, aký má vôbec význam zaoberať sa nediagonalizovateľnými maticami a lineárnymi operátormi. Dodajme teda, že na druhej strane typická matica $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ s aspoň jednou viacnásobnou vlastnou hodnotou je nediagonalizovateľná (pokúste sa samostatne upresniť význam tohto tvrdenia). Čo je však ešte dôležitejšie, s prirodzenými príkladmi takýchto operátorov sa možno stretnúť napr. v matematickej analýze.

[1] Dôkaz neriešiteľnosti rovníc piateho a vyššieho stupňa pomocou radikálov je súčasťou tzv. Galoisovej teórie a patrí k vrcholným výkonom algebry 19. storočia.

[2] Presnejšie, topologická dimenzia množiny všetkých matíc $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, ktoré nie sú podobné s diagonálnymi, je menšia než topologická dimenzia priestoru všetkých matíc $\mathbb{C}^{n\times n}$. Podobne, topologická dimenzia množiny všetkých matíc $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, ktoré nie sú podobné s diagonálnymi maticami z $\mathbb{C}^{n\times n}$, je menšia než topologická dimenzia priestoru všetkých matíc $\mathbb{R}^{n\times n}$. Pojem topologickej dimenzie je (zďaleka nie priamočiarym) zovšeobecnením dimenzie lineárnych a afinných podpriestorov nad $\mathbb{R}$.

20.1.4. Príklad. Lineárny operátor derivácie $D\colon \mathbb{R}^{(n)}[x] \to \mathbb{R}^{(n)}[x]$, kde $D(f) = f'(x) = \frac{\mathrm{d}f(x)}{\mathrm{d}x}$ pre $f(x) \in \mathbb{R}^{(n)}[x]$, na priestore všetkých reálnych polynómov stupňa $\leq n$ má v báze $\boldsymbol{\xi}^{(n)} = (1, x, x^2 \dots, x^n)$ maticu

$$(D)_{\boldsymbol{\xi}^{(n)}} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & n \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

Z toho už možno pomerne ľahko nahliadnuť (prípadne dopočítať), že maticou operátora D v báze $\left(1, x, \frac{x^2}{2}, \dots, \frac{x^n}{n!}\right)$ je Jordanova bunka $\boldsymbol{J}_{n+1} = \boldsymbol{J}_{n+1}(0)$.

20.2 Príklady úpravy matíc na Jordanov kanonický tvar

20.2.1. Príklad. Uvažujme maticu

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 2 & -2 \\ 1 & 3 & 1 & -1 \\ 1 & 0 & 0 & -1 \\ 3 & 4 & 3 & -4 \end{pmatrix} \in \mathbb{R}^{4\times 4}.$$

Jej charakteristický polynóm

$$\det(\boldsymbol{A} - x\boldsymbol{I}) = x^4 - 2x^2 + 1 = (x-1)^2(x+1)^2$$

má dva korene $x_{1,2} = 1$ a $x_{3,4} = -1$, oba dvojnásobné. Nájdeme k nim príslušné vlastné vektory. Podpriestor riešení homogénnej sústavy s maticou

$$\boldsymbol{A} - \boldsymbol{I} = \begin{pmatrix} 0 & 0 & 2 & -2 \\ 1 & 2 & 1 & -1 \\ 1 & 0 & -1 & -1 \\ 3 & 4 & 3 & -5 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

je jednorozmerný, generovaný vlastným vektorom $\boldsymbol{v}_1 = (2, -1, 1, 1)^{\mathsf{T}}$. Teda algebraicky dvojnásobné vlastné číslo 1 má geometrickú násobnosť 1. Ďalší vektor reťazca nájdeme ako nejaké riešenie $\boldsymbol{x} = \boldsymbol{v}_2$ sústavy $(\boldsymbol{A} - \boldsymbol{I}) \cdot \boldsymbol{x} = \boldsymbol{v}_1$ úpravou jej rozšírenej matice

$$\left(\begin{array}{cccc|c} 0 & 0 & 2 & -2 & 2 \\ 1 & 2 & 1 & -1 & -1 \\ 1 & 0 & -1 & -1 & 1 \\ 3 & 4 & 3 & -5 & 1 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & -2 & 2 \\ 0 & 1 & 0 & 1 & -2 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

teda napr. $\boldsymbol{v}_2 = (2, -2, 1, 0)^\mathsf{T}$.

Podpriestor riešení homogénnej sústavy s maticou

$$\boldsymbol{A} + \boldsymbol{I} = \begin{pmatrix} 2 & 0 & 2 & -2 \\ 1 & 4 & 1 & -1 \\ 1 & 0 & 1 & -1 \\ 3 & 4 & 3 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

má dimenziu 2 (a taká je tiež geometrická násobnosť algebraicky dvojnásobného vlastného čísla -1); jeho bázu tvoria vlastné vektory $\boldsymbol{v}_3 = (1, 0, -1, 0)^\mathsf{T}$, $\boldsymbol{v}_4 = (1, 0, 0, 1)^\mathsf{T}$.

Teda $\boldsymbol{A}$ je podobná matici v JKT

$$\boldsymbol{J} = \operatorname{diag}\big(\boldsymbol{J}_2(1), -1, -1\big) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

a príslušná Jordanova báza $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4)$ je tvorená stĺpcami matice prechodu

$$\boldsymbol{P} = \begin{pmatrix} 2 & 2 & 1 & 1 \\ -1 & -2 & 0 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Videli sme, že reťazce vektorov prislúchajúce rôznym vlastným číslam možno bez ťažkostí oddeliť. Preto sa odteraz sústredíme len na hľadanie a oddeľovanie reťazcov prislúchajúcich tomu istému vlastnému číslu a budeme sa zaoberať iba maticami s jednoprvkovým spektrom.

20.2.2. Príklad. Matica

$$\boldsymbol{A} = \begin{pmatrix} -2 & 1 & 0 \\ -3 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

má charakteristický polynóm

$$\det(\boldsymbol{A} - x\boldsymbol{I}) = -x^3$$

a jediné, algebraicky trojnásobné vlastné číslo $x_{1,2,3} = 0$. Vlastné vektory nájdeme riešením homogénnej sústavy s maticou

$$\boldsymbol{A} - 0\boldsymbol{I} = \boldsymbol{A} \sim \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Podpriestor riešení je jednorozmerný, generovaný vektorom $\boldsymbol{u} = (1,2,1)^\mathsf{T}$, teda geometrická násobnosť vlastného čísla 0 je 1. Hľadaná báza tak bude pozostávať z jediného reťazca prislúchajúceho vlastnému vektoru $\boldsymbol{u}_1 = \boldsymbol{u}$. Vektor $\boldsymbol{u}_2$ nájdeme ako nejaké riešenie $\boldsymbol{x} = \boldsymbol{u}_2$ sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{u}_1$ úpravou jej rozšírenej matice

$$(\boldsymbol{A} \,|\, \boldsymbol{u}_1) = \left(\begin{array}{ccc|c} -2 & 1 & 0 & 1 \\ -3 & 1 & 1 & 2 \\ -1 & 0 & 1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & -1 & -1 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right);$$

takže môžeme položiť napr. $\boldsymbol{u}_2 = (-1,-1,0)^\mathsf{T}$. Podobne, tretí vektor $\boldsymbol{u}_3$ nášho reťazca nájdeme ako nejaké riešenie $\boldsymbol{x} = \boldsymbol{u}_3$ sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{u}_2$ úpravou jej rozšírenej matice

$$(\boldsymbol{A} \,|\, \boldsymbol{u}_2) = \left(\begin{array}{ccc|c} -2 & 1 & 0 & -1 \\ -3 & 1 & 1 & -1 \\ -1 & 0 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

teda napr. $\boldsymbol{u}_3 = (0,-1,0)^\mathsf{T}$.

To znamená, že $\boldsymbol{A}$ je podobná priamo s Jordanovou bunkou

$$\boldsymbol{J}_3 = \boldsymbol{J}_3(0) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

prostredníctvom matice prechodu

$$\boldsymbol{P} = \begin{pmatrix} 1 & -1 & 0 \\ 2 & -1 & -1 \\ 1 & 0 & 0 \end{pmatrix},$$

tvorenej vektormi Jordanovej bázy $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$ ako stĺpcami.

20.2.3. Príklad. Matica

$$\boldsymbol{A} = \begin{pmatrix} 8 & 10 & -5 \\ -2 & -1 & 2 \\ 1 & 2 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

má charakteristický polynóm

$$\det(\boldsymbol{A} - x\boldsymbol{I}) = 27 - 27x + 9x^2 - x^3 = (3-x)^3$$

a algebraicky trojnásobné vlastné číslo $x_{1,2,3} = 3$. K nemu príslušné vlastné vektory nájdeme riešením homogénnej sústavy s maticou

$$\boldsymbol{A} - 3\boldsymbol{I} = \begin{pmatrix} 5 & 10 & -5 \\ -2 & -4 & 2 \\ 1 & 2 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Podpriestor riešení je dvojrozmerný, generovaný (napríklad) vektormi $\boldsymbol{u} = (2,-1,0)^{\mathsf{T}}$, $\boldsymbol{v} = (1,0,1)^{\mathsf{T}}$. Teda geometrická násobnosť vlastného čísla 3 je 2, takže k nemu budú prislúchať dva reťazce dĺžok 1 a 2. Keďže vopred nevieme, ku ktorému vlastnému vektoru $\boldsymbol{v}_1 \in [\boldsymbol{u}, \boldsymbol{v}]$ existuje vektor $\boldsymbol{v}_2$ taký, že $\boldsymbol{v}_1 = (\boldsymbol{A} - 3\boldsymbol{I}) \cdot \boldsymbol{v}_2$, musíme uvažovať ľubovoľnú lineárnu kombináciu $\boldsymbol{v}_1 = a\boldsymbol{u} + b\boldsymbol{v} = (2a+b, -a, b)^{\mathsf{T}}$, pričom parametre $a, b \in \mathbb{R}$ budeme voliť tak, aby sústava $(\boldsymbol{A} - 3\boldsymbol{I}) \cdot \boldsymbol{x} = \boldsymbol{v}_1$ mala nejaké riešenie $\boldsymbol{x} = \boldsymbol{v}_2$. Úpravou jej rozšírenej matice dostaneme

$$\left(\begin{array}{rrr|c} 5 & 10 & -5 & 2a+b \\ -2 & -4 & 2 & -a \\ 1 & 2 & -1 & b \end{array}\right) \sim \left(\begin{array}{rrr|c} 1 & 2 & -1 & b \\ 0 & 0 & 0 & a-2b \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Sústava má riešenie práve vtedy, keď $a = 2b$; volíme napr. $b = 1$, $a = 2$. Tomu zodpovedá $\boldsymbol{v}_1 = (5,-2,1)^{\mathsf{T}}$, $\boldsymbol{v}_2 = (1,0,0)^{\mathsf{T}}$. Za $\boldsymbol{v}_3$ možno zvoliť akýkoľvek vektor, ktorý spolu s $\boldsymbol{v}_1$ tvorí bázu vlastného podpriestoru $\mathcal{R}(\boldsymbol{A} - 3\boldsymbol{I}) = [\boldsymbol{u}, \boldsymbol{v}]$; vidíme, že vyhovujú obe voľby $\boldsymbol{v}_3 = \boldsymbol{u}$, resp. $\boldsymbol{v}_3 = \boldsymbol{v}$. Vyberme si napr. druhú možnosť $\boldsymbol{v}_3 = (1,0,1)^{\mathsf{T}}$.

Teda JKT matice $\boldsymbol{A}$ je

$$\boldsymbol{J} = \operatorname{diag}\big(\boldsymbol{J}_2(3), 3\big) = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

a príslušná matica prechodu tvorená stĺpcami Jordanovej bázy $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ vyzerá (napr.) takto

$$\boldsymbol{P} = \begin{pmatrix} 5 & 1 & 1 \\ -2 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

20.2.4. Príklad. Matica

$$\boldsymbol{A} = \begin{pmatrix} -4 & -6 & -19 & -1 \\ 2 & 4 & 3 & 1 \\ 1 & 1 & 6 & 0 \\ 5 & 5 & 20 & 2 \end{pmatrix} \in \mathbb{R}^{4\times 4}$$

má charakteristický polynóm

$$\det(\boldsymbol{A} - x\boldsymbol{I}) = x^4 - 8x^3 + 24x^2 - 32x + 16 = (x-2)^4$$

s jediným štvornásobným koreňom $x_{1\text{–}4} = 2$. Vlastné vektory nájdeme riešením homogénnej sústavy s maticou

$$\boldsymbol{A} - 2\boldsymbol{I} = \begin{pmatrix} -6 & -6 & -19 & -1 \\ 2 & 2 & 3 & 1 \\ 1 & 1 & 4 & 0 \\ 5 & 5 & 20 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 & 4/5 \\ 0 & 0 & 1 & -1/5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Podpriestor riešení je dvojrozmerný, generovaný vektormi $\boldsymbol{u} = (1, -1, 0, 0)^{\mathsf{T}}$, $\boldsymbol{v} = (4, 0, -1, -5)^{\mathsf{T}}$; všeobecné riešenie má tvar $a\boldsymbol{u}+b\boldsymbol{v} = (a+4b, -a, -b, -5b)^{\mathsf{T}}$, kde $a, b \in \mathbb{R}$. Vidíme, že hľadaná báza bude pozostávať z dvoch reťazcov, nepoznáme však ich dĺžky – sú totiž dve možnosti zodpovedajúce rozkladom $4 = 1 + 3$, resp. $4 = 2 + 2$. Ďalšie vektory reťazcov získame riešením sústavy $(\boldsymbol{A} - 2\boldsymbol{I}) \cdot \boldsymbol{x} = a\boldsymbol{u} + b\boldsymbol{v}$ s rozšírenou maticou

$$\left(\begin{array}{rrrr|c} -6 & -6 & -19 & -1 & a+4b \\ 2 & 2 & 3 & 1 & -a \\ 1 & 1 & 4 & 0 & -b \\ 5 & 5 & 20 & 0 & -5b \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 1 & 0 & 4/5 & -4a/5 + 3b/5 \\ 0 & 0 & 1 & -1/5 & a/5 - 2b/5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Ako vidno, sústava má riešenie pre ľubovoľné a, b; môžeme teda voliť $a = 1$, $b = 0$, čomu zodpovedá prvý reťazec $\boldsymbol{u}_1 = \boldsymbol{u} = (1, -1, 0, 0)^{\mathsf{T}}$, $\boldsymbol{u}_2 = (-4/5, 0, 1/5, 0)^{\mathsf{T}}$, resp. $a = 0$, $b = 1$, čomu zodpovedá druhý reťazec $\boldsymbol{v}_1 = \boldsymbol{v} = (4, 0, -1, -5)^{\mathsf{T}}$, $\boldsymbol{v}_2 = (3/5, 0, -2/5, 0)^{\mathsf{T}}$.

JKT matice $\boldsymbol{A}$ teda je

$$\boldsymbol{J} = \operatorname{diag}\big(\boldsymbol{J}_2(2), \boldsymbol{J}_2(2)\big) = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix},$$

a príslušná Jordanova báza $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{v}_1, \boldsymbol{v}_2)$ je tvorená stĺpcami matice prechodu

$$\boldsymbol{P} = \begin{pmatrix} 1 & -4/5 & 4 & 3/5 \\ -1 & 0 & 0 & 0 \\ 0 & 1/5 & -1 & -2/5 \\ 0 & 0 & -5 & 0 \end{pmatrix}.$$

20.2.5. Príklad. Uvažujme maticu

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & -2 & -1 & -1 \end{pmatrix} \in \mathbb{R}^{4\times 4}.$$

Jej charakteristický polynóm

$$\det(\boldsymbol{A} - x\boldsymbol{I}) = x^4$$

má jeden štvornásobný koreň $x_{1\text{-}4} = 0$. Vlastné vektory dostaneme ako riešenia homogénnej sústavy s maticou

$$\boldsymbol{A} - 0\boldsymbol{I} = \boldsymbol{A} \sim \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Podpriestor riešení je dvojrozmerný, generovaný vektormi $\boldsymbol{u} = (1,0,-1,0)^{\mathsf{T}}$, $\boldsymbol{v} = (1,0,0,-1)^{\mathsf{T}}$; všeobecné riešenie má tvar $a\boldsymbol{u} + b\boldsymbol{v} = (a+b,0,-a,-b)^{\mathsf{T}}$, kde $a, b \in \mathbb{R}$. Takže hľadaná báza bude opäť pozostávať z dvoch reťazcov dĺžok $1+3$ alebo $2+2$. Ďalšie vektory reťazcov získame riešením sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = a\boldsymbol{u} + b\boldsymbol{v}$ úpravou jej rozšírenej matice

$$\left(\begin{array}{cccc|c} 0 & 1 & 0 & 0 & a+b \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & -a \\ -1 & -2 & -1 & -1 & -b \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 1 & 1 & -2a-b \\ 0 & 1 & 0 & 0 & a+b \\ 0 & 0 & 0 & 0 & 2a+b \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Sústava má riešenie práve vtedy, keď $2a+b=0$. Z toho je jasné, že dĺžky hľadaných reťazcov budú 1 a 3. Zvolíme napr. $a=-1$, $b=2$. Dostaneme tak prvý vektor trojčlenného reťazca $\boldsymbol{v}_1 = -\boldsymbol{u}+2\boldsymbol{v} = (1,0,1,-2)^{\mathsf{T}}$. Druhý vektor zatiaľ ponecháme v tvare všeobecného riešenia $\boldsymbol{v}_2 = \boldsymbol{x} = (c+d,1,-c,-d)^{\mathsf{T}}$ sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{v}_1$ a parametre $c, d \in \mathbb{R}$ budeme voliť tak, aby existovalo nejaké riešenie $\boldsymbol{x} = \boldsymbol{v}_3$ sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{v}_2$. Úpravou jej rozšírenej matice dostaneme

$$\left(\begin{array}{cccc|c} 0 & 1 & 0 & 0 & c+d \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -c \\ -1 & -2 & -1 & -1 & -d \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 1 & 1 & -2c-d \\ 0 & 1 & 0 & 0 & c+d \\ 0 & 0 & 0 & 0 & 1+2c+d \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Sústava má riešenie práve vtedy, keď $1+2c+d=0$; zvoľme napr. $c=0$, $d=-1$. Tomu zodpovedá $\boldsymbol{v}_2 = (-1,1,0,1)^{\mathsf{T}}$ a (napr.) $\boldsymbol{v}_3 = (1,-1,0,0)^{\mathsf{T}}$. Jediný vektor druhého reťazca zvolíme tak, aby spolu s vektorom $\boldsymbol{v}_1$ tvorili bázu vlastného podpriestoru $[\boldsymbol{u}, \boldsymbol{v}]$; vyhovuje každý z vektorov $\boldsymbol{u}$, $\boldsymbol{v}$. Zvoľme napr. $\boldsymbol{v}_4 = \boldsymbol{u} = (1,0,-1,0)^{\mathsf{T}}$.

JKT matice $\boldsymbol{A}$ teda je

$$\boldsymbol{J} = \operatorname{diag}\bigl(\boldsymbol{J}_3(0), 0\bigr) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

a príslušná matica prechodu, tvorená stĺpcami Jordanovej bázy $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4)$, vyzerá takto:

$$\boldsymbol{P} = \begin{pmatrix} 1 & -1 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ -2 & 1 & 0 & 0 \end{pmatrix}.$$

20.3 Prípad viacnásobného komplexného vlastného čísla

V predchádzajúcom paragrafe sme sa úmyselne vyhli tak komplexným maticiam, ako aj reálnym maticiam s komplexnými vlastnými číslami. Viedla nás k tomu skôr pohodlnosť, než nejaké zásadné dôvody. Jednoducho sme sa snažili sústrediť len na samotnú úpravu matice na JKT a nechceli si zbytočne komplikovať život z tohto hľadiska nepodstatnými detailmi komplexnej aritmetiky. Úprava komplexných matíc (v ktorých sú už zahrnuté i reálne matice s komplexnými vlastnými hodnotami) na JKT sa od reálneho prípadu nijako zásadne nelíši. Len výsledná Jordanova matica môže mať na diagonále komplexné vlastné hodnoty a taktiež v príslušnej Jordanovej báze sa môžu objaviť komplexné vektory.

Z hľadiska aplikácií JKT najmä na riešenie diferenciálnych rovníc je však dôležité vedieť nahradiť komplexný JKT nejakej reálnej matice vhodným *reálnym* kanonickým tvarom, podobným s pôvodnou maticou prostredníctvom *reálnej* matice prechodu. To možno dosiahnuť rozvinutím už známych metód z paragrafu 19.5. Z toho dôvodu si dovolíme celý postup len stručne opísať a overenie detailov prenechať čitateľovi.

Ak $\lambda = a+ib$ je komplexné vlastné číslo matice $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, tak aj $\overline{\lambda} = a-ib$ je jej vlastné číslo, pričom λ a $\overline{\lambda}$ majú rovnakú algebraickú i geometrickú násobnosť. Ak navyše $\boldsymbol{\gamma} = (\boldsymbol{w}_1, \dots, \boldsymbol{w}_m)$ je časť Jordanovej bázy prislúchajúca k λ, zložená z reťazcov dĺžok $m_1 + \ldots + m_l = m$, tak $\overline{\boldsymbol{\gamma}} = (\overline{\boldsymbol{w}}_1, \dots, \overline{\boldsymbol{w}}_m)$ je časť Jordanovej bázy prislúchajúca k $\overline{\lambda}$, zložená v zodpovedajúcom poradí z komplexne združených reťazcov rovnakých dĺžok.

Z toho mimochodom vyplýva, že aj keby sme sa neusilovali o elimináciu komplexných čísel z JKT, možno vzťah medzi λ a $\overline{\lambda}$ a k nim prislušnými bázami výhodne využiť: stačí nájsť príslušnú časť Jordanovej bázy len pre jednu z vlastných hodnôt λ, $\overline{\lambda}$ – zodpovedajúcu časť bázy pre druhú vlastnú hodnotu už možno dostať čiste mechanicky „opruhovaním“. Podčiarkujeme však, že také niečo funguje len pre *reálne* a nie všeobecne pre komplexné matice.

Vráťme sa však k pôvodnej otázke. Pre $m \geq 1$ a $a, b \in \mathbb{R}$ označme

$$\boldsymbol{J}_m\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right) = \begin{pmatrix} a & -b & 1 & 0 & 0 & 0 & \dots & \dots & 0 & 0 \\ b & a & 0 & 1 & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & a & -b & 1 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & b & a & 0 & 1 & \dots & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \ddots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \dots & a & -b & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & \dots & b & a & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 & a & -b \\ 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 & b & a \end{pmatrix}$$

reálnu maticu rozmeru $2m \times 2m$, vytvorenú z m diagonálne umiestnených blokov $\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$ a $m-1$ blokov $\boldsymbol{I}_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ vedľa nich. Matice tvaru $\boldsymbol{J}_m\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$ nazývame *zovšeobecnenými Jordanovými bunkami.*

Práve zovšeobecnená Jordanova bunka $\boldsymbol{J}_m\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$ je tou reálnou maticou, ktorou možno v komplexnom JKT reálnej matice $\boldsymbol{A}$ nahradiť dvojicu Jordanových blokov $\boldsymbol{J}_m(\lambda)$, $\boldsymbol{J}_m(\overline{\lambda})$ prislúchajúcich komplexne združeným vlastným číslam $\lambda = a + \mathrm{i}b$, $\overline{\lambda} = a - \mathrm{i}b$ s nenulovým b. Presnejšie, ak $\boldsymbol{\gamma} = (\boldsymbol{w}_1, \dots, \boldsymbol{w}_m)$ je reťazec (komplexných) vektorov zodpovedajúci v príslušnej Jordanovej báze bloku $\boldsymbol{J}_m(\lambda)$, tak reálne vektory

$$\begin{aligned} \boldsymbol{u}_1 &= \operatorname{Re} \boldsymbol{w}_1 = \frac{1}{2}(\boldsymbol{w}_1 + \overline{\boldsymbol{w}}_1), & \boldsymbol{v}_1 &= -\operatorname{Im} \boldsymbol{w}_1 = \frac{1}{2\mathrm{i}}(\overline{\boldsymbol{w}}_1 - \boldsymbol{w}_1), \\ \boldsymbol{u}_2 &= \operatorname{Re} \boldsymbol{w}_2 = \frac{1}{2}(\boldsymbol{w}_2 + \overline{\boldsymbol{w}}_2), & \boldsymbol{v}_2 &= -\operatorname{Im} \boldsymbol{w}_2 = \frac{1}{2\mathrm{i}}(\overline{\boldsymbol{w}}_2 - \boldsymbol{w}_2), \\ &\vdots & &\vdots \\ \boldsymbol{u}_m &= \operatorname{Re} \boldsymbol{w}_m = \frac{1}{2}(\boldsymbol{w}_m + \overline{\boldsymbol{w}}_m), & \boldsymbol{v}_m &= -\operatorname{Im} \boldsymbol{w}_m = \frac{1}{2\mathrm{i}}(\overline{\boldsymbol{w}}_m - \boldsymbol{w}_m) \end{aligned}$$

tvoria úsek bázy $(\boldsymbol{u}_1, \boldsymbol{v}_1, \boldsymbol{u}_2, \boldsymbol{v}_2, \dots, \boldsymbol{u}_m, \boldsymbol{v}_m)$ zodpovedajúci bloku $\boldsymbol{J}_m\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$.

Ak teda v komplexnom JKT $\boldsymbol{J}$ štvorcovej reálnej matice $\boldsymbol{A}$ necháme reálne Jordanove bloky a im zodpovedajúce reťazce Jordanovej bázy na pokoji, ďalej každú dvojicu Jordanových blokov $\boldsymbol{J}_m(\lambda)$, $\boldsymbol{J}_m(\overline{\lambda})$ prislúchajúcich komplexne združeným imaginárnym vlastným hodnotám λ, $\overline{\lambda}$ nahradíme blokom $\boldsymbol{J}_m\left(\begin{smallmatrix} \operatorname{Re}\lambda & -\operatorname{Im}\lambda \\ \operatorname{Im}\lambda & \operatorname{Re}\lambda \end{smallmatrix}\right)$ a k nim zodpovedajúce reťazce $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_m)$, $(\overline{\boldsymbol{w}}_1, \dots, \overline{\boldsymbol{w}}_m)$ Jordanovej bázy nahradíme úsekom $(\operatorname{Re} \boldsymbol{w}_1, -\operatorname{Im} \boldsymbol{w}_1, \dots, \operatorname{Re} \boldsymbol{w}_m, -\operatorname{Im} \boldsymbol{w}_m)$, získame tak zovšeobecnenú Jordanovu maticu $\boldsymbol{J}' \in \mathbb{R}^{n \times n}$ a *reálnu* maticu prechodu $\boldsymbol{Q}$ (ktorej stĺpce sú vektory novej bázy) takú, že $\boldsymbol{J}' = \boldsymbol{Q}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{Q}$.

20.3.1. Príklad. Reálna matica

$$\begin{pmatrix} 5 & -5 & 5 & -4 & 4 & -2 \\ 5 & -5 & 7 & -7 & 8 & -4 \\ 5 & -8 & 12 & -12 & 13 & -6 \\ 5 & -9 & 12 & -12 & 15 & -7 \\ 5 & -9 & 12 & -15 & 20 & -9 \\ 5 & -9 & 12 & -16 & 20 & -8 \end{pmatrix}$$

má charakteristický polynóm

$$\det(\boldsymbol{A}-x\boldsymbol{I}) = x^6-12x^5+63x^4-184x^3+315x^2-300x+125 = (x^2-4x+5)^3,$$

s dvoma trojnásobnými komplexne združenými koreňmi $x_{1,2,3} = 2+\mathrm{i}$ a $x_{3,4,5} = \overline{x}_{1,2,3} = 2-\mathrm{i}$. Vlastný vektor prislúchajúci k vlastnému číslu $2+\mathrm{i}$ nájdeme ako riešenie homogénnej sústavy s maticou $\boldsymbol{A}-(2+\mathrm{i})\boldsymbol{I}$. Podpriestor riešení je jednorozmerný, generovaný vektorom $\boldsymbol{w}_1 = (3+\mathrm{i}, 5, 5, 5, 5, 5)^{\mathsf{T}}$. Vlastnej hodnote $2+\mathrm{i}$ tak zodpovedá jediný reťazec dĺžky 3. Ďalší vektor reťazca nájdeme ako riešenie sústavy $(\boldsymbol{A}-(2+\mathrm{i})\boldsymbol{I})\cdot\boldsymbol{w}_2 = \boldsymbol{w}_1$; vyhovuje $\boldsymbol{w}_2 = (-16/5-12\mathrm{i}/5, -5, -2+\mathrm{i}, 0, 0, 0)^{\mathsf{T}}$. Konečne tretí vektor dostaneme ako riešenie sústavy $(\boldsymbol{A}-(2+\mathrm{i})\boldsymbol{I})\cdot\boldsymbol{w}_3 = \boldsymbol{w}_2$; teda napr. $\boldsymbol{w}_3 = (-8/25 + 44\mathrm{i}/25, 0, -16/5-12\mathrm{i}/5, -5, -2+\mathrm{i}, 0)^{\mathsf{T}}$.

Komplexný JKT matice $\boldsymbol{A}$ teda je

$$\boldsymbol{J} = \operatorname{diag}(\boldsymbol{J}_3(2+\mathrm{i}), \boldsymbol{J}_3(2-\mathrm{i})) = \begin{pmatrix} 2+\mathrm{i} & 1 & 0 & 0 & 0 & 0 \\ 0 & 2+\mathrm{i} & 1 & 0 & 0 & 0 \\ 0 & 0 & 2+\mathrm{i} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2-\mathrm{i} & 1 & 0 \\ 0 & 0 & 0 & 0 & 2-\mathrm{i} & 1 \\ 0 & 0 & 0 & 0 & 0 & 2-\mathrm{i} \end{pmatrix}$$

a stĺpce príslušnej (komplexnej) matice prechodu (ktorú nevypisujeme z typografických dôvodov) tvoria vektory Jordanovej bázy $(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3, \overline{\boldsymbol{w}}_1, \overline{\boldsymbol{w}}_2, \overline{\boldsymbol{w}}_3)$.

Zovšeobecnený reálny JKT matice $\boldsymbol{A}$ už z toho možno dostať okamžite:

$$\boldsymbol{J}' = \boldsymbol{J}_3\left(\begin{smallmatrix} 2 & -1 \\ 1 & 2 \end{smallmatrix}\right) = \begin{pmatrix} 2 & -1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & -1 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix},$$

rovnako ako maticu prechodu

$$\boldsymbol{Q} = \begin{pmatrix} 3 & -1 & -16/5 & 12/5 & -8/25 & -44/25 \\ 5 & 0 & -5 & 0 & 0 & 0 \\ 5 & 0 & -2 & -1 & -16/5 & 12/5 \\ 5 & 0 & 0 & 0 & -5 & 0 \\ 5 & 0 & 0 & 0 & -2 & -1 \\ 5 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

tvorenú vektormi bázy $(\operatorname{Re}\boldsymbol{w}_1, -\operatorname{Im}\boldsymbol{w}_1, \operatorname{Re}\boldsymbol{w}_2, -\operatorname{Im}\boldsymbol{w}_2, \operatorname{Re}\boldsymbol{w}_3, -\operatorname{Im}\boldsymbol{w}_3)$.

V prípade zložitejšej štruktúry reťazcov prislúchajúcich komplexným vlastným hodnotám sa však už situácia stáva dosť neprehľadnou.

20.4 Rozklad na koreňové podpriestory*

Nech V je vektorový priestor nad poľom K, $\varphi\colon V \to V$ je lineárny operátor a $\lambda \in K$ je jeho vlastná hodnota. Hovoríme, že vektor $\boldsymbol{v} \in V$ je *koreňový vektor* lineárneho operátora φ prislúchajúci k jeho vlastnej hodnote λ, ak existuje prirodzené číslo k také, že $(\varphi - \lambda)^k(\boldsymbol{v}) = \mathbf{0}$. Najmenšie k s touto vlastnosťou nazývame *rádom koreňového vektora* $\boldsymbol{v}$ vzhľadom na φ.

Zafixujme na oba paragrafy 20.4 a 20.5 vektorový priestor V konečnej dimenzie n a lineárnu transformáciu $\varphi\colon V \to V$. Pre $k \in \mathbb{N}$, $\lambda \in \operatorname{Spec}\varphi$ označme

$$\operatorname{Ker}_\lambda^k = \operatorname{Ker}(\varphi - \lambda)^k = \big\{\boldsymbol{v} \in V;\ (\varphi - \lambda)^k(\boldsymbol{v}) = \mathbf{0}\big\}$$

lineárne podpriestory priestoru V. Zrejme platí

$$\{\mathbf{0}\} = \operatorname{Ker}_\lambda^0 \subseteq \operatorname{Ker}_\lambda^1 \subseteq \operatorname{Ker}_\lambda^2 \subseteq \ldots \subseteq \operatorname{Ker}_\lambda^k \subseteq \operatorname{Ker}_\lambda^{k+1} \subseteq \ldots,$$

a $\boldsymbol{v} \in V$ je koreňovým vektorom φ prislúchajúcim k λ práve vtedy, keď $\boldsymbol{v} \in \operatorname{Ker}_\lambda^k$ pre nejaké k. Keďže V je konečnorozmerný, uvedené inklúzie nemôžu byť všetky ostré, lebo inak by dimenzie podpriestorov $\operatorname{Ker}_\lambda^k \subseteq V$ rástli nad všetky medze. Preto existuje najmenšie také $r \in \mathbb{N}$ (závislé na λ), pre ktoré platí

$$\operatorname{Ker}_\lambda^r = \operatorname{Ker}_\lambda^{r+1};$$

potom už $\operatorname{Ker}_\lambda^r = \operatorname{Ker}_\lambda^k$ pre všetky $k \geq r$. Prirodzené číslo $r = r_\lambda$ nazývame *rádom vlastnej hodnoty* λ lineárnej transformácie φ. Lineárny podpriestor $\operatorname{Ker}_\lambda^r \subseteq V$ značíme $\operatorname{Ker}_\lambda$ a nazývame ho *koreňovým podpriestorom* operátora φ prislúchajúcim k vlastnej hodnote λ. Teda $\boldsymbol{v} \in \operatorname{Ker}_\lambda$ práve vtedy, keď $\boldsymbol{v}$ je koreňový vektor φ vzhľadom na λ.

Podobným spôsobom označme

$$\operatorname{Im}_\lambda^k = \operatorname{Im}(\varphi - \lambda)^k = \big\{(\varphi - \lambda)^k(\boldsymbol{u});\ \boldsymbol{u} \in V\big\}$$

členy postupnosti lineárnych podpriestorov

$$V = \mathrm{Im}_\lambda^0 \supseteq \mathrm{Im}_\lambda^1 \supseteq \mathrm{Im}_\lambda^2 \supseteq \ldots \supseteq \mathrm{Im}_\lambda^k \supseteq \mathrm{Im}_\lambda^{k+1} \supseteq \ldots$$

priestoru V. Podľa vety 6.2.3 o dimenzii jadra a obrazu pre každé k platí

$$\dim \mathrm{Ker}_\lambda^k + \dim \mathrm{Im}_\lambda^k = n.$$

Z toho vyplýva, že dimenzie podpriestorov Ker_λ^k prestanú rásť v tej istej chvíli, keď dimenzie podpriestorov Im_λ^k prestanú klesať. Preto $r = r_\lambda$ je taktiež najmenšie prirodzené číslo s vlastnosťou $\mathrm{Im}_\lambda^r = \mathrm{Im}_\lambda^{r+1}$. Položme $\mathrm{Im}_\lambda = \mathrm{Im}_\lambda^r$.

20.4.1. Lema. *Nech λ je vlastná hodnota operátora φ. Potom pre každé $k \in \mathbb{N}$ sú Ker_λ^k aj Im_λ^k invariantné podpriestory a platí*

$$V = \mathrm{Ker}_\lambda \oplus \mathrm{Im}_\lambda,$$

t. j. V je priamym súčtom invariantných podpriestorov Ker_λ a Im_λ.

Dôkaz. Pre $\boldsymbol{x} \in V$ platí

$$\varphi(\boldsymbol{x}) = (\varphi - \lambda)(\boldsymbol{x}) + \lambda\boldsymbol{x}.$$

Pre $\boldsymbol{x} \in \mathrm{Ker}_\lambda^k$ máme $(\varphi - \lambda)^k(\boldsymbol{x}) = \boldsymbol{0}$, a následne

$$(\varphi - \lambda)^k(\varphi\boldsymbol{x}) = (\varphi - \lambda)^{k+1}(\boldsymbol{x}) + \lambda(\varphi - \lambda)^k(\boldsymbol{x}) = \boldsymbol{0},$$

teda $\varphi(\boldsymbol{x}) \in \mathrm{Ker}_\lambda^k$. Podobne, ak $\boldsymbol{x} \in \mathrm{Im}_\lambda^k$, tak $\boldsymbol{x} = (\varphi - \lambda)^k(\boldsymbol{y})$ pre nejaké $\boldsymbol{y} \in V$. Potom

$$\varphi(\boldsymbol{x}) = (\varphi - \lambda)^{k+1}(\boldsymbol{y}) + \lambda(\varphi - \lambda)^k(\boldsymbol{y}) \in \mathrm{Im}_\lambda^k.$$

To dokazuje invariantnosť uvedených podpriestorov, špeciálne invariantnosť podpriestorov $\mathrm{Ker}_\lambda = \mathrm{Ker}_\lambda^r$ a $\mathrm{Im}_\lambda = \mathrm{Im}_\lambda^r$.

Keďže $(\varphi - \lambda)^k(\mathrm{Im}_\lambda) = \mathrm{Im}_\lambda$ pre všetky $k \in \mathbb{N}$, hodnosť lineárneho operátora $(\varphi - \lambda)^r \restriction \mathrm{Im}_\lambda \colon \mathrm{Im}_\lambda \to \mathrm{Im}_\lambda$ sa rovná dimenzii podpriestoru Im_λ. Preto podľa dôsledku 6.2.4 je to injektívny operátor a

$$\mathrm{Ker}_\lambda \cap \mathrm{Im}_\lambda = \mathrm{Ker}\big((\varphi - \lambda)^r \restriction \mathrm{Im}_\lambda\big) = \{\boldsymbol{0}\}.$$

Nakoľko súčet dimenzií oboch podpriestorov je n, z toho plynie $V = \mathrm{Ker}_\lambda \oplus \mathrm{Im}_\lambda$.

20.4.2. Lema. *Nech $\lambda \neq \mu$ sú vlastné hodnoty lineárneho operátora φ. Potom*

(a) Ker_λ je invariantný podpriestor operátora $\varphi - \mu$;

(b) lineárny operátor $(\varphi - \mu) \restriction \mathrm{Ker}_\lambda \colon \mathrm{Ker}_\lambda \to \mathrm{Ker}_\lambda$ *je bijektívny;*

(c) $\mathrm{Ker}_\lambda \subseteq \mathrm{Im}_\mu$.

Dôkaz. (a) Nech $\boldsymbol{u} \in \mathrm{Ker}_\lambda$, t. j. $(\varphi - \lambda)^k(\boldsymbol{u}) = \boldsymbol{0}$ pre nejaké k. Označme $\boldsymbol{v} = (\varphi - \mu)(\boldsymbol{u})$. Potom

$$\begin{aligned}(\varphi - \lambda)^k(\boldsymbol{v}) &= (\varphi - \lambda)^k(\varphi - \lambda + \lambda - \mu)(\boldsymbol{u}) \\ &= (\varphi - \lambda)^{k+1}(\boldsymbol{u}) + (\lambda - \mu)(\varphi - \lambda)^k(\boldsymbol{u}) = \boldsymbol{0},\end{aligned}$$

teda tiež $\boldsymbol{v} \in \mathrm{Ker}_\lambda$.

(b) Nech $\boldsymbol{u} \in \mathrm{Ker}\big((\varphi-\mu) \restriction \mathrm{Ker}_\lambda\big) = \mathrm{Ker}(\varphi-\mu) \cap \mathrm{Ker}_\lambda$. Potom $(\varphi-\mu)(\boldsymbol{u}) = \boldsymbol{0}$ a

$$(\varphi - \lambda)(\boldsymbol{u}) = (\varphi - \mu)(\boldsymbol{u}) + (\mu - \lambda)(\boldsymbol{u}) = (\mu - \lambda)(\boldsymbol{u}),$$

teda $(\varphi - \lambda)^k(\boldsymbol{u}) = (\mu - \lambda)^k(\boldsymbol{u})$ pre každé $k \in \mathbb{N}$. Pre $k = r_\lambda$ dostávame

$$\boldsymbol{0} = (\varphi - \lambda)^k(\boldsymbol{u}) = (\mu - \lambda)(\boldsymbol{u}).$$

Keďže $\lambda \neq \mu$, vyplýva z toho $\boldsymbol{u} = \boldsymbol{0}$. Teda $\mathrm{Ker}\big((\varphi - \mu) \restriction \mathrm{Ker}_\lambda\big) = \{\boldsymbol{0}\}$ a podľa dôsledku 6.2.4 je $(\varphi - \mu) \restriction \mathrm{Ker}_\lambda$ bijektívny lineárny operátor.

(c) Označme $\psi = (\varphi - \mu) \restriction \mathrm{Ker}_\lambda$. Z (b) vyplýva, že pre každé $k \in \mathbb{N}$ je lineárny operátor $\psi^k \colon \mathrm{Ker}_\lambda \to \mathrm{Ker}_\lambda$ bijektívny, teda pre $k = r_\mu$ dostávame

$$\mathrm{Ker}_\lambda = \psi^k(\mathrm{Ker}_\lambda) \subseteq \psi^k(V) = \mathrm{Im}_\mu\,.$$

20.4.3. Tvrdenie. *Nech φ má nad poľom K spektrum* $\mathrm{Spec}\,\varphi = \{\lambda_1, \dots, \lambda_k\}$ *plnej algebraickej váhy n, pričom $\lambda_i \neq \lambda_j$ pre $i \neq j$. Potom*

$$V = \mathrm{Ker}_{\lambda_1} \oplus \ldots \oplus \mathrm{Ker}_{\lambda_k},$$

t. j. V je priamym súčtom koreňových podpriestorov Ker_{λ_i} *prislúchajúcich jednotlivým vlastným hodnotám. Navyše, dimenzia každého koreňového podpriestoru* Ker_{λ_i} *sa rovná algebraickej násobnosti príslušnej vlastnej hodnoty λ_i.*

Dôkaz. Dôkaz prvej časti vykonáme indukciou podľa počtu prvkov spektra k. Pre $k = 1$ je tvrdenie dôsledkom liem 19.1.1 a 20.4.1. Nech teda $k \geq 2$ a predpokladajme, že tvrdenie platí pre lineárne operátory, ktorých spektrum má menej než k prvkov. Podľa lemy 20.4.1 možno V rozložiť na priamy súčet invariantných podpriestorov

$$V = \mathrm{Ker}_{\lambda_k} \oplus \mathrm{Im}_{\lambda_k}\,.$$

Na základe lemy 19.1.1 ľahko nahliadneme rovnosť $\mathrm{Spec}(\varphi \restriction \mathrm{Im}_{\lambda_k}) = \{\lambda_1, \dots, \lambda_{k-1}\}$. Podľa indukčného predpokladu Im_{λ_k} možno rozložiť na priamy súčet koreňových podpriestorov operátora $\varphi \restriction \mathrm{Im}_{\lambda_k}$. Tento rozklad má očividne tvar

$$\mathrm{Im}_{\lambda_k} = (\mathrm{Ker}_{\lambda_1} \cap \mathrm{Im}_{\lambda_k}) \oplus \ldots \oplus (\mathrm{Ker}_{\lambda_{k-1}} \cap \mathrm{Im}_{\lambda_k}) = \mathrm{Ker}_{\lambda_1} \oplus \ldots \oplus \mathrm{Ker}_{\lambda_{k-1}},$$

keďže podľa lemy 20.4.2 (c) pre $i < k$ platí $\mathrm{Ker}_{\lambda_i} \cap \mathrm{Im}_{\lambda_k} = \mathrm{Ker}_{\lambda_i}$. Z toho už dostávame požadovaný rozklad

$$V = \mathrm{Im}_{\lambda_k} \oplus \mathrm{Ker}_{\lambda_k} = \mathrm{Ker}_{\lambda_1} \oplus \ldots \oplus \mathrm{Ker}_{\lambda_{k-1}} \oplus \mathrm{Ker}_{\lambda_k} .$$

Rovnosť algebraickej násobnosti vlastnej hodnoty λ_i a dimenzie koreňového podpriestoru Ker_{λ_i} je dôsledkom uvedeného rozkladu a lemy 19.1.1. Podľa nich má totiž charakteristický polynóm operátora φ tvar

$$\mathrm{ch}_\varphi(x) = \mathrm{ch}_{\varphi_1}(x) \ldots \mathrm{ch}_{\varphi_k}(x),$$

kde $\varphi_i = \varphi \restriction \mathrm{Ker}_{\lambda_i}$. Na druhej strane, keďže $\mathrm{Spec}\,\varphi$ má plnú algebraickú váhu,

$$\mathrm{ch}_\varphi(x) = (\lambda_1 - x)^{m_i} \ldots (\lambda_k - x)^{m_k},$$

kde m_i je algebraická násobnosť λ_i. Nakoľko ch_{φ_i} je mocninou lineárneho faktora $\lambda_i - x$ a hodnoty λ_i sú navzájom rôzne, je $\mathrm{ch}_{\varphi_i}(x) = (\lambda_i - x)^{m_i}$ a $m_i = \dim \mathrm{Ker}_{\lambda_i}$.

20.5 Nilpotentné operátory*

Zúženie $\varphi_\lambda = (\varphi - \lambda) \restriction \mathrm{Ker}_\lambda$ lineárneho operátora $\varphi - \lambda$ na koreňový podpriestor Ker_λ lineárneho operátora φ vyhovuje podmienke $\varphi_\lambda^r = \mathbf{0}$, kde $r = r_\lambda$ je rád príslušnej vlastnej hodnoty λ a $\mathbf{0}\colon \mathrm{Ker}_\lambda \to \mathrm{Ker}_\lambda$ je identicky nulový lineárny operátor.

Hovoríme, že *lineárny operátor* $\varphi\colon V \to V$ je *nilpotentný*, ak pre niektoré prirodzené číslo r platí $\varphi^r = \mathbf{0}$. Najmenšie takéto r, t. j. vlastne rád vlastnej hodnoty 0 operátora φ, nazývame *rádom nilpotentného operátora*.

Celkom analogicky možno definovať aj pojem *nilpotentnej matice*. Príkladom nilpotentných matíc sú Jordanove bunky $\boldsymbol{J}_n = \boldsymbol{J}_n(0)$; platí totiž $\boldsymbol{J}_n^n = \mathbf{0}$. Keďže $\boldsymbol{J}_n^{n-1} \neq \mathbf{0}$, n je priamo rádom nilpotentnej matice $\boldsymbol{J}_n$.

Spektrum nilpotentného operátora $\varphi\colon V \to V$ zrejme pozostáva z jedinej vlastnej hodnoty 0; spektrum posunutého operátora $\varphi + \lambda$, daného predpisom $\boldsymbol{x} \mapsto \varphi(\boldsymbol{x}) + \lambda\boldsymbol{x}$, potom pozostáva z jediného skalára λ. Pritom spektrá oboch operátorov majú plnú algebraickú váhu rovnú $\dim V$.

V predchádzajúcom paragrafe sme vlastne ukázali, že každý lineárny operátor φ na konečnorozmernom vektorovom priestore V s k-prvkovým spektrom $\{\lambda_1, \ldots, \lambda_k\}$ plnej algebraickej váhy $n = \dim V$ možno rozložiť na priamy súčet vhodne posunutých nilpotentných operátorov. Presnejšie, ak $\boldsymbol{x} = \boldsymbol{x}_1 + \ldots + \boldsymbol{x}_k$ je (jednoznačne určený) rozklad vektora $\boldsymbol{x} \in V$ na zložky $\boldsymbol{x}_i \in \mathrm{Ker}_{\lambda_i}$, tak

$$\varphi(\boldsymbol{x}) = \varphi(\boldsymbol{x}_1) + \ldots + \varphi(\boldsymbol{x}_k) = (\varphi_{\lambda_1} + \lambda_1)(\boldsymbol{x}_1) + \ldots + (\varphi_{\lambda_k} + \lambda_1)(\boldsymbol{x}_k)$$

je rozklad vektora $\varphi(\boldsymbol{x})$ na zložky $\varphi(\boldsymbol{x}_i) = (\varphi_{\lambda_i} + \lambda_i)(\boldsymbol{x}_i) \in \mathrm{Ker}_{\lambda_i}$, pričom jednotlivé operátory $\varphi_{\lambda_i}\colon \mathrm{Ker}_{\lambda_i} \to \mathrm{Ker}_{\lambda_i}$ sú nilpotentné.

Na dovŕšenie dôkazu viet 20.1.1 a 20.1.2 o JKT tak v podstate stačí dokázať prvú z nich pre nilpotentné lineárne operátory. Ak sú totiž $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_k$ Jordanove bázy nilpotentných operátorov $\varphi_{\lambda_1}, \dots, \varphi_{\lambda_k}$, t. j. operátor φ_{λ_i} má vzhľadom na bázu $\boldsymbol{\beta}_i$ koreňového podpriestoru Ker_{λ_i} maticu v JKT

$$(\varphi_{\lambda_i})_{\boldsymbol{\beta}_i} = \mathrm{diag}\big(\boldsymbol{J}_{n_{i1}}, \dots, \boldsymbol{J}_{n_{iq_i}}\big),$$

tak φ má vzhľadom na bázu $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_k)$ priestoru V, ktorá vznikne ich spojením, maticu

$$\begin{aligned}(\varphi)_{\boldsymbol{\beta}} &= \mathrm{diag}\big((\varphi_{\lambda_1} + \lambda_1)_{\boldsymbol{\beta}_1}, \dots, (\varphi_{\lambda_k} + \lambda_k)_{\boldsymbol{\beta}_k}\big)\\ &= \mathrm{diag}\big(\boldsymbol{J}_{n_{11}}(\lambda_1), \dots, \boldsymbol{J}_{n_{1q_1}}(\lambda_1), \dots\dots, \boldsymbol{J}_{n_{k1}}(\lambda_k), \dots, \boldsymbol{J}_{n_{kq_k}}(\lambda_k)\big),\end{aligned}$$

a tá je opäť v JKT.

20.5.1. Tvrdenie. *Nech $\varphi\colon V \to V$ je nilpotentný lineárny operátor rádu r. Potom existujú prirodzené čísla $l_1, \dots, l_r$ také, že $\sum_{p=1}^{r} p\, l_p = n$, a vektory $\boldsymbol{u}_{pj} \in V$ rádu p také, že všetkých n vektorov $\varphi^i(\boldsymbol{u}_{pj})$, kde $1 \le p \le r$, $1 \le j \le l_p$, $0 \le i \le p-1$, je lineárne nezávislých, teda dohromady tvoria bázu priestoru V.*

Najprv niekoľko poznámok k zneniu tvrdenia. Pre pevné $1 \le p \le r$, $1 \le j \le l_p$ zakaždým dostávame reťazec

$$\varphi\colon \boldsymbol{u}_{pj} \mapsto \varphi(\boldsymbol{u}_{pj}) \mapsto \dots \mapsto \varphi^{p-2}(\boldsymbol{u}_{pj}) \mapsto \varphi^{p-1}(\boldsymbol{u}_{pj}) \mapsto \mathbf{0}.$$

Ak pre $1 \le i \le p$ položíme $\boldsymbol{v}_{pj}^i = \varphi^{p-i}(\boldsymbol{u}_{pj})$, t. j. očíslujeme vektory sprava doľava, prejde tento reťazec na tvar

$$\varphi\colon \boldsymbol{v}_{pj}^p \mapsto \boldsymbol{v}_{pj}^{p-1} \mapsto \dots \mapsto \boldsymbol{v}_{pj}^2 \mapsto \boldsymbol{v}_{pj}^1 \mapsto \mathbf{0}.$$

Spojením jednotlivých reťazcov $\boldsymbol{\beta}_{pj} = \big(\boldsymbol{v}_{pj}^1, \dots, \boldsymbol{v}_{pj}^p\big)$, dohromady dostaneme Jordanovu bázu

$$\boldsymbol{\beta} = \big(\boldsymbol{\beta}_{r1}, \dots, \boldsymbol{\beta}_{rl_r}, \boldsymbol{\beta}_{r-11}, \dots, \boldsymbol{\beta}_{r-1l_{r-1}}, \dots\dots, \boldsymbol{\beta}_{21}, \dots, \boldsymbol{\beta}_{2l_2}, \boldsymbol{\beta}_{11}, \dots, \boldsymbol{\beta}_{1l_1}\big)$$

operátora φ. Matica φ vzhľadom na ňu má tvar

$$\begin{aligned}(\varphi)_{\boldsymbol{\beta}} &= \mathrm{diag}\big(\boldsymbol{J}_r, \dots, \boldsymbol{J}_r, \boldsymbol{J}_{r-1}, \dots, \boldsymbol{J}_{r-1}, \dots\dots, \boldsymbol{J}_2, \dots, \boldsymbol{J}_2, \boldsymbol{J}_1, \dots, \boldsymbol{J}_1\big)\\ &= \mathrm{diag}\big(\boldsymbol{J}_r, \dots, \boldsymbol{J}_r, \boldsymbol{J}_{r-1}, \dots, \boldsymbol{J}_{r-1}, \dots\dots, \boldsymbol{J}_2, \dots, \boldsymbol{J}_2, 0, \dots, 0\big),\end{aligned}$$

v ktorom sa Jordanov blok $\boldsymbol{J}_p$ vyskytuje práve l_p-krát. Pre $p < r$ pritom nevylučujeme ani možnosť $l_p = 0$; vtedy sa v JKT matice operátora φ Jordanov blok $\boldsymbol{J}_p$ nevyskytuje.

Reťazcovú štruktúru Jordanovej bázy $\boldsymbol{\beta}$ možno vo všeobecnosti znázorniť schémou:

$$\begin{array}{cccccccccccc}
r & & r-1 & & & & 2 & & 1 & & & \\
\bullet & \mapsto & \bullet & \mapsto & \dots & \mapsto & \bullet & \mapsto & \bullet & \mapsto & \mathbf{0} & \\
\vdots & & \vdots & & & & \vdots & & \vdots & & \vdots & l_r \\
\bullet & \mapsto & \bullet & \mapsto & \dots & \mapsto & \bullet & \mapsto & \bullet & \mapsto & \mathbf{0} & \\
 & & \bullet & \mapsto & \dots & \mapsto & \bullet & \mapsto & \bullet & \mapsto & \mathbf{0} & \\
 & & \vdots & & & & \vdots & & \vdots & & \vdots & l_{r-1} \\
 & & \bullet & \mapsto & \dots & \mapsto & \bullet & \mapsto & \bullet & \mapsto & \mathbf{0} & \\
 & & & \ddots & & & \vdots & & \vdots & & \vdots & \\
 & & & & \ddots & & \vdots & & \vdots & & \vdots & \\
 & & & & & \ddots & \vdots & & \vdots & & \vdots & \\
 & & & & & & \bullet & \mapsto & \bullet & \mapsto & \mathbf{0} & \\
 & & & & & & \vdots & & \vdots & & \vdots & l_2 \\
 & & & & & & \bullet & \mapsto & \bullet & \mapsto & \mathbf{0} & \\
 & & & & & & & & \bullet & \mapsto & \mathbf{0} & \\
 & & & & & & & & \vdots & & \vdots & l_1 \\
 & & & & & & & & \bullet & \mapsto & \mathbf{0} &
\end{array}$$

kde znak $\bullet$ označuje jednotlivé bázické vektory a každá skupina je tvorená l_p reťazcami dĺžky p, pre $1 \leq p \leq r$. Čísla r, $r-1$, …, 2, 1 v záhlaví označujú rády vektorov v príslušných stĺpcoch – budeme ich jednoducho nazývať číslami týchto stĺpcov.

Napríklad zo schémy

$$\begin{array}{ccccccccccc}
\boldsymbol{t}_5 & \mapsto & \boldsymbol{t}_4 & \mapsto & \boldsymbol{t}_3 & \mapsto & \boldsymbol{t}_2 & \mapsto & \boldsymbol{t}_1 & \mapsto & \mathbf{0} \\
\boldsymbol{u}_5 & \mapsto & \boldsymbol{u}_4 & \mapsto & \boldsymbol{u}_3 & \mapsto & \boldsymbol{u}_2 & \mapsto & \boldsymbol{u}_1 & \mapsto & \mathbf{0} \\
 & & & & \boldsymbol{v}_3 & \mapsto & \boldsymbol{v}_2 & \mapsto & \boldsymbol{v}_1 & \mapsto & \mathbf{0} \\
 & & & & & & \boldsymbol{w}_2 & \mapsto & \boldsymbol{w}_1 & \mapsto & \mathbf{0} \\
 & & & & & & \boldsymbol{x}_2 & \mapsto & \boldsymbol{x}_1 & \mapsto & \mathbf{0} \\
 & & & & & & \boldsymbol{y}_2 & \mapsto & \boldsymbol{y}_1 & \mapsto & \mathbf{0} \\
 & & & & & & & & \boldsymbol{z}_1 & \mapsto & \mathbf{0}
\end{array}$$

možno vyčítať, že nilpotentný lineárny operátor φ má v Jordanovej báze

$$\boldsymbol{\beta} = (\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3, \boldsymbol{t}_4, \boldsymbol{t}_5, \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4, \boldsymbol{u}_5, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{z}_1)$$

maticu v JKT

$$\begin{aligned}
(\varphi)_{\boldsymbol{\beta}} &= \operatorname{diag}\big(\boldsymbol{J}_5, \boldsymbol{J}_5, \boldsymbol{J}_3, \boldsymbol{J}_2, \boldsymbol{J}_2, \boldsymbol{J}_2, \boldsymbol{J}_1\big) \\
&= \operatorname{diag}\big(\boldsymbol{J}_5, \boldsymbol{J}_5, \boldsymbol{J}_3, \boldsymbol{J}_2, \boldsymbol{J}_2, \boldsymbol{J}_2, 0\big)
\end{aligned}$$

rozmeru 20×20. (Všimnite si, že $l_4 = 0$, teda Jordanov blok $\boldsymbol{J}_4$ sa v matici $(\varphi)_{\boldsymbol{\beta}}$ nevyskytuje.)

Pri dôkaze tvrdenia 20.5.1 budeme potrebovať niekoľko nových pojmov a pomocných výsledkov.

Nech S je lineárny podpriestor vektorového priestoru V. Hovoríme, že *vektory* $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m \in V$ sú *lineárne nezávislé vzhľadom na podpriestor* S, ak pre všetky skaláry $c_1, \dots, c_m \in K$ platí

$$c_1\boldsymbol{x}_1 + \ldots + c_m\boldsymbol{x}_m \in S \;\Rightarrow\; c_1 = \ldots = c_m = 0.$$

Hovoríme, že vektory $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ tvoria *bázu priestoru* V *vzhľadom na podpriestor* S, ak sú lineárne nezávislé vzhľadom na S a $V = S + [\boldsymbol{x}_1, \dots, \boldsymbol{x}_m]$. Zrejme $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m \in V$ sú lineárne nezávislé (tvoria bázu V) práve vtedy, keď sú lineárne nezávislé (tvoria bázu V) vzhľadom na podpriestor $\{\boldsymbol{0}\}$.

Jednoduchý dôkaz nasledujúcej lemy prenechávame ako cvičenie čitateľovi (porovnajte s dôkazmi viet 5.4.1 o dimenzii súčtu a prieniku podpriestorov a 6.2.3 o dimenzii jadra a obrazu).

20.5.2. Lema. *Nech $S \subseteq V$ je lineárny podpriestor a $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m \in V$. Potom nasledujúce podmienky sú ekvivalentné:*

(i) vektory $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ sú lineárne nezávislé vzhľadom na S;

(ii) pre ľubovoľnú bázu $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ podpriestoru S vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ sú lineárne nezávislé;

(iii) existuje báza $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ podpriestoru S taká, že vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ sú lineárne nezávislé.

Na základe lemy 20.5.2 už ľahko odvodíme nasledujúce zovšeobecnenie vety 4.4.4.

20.5.3. Lema. *Nech $S \subseteq V$ je lineárny podpriestor a vektory $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m \in V$ sú lineárne nezávislé vzhľadom na podpriestor S. Potom existujú $\boldsymbol{y}_1, \dots, \boldsymbol{y}_l \in V$ také, že $l = \dim V - \dim S - m$ a vektory $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m, \boldsymbol{y}_1, \dots, \boldsymbol{y}_l$ tvoria bázu priestoru V vzhľadom na S.*

Dôkaz. Stačí zvoliť ľubovoľnú bázu $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ podpriestoru S a lineárne nezávislý systém $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ doplniť vektormi $\boldsymbol{y}_1, \dots, \boldsymbol{y}_l$ na bázu V.

20.5.4. Lema. *Nech $S \subseteq T \subseteq V$ sú lineárne podpriestory také, že $\varphi^{-1}(S) \subseteq T$. Potom pre ľubovoľné vektory $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m \in V$ lineárne nezávislé vzhľadom na T vektory $\varphi(\boldsymbol{x}_1), \dots, \varphi(\boldsymbol{x}_m)$ sú lineárne nezávislé nad S.*

Dôkaz. Nech platí $c_1\varphi(\boldsymbol{x}_1) + \cdots + c_m\varphi(\boldsymbol{x}_m) = \varphi(c_1\boldsymbol{x}_1 + \ldots + c_m\boldsymbol{x}_m) \in S$. Potom $c_1\boldsymbol{x}_1 + \ldots c_m\boldsymbol{x}_m \in \varphi^{-1}(S) \subseteq T$ a z nezávislosti vektorov $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$

nad T vyplýva $c_1 = \ldots = c_m = 0$, teda $\varphi(\boldsymbol{x}_1), \ldots, \varphi(\boldsymbol{x}_m)$ sú nezávislé nad S.

Dôkaz tvrdenia 20.5.1. Operátor φ má jediné vlastné číslo $\lambda = 0$ a V je jeho jediný koreňový podpriestor. Využijúc označenie z predchádzajúceho paragrafu položme $S_p = \operatorname{Ker}_0^p = \operatorname{Ker}\varphi^p$ pre $0 \leq p \leq r$. Dostaneme tak ostro klesajúcu postupnosť lineárnych podpriestorov

$$V = S_r \supset S_{r-1} \supset \ldots \supset S_1 \supset S_0 = \{\mathbf{0}\}$$

usporiadanú inklúziou.

Hľadané vektory $\boldsymbol{u}_{pj}$, zostrojíme rekurziou v r krokoch, v ktorých postupne, zľava doprava, dopĺňame spodnú časť jednotlivých nenulových stĺpcov uvedenej „veľkej schémy“.

Vektory $\boldsymbol{u}_{r1}, \ldots, \boldsymbol{u}_{rl_r} \in V$ vyberieme tak, aby tvorili bázu priestoru $V = S_r$ vzhľadom na podpriestor S_{r-1}.

Nech $r > p \geq 1$. Predpokladajme, že sme už zostrojili všetky stĺpce schémy s číslami $r, r-1, \ldots, p+1$ a vektory každého stĺpca q, kde $r \geq q \geq p+1$, tvoria bázu priestoru S_q vzhľadom na jeho podpriestor S_{q-1}. Označme $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ vektory stĺpca $p+1$. Keďže $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ sú nezávislé nad $S_p = \varphi^{-1}(S_{p-1})$, podľa lemy 20.5.4 sú vektory $\varphi(\boldsymbol{x}_1), \ldots, \varphi(\boldsymbol{x}_m) \in S_p$ nezávislé nad S_{p-1}. Podľa (dôkazu) lemy 20.5.3 ich možno vhodnými vektormi $\boldsymbol{u}_{p1}, \ldots, \boldsymbol{u}_{p\,l_p} \in S_p$ doplniť do bázy podpriestoru S_p vzhľadom na podpriestor S_{p-1}. Stĺpec p je potom tvorený vektormi $\varphi(\boldsymbol{x}_1), \ldots, \varphi(\boldsymbol{x}_m), \boldsymbol{u}_{p1}, \ldots, \boldsymbol{u}_{pl_p}$.

Na dokončenie dôkazu treba ešte overiť, že systém takto zostrojených vektorov $\boldsymbol{v}_{pj}^i = \varphi^{p-i}(\boldsymbol{u}_{pj})$, $1 \leq p \leq r$, $1 \leq j \leq l_p$, $1 \leq i \leq p$, naozaj tvorí bázu priestoru V.

Z konštrukcie vyplýva, že vektory v stĺpci 1 sú lineárne nezávislé. Nech teda $1 \leq p \leq r$ je najväčšie číslo také, že vektory v stĺpcoch $1, \ldots, p$ sú lineárne nezávislé. Stačí ukázať rovnosť $p = r$. V opačnom prípade by vektory v stĺpcoch $1, \ldots, p, p+1$ boli lineárne závislé. To by však znamenalo, že vektory v stĺpci $p+1$ sú lineárne závislé vzhľadom na podpriestor S_p (podrobne si rozmyslite prečo). Ale to odporuje podmienkam našej konštrukcie. Teda všetky vektory „veľkej schémy“ sú lineárne nezávislé.

Konečne ukážeme, že zostrojené vektory generujú celé V. Na to stačí overiť, že ich počet, ktorý je zrejme $\sum_{p=1}^r p\, l_p$, sa naozaj rovná $n = \dim V$. Pri konštrukcii p-teho stĺpca sme k φ-obrazom vektorov predošlého stĺpca vždy pridávali

$$l_p = \dim S_p - \dim S_{p-1} - \sum_{i=p+1}^{r} l_i$$

vektorov. Sčítaním uvedených rovností pre $1 \leq p \leq r$ dostaneme

$$\begin{aligned}\sum_{p=1}^{r} l_p &= \sum_{p=1}^{r}(\dim S_p - \dim S_{p-1}) - \sum_{p=1}^{r}\sum_{i=p+1}^{r} l_i \\ &= \dim S_r - \dim S_0 - \sum_{i=1}^{r}\sum_{p=1}^{i-1} l_i = n - \sum_{i=1}^{r}(i-1)\,l_i,\end{aligned}$$

teda $\sum_{p=1}^{r} p\,l_p = n$.

Dokončenie dôkazu viet 20.1.1 a 20.1.2 si už vyžaduje len ukázať jednoznačnosť Jordanovho kanonického tvaru.

Predovšetkým spektrum $\{\lambda_1, \ldots, \lambda_k\}$ lineárneho operátora φ (vrátane algebraických i geometrických násobností jednotlivých vlastných hodnôt), ako aj jeho koreňové podpriestory Ker_{λ_i} (a tým aj ich dimenzie r_{λ_i}) sú definované bez akéhokoľvek odkazu na maticu φ v nejakej báze. Zúženia $\varphi_{\lambda_i} = (\varphi - \lambda_i) \restriction \mathrm{Ker}_{\lambda_i}$ sú nilpotentné operátory. Takže stačí dokázať jednoznačnosť JKT matice nilpotentného operátora φ, t. j. vlastne počtov l_p jednotlivých reťazcov dĺžok p, $1 \leq p \leq r$. Ako sme však videli, tie sú definované rekurzívne pomocou dimenzií podpriestorov $S_p = \mathrm{Ker}_0^p$, ktoré opäť závisia len na operátore φ.

20.6 Ešte raz úprava na JKT

V tvrdeniach 20.4.3 a 20.5.1 je priamo obsiahnutý opis druhej metódy úpravy matice na JKT, pri ktorej budujeme reťazce Jordanovej bázy „zľava doprava“. Keďže aj lineárny operátor na konečnorozmernom priestore býva väčšinou zadaný maticou v nejakej báze, možno návod očividným spôsobom aplikovať aj na tento prípad. Za predpokladu znalosti plného spektra matice $\boldsymbol{A} \in K^{n\times n}$ postupujeme podľa nasledujúcich bodov:

(1) Pre každé $\lambda \in \operatorname{Spec}\boldsymbol{A}$ osobitne vypočítame mocniny matice $\boldsymbol{A}_\lambda = \boldsymbol{A} - \lambda\boldsymbol{I}$ a každú z nich upravíme na redukovaný stupňovitý tvar. Dostávame tak postupnosť dvojíc matíc

$$\boldsymbol{A}_\lambda \sim \boldsymbol{B}_1,\ \boldsymbol{A}_\lambda^2 \sim \boldsymbol{B}_2,\ \ldots,\ \boldsymbol{A}_\lambda^r \sim \boldsymbol{B}_r,\ \boldsymbol{A}_\lambda^{r+1} \sim \boldsymbol{B}_{r+1},$$

ktorú ukončíme, v prvom kroku $r = r_\lambda$ takom, že $\boldsymbol{B}_r = \boldsymbol{B}_{r+1}$ (čo môže nastať aj keď $\boldsymbol{A}_\lambda^r \neq \boldsymbol{A}_\lambda^{r+1}$). Potom podpriestor $\mathcal{R}(\boldsymbol{A}_\lambda^r) = \mathcal{R}(\boldsymbol{B}_r)$ riešení homogénnej sústavy $\boldsymbol{A}_\lambda^r \cdot \boldsymbol{x} = \boldsymbol{0}$ je koreňový podpriestor matice $\boldsymbol{A}$ prislúchajúci k λ. Špeciálne, v prípade jednobodového spektra je $\boldsymbol{A}_\lambda$ nilpotentná a r je prvý krok, pre ktorý $\boldsymbol{A}_\lambda^r = \boldsymbol{0}$.

(2) Na základe každej z matíc $\boldsymbol{B}_p$, $1 \le p \le r$, nájdeme bázu podpriestoru riešení sústavy $\boldsymbol{A}_\lambda^p \cdot \boldsymbol{x} = \boldsymbol{0}$, ktorú zapíšeme ako stĺpce matice $\boldsymbol{C}_p$.

(3) Vektory $\boldsymbol{u}_{r1}, \dots, \boldsymbol{u}_{rl_r}$ ľavého, t. j. r-tého stĺpca „veľkej schémy" vyberieme zo stĺpcov matice $\boldsymbol{C}_r$ podľa vety 4.4.4 tak, aby dopĺňali bázu podpriestoru $\mathcal{R}\big(\boldsymbol{A}_\lambda^{r-1}\big)$, t. j. stĺpce matice $\boldsymbol{C}_{r-1}$, do bázy podpriestoru $\mathcal{R}(\boldsymbol{A}_\lambda^r)$.

(4) Keď už máme zostrojené stĺpce $r, \dots, p+1$, kde $r > p \ge 1$, pričom stĺpec $p+1$ je tvorený vektormi $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$, tak p-ty stĺpec bude pozostávať z vektorov $\boldsymbol{A}_\lambda \cdot \boldsymbol{x}_1, \dots, \boldsymbol{A}_\lambda \cdot \boldsymbol{x}_m$, $\boldsymbol{u}_{p1}, \dots, \boldsymbol{u}_{p\,l_p}$, pričom vektory $\boldsymbol{u}_{p1}, \dots, \boldsymbol{u}_{p\,l_p}$ získame tak, že bázu $\boldsymbol{C}_{p-1}$ podpriestoru $\mathcal{R}\big(\boldsymbol{A}_\lambda^{p-1}\big)$ rozšírenú o vektory $\boldsymbol{A}_\lambda \cdot \boldsymbol{x}_1, \dots, \boldsymbol{A}_\lambda \cdot \boldsymbol{x}_m$ (ktoré, podľa lemy 20.5.4 tvoria spolu lineárne nezávislý systém) doplníme do bázy podpriestoru $\mathcal{R}(\boldsymbol{A}_\lambda^p)$ vhodnými stĺpcami matice $\boldsymbol{C}_p$ podľa vety 4.4.4.

(5) Vlastnej hodnote λ potom zodpovedá Jordanova báza zložená z takto zostrojených reťazcov a matica v JKT

$$\operatorname{diag}\big(\boldsymbol{J}_r(\lambda), \dots, \boldsymbol{J}_r(\lambda), \dots\dots, \boldsymbol{J}_1(\lambda), \dots, \boldsymbol{J}_1(\lambda)\big)$$

s l_p blokmi $\boldsymbol{J}_p(\lambda)$.

(6) Nakoniec zoradíme vektory Jordanových báz pre jednotlivé vlastné hodnoty pekne za sebou a príslušné JKT blokovo diagonálne. Tým získame výslednú Jordanovu bázu (maticu prechodu) a JKT pôvodnej matice $\boldsymbol{A}$.

Poznámka. 1. S trochou skúsenosti a šikovnosti možno obe metódy úpravy na JKT výhodne kombinovať a budovať reťazce Jordanových báz zároveň zľava i sprava.

2. Nad nekonečným poľom K (najmä v typických prípadoch polí $\mathbb{R}$ a $\mathbb{C}$) možno nové vektory $\boldsymbol{u}_{p1}, \dots, \boldsymbol{u}_{p\,l_p}$ v bodoch (3), (4) voliť v podstate náhodne. Treba si len vopred zistiť ich počet, t. j. $l_r = h(\boldsymbol{B}_{r-1}) - h(\boldsymbol{B}_r)$ a $l_p = h(\boldsymbol{B}_{p-1}) - h(\boldsymbol{B}_p) - m$ pre $r > p \ge 1$. Za predpokladu $l_p > 0$ je totiž nekonečne málo pravdepodobné, že pri skutočne náhodnej voľbe l_p vektorov z podpriestoru $\mathcal{R}\big(\boldsymbol{A}_\lambda^p\big)$ budú vektory $\boldsymbol{A}_\lambda \cdot \boldsymbol{x}_1, \dots, \boldsymbol{A}_\lambda \cdot \boldsymbol{x}_m$, $\boldsymbol{u}_{p1}, \dots, \boldsymbol{u}_{p\,l_p}$ lineárne závislé vzhľadom na podpriestor $\mathcal{R}\big(\boldsymbol{A}_\lambda^{p-1}\big)$. Problémom však môže byť zvoliť vektory $\boldsymbol{u}_{pj}$ „naozaj náhodne". Často práve v snahe o to – spolu s mimovoľnou tendenciou voliť „čo najjednoduchšie" či „čo najkrajšie" vektory – môžeme nakoniec dostať vektory závislé nad $\mathcal{R}\big(\boldsymbol{A}_\lambda^{p-1}\big)$. Podobne, ako je niekedy ťažké nakresliť „naozaj všeobecný" trojuholník.

20.6.1. Príklad. Matica

$$\boldsymbol{A} = \begin{pmatrix} -5 & 3 & 1 & -1 & 6 \\ -3 & 0 & 1 & -5 & 3 \\ -4 & 4 & 1 & 4 & 4 \\ -1 & 1 & 0 & 2 & 1 \\ -6 & 3 & 1 & -1 & 7 \end{pmatrix} \in \mathbb{R}^{5\times 5}$$

má charakteristický polynóm

$$\det(\boldsymbol{A} - x\boldsymbol{I}) = 1 - 5x + 10x^2 - 10x^3 + 5x^4 - x^5 = (1-x)^5$$

a jediné vlastné číslo $x_{1\text{–}5} = 1$ s algebraickou násobnosťou 5.

Postupne vypočítame mocniny matice $\boldsymbol{A} - \boldsymbol{I}$, ich redukované stupňovité tvary a bázy podpriestorov riešení sústav $(\boldsymbol{A} - \boldsymbol{I})^p \cdot \boldsymbol{x} = \boldsymbol{0}$:

$$\boldsymbol{A} - \boldsymbol{I} = \begin{pmatrix} -6 & 3 & 1 & -1 & 6 \\ -3 & -1 & 1 & -5 & 3 \\ -4 & 4 & 0 & 4 & 4 \\ -1 & 1 & 0 & 1 & 1 \\ -6 & 3 & 1 & -1 & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

fundamentálny systém riešení tvoria dva lineárne nezávislé vlastné vektory, ktoré zapíšeme ako stĺpce matice

$$\boldsymbol{C}_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \\ 4 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Druhá mocnina je

$$(\boldsymbol{A} - \boldsymbol{I})^2 = \begin{pmatrix} -12 & 0 & 3 & -12 & 12 \\ 4 & 0 & -1 & 4 & -4 \\ -16 & 0 & 4 & -16 & 16 \\ -4 & 0 & 1 & -4 & 4 \\ -12 & 0 & 3 & -12 & 12 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1/4 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

s fundamentálnym systémom riešení tvoreným stĺpcami matice

$$\boldsymbol{C}_2 = \begin{pmatrix} 0 & 1 & -1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Konečne $(\boldsymbol{A}-\boldsymbol{I})^3=\mathbf{0}$ (takže $\boldsymbol{B}_3=\mathbf{0}$, $\boldsymbol{C}_3=\boldsymbol{I}_5$), a ďalšie mocniny nemusíme počítať.

Keďže máme dva lineárne nezávislé vlastné vektory, Jordanova báza bude mať dva reťazce. Dlhší z nich bude mať dĺžku 3, teda ten kratší musí mať nutne dĺžku $5-3=2$.

Počiatočný vektor dlhšieho reťazca zvolíme tak, aby dopĺňal stĺpce matice $\boldsymbol{C}_2$ na bázu priestoru $\mathbb{R}^5=\mathcal{R}\big((\boldsymbol{A}-\boldsymbol{I})^3\big)$; vyhovuje napr. $\boldsymbol{u}_3=(1,0,0,0,0)^{\mathsf{T}}$. Druhý a tretí vektor reťazca potom sú $\boldsymbol{u}_2=(\boldsymbol{A}-\boldsymbol{I})\cdot\boldsymbol{u}_3=(-6,-3,-4,-1,-6)^{\mathsf{T}}$, $\boldsymbol{u}_1=(\boldsymbol{A}-\boldsymbol{I})\cdot\boldsymbol{u}_2=(\boldsymbol{A}-\boldsymbol{I})^2\cdot\boldsymbol{u}_3=(-12,4,-16,-4,-12)^{\mathsf{T}}$.

Počiatočný vektor kratšieho reťazca vyberieme spomedzi stĺpcov matice $\boldsymbol{C}_2$ tak, aby dopĺňal stĺpce matice $(\boldsymbol{C}_1,\boldsymbol{u}_2)$ na bázu priestoru $\mathcal{R}\big((\boldsymbol{A}-\boldsymbol{I})^2\big)$. Vyhovuje druhý stĺpec $\boldsymbol{v}_2=(1,0,4,0,0)^{\mathsf{T}}$. Druhý vektor kratšieho reťazca potom je $\boldsymbol{v}_1=(\boldsymbol{A}-\boldsymbol{I})\cdot\boldsymbol{v}_2=(-2,1,-4,-1,-2)^{\mathsf{T}}$.

Lineárne nezávislé vektory $\boldsymbol{u}_1,\boldsymbol{v}_1$ tvoria bázu vlastného podpriestoru $\mathcal{R}(\boldsymbol{A}-\boldsymbol{I})$, takže ďalej už nemusíme nič dopĺňať.

Jordanov tvar matice $\boldsymbol{A}$ teda je

$$\boldsymbol{J}=\operatorname{diag}\big(\boldsymbol{J}_3(1),\boldsymbol{J}_2(1)\big)=\begin{pmatrix}1&1&0&0&0\\0&1&1&0&0\\0&0&1&0&0\\0&0&0&1&1\\0&0&0&0&1\end{pmatrix}$$

a stĺpce matice prechodu $\boldsymbol{P}$ tvoria vektory Jordanovej bázy $\boldsymbol{\beta}=(\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3,\boldsymbol{v}_1,\boldsymbol{v}_2)$, t. j.

$$\boldsymbol{P}=\begin{pmatrix}-12&-6&1&-2&1\\4&-3&0&1&0\\-16&-4&0&-4&4\\-4&-1&0&-1&0\\-12&-6&0&-2&0\end{pmatrix}.$$

Cvičenia

20.1. Reálna matica $\boldsymbol{A}=\begin{pmatrix}a&b\\c&d\end{pmatrix}$ má dvojnásobné vlastné číslo práve vtedy, keď $(a-d)^2+4bc=0$ (pozri cvičenie 18.7).

(a) Vysvetlite, v akom zmysle je nekonečne málo pravdepodobné, že náhodne vybraná matica $\boldsymbol{A}\in\mathbb{R}^{n\times n}$ spĺňa túto podmienku.

(b) Ak to však nastane, tak $\boldsymbol{A}$ je diagonalizovateľná práve vtedy, keď má tvar $\boldsymbol{A}=\lambda\boldsymbol{I}_2$. Opäť vysvetlite, v akom zmysle je to nekonečne málo pravdepodobné aj za uvedeného predpokladu.

(c) Urobte podobné úvahy aj pre maticu $\boldsymbol{A} \in \mathbb{C}^{2\times 2}$.

20.2. (a) Matice $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$ sú podobné nad poľom $\mathbb{R}$ práve vtedy, keď sú podobné nad poľom $\mathbb{C}$. Dokážte.

(b) Skúste zovšeobecniť tvrdenie (a) na prípad ľubovoľného poľa K a jeho rozšírenia L. (Ak sa vám to nedarí, budete mať možnosť vrátiť sa k tejto otázke po preštudovaní paragrafu 21.4 v cvičení 21.20.)

20.3. (a) V každom z príkladov 20.2.1–5 doplňte vynechané výpočty.

(b) Nájdite v uvedených príkladoch aj nejaké iné Jordanove bázy (matice prechodu), než sme uviedli pri ich riešení.

(c) V príklade 20.2.4 nájdite celočíselnú maticu prechodu.

(d) Ako sa v uvedených príkladoch zmení výsledný JKT, ak zmeníte poradie niektorých reťazcov v príslušnej Jordanovej báze?

20.4. Metódou z paragrafu 20.2 upravte na Jordanov kanonický tvar nasledujúce reálne matice:

$$\text{(a)}\ \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ -1 & 1 & -1 & 3 \end{pmatrix}, \qquad \text{(b)}\ \begin{pmatrix} 2 & 1 & 0 & 0 \\ -4 & -2 & 1 & 0 \\ -5 & -3 & 3 & 0 \\ -3 & -2 & 1 & 1 \end{pmatrix}, \qquad \text{(c)}\ \begin{pmatrix} 2 & 1 & 0 & 0 \\ -1 & 4 & 0 & 0 \\ 0 & 2 & 2 & 1 \\ 2 & 0 & -1 & 4 \end{pmatrix},$$

$$\text{(d)}\ \begin{pmatrix} -6 & 4 & -2 & 1 & 1 & 0 \\ -12 & 8 & -4 & 1 & 1 & 1 \\ -2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 & -1 \\ -3 & 2 & -1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & -2 & 2 \end{pmatrix}, \qquad \text{(e)}\ \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ -4 & 3 & 0 & 0 & 0 & 0 \\ 3 & -2 & -1 & 1 & 0 & 0 \\ -3 & 1 & -2 & 2 & 1 & 0 \\ -7 & 4 & 2 & -1 & 2 & 0 \\ 1 & -1 & -2 & 1 & 0 & 1 \end{pmatrix}.$$

20.5. Matica $\boldsymbol{A} \in \mathbb{R}^{4\times 4}$ má komplexné vlastné číslo i s algebraickou násobnosťou 2 a geometrickou násobnosťou 1.

(a) Nájdite jej spektrum a určte algebraickú a geometrickú násobnosť zvyšných vlastných čísel.

(b) Nájdite JKT matice $\boldsymbol{A}$ ako aj jej reálny zovšeobecnený JKT.

(c) Predpokladajte, že $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^4$ sú také vektory, že $\boldsymbol{A}(\boldsymbol{u} - \mathrm{i}\boldsymbol{v})$ je vlastný vektor matice $\boldsymbol{A}$ prislúchajúci k jej vlastnému číslu i. Možno už z tejto informácie jednoznačne zrekonštruovať maticu $\boldsymbol{A}$? Ako vyzerá matica $\boldsymbol{A}$, ak $\boldsymbol{u} = (1,0,1,0)^\mathsf{T}$, $\boldsymbol{v} = (1,-1,0,-1)^\mathsf{T}$?

20.6. Dokážte nasledujúce tvrdenia:
(a) Matica $\boldsymbol{A} \in K^{n\times n}$ je nilpotentná práve vtedy, keď $\mathrm{ch}_{\boldsymbol{A}}(x) = (-x)^n$, t. j. práve vtedy, keď $\operatorname{Spec} \boldsymbol{A} = \{0\}$.

(b) Rád vlastnej hodnoty λ lineárneho operátora φ na konečnorozmernom vektorovom priestore sa rovná rozmeru najväčšej Jordanovej bunky $\boldsymbol{J}_m(\lambda)$ v jeho matici v JKT.

20.7. (a) Doplňte vynechané výpočty v príklade 20.6.1.

(b) Upravte maticu z príkladu 20.6.1 na JKT „pôvodnou" metódou z paragrafu 20.2.

(c) Upravte matice z príkladov 20.2.1–5 a z cvičenia 4 na JKT metódou opísanou v paragrafe 20.6.

(d) Pre každú z matíc porovnajte obe metódy jej úpravy na JKT a skúste rozhodnúť, ktorá je výhodnejša.

(e) Ak má matica $\boldsymbol{A} \in K^{n\times n}$ viacprvkové spektrum, tak pri jej úprave na JKT podľa návodu z paragrafu 20.6 sa môže stať, že $\boldsymbol{A}_\lambda^r \neq \boldsymbol{A}_\lambda^{r+1}$, aj keď $\boldsymbol{B}_r = \boldsymbol{B}_{r+1}$. Presvedčte sa o tom pri úprave matíc z príkladu 20.2.1 a cvičenia 20.4 (d).

21. Polynomické invarianty podobnosti matíc

V tejto kapitole najprv prirodzene rozšírime oblasť pôsobnosti polynomických funkcií aj na maticové resp. operátorové argumenty. To nám umožní dokázať kľúčovú vlastnosť charakteristického polynómu, známu pod názvom *Cayleyho-Hamiltonova veta*, a zaviesť ďalší dôležitý polynomický invariant podobnosti matíc – tzv. *minimálny polynóm.*

V druhej polovici sa informatívne oboznámime s dvoma typmi sústav polynómov, z ktorých každá úplným a jednoznačným spôsobom charakterizuje triedu všetkých matíc daného lineárneho operátora resp. triedu všetkých matíc podobných s danou štvorcovou maticou. Na ich základe potom zostrojíme dva nové typy kanonických tvarov matíc lineárnych operátorov, ktorých určenie si – na rozdiel od Jordanovho kanonického tvaru – nevyžaduje znalosť spektra a nevybočuje z pôvodného poľa. Vopred prezraďme aspoň ich názvy: pôjde o tzv. *sústavu elementárnych deliteľov* a *sústavu invariantných faktorov*, ku ktorým prislúcha tzv. *primárny kanonický tvar* resp. *racionálny kanonický tvar matice.*

21.1 Polynomické maticové funkcie

Každý polynóm $f(x) = c_0 + c_1x + \ldots + c_mx^m = \sum_{i=0}^m c_ix^i$ nad poľom K prirodzene definuje (rovnako značenú) funkciu $f\colon K \to K$ danú dosadením hodnoty $a \in K$ za premennú x, t. j. $f(a) = c_0 + c_1a + \ldots + c_ma^m = \sum_{i=0}^m c_ia^i$. Tak isto však môžeme do polynómu $f(x)$ dosadiť za premennú x ľubovoľnú štvorcovú maticu $\boldsymbol{A} \in K^{n\times n}$. Tým dostaneme *polynomickú maticovú funkciu* $f\colon K^{n\times n} \to K^{n\times n}$ danú predpisom

$$f(\boldsymbol{A}) = c_0\boldsymbol{I} + c_1\boldsymbol{A} + \ldots + c_m\boldsymbol{A}^m = \sum_{i=0}^m c_i\boldsymbol{A}^i.$$

Podobne definuje polynóm $f(x)$ funkciu $f\colon \mathcal{L}(V,V) \to \mathcal{L}(V,V)$ na vektorovom priestore $\mathcal{L}(V,V)$ všetkých lineárnych operátorov $\varphi\colon V \to V$. Stačí položiť

$$f(\varphi) = c_0\,\mathrm{id}_V + c_1\varphi + \ldots + c_m\varphi^m = \sum_{i=0}^m c_i\varphi^i,$$

kde φ^i je i-ta iterácia operátora φ, t.j. $\varphi^0 = \mathrm{id}_V$, $\varphi^1 = \varphi$, $\varphi^2 = \varphi\circ\varphi$, atď.

Keďže v centre našej našej pozornosti zostávajú i naďalej konečnorozmerné vektorové priestory, obmedzíme sa na štúdium maticových funkcií, ktoré sú o niečo názornejšie, a uspokojíme sa s poznámkou, že príslušné výsledky možno na operátorové funkcie jednoducho preniesť na základe vzájomne jednoznačnej korešpondencie medzi operátormi $\varphi \in \mathcal{L}(V, V)$ a maticami $\boldsymbol{A} \in K^{n\times n}$ (danej voľbou vhodnej bázy n-rozmerného priestoru V).

Naše štúdium polynomických maticových funkcií začneme niekoľkými jednoduchými pozorovaniami, ktoré navyše plne oprávňujú takýto prístup.

21.1.1. Tvrdenie. *Nech $f(x) \in K[x]$, $\boldsymbol{A}, \boldsymbol{P} \in K^{n\times n}$, pričom $\boldsymbol{P}$ je regulárna. Potom*

$$f\big(\boldsymbol{P}\cdot\boldsymbol{A}\cdot\boldsymbol{P}^{-1}\big) = \boldsymbol{P}\cdot f(\boldsymbol{A})\cdot\boldsymbol{P}^{-1}.$$

Dôkaz. Stačí si uvedomiť, že

$$\big(\boldsymbol{P}\cdot\boldsymbol{A}\cdot\boldsymbol{P}^{-1}\big)^2 = \boldsymbol{P}\cdot\boldsymbol{A}\cdot\boldsymbol{P}^{-1}\cdot\boldsymbol{P}\cdot\boldsymbol{A}\cdot\boldsymbol{P}^{-1} = \boldsymbol{P}\cdot\boldsymbol{A}^2\cdot\boldsymbol{P}^{-1},$$

pričom tento vzťah možno zrejmým spôsobom zovšeobecniť aj na vyššie mocniny. Dokončenie dôkazu už možno prenechať čitateľovi.

21.1.2. Dôsledok. *Nech $f(x) \in K[x]$, $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$. Potom*

$$\boldsymbol{A} \approx \boldsymbol{B} \;\Rightarrow\; f(\boldsymbol{A}) \approx f(\boldsymbol{B}).$$

Taktiež jednoduchý dôkaz nasledujúcej lemy prenechávame ako cvičenie čitateľovi.

21.1.3. Lema. *Nech $f(x) \in K[x]$, $\boldsymbol{A}_1 \in K^{n_1\times n_1}$, …, $\boldsymbol{A}_k \in K^{n_k\times n_k}$. Potom*

$$f\big(\operatorname{diag}(\boldsymbol{A}_1,\dots,\boldsymbol{A}_k)\big) = \operatorname{diag}\big(f(\boldsymbol{A}_1),\dots,f(\boldsymbol{A}_k)\big).$$

Teraz už môžeme uviesť nasledujúci efektný výsledok.

21.1.4. Veta. (Cayley-Hamilton) *Pre ľubovoľnú maticu $\boldsymbol{A} \in K^{n\times n}$ platí*

$$\operatorname{ch}_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{0},$$

t. j. matica $\boldsymbol{A}$ vyhovuje svojej charakteristickej rovnici.

V dôkaze využijeme poznatky o Jordanovom kanonickom tvare z predchádzajúcej kapitoly.

Nech L je konečné rozšírenie poľa K, v ktorom má $\boldsymbol{A}$ spektrum plnej algebraickej váhy n. Označme $\boldsymbol{J} = \operatorname{diag}\big(\boldsymbol{J}_{n_1}(\lambda_1),\dots,\boldsymbol{J}_{n_k}(\lambda_k)\big) \in L^{n\times n}$ Jordanov kanonický tvar matice $\boldsymbol{A}$ a $\boldsymbol{P} \in L^{n\times n}$ regulárnu maticu, ktorej stĺpce tvoria vektory príslušnej Jordanovej bázy. Potom $\boldsymbol{A} = \boldsymbol{P}\cdot\boldsymbol{J}\cdot\boldsymbol{P}^{-1}$, $\operatorname{ch}_{\boldsymbol{A}}(x) = \operatorname{ch}_{\boldsymbol{J}}(x)$ a podľa tvrdenia 21.1.1

$$\operatorname{ch}_{\boldsymbol{A}}(\boldsymbol{A}) = \operatorname{ch}_{\boldsymbol{J}}(\boldsymbol{A}) = \boldsymbol{P}\cdot\operatorname{ch}_{\boldsymbol{J}}(\boldsymbol{J})\cdot\boldsymbol{P}^{-1}.$$

Stačí teda dokázať rovnosť $\mathrm{ch}_{\boldsymbol{J}}(\boldsymbol{J}) = \mathbf{0}$ pre matice v JKT. Z lemy 19.1.1 vyplýva

$$\mathrm{ch}_{\boldsymbol{J}}(x) = \mathrm{ch}_{\boldsymbol{J}_1}(x) \dots \mathrm{ch}_{\boldsymbol{J}_k}(x) = (\lambda_1 - x)^{n_1} \dots (\lambda_k - x)^{n_k},$$

kde $\boldsymbol{J}_i = \boldsymbol{J}_{n_i}(\lambda_i)$ pre $1 \leq i \leq k$. Teda podľa lemy 21.1.3 je $\mathrm{ch}_{\boldsymbol{J}}(\boldsymbol{J})$ blokovo diagonálna matica s diagonálnymi blokmi

$$\mathrm{ch}_{\boldsymbol{J}}(\boldsymbol{J}_i) = (\lambda_1 \boldsymbol{I}_{n_i} - \boldsymbol{J}_i)^{n_1} \dots (\lambda_k \boldsymbol{I}_{n_i} - \boldsymbol{J}_i)^{n_k}$$

pre $1 \leq i \leq k$. V i-tom bloku sa ako i-ty činiteľ nachádza mocnina $(\lambda_i \boldsymbol{I}_{n_i} - \boldsymbol{J}_i)^{n_i} = \mathbf{0}$ nilpotentnej matice $\lambda_i \boldsymbol{I}_{n_i} - \boldsymbol{J}_i = -\boldsymbol{J}_{n_i}(0)$ rádu n_i (pozri začiatok paragrafu 20.5). Preto každý blok i celý výraz $\mathrm{ch}_{\boldsymbol{J}}(\boldsymbol{J})$ sú nevyhnutne nulové matice.

Caylyeho-Hamiltonovu vetu možno využiť na výpočet mocnín štvorcových matíc.

21.1.5. Príklad. Matica

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & -2 \\ -1 & 3 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

má charakteristický polynóm

$$\mathrm{ch}_{\boldsymbol{A}}(x) = \det(\boldsymbol{A} - x\boldsymbol{I}) = 8 - 7x + 3x^2 - x^3.$$

Preto

$$\begin{aligned} \boldsymbol{A}^3 &= 8\boldsymbol{I} - 7\boldsymbol{A} + 3\boldsymbol{A}^2, \\ \boldsymbol{A}^4 &= 8\boldsymbol{A} - 7\boldsymbol{A}^2 + 3\boldsymbol{A}^3 = 8\boldsymbol{A} - 7\boldsymbol{A}^2 + 24\boldsymbol{I} - 21\boldsymbol{A} + 9\boldsymbol{A}^2 \\ &= 24\boldsymbol{I} - 13\boldsymbol{A} + 2\boldsymbol{A}^2, \\ \boldsymbol{A}^5 &= 24\boldsymbol{A} - 13\boldsymbol{A}^2 + 2\boldsymbol{A}^3 = 24\boldsymbol{A} - 13\boldsymbol{A}^2 + 16\boldsymbol{I} - 14\boldsymbol{A} + 6\boldsymbol{A}^2 \\ &= 16\boldsymbol{I} + 10\boldsymbol{A} - 7\boldsymbol{A}^2, \qquad \text{atď.} \end{aligned}$$

Všetky vyššie mocniny matice $\boldsymbol{A}$ teda možno vyjadriť pomocou jej nultej, prvej a druhej mocniny, presnejšie, ako hodnoty vhodných polynómov nanajvýš druhého stupňa pre maticu $\boldsymbol{A}$.

V prípade, že poznáme JKT matice $\boldsymbol{A} \in K^{n\times n}$ a príslušnú maticu prechodu, máme k dispozícii ešte efektívnejší spôsob výpočtu hodnôt $f(\boldsymbol{A})$ pre polynómy $f(x) = \sum_{i=0}^{m} c_i x^i \in K[x]$. Ak $\boldsymbol{J} = \mathrm{diag}\big(\boldsymbol{J}_{n_1}(\lambda_1), \dots, \boldsymbol{J}_{n_k}(\lambda_k)\big)$ je

JKT matice $\boldsymbol{A}$ a $\boldsymbol{P}$ je príslušná matica prechodu, t. j. $\boldsymbol{A} = \boldsymbol{P} \cdot \boldsymbol{J} \cdot \boldsymbol{P}^{-1}$, tak podľa tvrdenia 21.1.1 a lemy 21.1.3 platí

$$f(\boldsymbol{A}) = \boldsymbol{P} \cdot f(\boldsymbol{J}) \cdot \boldsymbol{P}^{-1} = \boldsymbol{P} \cdot \operatorname{diag}\big(f\big(\boldsymbol{J}_{n_1}(\lambda_1)\big), \dots, f\big(\boldsymbol{J}_{n_k}(\lambda_k)\big)\big) \cdot \boldsymbol{P}^{-1}.$$

Preto sa stačí naučiť počítať hodnoty polynómov pre Jordanove bunky $\boldsymbol{J}_n(\lambda)$. Samozrejme, začneme polynómami x^m. Na výpočet mocniny $\boldsymbol{J}_n(\lambda)^m$ sa nám zíde nasledujúca maticová verzia *binomickej vety*, ktorú možno dokázať indukciou rovnako ako binomickú vetu v príslušnom poli. Hovoríme, že *matice* $\boldsymbol{A}$, $\boldsymbol{B} \in K^{n\times n}$ *komutujú*, ak $\boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{B} \cdot \boldsymbol{A}$.

21.1.6. Lema. *Nech matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{n\times n}$ komutujú. Potom pre ľubovoľné $m \in \mathbb{N}$ platí*

$$(\boldsymbol{A} + \boldsymbol{B})^m = \sum_{i=0}^{m} \binom{m}{i} \boldsymbol{A}^{m-i} \boldsymbol{B}^i.$$

Ak si ešte uvedomíme, že $\boldsymbol{J}_n(\lambda) = \lambda \boldsymbol{I}_n + \boldsymbol{J}_n(0) = \lambda \boldsymbol{I}_n + \boldsymbol{J}_n$, pričom $(\lambda \boldsymbol{I}_n) \cdot \boldsymbol{A} = \lambda \boldsymbol{A} = \boldsymbol{A} \cdot (\lambda \boldsymbol{I}_n)$, teda matica $\lambda \boldsymbol{I}_n$ komutuje s každou maticou $\boldsymbol{A} \in \boldsymbol{K}^{n\times n}$, okamžite z toho dostávame

$$\boldsymbol{J}_n(\lambda)^m = \sum_{i=0}^{m} \binom{m}{i} \lambda^{m-i} \boldsymbol{J}_n^i.$$

Mocniny nilpotentej matice $\boldsymbol{J}_n$ už vypočítame ľahko na základe vzťahov

$$\boldsymbol{J}_n^m \cdot \boldsymbol{e}_i = \begin{cases} \boldsymbol{0}, & \text{pre } 1 \le i \le m, \\ \boldsymbol{e}_{i+m}, & \text{pre } m \le i \le n. \end{cases}$$

Takže pre $m \ge n$ je $\boldsymbol{J}_n^m = \boldsymbol{0}$, a pre $0 \le m < n$ je to bloková matica

$$\boldsymbol{J}_n^m = \begin{pmatrix} \boldsymbol{0}_{n-m\,m} & \boldsymbol{I}_{n-m} \\ \boldsymbol{0}_{m\,m} & \boldsymbol{0}_{m\,n-m} \end{pmatrix}.$$

Po dosadení do binomickej vety a sčítaní všetkých členov dostávame hľadaný vzorec.

21.1.7. Lema. *Pre ľubovoľné $m, n \in \mathbb{N}$, $n \ge 1$, $\lambda \in K$ platí*

$$\boldsymbol{J}_n(\lambda)^m = \begin{pmatrix} \lambda^m & \binom{m}{1}\lambda^{m-1} & \dots & \binom{m}{n-1}\lambda^{m-n+1} \\ 0 & \lambda^m & \dots & \binom{m}{n-2}\lambda^{m-n+2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda^m \end{pmatrix},$$

pričom pre $j > m$ definitoricky kladieme $\binom{m}{j}\lambda^{m-j} = 0$.

Pre polynóm $f(x) = \sum_{i=0}^{m} c_i x^i \in K[x]$ definujeme jeho *formálnu p-tu deriváciu* ako polynóm

$$f^{(p)}(x) = \sum_{i=p}^{m} i(i-1)\ldots(i-p+1)c_i x^{i-p} = \sum_{i=0}^{m-p} \frac{(i+p)!}{i!} c_{i+p} x^i;$$

pre $i = 1$ samozrejme píšeme $f^{(1)}(x) = f'(x)$. Pre $K = \mathbb{R}$ formálna derivácia splýva s obvyklou deriváciou $f^{(p)}(x) = \mathrm{d}^p f(x)/\,\mathrm{d}x^p$, definovanou pomocou známej limity.

Všimnime si, že prvky matice $\boldsymbol{J}_n(\lambda)^m$ možno vyjadriť aj pomocou formálnych derivácií polynómu $f(x) = x^m$ v bode $x = \lambda$:

$$\binom{m}{j}\lambda^{m-j} = \frac{1}{j!} f^{(j)}(\lambda).$$

Z toho a lineárnosti operátora derivácie už priamo vyplýva nasledujúci všeobecný vzorec.

21.1.8. Veta. *Nech $f(x) \in K[x]$ je ľubovoľný polynóm, $\lambda \in K$ a $1 \leq n \in \mathbb{N}$. Potom*

$$f\big(\boldsymbol{J}_n(\lambda)\big) = \begin{pmatrix} f(\lambda) & \frac{1}{1!}f'(\lambda) & \ldots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda) \\ 0 & f(\lambda) & \ldots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & f(\lambda) \end{pmatrix}.$$

Keby sa niekomu zdal tento vzorec príliš komplikovaný, snáď ho uteší, ak mu pripomenieme, že „typická“ štvorcová matica nad $\mathbb{R}$ či $\mathbb{C}$ je podobná s diagonálnou. Pre $\boldsymbol{A} = \boldsymbol{P} \cdot \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \cdot \boldsymbol{P}^{-1}$, máme $f(\boldsymbol{A}) = \boldsymbol{P} \cdot \mathrm{diag}\big(f(\lambda_1), \ldots, f(\lambda_n)\big) \cdot \boldsymbol{P}^{-1}$, a basta.

21.2 Minimálny polynóm

Charakteristický polynóm štvorcovej matice $\boldsymbol{A}$ nie je nevyhnutne polynómom najnižšieho možného stupňa, ktorý anuluje maticu $\boldsymbol{A}$. Napríklad charakteristický polynóm matice $\boldsymbol{J} = \mathrm{diag}\big(\boldsymbol{J}_m(\lambda), \boldsymbol{J}_n(\lambda)\big)$ v JKT je zrejme $(\lambda - x)^{m+n}$. Avšak už polynóm $f(x) = (\lambda - x)^{\max(m,n)}$ ju anuluje, t. j. platí $f(\boldsymbol{J}) = \boldsymbol{0}$.

Minimálnym polynómom matice $\boldsymbol{A} \in K^{n \times n}$ nazývame polynóm nad K najnižšieho možného stupňa, ktorý anuluje maticu $\boldsymbol{A}$ a je navyše *normovaný* (t. j. rôzny od 0 a s koeficientom 1 pri najvyššej mocnine x).

Keďže napríklad $(-1)^n \mathrm{ch}_{\boldsymbol{A}}(x)$ je normovaný polynóm anulujúci maticu $\boldsymbol{A} \in K^{n \times n}$, minimálny polynóm matice $\boldsymbol{A}$ určite existuje a má stupeň $\leq$

n. Taktiež ľahko nahliadneme, že je určený jednoznačne. Keby totiž $f(x)$, $g(x)$ boli dva rôzne minimálne polynómy matice $\boldsymbol{A}$, tak by museli mať rovnaký stupeň. Potom by však vhodný skalárny násobok nenulového polynómu $f(x) - g(x)$ bol normovaný polynóm nižšieho stupňa, ktorý tiež anuluje maticu $\boldsymbol{A}$.

To nás oprávňuje zaviesť pre minimálny polynóm matice $\boldsymbol{A}$ označenie $\mu_{\boldsymbol{A}}(x)$. Celkom obdobne možno definovať aj *minimálny polynóm $\mu_\varphi(x)$ lineárneho operátora φ* na konečnorozmernom vektorovom priestore V.

Z tvrdenia 21.1.1 okamžite vyplýva očakávaný výsledok.

21.2.1. Tvrdenie. *Podobné matice majú rovnaký minimálny polynóm. Minimálny polynóm lineárneho operátora na konečnorozmernom vektorovom priestore sa rovná minimálnemu polynómu jeho matice vzhľadom na ľubovoľnú bázu.*

Minimálny polynóm je tak popri charakteristickom polynóme ďalším dôležitým invariantom podobnosti matíc.

21.2.2. Tvrdenie. *Nech $\boldsymbol{A} \in K^{n\times n}$, $f(x) \in K[x]$. Potom $f(\boldsymbol{A}) = \boldsymbol{0}$ práve vtedy, keď minimálny polynóm $\mu_{\boldsymbol{A}}(x)$ delí polynóm $f(x)$. V dôsledku toho minimálny polynóm matice $\boldsymbol{A}$ delí jej charakteristický polynóm.*

Dôkaz. Zrejme stačí dokázať prvú časť tvrdenia, a v tej je netriválna len jedna implikácia. Nech teda $f(\boldsymbol{A}) = \boldsymbol{0}$. Označme $r(x) \in K[x]$ zvyšok po delení polynómu $f(x)$ minimálnym polynómom $\mu_{\boldsymbol{A}}(x)$. Teda stupeň $r(x)$ je menší ako stupeň $\mu_{\boldsymbol{A}}(x)$ a existuje čiastočný podiel $q(x) \in K[x]$ taký, že $f(x) = q(x)\,\mu_{\boldsymbol{A}}(x) + r(x)$. Potom $r(\boldsymbol{A}) = f(\boldsymbol{A}) - q(\boldsymbol{A})\,\mu_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{0}$, takže nevyhnutne $r(x) = 0$, čiže $\mu_{\boldsymbol{A}}(x)$ delí $f(x)$. V opačnom prípade by totiž vhodný skalárny násobok nenulového zvyšku $r(x)$ bol normovaný polynóm nižšieho stupňa ako $\mu_{\boldsymbol{A}}(x)$ anulujúci maticu $\boldsymbol{A}$.

Minimálny polynóm matice možno určiť z jej Jordanovho kanonického tvaru. Na to stačí poznať minimálny polynóm matíc v JKT. Jednoduchú odpoveď na túto otázku dáva nasledujúce tvrdenie. Jedna časť jeho dôkazu je v podstate zahrnutá v dôkaze vety 21.1.4, zvyšok vyplýva z tvrdenia 21.2.2 a z faktu, že polynóm $(x-\lambda)^m$ pre $m < n$ neanuluje Jordanovu bunku $\boldsymbol{J}_n(\lambda)$.

21.2.3. Tvrdenie. *Nech $\boldsymbol{J} = \operatorname{diag}\big(\boldsymbol{J}_{n_1}(\lambda_1), \dots, \boldsymbol{J}_{n_k}(\lambda_k)\big)$ je JKT matice $\boldsymbol{A} \in K^{n\times n}$. Potom minimálny polynóm $\mu_{\boldsymbol{A}}(x) = \mu_{\boldsymbol{J}}(x)$ je súčinom mocnín $(x-\lambda)^{m(\lambda)}$ lineárnych faktorov $x - \lambda$ s exponentmi $m(\lambda) = \max\{n_i;\ \lambda_i = \lambda\}$ pre $\lambda \in \operatorname{Spec}\boldsymbol{A}$.*

Na druhej strane, charakteristický polynóm $\operatorname{ch}_{\boldsymbol{A}}(x) = \operatorname{ch}_{\boldsymbol{J}}(x)$ je v takom prípade súčinom mocnín $(\lambda - x)^{n(\lambda)}$ (až na znamienko) rovnakých lineárnych faktorov s exponentmi $n(\lambda) = \sum_{\lambda_i=\lambda} n_i$.

21.2.4. Dôsledok. *Pre maticu $\boldsymbol{A} \in K^{n\times n}$ platí $\mu_{\boldsymbol{A}}(x) = (-1)^n \operatorname{ch}_{\boldsymbol{A}}(x)$ práve vtedy, keď sa v JKT matice $\boldsymbol{A}$ každá vlastná hodnota $\lambda \in \operatorname{Spec}\boldsymbol{A}$ vyskytuje v práve jednej Jordanovej bunke $\boldsymbol{J}_{n_i}(\lambda_i)$.*

Napr. charakteristický polynóm matice $\boldsymbol{I}_4$ je $(x-1)^4$, kým jej minimálny polynóm je „len“ $x-1$. Na druhej strane, pre diagonálnu maticu $\boldsymbol{A} = \operatorname{diag}(0,1,2,3)$ platí $\operatorname{ch}_{\boldsymbol{A}}(x) = x(x-1)(x-2)(x-3) = \mu_{\boldsymbol{A}}(x)$. Ukážeme si, že každý normovaný polynóm

$$f(x) = x^n - c_1x^{n-1} - \ldots - c_{n-1}x - c_n \in K[x]$$

je minimálnym polynómom vhodnej matice $\boldsymbol{A} \in K^{n\times n}$. Označme

$$\boldsymbol{M}_f = \begin{pmatrix} c_1 & 1 & 0 & \ldots & 0 \\ c_2 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & 0 & 0 & \ldots & 1 \\ c_n & 0 & 0 & \ldots & 0 \end{pmatrix};$$

maticu $\boldsymbol{M}_f$ nazývame *pridruženou maticou polynómu $f(x)$*.

Napríklad pridruženou maticou polynómu x^n je Jordanova bunka $\boldsymbol{J}_n(0)$.

21.2.5. Tvrdenie. *Nech $f(x) \in K^{(n)}[x]$ je normovaný polynóm. Potom $f(x)$ je minimálnym polynómom svojej pridruženej matice $\boldsymbol{M}_f$; jej charakteristický polynóm je $(-1)^n f(x)$.*

Dôkaz. Nech $f(x) = x^n - \sum_{i=1}^n c_i x^{n-i}$. Matica $\boldsymbol{M}_f$ zobrazuje vektory kanonickej bázy $\boldsymbol{\varepsilon}$ v K^n podľa schémy

$$\boldsymbol{e}_n \mapsto \boldsymbol{e}_{n-1} \mapsto \ldots \mapsto \boldsymbol{e}_2 \mapsto \boldsymbol{e}_1 \mapsto (c_1, c_2 \ldots, c_{n-1}, c_n)^{\mathsf{T}},$$

z ktorej vyplývajú rovnosti $\boldsymbol{e}_i = \boldsymbol{M}_f^{n-i} \cdot \boldsymbol{e}_n$, pre $1 \le i \le n$, a tiež

$$\boldsymbol{M}_f^n \cdot \boldsymbol{e}_n = \boldsymbol{M}_f \cdot \boldsymbol{e}_1 = \sum_{i=1}^n c_i \boldsymbol{e}_i = \sum_{i=1}^n c_i \boldsymbol{M}_f^{n-i} \cdot \boldsymbol{e}_n.$$

Keďže matice $\boldsymbol{M}_f$ a $f(\boldsymbol{M}_f)$ zrejme komutujú, platí

$$\begin{aligned} f(\boldsymbol{M}_f) \cdot \boldsymbol{e}_i &= f(\boldsymbol{M}_f) \cdot \boldsymbol{M}_f^{n-i} \cdot \boldsymbol{e}_n = \boldsymbol{M}_f^{n-i} \cdot f(\boldsymbol{M}_f) \cdot \boldsymbol{e}_n \\ &= \boldsymbol{M}_f^{n-i} \cdot \left(\boldsymbol{M}_f^n \cdot \boldsymbol{e}_n - \sum_{i=1}^n c_i \boldsymbol{M}_f^{n-i} \cdot \boldsymbol{e}_n \right) = \boldsymbol{0} \end{aligned}$$

pre každé $i \le n$, v dôsledku čoho $f(\boldsymbol{M}_f) = \boldsymbol{0}$.

Ak $g(x) = \sum_{i=0}^{m} d_i x^i$ je polynóm stupňa $m < n$, tak

$$g(\boldsymbol{M}_f) \cdot \boldsymbol{e}_n = \sum_{i=0}^{m} d_i \boldsymbol{M}_f^i \cdot \boldsymbol{e}_n = d_0 \boldsymbol{e}_n + d_1 \boldsymbol{e}_{n-1} + \ldots + d_m \boldsymbol{e}_{n-m}.$$

Z lineárnej nezávislosti vektorov $\boldsymbol{e}_{n-m}, \ldots, \boldsymbol{e}_n$ však vyplýva, že uvedená lineárna kombinácia sa môže rovnať $\mathbf{0}$, len ak sú všetky koeficienty d_i rovné 0, t. j. keď $g(x) = 0$. Teda $f(x)$ je naozaj minimálny polynóm matice $\boldsymbol{M}_f$.

Kedže charakteristický polynóm matice $\boldsymbol{M}_f$ je stupňa n s koeficientom $(-1)^n$ pri najvyššej mocnine premennej x, a je deliteľný jej minimálnym polynómom $\mu_{\boldsymbol{M}_f}(x) = f(x)$, ktorý je rovnakého stupňa, nevyhnutne platí $\mathrm{ch}_{\boldsymbol{M}_f}(x) = (-1)^n f(x)$.

21.3 Cyklické podpriestory

V tomto paragrafe dáme pre zmenu prednosť reči lineárnych operátorov. Je na čitateľovi, aby si v prípade potreby sformuloval príslušné pojmy a výsledky v jazyku matíc.

Nech $S \subseteq V$ je invariantný podpriestor lineárneho operátora $\varphi\colon V \to V$. Hovoríme, že S je *cyklický podpriestor* operátora φ, ak existuje vektor $\boldsymbol{v} \in S$ taký, že nejaká konečná postupnosť $\boldsymbol{v}, \varphi(\boldsymbol{v}), \ldots, \varphi^{k-1}(\boldsymbol{v})$ jeho iterovaných obrazov tvorí bázu podpriestoru S. Vektor $\boldsymbol{v}$ nazývame *cyklickým generátorom* podpriestoru S a bázu $\big(\boldsymbol{v}, \varphi(\boldsymbol{v}), \ldots, \varphi^{k-1}(\boldsymbol{v})\big)$, prípadne v obrátenom poradí $\big(\varphi^{k-1}(\boldsymbol{v}), \ldots, \varphi(\boldsymbol{v}), \boldsymbol{v}\big)$, nazývame *cyklickou bázou* podpriestoru S (vzhľadom na φ).

Každý nenulový vektor $\boldsymbol{v} \in V$ je cyklickým generátorom práve jedného cyklického podpriestoru operátora φ. Keďže V je konečnorozmerný, v postupnosti obrazov $\boldsymbol{v}, \varphi(\boldsymbol{v}), \varphi^2(\boldsymbol{v}), \ldots$ sa raz musí vyskytnúť člen, ktorý je lineárnou kombináciou predchádzajúcich. Označme k najmenšie prirodzené číslo, pre ktoré to nastane. Potom vektory $\boldsymbol{v}, \varphi(\boldsymbol{v}), \ldots, \varphi^{k-1}(\boldsymbol{v})$ sú lineárne nezávislé a

$$\varphi^k(\boldsymbol{v}) = c_1 \varphi^{k-1}(\boldsymbol{v}) + \ldots + c_{k-1}\varphi(\boldsymbol{v}) + c_k \boldsymbol{v}$$

pre nejaké $c_1, \ldots, c_k \in K$. Zrejme všetky ďalšie obrazy $\varphi^l(\boldsymbol{v})$, $l \geq k$, sú takisto v lineárnom obale $S = \big[\boldsymbol{v}, \varphi(\boldsymbol{v}), \ldots, \varphi^{k-1}(\boldsymbol{v})\big]$, takže S je cyklický podpriestor operátora φ s cyklickým generátorom $\boldsymbol{v}$. Maticou zúženého operátora $\varphi \restriction S$ vzhľadom na cyklickú bázu $\big(\varphi^{k-1}(\boldsymbol{v}), \ldots, \varphi(\boldsymbol{v}), \boldsymbol{v}\big)$ je pridružená matica $\boldsymbol{M}_f$ polynómu $f(x) = x^k - c_1 x^{k-1} - \ldots - c_{k-1}x - c_k$. Tento polynóm nazývame *cyklickým rádom* vektora $\boldsymbol{v}$ (vzhľadom na φ).

Ako uvidíme v nasledujúcej kapitole, v aplikáciách lineárnej algebry na diferenciálne rovnice hrajú významnú úlohu lineárne operátory $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ (teda vlastne matice $\boldsymbol{A} \in \mathbb{R}^{n \times n}$), vzhľadom na ktoré je celý priestor

$\mathbb{R}^n$ cyklický. Tým nadobúda na dôležitosti otázka charakterizácie takýchto operátorov a matíc, ktorá nás vracia späť k dôsledku 21.2.4.

21.3.1. Tvrdenie. *Nech $\varphi\colon V \to V$ je lineárny operátor na n-rozmernom vektorovom priestore V. Potom nasledujúce podmienky sú ekvivalentné:*

(i) *V je cyklickým podpriestorom operátora φ;*

(ii) *maticou φ vo vhodnej báze priestoru V je pridružená matica $\boldsymbol{M}_f$ nejakého normovaného polynómu $f(x)$ stupňa n;*

(iii) *minimálny polynóm $\mu_\varphi(x)$ má stupeň n;*

(iv) *$\mu_\varphi(x) = (-1)^n \operatorname{ch}_\varphi(x)$;*

(v) *v Jordanovej matici operátora φ (nad nejakým konečným rozšírením poľa K) sa každá vlastná hodnota $\lambda \in \operatorname{Spec}\varphi$ vyskytuje v práve jednej Jordanovej bunke.*

Ešte inak môžeme povedať, že vektor $\boldsymbol{v} \in V$ je cyklickým generátorom priestoru V vzhľadom na lineárny operátor φ práve vtedy, keď jeho cyklickým rádom je polynóm $(-1)^n \operatorname{ch}_\varphi(x)$.

21.4 Primárny a racionálny kanonický tvar*

Tento paragraf má prevažne informatívny charakter a v našom kurze naň nebudeme nadväzovať. Obsahuje stručný prehľad niektorých základných pojmov a výsledkov o ďalších kanonických tvaroch matíc lineárnych operátorov a invariantoch podobnosti štvorcových matíc, ktoré uvádzame bez dôkazov. Čitateľ s hlbším záujmom o túto problematiku, môže siahnuť po inej dostupnej literatúre. [1]

Jordanov kanonický tvar zostáva i naďalej našim „hlavným“ kanonickým tvarom. Jeho prednosti sa zakladajú na jednoduchých pravidlách výpočtu mocnín a – v dôsledku toho – aj hodnôt iných funkcií pre Jordanove bunky $\boldsymbol{J}_n(\lambda)$. Popri tom však Jordanov tvar má i dve nezanedbateľné nevýhody. Predovšetkým, na jeho určenie je potrebná znalosť spektra príslušného lineárneho operátora alebo matice, čo môže byť problém najmä pri vyšších dimenziách. Po druhé, JKT je dosiahnuteľný len pre operátory či matice s plnou algebraickou váhou spektra; v prípade, že tomu tak nie je, musíme napr. JKT matice $\boldsymbol{A} \in K^{n\times n}$ hľadať až v nejakom rozšírení L pôvodného poľa K.

V tomto paragrafe sa zoznámime s dvoma typmi kanonických tvarov matíc, ktorých určenie si nevyžaduje znalosť spektra a – ani v prípade jeho neúplnej algebraickej váhy – nevybočuje z pôvodného poľa. Začneme však

[1]Pozri napr. G. Birkhoff, S. Mac Lane, *Prehľad modernej algebry*, ako aj S. Mac Lane, G. Birkhoff, *Algebra*.

ďalším malým prídavkom k našim mimovoľne sa rozrastajúcim vedomostiam o polynómoch.

Hovoríme, že *polynómy* $f(x), g(x) \in K[x]$ stupňa aspoň 1 sú *nesúdeliteľné*, ak okrem konštantných polynómov $0 \neq c \in K$ nemajú v $K[x]$ iné spoločné delitele. Hovoríme, že *polynóm* $p(x) \in K[x]$ stupňa aspoň 1 je *ireducibilný*, ak každý jeho deliteľ je alebo konštantný polynóm alebo skalárny násobok polynómu $p(x)$. Ireducibilné polynómy hrajú v obore $K[x]$ podobnú úlohu ako prvočísla v obore $\mathbb{N}$ alebo $\mathbb{Z}$. Obdobou vety o jednoznačnom rozklade celých čísel na prvočinitele je nasledujúce tvrdenie, ktoré možno pomerne jednoducho dokázať indukciou podľa stupňa polynómu $f(x)$.

21.4.1. Tvrdenie. *Každý polynóm $f(x) \in K[x]$ stupňa aspoň 1 možno rozložiť na súčin tvaru*

$$f(x) = a\, p_1(x)^{\alpha_1} \dots p_k(x)^{\alpha_k}$$

kde $0 \neq a \in K$, $\alpha_1, \dots, \alpha_k$ sú kladné celé čísla a $p_1(x), \dots, p_k(x)$ sú navzájom rôzne normované ireducibilné polynómy nad K. Tento rozklad je jednoznačný až na poradie mocnín $p_i(x)^{\alpha_i}$ jednotlivých ireducibilných faktorov $p_i(x)$.

Teraz už môžeme vysloviť vetu o tzv. *primárnom kanonickom tvare.*

21.4.2. Veta. *Nech $\varphi\colon V \to V$ je lineárny operátor na n-rozmernom vektorovom priestore V, ktorého minimálny polynóm má v $K[x]$ rozklad*

$$\mu_\varphi(x) = p_1(x)^{\alpha_1} \dots p_k(x)^{\alpha_k}$$

na súčin mocnín rôznych normovaných ireducibilných polynómov $p_1(x), \dots, p_k(x)$ stupňov $n_1, \dots, n_k$. Potom existujú postupnosti prirodzených čísel $\alpha_i = \alpha_{i1} \geq \alpha_{i2} \geq \dots \geq \alpha_{ir_i} \geq 1$ také, že $\sum_{i=1}^k n_i \sum_{j=1}^{r_i} \alpha_{ij} = n$, a báza priestoru V, vzhľadom na ktorú má matica lineárneho operátora φ blokovo diagonálny tvar

$$\operatorname{diag}\bigl(\boldsymbol{M}_{p_{11}}, \dots, \boldsymbol{M}_{p_{1r_1}}, \dots\dots, \boldsymbol{M}_{p_{k1}}, \dots, \boldsymbol{M}_{p_{kr_k}}\bigr)$$

pozostávajúci z pridružených matíc polynómov $p_{ij}(x) = p_i(x)^{\alpha_{ij}}$, kde $1 \leq i \leq k$, $1 \leq j \leq r_j$.

Sústava polynómov $p_{ij}(x) = p_i(x)^{\alpha_{ij}}$ $(1 \leq i \leq k,\ 1 \leq j \leq r_j)$ sa nazýva *sústava elementárnych deliteľov* lineárneho operátora φ, a je až na poradie podsústav $p_{i1}(x), \dots, p_{ir_i}(x)$ určená jednoznačne. Uvedený tvar matice operátora φ sa nazýva *primárny kanonický tvar.* I tento tvar je jednoznačne určený až na poradie sústav blokov $\boldsymbol{M}_{p_{i1}}, \dots, \boldsymbol{M}_{p_{ir_i}}$, pridružených k mocninám toho istého ireducibilného polynómu $p_i(x)$. Charakteristický polynóm operátora φ je (až na znamienko) súčinom všetkých elementárnych deliteľov, t. j.

$$\begin{aligned}\operatorname{ch}_\varphi(x) &= (-1)^n p_{11}(x) \dots p_{1r_1}(x) \dots\dots p_{k1}(x) \dots p_{kr_k}(x)\\ &= (-1)^n p_1(x)^{\alpha_{11}+\dots+\alpha_{1r_1}} \dots\dots p_k(x)^{\alpha_{k1}+\dots+\alpha_{kr_k}}\end{aligned}$$

Primárny kanonický tvar matice lineárneho operátora φ zodpovedá rozkladu priestoru V najprv na priamy súčet invariantných podpriestorov $S_i = \operatorname{Ker} p_i^{\alpha_i}(\varphi)$, o ktorom hovorí tzv. *prvá veta o rozklade*. Podľa tzv. *druhej vety o rozklade* sa každý podpriestor S_i ďalej rozpadá na priamy súčet cyklických podpriestorov $S_{i1}, \dots, S_{ir_i}$, ktorých cyklické rády vzhľadom na φ sú polynómy $p_{ij}(x) = p_i(x)^{\alpha_{ij}}$.

S ohľadom na spomínanú jednoznačnosť možno teraz vetu 21.4.2 vysloviť v reči matíc takto:

21.4.3. Veta. *Každá matica $\boldsymbol{A} \in K^{n\times n}$ je podobná s nejakou s maticou $\boldsymbol{B} \in K^{n\times n}$ v primárnom kanonickom tvare. Matica $\boldsymbol{B}$ je určená jednoznačne až na poradie sústav blokov $\boldsymbol{M}_{p_{i1}}, \dots, \boldsymbol{M}_{p_{ir_i}}$, pridružených k mocninám jednotlivých ireducibilných polynómov $p_i(x) \in K[x]$, ktorých najvyššie mocniny tvoria rozklad minimálneho polynómu $\mu_{\boldsymbol{A}}(x) = p_1(x)^{\alpha_1} \dots p_k(x)^{\alpha_k}$ na po dvoch nesúdeliteľné faktory.*

Ak v uvedenej sústave elementárnych deliteľov položíme $r = \max\{r_i;\ 1 \le i \le k\}$, ďalej $\alpha_{ij} = 0$, teda $p_{ij}(x) = p_i(x)^{\alpha_{ij}} = 1$ pre $r_i < j \le r$, tak po prenásobení „po stĺpcoch“, obdržíme sústavu polynómov

$$h_j(x) = p_{1j}(x) \dots p_{kj}(x) = p_1(x)^{\alpha_{1j}} \dots p_k(x)^{\alpha_{kj}},$$

pre $1 \le j \le r$. Prechod od primárneho k tzv. *racionálnemu kanonickému tvaru* sa zakladá na pozorovaní, že priamy súčet cyklických podpriestorov $S_{1j}, \dots, S_{kj}$ s po dvoch nesúdeliteľnými cyklickými rádmi $p_{1j}(x), \dots, p_{kj}(x)$ je opäť cyklický podpriestor s rádom $h_j(x) = p_{1j}(x) \dots p_{kj}(x)$.

21.4.4. Veta. *Nech $\varphi\colon V \to V$ je lineárny operátor na n-rozmernom vektorovom priestore V. Potom existuje jednoznačne určená postupnosť nekonštantných normovaných polynómov $h_1(x) = \mu_\varphi(x)$, $h_2(x), \dots, h_r(x) \in K[x]$, v ktorej každý nasledujúci člen delí predchádzajúci a súčet ich stupňov je n, ako aj báza priestoru V, vzhľadom na ktorú má matica lineárneho operátora φ blokovo diagonálny tvar*

$$\operatorname{diag}\big(\boldsymbol{M}_{h_1}, \boldsymbol{M}_{h_2}, \dots, \boldsymbol{M}_{h_r}\big),$$

pozostávajúci z pridružených matíc jednotlivých polynómov $h_j(x)$ pre $1 \le j \le r$.

Jednotlivé polynómy $h_j(x)$ sa nazývajú *invariantné faktory* lineárneho operátora φ a celá postupnosť $h_1(x) = \mu_\varphi(x)$, $h_2(x), \dots, h_r(x)$ jeho *sústavou invariantných faktorov*. Uvedený tvar matice operátora φ sa nazýva *racionálny kanonický tvar*. Sústava invariantných faktorov, rovnako ako racionálny kanonický tvar matice operátora φ sú určené jednoznačne, dokonca vrátane poradia jednotlivých polynómov resp. k nim pridružených maticových blokov. Rozkladom každého invariantného faktora na mocniny ireducibilných polynómov možno z ich sústavy spätne získať sústavu elementárnych deliteľov. Charakteristický polynóm operátora φ je (až na znamienko) súčinom všetkých jeho invariantných faktorov, t. j.

$$\mathrm{ch}_\varphi(x) = (-1)^n h_1(x) h_2(x) \dots h_r(x).$$

Vetu 21.4.4 možno v jazyku matíc vysloviť v nasledujúcej podobe.

21.4.5. Veta. *Každá matica* $\boldsymbol{A} \in K^{n\times n}$ *je podobná s jednoznačne určenou maticou v racionálnom kanonickom tvare* $\mathrm{diag}\big(\boldsymbol{M}_{h_1}, \boldsymbol{M}_{h_2}, \dots, \boldsymbol{M}_{h_r}\big)$*, pozostávajúcom z matíc pridružených k jej invariantným faktorom* $h_1(x) = \mu_{\boldsymbol{A}}(x)$*,* $h_2(x), \dots, h_r(x)$*.*

Poznámka. Výslovne zdôrazňujeme, že toľkokrát spomínaná jednoznačnosť sa vzťahuje len na primárny resp. racionálny kanonický tvar matice a nie na príslušnú bázu prípadne maticu prechodu. Napokon, rovnako je tomu aj v prípade JKT.

21.4.6. Príklad. Matica $\boldsymbol{A} \in \mathbb{R}^{6\times 6}$ má minimálny polynóm $\mu_{\boldsymbol{A}}(x) = (x^2 + 3)(x-1)^2$. Potom jej charakteristický polynóm je $\mathrm{ch}_{\boldsymbol{A}}(x) = (x^2+3)^2(x-1)^2$ alebo $\mathrm{ch}_{\boldsymbol{A}}(x) = (x^2+3)(x-1)^4$.

V prvom prípade sústavu elementárnych deliteľov matice $\boldsymbol{A}$ tvoria polynómy x^2+3, x^2+3, $(x-1)^2$ a sústavu invariantných faktorov polynómy $(x^2+3)(x-1)^2$, x^2+3.

V druhom prípade sú dve možnosti: Sústava elementárnych deliteľov má tvar x^2+3, $(x-1)^2$, $(x-1)^2$, so sústavou invariantných faktorov $(x^2+3)(x-1)^2$, $(x-1)^2$; alebo sústavu elementárnych deliteľov tvoria polynómy x^2+3, $(x-1)^2$, $x-1$, $x-1$ a sústavu invariantných faktorov polynómy $(x^2+3)(x-1)^2$, $x-1$, $x-1$.

21.5 Výpočet invariantných faktorov*

V tomto paragrafe predvedieme, ako možno k danej matici $\boldsymbol{A} \in K^{n\times n}$ zostrojiť sústavu invariantných faktorov. Z nej už vieme priamo napísať racionálny kanonický tvar (ako aj sústavu elementárnych deliteľov a primárny kanonický tvar) tejto matice. Taktiež si ukážeme, ako dostaneme príslušné matice prechodu. Pritom sa opäť sústredíme len na informáciu o podstatných súvislostiach a výpočtových postupoch, a príslušné dôkazy iba naznačíme alebo celkom vynecháme.

Najprv si uvedomme, že prvkami charakteristickej matice matice $\boldsymbol{A} \in K^{n\times n}$ sú polynómy s koeficientmi z poľa K, čiže $\boldsymbol{A} - x\boldsymbol{I} \in K[x]^{n\times n}$ je vlastne maticou nad množinou $K[x]$ všetkých polynómov v premennej x nad poľom K. Tradične sa v tejto súvislosti miesto premennej x zvykne používať premenná λ a maticiam nad množinou polynómov $K[\lambda]$ sa zvykne hovoriť *λ-matice*. Aj my sa prispôsobíme tejto tradícii. Aby sme vyznačili výskyt premennej λ, budeme λ-matice značiť $\boldsymbol{A}(\lambda)$, $\boldsymbol{B}(\lambda)$ a pod. Každú maticu $\boldsymbol{A}$ nad poľom K možno zároveň považovať za λ-maticu, ktorej prvky sú konštantné polynómy; hovoríme, že $\boldsymbol{A} = \boldsymbol{A}(\lambda)$ je *konštantná λ-matica*. Zrejme λ-matice (pokiaľ to dovolia ich rozmery) možno sčítať i násobiť celkom analogicky ako matice nad poľom K. Podobne ako v prípade charakteristickej matice, determinant štvorcovej λ-matice $\boldsymbol{A}(\lambda)$ je vždy polynóm $\det \boldsymbol{A}(\lambda) \in K[\lambda]$.

Elementárnou riadkovou operáciou, krátko λ-ERO, *na λ-matici* $\boldsymbol{A}(\lambda)$ rozumieme

I. výmenu dvoch riadkov matice $\boldsymbol{A}(\lambda)$;

II. vynásobenie niektorého riadku matice $\boldsymbol{A}(\lambda)$ nenulovým skalárom z poľa K;

III. pripočítanie niektorého riadku matice $\boldsymbol{A}(\lambda)$ vynásobeného ľubovoľným polynómom z $K[\lambda]$ k jej inému riadku.

Celkom analogicky definujeme aj *elementárne stĺpcové operácie na λ-maticiach*, krátko λ-ESO. Dve *λ-matice* $\boldsymbol{A}(\lambda)$, $\boldsymbol{B}(\lambda)$ rovnakých rozmerov sa nazývajú *λ-ekvivalentné*, označenie $\boldsymbol{A}(\lambda) \sim_\lambda \boldsymbol{B}(\lambda)$, ak jednu z nich možno dostať z druhej konečným počtom λ-ERO a λ-ESO. Zrejme pre každé $m, n \in \mathbb{N}$ je vzťah $\sim_\lambda$ naozaj ekvivalenciou na množine λ-matíc $K[\lambda]^{m\times n}$.

λ-matica $\boldsymbol{P}(\lambda) \in K[\lambda]^{n\times n}$ sa nazýva *invertibilná*, ak existuje $\boldsymbol{Q}(\lambda) \in K[\lambda]^{n\times n}$ taká, že $\boldsymbol{P}(\lambda)\cdot\boldsymbol{Q}(\lambda) = \boldsymbol{I}_n = \boldsymbol{Q}(\lambda)\cdot\boldsymbol{P}(\lambda)$. Zrejme λ-matica $\boldsymbol{Q}(\lambda) = \boldsymbol{P}(\lambda)^{-1}$ je k danej invertibilnej $\boldsymbol{P}(\lambda)$ určená jednoznačne už jednou z oboch požadovaných rovností.

Čitateľ by si mal sám premyslieť, ako možno λ-ERO a λ-ESO realizovať pomocou násobenia vhodnými invertibilnými λ-maticami. Na základe toho si už ľahko prispôsobí postup z paragrafu 7.4 na výpočet výrazov tvaru $\boldsymbol{P}(\lambda)^{-1}$, $\boldsymbol{P}(\lambda)^{-1}\cdot\boldsymbol{A}(\lambda)$, $\boldsymbol{A}(\lambda)\cdot\boldsymbol{P}(\lambda)^{-1}$ s invertibilnou λ-maticou $\boldsymbol{P}(\lambda)$.

Kľúčom k výpočtu invariantných faktorov je nasledujúca veta.

21.5.1. Veta. *Matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{n\times n}$ sú podobné práve vtedy, keď ich charakteristické λ-matice $\boldsymbol{A}-\lambda\boldsymbol{I}$, $\boldsymbol{B}-\lambda\boldsymbol{I}$ sú λ-ekvivalentné; presnejšie, ak $\boldsymbol{P}(\lambda), \boldsymbol{Q}(\lambda) \in K[\lambda]^{n\times n}$ sú invertibilné matice také, že $\boldsymbol{B} - \lambda\boldsymbol{I} = \boldsymbol{Q}(\lambda)\cdot(\boldsymbol{A}-\lambda\boldsymbol{I})\cdot\boldsymbol{P}(\lambda)$, tak $\boldsymbol{P}(\lambda) = \boldsymbol{P}$, $\boldsymbol{Q}(\lambda) = \boldsymbol{Q}$ sú konštantné matice, pre ktoré platí $\boldsymbol{Q} = \boldsymbol{P}^{-1}$ a $\boldsymbol{B} = \boldsymbol{Q}\cdot\boldsymbol{A}\cdot\boldsymbol{P}$.*

Náčrt dôkazu. Ak $\boldsymbol{A} \approx \boldsymbol{B}$ a $\boldsymbol{P} \in K^{n\times n}$ je regulárna taká, že $\boldsymbol{B} = \boldsymbol{P}^{-1}\cdot\boldsymbol{A}\cdot\boldsymbol{P}$, tak zrejme $\boldsymbol{B}-\lambda\boldsymbol{I} = \boldsymbol{P}^{-1}\cdot(\boldsymbol{A}-\lambda\boldsymbol{I})\cdot\boldsymbol{P}$, z čoho už vyplýva $\boldsymbol{A}-\lambda\boldsymbol{I} \sim_\lambda \boldsymbol{B}-\lambda\boldsymbol{I}$.

Naopak, pre $\boldsymbol{A}-\lambda\boldsymbol{I} \sim_\lambda \boldsymbol{B}-\lambda\boldsymbol{I}$ existujú invertibilné λ-matice $\boldsymbol{P}(\lambda)$, $\boldsymbol{Q}(\lambda)$ také, že $\boldsymbol{B}-\lambda\boldsymbol{I} = \boldsymbol{Q}(\lambda)\cdot(\boldsymbol{A}-\lambda\boldsymbol{I})\cdot\boldsymbol{P}(\lambda)$. Netriválnym faktom je, že za týchto okolností už $\boldsymbol{P}(\lambda) = \boldsymbol{P}$, $\boldsymbol{Q}(\lambda) = \boldsymbol{Q}$ musia byť konštantné matice, pre ktoré navyše platí $\boldsymbol{Q} \cdot \boldsymbol{P} = \boldsymbol{I}$ a $\boldsymbol{B} = \boldsymbol{Q} \cdot \boldsymbol{A} \cdot \boldsymbol{P}$, teda $\boldsymbol{A} \approx \boldsymbol{B}$.

Hovoríme, že λ-*matica* $\boldsymbol{H}(\lambda) \in K[\lambda]^{n \times n}$ je *v kanonickom tvare*, ak $\boldsymbol{H}(\lambda) = \operatorname{diag}(h_1(\lambda), \ldots, h_n(\lambda))$ je diagonálna matica, ktorej diagonálne prvky sú normované polynómy alebo 0 a každý nasledujúci člen delí predchádzajúci (pritom nulový polynóm delí len nulový polynóm a je deliteľný každým polynómom, teda v kanonickom tvare z $h_j(\lambda) = 0$ vyplýva $h_i(\lambda) = 0$ pre každé $i \le j$).

Polynóm $d(\lambda)$ je *najväčším spoločným deliteľom polynómov* $f_1(\lambda), \ldots, f_k(\lambda)$, ak $d(\lambda)$ delí každý z polynómov $f_i(\lambda)$ a je deliteľný každým polynómom $g(x)$, ktorý delí všetky polynómy $f_i(\lambda)$. Hovoríme, že *polynómy* $f_1(\lambda), \ldots, f_k(\lambda)$ sú *nesúdeliteľné*, ak ich najväčším spločným deliteľom je konštantný polynóm 1. Pripomeňme, že minorom matice $\boldsymbol{A}$ rozumieme determinant ľubovoľnej štvorcovej matice, ktorá vznikne z matice $\boldsymbol{A}$ vynechaním niektorých riadkov a stĺpcov.

21.5.2. Veta. *Každá štvorcová λ-matica $\boldsymbol{A}(\lambda)$ je λ-ekvivalentná s jednoznačne určenou λ-maticou $\boldsymbol{H}(\lambda) = \operatorname{diag}(h_1(\lambda), \ldots, h_n(\lambda))$ v kanonickom tvare. Pre jej diagonálne prvky platí*

$$h_n(\lambda) = d_1(\lambda), \qquad h_{n-j}(\lambda) = \begin{cases} \frac{d_{j+1}(\lambda)}{d_j(\lambda)}, & \text{ak } d_j(\lambda) \neq 0, \\ 0, & \text{ak } d_j(\lambda) = 0, \end{cases}$$

pre $1 \le j \le n-1$, kde $d_k(\lambda)$ je najväčší spoločný deliteľ všetkých minorov rádu k matice $\boldsymbol{A}(\lambda)$, a ak $d_k(\lambda) \neq 0$, tak je to normovaný polynóm.

Poznámka. (a) Polynómy $d_1(\lambda), \ldots, d_n(\lambda)$ sa nazývajú *determinantové delitele λ-matice* $\boldsymbol{A}(\lambda)$. Zrejme $d_n(\lambda)$ sa až na skalárny násobok rovná determinantu $\det \boldsymbol{A}(\lambda)$.

(b) Pokiaľ je $\boldsymbol{H}(\lambda)$ kanonickým tvarom *charakteristickej matice* $\boldsymbol{A} - \lambda \boldsymbol{I}$, tak (z dôvodu zachovania hodnosti pri λ-ERO a λ-ESO) $\boldsymbol{H}(\lambda)$ nemôže obsahovať nulové prvky na diagonále. Preto $\boldsymbol{H}(\lambda) = \operatorname{diag}\big(h_1(\lambda), \ldots, h_r(\lambda), 1, \ldots, 1\big)$ a jej nekonštantné diagonálne prvky tvoria sústavu invariantných faktorov matice $\boldsymbol{A}$.

Miesto dôkazu len opíšeme algoritmus, ako nájsť kanonický tvar štvorcovej λ-matice $\boldsymbol{A}(\lambda) = \big(a_{ij}(\lambda)\big)_{n \times n}$. Jedna možnosť (tá pracnejšia) spočíva vo výpočte všetkých minorov matice $\boldsymbol{A}(\lambda)$ všetkých možných rádov a ich spoločných deliteľov $d_k(\lambda)$. Efektívnejšia je úprava pomocou λ-ERO a λ-ESO. Kľúčovým faktom je pozorovanie, že determinantové delitele $d_k(\lambda)$ sa pri úprave λ-matice $\boldsymbol{A}(\lambda)$ pomocou λ-ERO a λ-ESO nemenia.

Ak $\boldsymbol{A}(\lambda)$ je nulová matica, sme hotoví. V opačnom prípade možno výmenou riadkov a stĺpcov zariadiť, aby $a_{nn}(\lambda) \neq 0$. Teraz postupne dosiahneme, aby $a_{nn}(\lambda)$ delil všetky prvky v n-tom stĺpci i riadku. Ak napr. $a_{in}(\lambda) =$

$a_{nn}(\lambda)\,p(\lambda)+r(\lambda)$, pričom stupeň zvyšku $r(\lambda)\neq 0$ je menší než stupeň $a_{nn}(\lambda)$, odpočítame od i-teho riadku $p(\lambda)$-násobok n-tého riadku a vymeníme i-ty a n-tý riadok. Tým zakaždým znížime stupeň $a_{nn}(\lambda)$. Pretože konštantný nenulový polynóm delí každý polynóm, po konečnom počte krokov bude $a_{nn}(\lambda)$ deliť všetky prvky $a_{in}(\lambda)$, $a_{nj}(\lambda)$; v tej chvíli môžeme pripočítaním vhodných násobkov posledného riadku resp. stĺpca vynulovať všetky prvky v poslednom stĺpci aj riadku okrem $a_{nn}(\lambda)$.

Pokiaľ $a_{nn}(\lambda)$ nedelí všetky prvky v matici, možno opäť znížiť jeho stupeň. Ak totiž $a_{ij}(\lambda) = a_{nn}(\lambda)\,q(\lambda) + s(\lambda)$, kde zvyšok $s(\lambda)\neq 0$ má menší stupeň, pripočítame i-ty riadok k n-tému a opäť postupujeme podľa predchádzajúceho odstavca.

Po konečnom počte krokov dosiahneme, že $a_{nn}(\lambda)$ delí všetky polynómy v matici a je jediným nenulovým prvkom v n-tom riadku i stĺpci. Navyše, vynásobením vhodným skalárom možno dosiahnuť, aby to bol normovaný polynóm. Teraz aplikujeme rovnaký postup na maticu tvorenú prvými $n-1$ riadkami a stĺpcami takto upravenej matice a postupujeme tak dlho, až kým po konečnom počte krokov nedostaneme kanonický tvar.

Samozrejme, v konkrétnych prípadoch sa nemusíme otrocky držať uvedeného postupu, ale môžeme výhodne využívať podobu upravovanej matice.

21.5.3. Príklad. Nájdeme sústavu invariantných faktorov matice

$$\boldsymbol{A} = \begin{pmatrix} -3 & 9 & -8 & 3 \\ -1 & 3 & -2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{4\times 4}$$

úpravou jej charakteristickej matice

$$\boldsymbol{A}-\lambda\boldsymbol{I} = \begin{pmatrix} -3-\lambda & 9 & -8 & 3 \\ -1 & 3-\lambda & -2 & 1 \\ 0 & 0 & 1-\lambda & 0 \\ 0 & 0 & 0 & 1-\lambda \end{pmatrix}$$

na kanonický tvar. V prvom kroku vymeníme druhý a štvrtý riadok, v druhom pomocou jednotky na mieste $(4,4)$ vynulujeme zvyšné prvky štvrtého riadku i stĺpca. $\boldsymbol{A}-\lambda\boldsymbol{I}$ je tak postupne λ-ekvivalentná s λ-maticami

$$\begin{pmatrix} -3-\lambda & 9 & -8 & 3 \\ 0 & 0 & 0 & 1-\lambda \\ 0 & 0 & 1-\lambda & 0 \\ -1 & 3-\lambda & -2 & 1 \end{pmatrix} \sim_\lambda \begin{pmatrix} -\lambda & 3\lambda & -2 & 0 \\ 1-\lambda & (1-\lambda)(\lambda-3) & 2(1-\lambda) & 0 \\ 0 & 0 & 1-\lambda & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Štvrtý riadok a stĺpec sa pri ďalších úpravách už nebudú meniť, preto ich jednoducho „odpílime“. Teraz odpočítame od druhého riadku tretí, potom vymeníme $(-1/2)$-násobok prvého riadku s tretím. V ďalšom kroku pomocou prvku 1 v pozícii $(3,3)$ vynulujeme zvyšné prvky tretieho stĺpca i riadku a vo výsledku vynásobíme prvý aj druhý riadok dvomi. Tým dostaneme

$$\begin{pmatrix} 0 & 0 & 1-\lambda \\ 1-\lambda & (1-\lambda)(\lambda-3) & 1-\lambda \\ \lambda/2 & -3\lambda/2 & 1 \end{pmatrix} \sim_\lambda \begin{pmatrix} \lambda(\lambda-1) & 3\lambda(\lambda-1) & 0 \\ (1-\lambda)(\lambda-2) & (1-\lambda)(5\lambda-6) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Zase odpílime tretí riadok i stĺpec. Teraz od druhého stĺpca odpočítame trojnásobok prvého, potom k druhému riadku pripočítame prvý a výsledok vynásobime $-1/2$. Keďže prvok na mieste $(2,1)$ delí všetky ostatné prvky, môžeme pomocou neho vynulovať oba diagonálne členy. Nakoniec stačí vymeniť prvý a druhý stĺpec a prenásobením riadkov vhodnými skalármi znormovať oba polynómy na diagonále. Postupne tak dostávame

$$\begin{pmatrix} \lambda(\lambda-1) & 0 \\ 1-\lambda & \lambda(\lambda-1) \end{pmatrix} \sim_\lambda \begin{pmatrix} 0 & 2\lambda^2(\lambda-1) \\ 1-\lambda & 0 \end{pmatrix} \sim_\lambda \begin{pmatrix} \lambda^2(\lambda-1) & 0 \\ 0 & \lambda-1 \end{pmatrix}.$$

Po opätovnom pripojení odpílených riadkov a stĺpcov máme

$$\boldsymbol{A} - \lambda \boldsymbol{I} \sim_\lambda \operatorname{diag}\bigl(\lambda^2(\lambda-1), \lambda-1, 1, 1\bigr),$$

pričom posledná λ-matica už je v kanonickom tvare. Nekonštantné polynómy v nej, t. j. $\lambda^2(\lambda-1) = \lambda^3 - \lambda^2$, $\lambda - 1$, tvoria sústavu invariantných faktorov matice $\boldsymbol{A}$. Jej sústava elementárnych deliteľov preto je λ^2, $\lambda-1$, $\lambda-1$. Na základe toho už môžeme napísať racionálny, primárny i Jordanov kanonický tvar matice $\boldsymbol{A}$:

$$\begin{aligned} \boldsymbol{A} &\approx \operatorname{diag}\bigl(\boldsymbol{M}_{\lambda^2(\lambda-1)}, \boldsymbol{M}_{\lambda-1}\bigr) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &\approx \operatorname{diag}\bigl(\boldsymbol{M}_{\lambda^2}, \boldsymbol{M}_{\lambda-1}, \boldsymbol{M}_{\lambda-1}\bigr) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \operatorname{diag}\bigl(\boldsymbol{J}_2(0), 1, 1\bigr). \end{aligned}$$

Všimnite si, že primárny a Jordanov kanonický tvar matice $\boldsymbol{A}$ (náhodou) splývajú.

21.5.4. Príklad. Keby nám v predchádzajúcom príklade napr. vyšlo

$$\boldsymbol{A} - \lambda \boldsymbol{I} \sim_\lambda \operatorname{diag}\bigl((\lambda^2 - \lambda + 3)(\lambda-2)^2, 1, 1, 1\bigr),$$

znamenalo by to, že sústava invariantných faktorov pozostáva z jediného polynómu $(\lambda^2-\lambda+3)(\lambda-2)^2 = \lambda^4-5\lambda^3+11\lambda^2-16\lambda+12$, sústava elementárnych deliteľov z dvoch polynómov $\lambda^2 - \lambda + 3$, $(\lambda-2)^2 = \lambda^2 - 4\lambda + 4$ a racionálny resp. primárny kanonický tvar matice $\boldsymbol{A}$ je

$$\begin{aligned}
\boldsymbol{A} \approx \boldsymbol{M}_{(\lambda^2-\lambda+3)(\lambda-2)^2} &= \begin{pmatrix} 5 & 1 & 0 & 0 \\ -11 & 0 & 1 & 0 \\ 16 & 0 & 0 & 1 \\ -12 & 0 & 0 & 0 \end{pmatrix} \\
\approx \operatorname{diag}\big(\boldsymbol{M}_{\lambda^2-\lambda+3}, \boldsymbol{M}_{(\lambda-2)^2}\big) &= \begin{pmatrix} 1 & 1 & 0 & 0 \\ -3 & 0 & 0 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & -4 & 0 \end{pmatrix}.
\end{aligned}$$

Keďže polynóm $\lambda^2-\lambda+3$ je ireducibilný nad $\mathbb{R}$ (má dva korene $(1\pm \mathrm{i}\sqrt{11})/2 \in \mathbb{C}$), JKT matice $\boldsymbol{A}$ nad $\mathbb{R}$ neexistuje. Jej komplexný resp. zovšeobecnený JKT (pozri paragraf 20.3) by bol

$$\begin{aligned}
\boldsymbol{A} &\approx \operatorname{diag}\big((1+\mathrm{i}\sqrt{11})/2, (1-\mathrm{i}\sqrt{11})/2, \boldsymbol{J}_2(2)\big) \\
&\approx \operatorname{diag}\left(\boldsymbol{J}_1\!\begin{pmatrix} 1/2 & -\sqrt{11}/2 \\ \sqrt{11}/2 & 1/2 \end{pmatrix}, \boldsymbol{J}_2(2)\right) = \begin{pmatrix} 1/2 & -\sqrt{11}/2 & 0 & 0 \\ \sqrt{11}/2 & 1/2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.
\end{aligned}$$

K úprave na kanonický tvar však navyše patrí nájdenie príslušnej bázy resp. matice prechodu. Nech teda $\boldsymbol{A} - \lambda\boldsymbol{I} \sim_\lambda \boldsymbol{H}(\lambda)$, kde λ-matica $\boldsymbol{H}(\lambda)$ je v kanonickom tvare a $\boldsymbol{C} \approx \boldsymbol{A}$ je niektorý z kanonických tvarov (racionálny, primárny či Jordanov) matice $\boldsymbol{A}$, zostrojený na základe $\boldsymbol{H}(\lambda)$. Potom tiež $\boldsymbol{C} - \lambda\boldsymbol{I} \sim_\lambda \boldsymbol{H}(\lambda)$. Preto existujú invertibilné λ-matice $\boldsymbol{P}_1(\lambda)$, $\boldsymbol{Q}_1(\lambda)$, $\boldsymbol{P}_2(\lambda)$, $\boldsymbol{Q}_2(\lambda)$ také, že

$$\boldsymbol{Q}_1(\lambda)\cdot(\boldsymbol{A}-\lambda\boldsymbol{I})\cdot\boldsymbol{P}_1(\lambda) = \boldsymbol{H}(\lambda) = \boldsymbol{Q}_2(\lambda)\cdot(\boldsymbol{C}-\lambda\boldsymbol{I})\cdot\boldsymbol{P}_2(\lambda).$$

V dôsledku toho

$$\boldsymbol{Q}_2(\lambda)^{-1}\cdot\boldsymbol{Q}_1(\lambda)\cdot(\boldsymbol{A}-\lambda\boldsymbol{I})\cdot\boldsymbol{P}_1(\lambda)\cdot\boldsymbol{P}_2(\lambda)^{-1} = \boldsymbol{C}-\lambda\boldsymbol{I},$$

takže podľa vety 21.5.1 sú $\boldsymbol{P} = \boldsymbol{P}_1(\lambda)\cdot\boldsymbol{P}_2(\lambda)^{-1}$, $\boldsymbol{Q} = \boldsymbol{Q}_2(\lambda)^{-1}\cdot\boldsymbol{Q}_1(\lambda)$ konštantné matice a platí $\boldsymbol{Q} = \boldsymbol{P}^{-1}$ a $\boldsymbol{C} = \boldsymbol{P}^{-1}\cdot\boldsymbol{A}\cdot\boldsymbol{P}$. To znamená, že lineárny operátor $\boldsymbol{x} \mapsto \boldsymbol{A}\cdot\boldsymbol{x}$ má v báze priestoru K^n tvorenej stĺpcami matice $\boldsymbol{P}$ maticu v príslušnom kanonickom tvare $\boldsymbol{C}$.

Stačí teda pri úprave $\boldsymbol{A}-\lambda\boldsymbol{I} \longrightarrow \boldsymbol{H}(\lambda)$ pomocou λ-ERO a λ-ESO zaznamenať *len stĺpcové* operácie rovnakými λ-ESO na jednotkovej matici $\boldsymbol{I}$, čím získame $\boldsymbol{P}_1(\lambda)$. Ďalej treba vykonať úpravu $\boldsymbol{C} - \lambda\boldsymbol{I} \longrightarrow \boldsymbol{H}(\lambda)$ (čo vzhľadom na špeciálny tvar matice $\boldsymbol{C}$ býva podstatne jednoduchšie) a opäť zaznamenať *len stĺpcové* operácie zodpovedajúcimi λ-ESO na matici $\boldsymbol{I}$. Tým získame maticu $\boldsymbol{P}_2(\lambda)$. Hľadaná matica prechodu potom je $\boldsymbol{P} = \boldsymbol{P}_1(\lambda) \cdot \boldsymbol{P}_2(\lambda)^{-1}$. Na ostatok si ešte uvedomme, že (po doplnení vynechaných tvrdení a dôkazov) obsahujú paragrafy 21.4 a 21.5 návod na nový dôkaz vety o Jordanovom kanonickom tvare.

Cvičenia

21.1. Dokončite dôkaz tvrdenia 21.1.1, odvoďte z neho dôsledok 21.1.2 a dokážte lemu 21.1.3.

21.2. Bez Cayleyho-Hamiltonovej vety dokážte, že pre danú maticu $\boldsymbol{A} \in K^{n\times n}$ existuje polynóm $f(x) \in K[x]$ stupňa $\leq n^2$ taký, že $f(\boldsymbol{A}) = \boldsymbol{0}$. (*Návod*: Uvažujte postupnosť mocnín $\boldsymbol{I}, \boldsymbol{A}, \boldsymbol{A}^2, \ldots, \boldsymbol{A}^k, \ldots$, matice $\boldsymbol{A}$ a spomeňte si, aká je dimenzia priestoru $K^{n\times n}$.) Uvedomte si, ako výrazne oproti tomu znižuje Cayleyho-Hamiltonova veta stupeň polynómu potrebného na anuláciu matice $\boldsymbol{A}$.

21.3. Vysvetlite, prečo nie je v poriadku nasledujúci „dôkaz" Cayleyho-Hamiltonovej vety: $\operatorname{ch}_{\boldsymbol{A}}(x) = \det(\boldsymbol{A} - x\boldsymbol{I})$, preto $\operatorname{ch}_{\boldsymbol{A}}(\boldsymbol{A}) = \det(\boldsymbol{A} - \boldsymbol{A}\cdot\boldsymbol{I}) = \det \boldsymbol{0} = 0$.

21.4. Vypočítajte maticu $\boldsymbol{A}^2$ z príkladu 21.1.5. Na základe toho dopočítajte matice $\boldsymbol{A}^3$, $\boldsymbol{A}^4$, $\boldsymbol{A}^5$.

21.5. Nech $f(x), g(x) \in K[x]$, $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in K^{n\times n}$. Dokážte nasledujúce tvrdenia:
(a) Matice $f(\boldsymbol{A})$, $g(\boldsymbol{A})$ komutujú

(b) Ak $\boldsymbol{A}$ komutuje s $\boldsymbol{B}$, tak $f(\boldsymbol{A})$ komutuje s $g(\boldsymbol{B})$

(c) Ak $\boldsymbol{A}$ komutuje s $\boldsymbol{B}$ aj s $\boldsymbol{C}$, tak $\boldsymbol{A}$ komutuje s $\boldsymbol{B}\cdot\boldsymbol{C}$.

21.6. (a) Dokážte lemu 21.1.6. Rozmyslite si, prečo je v nej potrebný predpoklad $\boldsymbol{A}\cdot\boldsymbol{B} = \boldsymbol{B}\cdot\boldsymbol{A}$.

(b) Nájdite príklad matíc $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{2\times 2}$, pre ktoré $(\boldsymbol{A}+\boldsymbol{B})^2 \neq \boldsymbol{A}^2 + 2\boldsymbol{A}\cdot\boldsymbol{B} + \boldsymbol{B}^2$.

(c) Nájdite správne vzorce na výpočet mocnín $(\boldsymbol{A}+\boldsymbol{B})^2$, $(\boldsymbol{A}+\boldsymbol{B})^3$ platné pre ľubovoľné (aj nekomutujúce) matice $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$. Zovšeobecnite na ľubovoľný exponent $m \in \mathbb{N}$. Kde sa v nich skrývajú binomické koeficienty $\binom{m}{i}$?

21.7. Nájdite minimálne polynómy matíc z príkladov 20.2.1–5 a cvičenia 20.4.

21.8. (a) Ako vyzerajú minimálny a charakteristický polynóm zovšeobecnenej Jordanovej bunky $\boldsymbol{J}_n\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$?

(b) Nájdite minimálne polynómy matíc z príkladu 20.3.1 a cvičenia 20.5.

21.9. Podrobne dokážte tvrdenie 21.2.3.

21.10. Nech $\boldsymbol{A}$, $\boldsymbol{B}$ sú štvorcové matice rádu n nad poľom K, také, že $\operatorname{ch}_{\boldsymbol{A}}(x) = \operatorname{ch}_{\boldsymbol{B}}(x)$ a $\mu_{\boldsymbol{A}}(x) = \mu_{\boldsymbol{B}}(x)$.
(a) Dokážte, že pre $n \leq 3$ z toho už vyplýva $\boldsymbol{A} \approx \boldsymbol{B}$.

(b) Nájdite príklad takých matíc rádu 4, ktoré napriek uvedeným predpokladom nie sú podobné.

21.11. Priamym výpočtom príslušného determinantu overte, že charakterický polynóm pridruženej matice $\boldsymbol{M}_f$ normovaného polynómu $f(x) \in K[x]$ stupňa n je $(-1)^n f(x)$. (*Návod*: Odvoďte rekurzívny vzorec pomocou vhodného Laplaceovho rozvoja.)

21.12. Nech φ je lineárny operátor na vektorovom priestore V nad poľom K a $\mathbf{0} \neq \boldsymbol{v} \in V$. Dokážte nasledujúce tvrdenia:

(a) $\boldsymbol{v} \in \operatorname{Ker} \varphi$ práve vtedy, keď cyklickým rádom vektora $\boldsymbol{v}$ je polynóm x.

(b) $\boldsymbol{v}$ je vlastný vektor operátora φ prislúchajúci k jeho vlastnej hodnote $\lambda \in K$ práve vtedy, keď jeho cyklickým rádom je polynóm $x - \lambda$.

21.13. Nech $S \subseteq V$ je cyklický podpriestor lineárneho operátora $\varphi\colon V \to V$, ktorého cyklický generátor $\boldsymbol{v}$ má cyklický rád $f(x) = x^k - \sum_{i=1}^{k} c_i x^{k-i}$. Napíšte maticu lineárneho operátora $\varphi \restriction S$ v cyklickej báze $(\boldsymbol{v}, \varphi(\boldsymbol{v}), \ldots \varphi^{k-1}(\boldsymbol{v}))$ podpriestoru S. Niektorí autori uvádzajú pridruženú maticu normovaného polynómu $f(x)$ práve v takomto tvare, prípadne v tvare transponovanej matice k tejto matici.

21.14. Dokážte tvrdenie 21.3.1.

21.15. Overte priamym výpočtom, že pre ľubovoľný skalár λ matice $\left(\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 2\lambda & 1 \\ -\lambda^2 & 0 \end{smallmatrix}\right)$, sú podobné. Nájdite *všetky* matice prechodu (porovnajte s cvičením 19.3 (c)). Zovšeobecnite na ľubovoľné n a matice $\boldsymbol{J}_n(\lambda)$, $\boldsymbol{M}_{(x-\lambda)^n}$.

21.16. Nech $\boldsymbol{A} \in K^{n\times n}$, $f(x), g(x) \in K[x]$ sú nesúdeliteľné polynómy a a $\boldsymbol{u}, \boldsymbol{v} \in K^n$ sú vektory s cyklickými rádmi $f(x)$ resp. $g(x)$ vzhľadom na $\boldsymbol{A}$. Dokážte postupne nasledujúce tvrdenia:

(a) Cyklický rád vektora $\boldsymbol{u} + \boldsymbol{v}$ je polynóm $f(x)\, g(x)$.

(b) Existuje vektor $\boldsymbol{w} \in K^n$, ktorého cyklický rád vzhľadom na $\boldsymbol{A}$ je práve minimálny polynóm $\mu_{\boldsymbol{A}}(x)$. (Riešte najprv pre prípad, že $\mu_{\boldsymbol{A}}(x) = p(x)^\alpha$, kde $p(x)$ je ireducibilný polynóm.)

21.17. (a) Dokážte, že minimálny polynóm matice $\boldsymbol{C} = \operatorname{diag}(\boldsymbol{A}, \boldsymbol{B})$, kde $\boldsymbol{A}$, $\boldsymbol{B}$ sú štvorcové matice nad poľom K, ktorých minimálne polynómy sú nesúdeliteľné, je $\mu_{\boldsymbol{C}}(x) = \mu_{\boldsymbol{A}}(x)\, \mu_{\boldsymbol{B}}(x)$.

(b) Polynóm najnižšieho možného stupňa, ktorý je deliteľný oboma nenulovými polynómami $f(x), g(x) \in K[x]$ sa nazýva ich *najmenší spoločný násobok*. Dokážte, že vo všeobecnosti je minimálny polynóm matice $\operatorname{diag}(\boldsymbol{A}, \boldsymbol{B})$ najmenším spoločným násobkom minimálnych polynómov $\mu_{\boldsymbol{A}}(x)$, $\mu_{\boldsymbol{B}}(x)$.

(c) Nech $f(x), g(x) \in K[x]$ sú nesúdeliteľné normované polynómy stupňov n resp. m. Dokážte, že K^{m+n} je cyklickým podpriestorom matice $\operatorname{diag}(\boldsymbol{M}_f, \boldsymbol{M}_g)$. Nájdite jeho cyklický generátor.

21.18. (a) Sformulujte vetu o prvočíselnom rozklade v obore celých čísel (prípadne si ju vyhľadajte v literatúre) a dokážte ju.

(b) Dokážte tvrdenie 21.4.1.

21.19. Podrobne popíšte, ako možno zo sústavy invariantných faktorov matice $\boldsymbol{A} \in K^{n\times n}$ získať sústavu jej elementárnych deliteľov.

21.20. Nech L je ľubovoľné rozšírenie poľa K. Potom matice $\boldsymbol{A}$, $\boldsymbol{B} \in K^{n\times n}$ sú podobné nad K práve vtedy, keď sú podobné nad L. Dokážte. (Porovnajte svoje súčasné možnosti s možnosťami, ktoré ste mali k dispozícii v cvičení 20.2.)

21.21. (a) Matica $\boldsymbol{A} \in K^{n\times n}$ je podobná s diagonálnou maticou nad poľom K práve vtedy, keď jej mininálny polynóm je súčinom lineárnych faktorov z $K[x]$. Dokážte

(b) Za predpokladu, že všetky elementárne delitele matice $\boldsymbol{A}$ sú mocninami lineárnych faktorov, vyjasnite vzťah medzi jej primárnym a Jordanovým kanonickým tvarom.

21.22. Ku každej možnosti z príkladu 21.4.6 napíšte primárny, racionálny aj Jordanov kanonický tvar matice $\boldsymbol{A}$.

21.23. *Doplňte medzery v dôkaze vety 21.5.1.

21.24. Pre všetky tri kanonické tvary matice $\boldsymbol{A}$ z príkladu 21.5.3 nájdite príslušnú maticu prechodu.

21.25. Determinantové delitele charakteristickej matice tvoria úplný systém invariantov podobnosti danej štvorcovej matice $\boldsymbol{A}$. To znamená, že matice $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$ sú podobné práve vtedy, keď ich charakteristické matice majú rovnaký systém determinantových deliteľov. Dokážte.

21.26. Nájdite sústavy invariantných faktorov a sústavy elementárnych deliteľov ako aj racionálne a primárne kanonické tvary matíc z príkladov 20.2.1–5, 20.3.1, 20.6.1 a cvičení 20.4, 20.5 (c) úpravou ich charakteristických matíc pomocou λ-ERO a λ-ESO. Vysvetlite, ako z nich možno získať JKT pôvodnej matice a príslušnú Jordanovu bázu (maticu prechodu).

22. Maticové funkcie

V tejto kapitole najprv preskúmame funkcie, ktoré možno získať dosadením matíc $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ do potenčných radov s komplexnými koeficientmi, t. j. do funkcií v istom zmysle blízkych polynómom. Teda na rozdiel od predchádzajúcej kapitoly sa obmedzíme len na pole $\mathbb{C}$ všetkých komplexných čísel, pričom reálne matice i funkcie budeme chápať ako špeciálny prípad komplexných. Najdôležitejšou funkciou, ktorú takto získame, bude exponenciála $\mathrm{e}^{\boldsymbol{A}}$ ľubovoľnej matice $\boldsymbol{A} \in \mathbb{C}^{n \times n}$. Potom stručne preskúmame možnosť rozšíriť na maticové argumenty aj iné funkcie komplexnej alebo reálnej premennej.

Ďalej sa budeme zaoberať maticovými funkciami reálnej premennej, pre ktoré odvodíme niektoré analógy základných formúl diferenciálneho a integrálneho počtu. Tie sa nám budú hodiť pri riešení sústav lineárnych diferenciálnych rovníc. Podrobnejšie sa budeme venovať sústavám homogénnych rovníc s konštantnými koeficientmi, ktoré predstavujú vďačnú oblasť aplikácií exponenciály a Jordanovho kanonického tvaru matíc nad $\mathbb{C}$.

U čitateľa predpokladáme základné znalosti z matematickej analýzy. Časť z nich pripomenieme v poznámkach pod čiarou prípadne zhrnieme do ucelených tvrdení, na ktoré sa budeme (bez dôkazov) odvolávať. Poznamenajme, že k deriváciám funkcií komplexnej premennej budeme pristupovať skôr algebraicky než analyticky – v podstate vystačíme s formálnymi deriváciami polynómov a potenčných radov. To nám umožní vedome prehliadnuť istý zásadný rozdiel medzi deriváciami funkcií reálnej a komplexnej premennej.

22.1 Mocninné rady maticovej premennej

Po polynómoch najjednoduchšie reálne či komplexné funkcie sú definované *mocninnými* alebo tiež *potenčnými radmi*, t. j. formálnymi výrazmi tvaru

$$f(x) = \sum_{k=0}^{\infty} c_k x^k = \sum c_k x^k,$$

kde $(c_k)_{k=0}^{\infty}$ je ľubovoľná postupnosť reálnych alebo komplexných čísel. [1] Množinu všetkých potenčných radov v premennej x s koeficientmi z poľa K budeme značiť $K[[x]]$. Zrejme $K[[x]]$ je nekonečnorozmerný vektorový priestor nad poľom K.

[1] V analýze sa uvažujú potenčné rady $\sum_{k=0}^{\infty} c_k (x - w)^k$ so *stredom* v ľubovoľnom bode $w \in \mathbb{C}$. Pre naše účely však celkom postačí, ak sa obmedzíme na potenčné rady so stredom $w = 0$.

Pod *formálnou p-tou deriváciou potenčného radu* $f(x) = \sum_{k=0}^{\infty} \in K[[x]]$ rozumieme potenčný rad

$$f^{(p)}(x) = \sum_{k=p}^{\infty} k(k-1)\dots(k-p+1)c_k x^{k-p} = \sum_{k=0}^{\infty} \frac{(k+p)!}{k!} c_{k+p} x^k \in K[[x]];$$

pre $p = 0$ je $f^{(0)}(x) = f(x)$ a pre $p = 1$ píšeme $f^{(1)}(x) = f'(x)$.

Každý potenčný rad $f(x) = \sum c_k x^k$ nad $\mathbb{R}$ alebo $\mathbb{C}$ definuje predpisom $f(a) = \sum_{k=0}^{\infty} c_k a^k$ (rovnako značenú) funkciu na množine všetkých tých reálnych alebo komplexných čísel a, pre ktoré uvedený rad konverguje. [2] Táto množina, ktorú nazývame *definičný obor* alebo *obor konvergencie potenčného radu* $f(x)$ a značíme $\mathrm{Dom}(f)$ (z anglického *domain*), je (až na hraničné body) charakterizovaná svojim polomerom konvergencie.

Označme $s = \limsup_{k\to\infty} |c_k|^{1/k}$ a $r = s^{-1}$, ak $0 < s < \infty$, resp. $r = \infty$, ak $s = 0$, resp. $r = 0$, ak $s = \infty$. [3] Potom r nazývame *polomerom konvergencie mocninného radu* $\sum c_k x^k$. Pre nás podstatné poznatky matematickej analýzy o mocninných radoch sú zhrnuté v nasledujúcej vete.

22.1.1. Veta. *Nech mocninný rad $f(x) = \sum_{k=0}^{\infty} c_k x^k \in \mathbb{C}[[x]]$ má polomer konvergencie r. Potom pre každé $a \in \mathbb{C}$ platí*

(a) ak $|a| < r$, tak rad $\sum c_k a^k$ absolútne konverguje;

(b) ak $|a| > r$, tak rad $\sum c_k a^k$ diverguje.

Ak navyše $r > 0$, tak funkcia f je spojitá na celom otvorenom kruhu $\{a \in \mathbb{C};\ |a| < r\}$ a má tam spojité derivácie všetkých rádov dané potenčnými radmi $f^{(p)}(x)$, z ktorých každý má rovnaký polomer konvergencie ako pôvodný rad a konverguje rovnomerne na každom uzavretom kruhu $\{a \in \mathbb{C};\ |a| \leq q\}$ pre

[2] Pre pohodlie čitateľa pripomíname, že súčet radu $\sum_{k=0}^{\infty} a_k$ komplexných čísel je definovaný ako limita jeho čiastočných súčtov, t. j. (ak prijmeme dohodu, že súčet radu značíme rovnako ako príslušný rad)

$$\sum_{k=0}^{\infty} a_k = \lim_{m\to\infty} \sum_{k=0}^{m} a_k.$$

Ak táto limita existuje (a je to komplexné číslo), hovoríme, že príslušný rad *konverguje*, v opačnom prípade hovoríme, že rad *diverguje*. Rad $\sum a_k$ *konverguje absolútne*, ak konverguje rad $\sum |a_k|$.

[3] *Limes superior* $\limsup a_k$ *postupnosti reálnych čísel* (a_k) je definované ako supremum množiny všetkých hromadných bodov tejto postupnosti. Pritom $b \in \mathbb{R} \cup \{-\infty, \infty\}$ je *hromadný bod postupnosti* (a_k), ak existuje z nej vybraná podpostupnosť $\left(a_{k_n}\right)$, kde (k_n) je rastúca postupnosť prirodzených čísel, taká, že $\lim_{n\to\infty} a_{k_n} = b$.

$0 < q < r$.[4] *Naopak, pre koeficienty pôvodného potenčného radu platí*

$$c_k = \frac{1}{k!} f^{(k)}(0).$$

Špeciálne, ak $r = \infty$, tak funkcia f je definovaná v celej komplexnej rovine $\mathbb{C}$ a má tam všetky „príjemné" vlastnosti uvedené v druhej časti vety. Ak $r = 0$, tak f je definovaná v jedinom bode $x = 0$. Ak $0 < r < \infty$, tak funkcia f môže no nemusí byť definovaná aj v niektorých bodoch hraničnej kružnice $\{a \in \mathbb{C};\ |a| = r\}$ svojho oboru konvergencie. Platí teda

$$\{a \in \mathbb{C};\ |a| < r\} \subseteq \mathrm{Dom}(f) \subseteq \{a \in \mathbb{C};\ |a| \le r\}.$$

Podrobnejší rozbor hraničných situácií však presahuje rámec nášho kurzu. (Pozri tiež cvičenie 22.3.)

Práve rovnomerná konvergencia vo vnútri definičného oboru potenčného radu zabezpečuje nielen spojitosť jeho súčtu a derivácií, ale tiež umožňuje tento súčet derivovať alebo integrovať formálnym derivovaním resp. integrovaním pôvodného radu člen za členom. Vo vnútri kruhu konvergencie tak nemusíme rozlišovať medzi radom a funkciou definovanou jeho súčtom, a rovnako ani medzi p-tou formálnou deriváciou radu a p-tou deriváciou tejto funkcie.

Formálne možno do mocninného radu $f(x) = \sum_{k=0}^{\infty} c_k x^k$ dosadiť za premennú x nielen reálne či komplexné číslo, ale aj ľubovoľnú štvorcovú maticu $\boldsymbol{A} = (a_{ij})_{n\times n}$ nad $\mathbb{R}$ alebo $\mathbb{C}$. Aby sme však mohli bližšie preskúmať maticové funkcie, ktoré takto dostaneme, musíme si najprv ujasniť niektoré základné otázky týkajúce sa konvergencie postupností a radov komplexných matíc.

Konvergenciu postupnosti matíc definujeme po zložkách. Teda postupnosť matíc $(\boldsymbol{A}_k)_{k=0}^{\infty}$, kde $\boldsymbol{A}_k = (a_{kij}) \in \mathbb{C}^{m\times n}$, konverguje k matici $\boldsymbol{A} = (a_{ij}) \in \mathbb{C}^{m\times n}$, ak pre všetky $i \le m$, $j \le n$ postupnosť $(a_{kij})_{k=0}^{\infty}$ konverguje k prvku a_{ij}, čiže $\lim\limits_{k\to\infty} a_{kij} = a_{ij}$. V takom prípade píšeme

$$\lim_{k\to\infty} \boldsymbol{A}_k = \boldsymbol{A}.$$

V dôsledku spojitosti sčítania a násobenia v poli $\mathbb{C}$ možno pre súčty a súčiny konvergentných maticových postupností $(\boldsymbol{A}_k)$, $(\boldsymbol{B}_k)$ vhodných rozmerov dokázať obdobné vzťahy ako pre číselné postupnosti:

$$\lim_{k\to\infty} (\boldsymbol{A}_k + \boldsymbol{B}_k) = \lim_{k\to\infty} \boldsymbol{A}_k + \lim_{k\to\infty} \boldsymbol{B}_k,$$
$$\lim_{k\to\infty} (\boldsymbol{A}_k \cdot \boldsymbol{B}_k) = \lim_{k\to\infty} \boldsymbol{A}_k \cdot \lim_{k\to\infty} \boldsymbol{B}_k.$$

[4]Funkcionálny rad $\sum_{k=0}^{\infty} g_k(x)$ *konverguje rovnomerne na množine* $X \subseteq \mathbb{C}$ k funkcii $g(x)$, ak $(\forall\, \varepsilon > 0)(\exists\, m_0)(\forall\, m \ge m_0)(\forall\, x \in X)\big(\big|\sum_{k=0}^{m} g_k(x) - g(x)\big| < \varepsilon\big)$. Vo všeobecnosti súčet konvergentného funkcionálneho radu funkcií spojitých na množine X nemusí byť spojitá funkcia na X. Ak však rad konverguje k svojmu súčtu na tejto množine *rovnomerne*, tak aj jeho súčet je spojitá funkcia na X.

Tieto pravidlá nám umožňujú počítať s limitami maticových postupností do značnej miery podobne ako s limitami číselných postupností.

Súčet maticového radu $\sum_{k=0}^{\infty} \boldsymbol{A}_k$, kde $(\boldsymbol{A}_k)_{k=0}^{\infty}$ je nejaká postupnosť komplexných matíc rovnakého rozmeru $m \times n$, potom definujeme ako limitu postupnosti jeho čiastočných súčtov, t. j.

$$\sum_{k=0}^{\infty} \boldsymbol{A}_k = \lim_{p\to\infty} \sum_{k=0}^{p} \boldsymbol{A}_k,$$

samozrejme, ak uvedená limita (t. j. matica $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ príslušných vlastností) existuje. V takom prípade hovoríme, že *maticový rad* $\sum \boldsymbol{A}_k$ *konverguje*, v opačnom prípade hovoríme, že *diverguje*.

Dosadením konkrétnej matice $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ do mocninného radu $f(x) = \sum_{k=0}^{\infty} c_k x^k$ tak dostaneme maticový rad $\sum_{k=0}^{\infty} c_k \boldsymbol{A}^k$, ktorého súčet je definovaný ako limita

$$f(\boldsymbol{A}) = \sum_{k=0}^{\infty} c_k \boldsymbol{A}^k = \lim_{m\to\infty} \sum_{k=0}^{m} c_k \boldsymbol{A}^k$$

postupnosti hodnôt $f_m(\boldsymbol{A})$ polynómov $f_m(x) = \sum_{k=0}^{m} c_k x^k$ pre maticu $\boldsymbol{A}$. Definičným oborom takejto funkcie f je množina tých matíc $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, pre ktoré uvedený rad, t. j. postupnosť $\big(f_m(\boldsymbol{A})\big)$, konverguje. Túto množinu (pri pevnom n) budeme značiť $\mathrm{Dom}_n(f)$.

Limitným prechodom pre $m \to \infty$ dostávame z výsledkov 21.1.1–3 obdobné tvrdenia aj pre maticové potenčné rady.

22.1.2. Tvrdenie. *Nech $f(x) \in \mathbb{C}[[x]]$, $\boldsymbol{A}$, $\boldsymbol{P} \in \mathbb{C}^{n\times n}$, pričom $\boldsymbol{P}$ je regulárna. Potom $\boldsymbol{A} \in \mathrm{Dom}_n(f)$ práve vtedy, keď $\boldsymbol{P}\cdot\boldsymbol{A}\cdot\boldsymbol{P}^{-1} \in \mathrm{Dom}_n(f)$, a v tom prípade platí*

$$f\big(\boldsymbol{P}\cdot\boldsymbol{A}\cdot\boldsymbol{P}^{-1}\big) = \boldsymbol{P}\cdot f(\boldsymbol{A})\cdot\boldsymbol{P}^{-1}.$$

22.1.3. Dôsledok. *Nech $f(x) \in \mathbb{C}[[x]]$, $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$, pričom $\boldsymbol{A} \approx \boldsymbol{B}$. Potom $\boldsymbol{A} \in \mathrm{Dom}_n(f)$ práve vtedy, keď $\boldsymbol{B} \in \mathrm{Dom}_n(f)$, a v tom prípade $f(\boldsymbol{A}) \approx f(\boldsymbol{B})$.*

22.1.4. Lema. *Nech $f(x) \in \mathbb{C}[[x]]$, $\boldsymbol{A}_1 \in \mathbb{C}^{n_1\times n_1}, \ldots, \boldsymbol{A}_p \in \mathbb{C}^{n_p\times n_p}$. Potom $\mathrm{diag}(\boldsymbol{A}_1,\ldots,\boldsymbol{A}_p) \in \mathrm{Dom}_n(f)$ práve vtedy, keď $\boldsymbol{A}_j \in \mathrm{Dom}_{n_j}(f)$ pre každé $j \le p$, a v tom prípade platí*

$$f\big(\mathrm{diag}(\boldsymbol{A}_1,\ldots,\boldsymbol{A}_p)\big) = \mathrm{diag}\big(f(\boldsymbol{A}_1),\ldots,f(\boldsymbol{A}_p)\big).$$

Ľubovoľný polynóm, ktorý anuluje maticu $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, nám umožňuje zjednodušiť výpočet súčtu potenčného radu $f(\boldsymbol{A}) = \sum c_k \boldsymbol{A}^k$, a to tým väčšmi, čím nižší je jeho stupeň m. Pomocou neho možno totiž všetky mocniny $\boldsymbol{A}^k$,

$k \geq m$, vyjadriť ako hodnoty polynómov stupňa $< m$ pre maticu $\boldsymbol{A}$ (porovnajte s príkladom 21.1.5). Najvýhodnejší na ten účel preto je minimálny polynóm matice $\boldsymbol{A}$. Ak totiž $\mu_{\boldsymbol{A}}(x) = x^m - \sum_{j=1}^{m} d_j x^{m-j}$, tak $\boldsymbol{A}^m = \sum_{j=1}^{m} d_j \boldsymbol{A}^{m-j}$. Potom

$$\boldsymbol{A}^{m+k} = \sum_{j=1}^{m} d_j^{(k)} \boldsymbol{A}^{m-j},$$

kde koeficienty $d_j^{(k)}$ možno vypočítať z rekurentného vzťahu

$$d_j^{(0)} = d_j, \qquad d_j^{(k+1)} = d_j d_1^{(k)} + d_{j+1}^{(k)},$$

pričom pre $j > m$ definitoricky kladieme $d_j^{(k)} = 0$. Ak teda označíme

$$\nu_{\boldsymbol{A}}^{(k)}(x) = d_1^{(k)} x^{m-1} + \ldots + d_{m-1}^{(k)} x + d_m^{(k)} = \sum_{j=1}^{m} d_j^{(k)} x^{m-j},$$

tak všetky tieto polynómy majú stupeň $< m$ a pre $\boldsymbol{A} \in \mathrm{Dom}_n(f)$ platí

$$f(\boldsymbol{A}) = \sum_{k=0}^{\infty} c_k \boldsymbol{A}^k = c_0 \boldsymbol{I} + \ldots + c_{m-1} \boldsymbol{A}^{m-1} + \sum_{k=0}^{\infty} c_{m+k} \nu_{\boldsymbol{A}}^{(k)}(\boldsymbol{A}).$$

Iný spôsob výpočtu potenčného radu $f(\boldsymbol{A}) = \sum c_k \boldsymbol{A}^k$, pri ktorom zároveň možno rozhodnúť otázku, či $\boldsymbol{A} \in \mathrm{Dom}_n(f)$, sa zakladá na znalosti Jordanovho kanonického tvaru matice $\boldsymbol{A} \in \mathbb{C}^{n\times n}$. Keďže pole $\mathbb{C}$ je algebraicky uzavreté, existuje matica $\boldsymbol{J} = \mathrm{diag}\big(\boldsymbol{J}_{n_1}(\lambda_1), \ldots, \boldsymbol{J}_{n_p}(\lambda_p)\big) \in \mathbb{C}^{n\times n}$ v JKT a regulárna matica $\boldsymbol{P} \in \mathbb{C}^{n\times n}$ také, že $\boldsymbol{A} = \boldsymbol{P} \cdot \boldsymbol{J} \cdot \boldsymbol{P}^{-1}$. Vďaka tvrdeniu 22.1.2 a leme 22.1.4 potom

$$\boldsymbol{A} \in \mathrm{Dom}_n(f) \;\Leftrightarrow\; \boldsymbol{J} \in \mathrm{Dom}_n(f) \;\Leftrightarrow\; (\forall\, i \leq p)\big(\boldsymbol{J}_{n_i}(\lambda_i) \in \mathrm{Dom}_{n_i}(f)\big),$$

a – podobne ako pre polynómy – v tom prípade platí

$$f(\boldsymbol{A}) = \boldsymbol{P} \cdot f(\boldsymbol{J}) \cdot \boldsymbol{P}^{-1} = \boldsymbol{P} \cdot \mathrm{diag}\big(f\big(\boldsymbol{J}_{n_1}(\lambda_1)\big), \ldots, f\big(\boldsymbol{J}_{n_p}(\lambda_p)\big)\big) \cdot \boldsymbol{P}^{-1}.$$

Tým sme otázku konvergencie potenčných radov pre všeobecné matice $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ zredukovali na otázku konvergencie takýchto radov pre Jordanove bunky $\boldsymbol{J}_n(\lambda)$, kde $\lambda \in \mathbb{C}$.

Použitím viet 22.1.1, 21.1.8 a limitným prechodom $m \to \infty$ na postupnosť čiastočných súčtov $f_m(x) = \sum_{k=0}^{m} c_k x^k$, umožneným rovnomernou konvergenciou, dostávame

22.1.5. Veta. *Nech $f(x) = \sum_{k=0}^{\infty} c_k x^k \in \mathbb{C}[[x]]$ je potenčný rad s polomerom konvergencie r a $\boldsymbol{J}_n(\lambda)$ je komplexná Jordanova bunka. Potom platí*

(a) ak $|\lambda| < r$, *tak* $\boldsymbol{J}_n(\lambda) \in \mathrm{Dom}_n(f)$;
(b) ak $|\lambda| > r$, *tak* $\boldsymbol{J}_n(\lambda) \notin \mathrm{Dom}_n(f)$.
V prípade (a) navyše platí

$$f\big(\boldsymbol{J}_n(\lambda)\big) = \begin{pmatrix} f(\lambda) & \frac{1}{1!}f'(\lambda) & \dots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda) \\ 0 & f(\lambda) & \dots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f(\lambda) \end{pmatrix}.$$

K hraničnému prípadu $|\lambda| = r$ poznamenajme len toľko, že rad $f\big(\boldsymbol{J}_n(\lambda)\big)$ konverguje práve vtedy, keď konverguje rad $f^{(n-1)}(\lambda)$. Potom konvergujú i všetky ostatné rady $f^{(j)}(\lambda)$, $0 \le j \le n-1$, a vyššie uvedený vzorec zostáva v platnosti.

Vzorec pre $f\big(\boldsymbol{J}_n(\lambda)\big)$ je cenný najmä vtedy, keď máme k dispozícii kompaktnú formulu pre funkciu f, z ktorej vieme priamo vypočítať aj jej derivácie. Ako uvidíme v paragrafe 22.3, taktiež ho možno použiť na *definíciu* hodnôt $f\big(\boldsymbol{J}_n(\lambda)\big)$ aj pre iné funkcie, než len dané potenčnými radmi. Stačí, aby funkcia f mala v bode λ všetky derivácie až do rádu $n-1$. V ďalšom kroku možno definíciu $f(\boldsymbol{A})$ rozšíriť na všetky matice $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ za predpokladu, že ich spektrum spĺňa istú podmienku, ktorú presnejšie sformulujeme v 22.3.

Spektrálnym polomerom komplexnej (a tým i reálnej) *štvorcovej matice* $\boldsymbol{A}$ nazývame maximum absolútnych hodnôt jej vlastných čísel, t. j. číslo

$$\rho(\boldsymbol{A}) = \max\{|\lambda|;\ \lambda \in \mathrm{Spec}\,\boldsymbol{A}\}.$$

Zrejme $\rho(\boldsymbol{A}) \ge 0$ je reálne číslo.

Otázku konvergencie potenčných radov pre všeobecné matice možno vyjasniť na základe vzťahu medzi polomerom konvergencie radu a spektrálnym polomerom matice. Keďže každá komplexná štvorcová matica je podobná s maticou v JKT, z tvrdení 22.1.2, 22.1.4 a 22.1.5 vyplýva

22.1.6. Veta. *Nech* $f(x) = \sum c_k x^k \in \mathbb{C}[[x]]$ *je potenčný rad s polomerom konvergencie* r. *Potom pre ľubovoľnú maticu* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ *platí*
(a) ak $\rho(\boldsymbol{A}) < r$, *tak* $\boldsymbol{A} \in \mathrm{Dom}_n(f)$;
(b) ak $\rho(\boldsymbol{A}) > r$, *tak* $\boldsymbol{A} \notin \mathrm{Dom}_n(f)$.

22.1.7. Príklad. Funkcie $(1-x)^{-1}$, $(1+x)^{-1}$ možno pre $|x| < 1$ vyjadriť Mac Laurinovými radmi

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k, \qquad \frac{1}{1+x} = \sum_{k=0}^{\infty} (-1)^k x^k,$$

s rovnakým polomerom konvergencie $r = 1$. Preto i maticové rady $\sum \boldsymbol{A}^k$, $\sum (-1)^k \boldsymbol{A}^k$ konvergujú pre každé $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, $\rho(\boldsymbol{A}) < 1$, a platí

$$\sum_{k=0}^{\infty} \boldsymbol{A}^k = (\boldsymbol{I} - \boldsymbol{A})^{-1}, \qquad \sum_{k=0}^{\infty} (-1)^k \boldsymbol{A}^k = (\boldsymbol{I} + \boldsymbol{A})^{-1};$$

pozri cvičenie 22.4. Rad $\sum \boldsymbol{A}^k = (\boldsymbol{I} - \boldsymbol{A})^{-1}$ sa zvykne tiež nazývať *von Neumannovým radom* matice $\boldsymbol{A}$.

22.1.8. Príklad. Funkcie $\ln(1-x)$, $\ln(1+x)$ možno pre $|x| < 1$ rozvinúť do Mac Laurinových radov

$$\ln(1-x) = -\sum_{k=1}^{\infty} \frac{1}{k}\, x^k, \qquad \ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}\, x^k,$$

opäť s polomerom konvergencie $r = 1$, ktoré sme získali z radov pre $(1 \pm x)^{-1}$ formálnou integráciou člen po člene a využitím rovnosti $\ln 1 = 0$. Uvedené rady teda môžeme použiť na *definíciu* funkcií

$$\ln(\boldsymbol{I} - \boldsymbol{A}) = -\sum_{k=1}^{\infty} \frac{1}{k} \boldsymbol{A}^k, \qquad \ln(\boldsymbol{I} + \boldsymbol{A}) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \boldsymbol{A}^k,$$

pre $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, $\rho(\boldsymbol{A}) < 1$. Oprávnenosťou takýchto definícií sa budeme zaoberať v cvičení 22.5.

Uvedomme si, že $\operatorname{Spec}(\boldsymbol{A} - \boldsymbol{I}) = \{\lambda - 1;\ \lambda \in \operatorname{Spec} \boldsymbol{A}\}$, a dosaďme do druhého radu maticu $\boldsymbol{A} - \boldsymbol{I}$ miesto matice $\boldsymbol{A}$ (prípadne do prvého radu maticu $\boldsymbol{I} - A$). Maticový rad

$$\ln \boldsymbol{A} = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}\, (\boldsymbol{A} - \boldsymbol{I})^k = -\sum_{k=1}^{\infty} \frac{1}{k}\, (\boldsymbol{I} - \boldsymbol{A})^k,$$

ktorý takto dostaneme, konverguje pre všetky $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ také, že $|\lambda - 1| < 1$ pre každé $\lambda \in \operatorname{Spec} \boldsymbol{A}$; pre takéto matice ho teda možno použiť na definíciu funkcie $\ln \boldsymbol{A}$.

22.2 Exponenciála matice

Vari najdôležitejšou funkciou v matematickej analýze je *exponenciála* e^x, ktorú možno pre každé $x \in \mathbb{C}$ definovať potenčným radom

$$\mathrm{e}^x = \exp x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Z uvedeného rozvoja už možno odvodiť všetky dôležité vlastnosti exponenciálnej funkcie, vrátane kľúčového vzťahu $\left(\mathrm{e}^x\right)' = \mathrm{e}^x$.

S exponenciálnou funkciou úzko súvisia ďalšie dve funkcie, kosínus a sínus, ktoré je pre naše účely najvýhodnejšie definovať pre každé $x \in \mathbb{C}$ mocninnými radmi

$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}, \qquad \sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}.$$

Presvedčte sa samostatne, že polomer konvergencie všetkých troch uvedených radov je $r = \infty$.

To nám umožňuje definovať exponenciálu, kosínus a sínus pre ľubovoľnú maticu $\boldsymbol{A} \in \mathbb{C}^{n\times n}$:

$$\mathrm{e}^{\boldsymbol{A}} = \exp \boldsymbol{A} = \sum_{k=0}^{\infty} \frac{1}{k!} \boldsymbol{A}^k,$$

$$\cos \boldsymbol{A} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \boldsymbol{A}^{2k}, \qquad \sin \boldsymbol{A} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \boldsymbol{A}^{2k+1}.$$

Jednoduchým výpočtom možno overiť maticové zovšeobecnenie slávneho *Eulerovho vzťahu* $\mathrm{e}^{\mathrm{i}x} = \cos x + \mathrm{i} \sin x$.

22.2.1. Tvrdenie. *Pre ľubovoľnú maticu* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ *platí*

$$\mathrm{e}^{\mathrm{i}\boldsymbol{A}} = \cos \boldsymbol{A} + \mathrm{i} \sin \boldsymbol{A}.$$

Dôkaz. Sčítaním zvlášť cez párne a nepárne členy radu dostaneme

$$\begin{aligned} \mathrm{e}^{\mathrm{i}\boldsymbol{A}} &= \sum_{k=0}^{\infty} \frac{1}{k!} (\mathrm{i}\boldsymbol{A})^k = \sum_{k=0}^{\infty} \frac{\mathrm{i}^{2k}}{(2k)!} \boldsymbol{A}^{2k} + \sum_{k=0}^{\infty} \frac{\mathrm{i}^{2k+1}}{(2k+1)!} \boldsymbol{A}^{2k+1} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \boldsymbol{A}^{2k} + \mathrm{i} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \boldsymbol{A}^{2k+1} = \cos \boldsymbol{A} + \mathrm{i} \sin \boldsymbol{A}. \end{aligned}$$

Známy vzťah $\mathrm{e}^{x+y} = \mathrm{e}^x\,\mathrm{e}^y$ má tiež maticové zovšeobecnenie.

22.2.2. Tvrdenie. *Ak matice* $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$ *komutujú, tak*

$$\mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}} = \mathrm{e}^{\boldsymbol{A}} \cdot \mathrm{e}^{\boldsymbol{B}}.$$

Dôkaz. S použitím rovnosti $\binom{k}{j} = \frac{k!}{j!(k-j)!}$ a binomickej vety (lema 21.1.6) postupnými úpravami vypočítame

$$\begin{aligned}\mathrm{e}^{\boldsymbol{A}} \cdot \mathrm{e}^{\boldsymbol{B}} &= \sum_{i=0}^{\infty} \frac{1}{i!}\boldsymbol{A}^i \cdot \sum_{j=0}^{\infty} \frac{1}{j!}\boldsymbol{B}^j = \sum_{k=0}^{\infty} \sum_{i+j=k} \frac{1}{i!\,j!}\boldsymbol{A}^i\boldsymbol{B}^j \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j} \boldsymbol{A}^{k-j}\boldsymbol{B}^j = \sum_{k=0}^{\infty} \frac{1}{k!}(\boldsymbol{A}+\boldsymbol{B})^k = \mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}\,.\end{aligned}$$

Ak matice $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$ komutujú, tak komutujú aj matice $\boldsymbol{A}$, $\mathrm{i}\boldsymbol{B}$. Z tvrdení 22.2.1, 22.2.2 tak dostávame

22.2.3. Dôsledok. *Ak matice $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$ komutujú, tak*

$$\mathrm{e}^{\boldsymbol{A}+\mathrm{i}\boldsymbol{B}} = \mathrm{e}^{\boldsymbol{A}}(\cos \boldsymbol{B} + \mathrm{i}\sin \boldsymbol{B}),$$

pričom $\mathrm{e}^{\boldsymbol{A}}, \cos \boldsymbol{B}, \sin \boldsymbol{B} \in \mathbb{R}^{n\times n}$.

Uvedomte si, že pre nulovú maticu $\mathbf{0}_{n,n} = \mathrm{diag}(0, \dots, 0)$ podľa lemy 21.1.3 platí $\exp \mathbf{0}_{n,n} = \mathrm{diag}(\mathrm{e}^0, \dots, \mathrm{e}^0) = \boldsymbol{I}_n$. Keďže každá štvorcová matica $\boldsymbol{A}$ komutuje s maticou $-\boldsymbol{A}$, z tvrdenia 22.2.2 okamžite vyplýva ďalší dôsledok.

22.2.4. Dôsledok. *Exponenciála $\mathrm{e}^{\boldsymbol{A}}$ ľubovoľnej matice $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ je regulárna a platí*

$$\left(\mathrm{e}^{\boldsymbol{A}}\right)^{-1} = \mathrm{e}^{-\boldsymbol{A}}\,.$$

Kľúčom k výpočtu exponenciály matice je opäť znalosť exponenciály Jordanových buniek. Vďaka rovnosti $(\mathrm{e}^x)' = \mathrm{e}^x$ podľa vety 22.1.5 dostávame

22.2.5. Tvrdenie. *Pre $n \geq 1$, $\lambda \in \mathbb{C}$ platí*

$$\exp \boldsymbol{J}_n(\lambda) = \begin{pmatrix} \mathrm{e}^\lambda & \frac{\mathrm{e}^\lambda}{1!} & \dots & \frac{\mathrm{e}^\lambda}{(n-1)!} \\ 0 & \mathrm{e}^\lambda & \dots & \frac{\mathrm{e}^\lambda}{(n-2)!} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathrm{e}^\lambda \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{1!} & \dots & \frac{1}{(n-1)!} \\ 0 & 1 & \dots & \frac{1}{(n-2)!} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \mathrm{e}^\lambda\,.$$

Na záver ešte zaznamenáme jeden užitočný a zaujímavý vzťah, umožňujúci jednoduchý vypočet determinantu exponenciály matice len na základe jej stopy.

22.2.6. Veta. (Liouvilleova formula) *Pre ľubovoľnú maticu $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ platí*

$$\det \mathrm{e}^{\boldsymbol{A}} = \mathrm{e}^{\mathrm{tr}\,\boldsymbol{A}}\,.$$

Dôkaz. Pre $\boldsymbol{J} = \operatorname{diag}\big(\boldsymbol{J}_{n_1}(\lambda_1)\dots,\boldsymbol{J}_{n_p}(\lambda_p)\big)$ podľa lemy 22.1.4 máme

$$\mathrm{e}^{\boldsymbol{J}} = \operatorname{diag}\big(\exp\boldsymbol{J}_{n_1}(\lambda_1),\dots,\exp\boldsymbol{J}_{n_p}(\lambda_p)\big).$$

Takže v dôsledku tvrdenia 22.2.5 platí

$$\det\mathrm{e}^{\boldsymbol{J}} = \big(\mathrm{e}^{\lambda_1}\big)^{n_1}\dots\big(\mathrm{e}^{\lambda_p}\big)^{n_p} = \exp(n_1\lambda_1+\dots+n_p\lambda_p) = \mathrm{e}^{\operatorname{tr}\boldsymbol{J}}\,.$$

Nech teraz $\boldsymbol{J}\approx\boldsymbol{A}$ je JKT matice $\boldsymbol{A}$. Podľa dôsledku 22.1.3 platí $\mathrm{e}^{\boldsymbol{A}}\approx\mathrm{e}^{\boldsymbol{J}}$. Keďže podobné matice majú rovnaký determinant aj stopu (pozri dôsledok 18.1.4), z prvej časti dôkazu vyplýva

$$\det\mathrm{e}^{\boldsymbol{A}} = \det\mathrm{e}^{\boldsymbol{J}} = \mathrm{e}^{\operatorname{tr}\boldsymbol{J}} = \mathrm{e}^{\operatorname{tr}\boldsymbol{A}}\,.$$

Keďže $\det\mathrm{e}^{\boldsymbol{A}} = \mathrm{e}^{\operatorname{tr}\boldsymbol{A}}\neq 0$, dostávame tak iný dôkaz regularity matice $\mathrm{e}^{\boldsymbol{A}}$.

22.3 Skalárne funkcie maticového argumentu

Tvrdenie 21.1.1, lema 21.1.3 a veta 21.1.8, spolu s vetou 20.1.2 o Jordanovom kanonickom tvare, nám umožňujú vypočítať hodnotu $f(\boldsymbol{A})$ pre ľubovoľný polynóm $f(x)$ a štvorcovú maticu $\boldsymbol{A}$ nad poľom komplexných čísel. Vďaka tvrdeniu 22.1.2, leme 22.1.4 a vete 22.1.5 môžeme tú istú schému uplatniť aj pre potenčný rad $f(x)\in\mathbb{C}[[x]]$ a maticu $\boldsymbol{A}\in\mathbb{C}^{n\times n}$, ktorej spektrálny polomer je menší ako polomer konvergencie radu $f(x)$. To navodzuje myšlienku použiť rovnakú schému na *definíciu* rozšírenia komplexných funkcií komplexnej premennej (prípadne reálnych funkcií reálnej premennej) na maticové argumenty. Takto získaným funkciám budeme hovoriť *skalárne funkcie maticového argumentu.*

Pripomeňme, že *rádom vlastnej hodnoty* λ matice $\boldsymbol{A}\in K^{n\times n}$ nad ľubovoľným poľom K nazývame rád r najväčšej Jordanovej bunky $\boldsymbol{J}_r(\lambda)$ v JKT matice $\boldsymbol{A}$ (pozri paragraf 20.4 a cvičenie 20.6).

Nech f je komplexná funkcia komplexnej premennej a $\boldsymbol{A}\in\mathbb{C}^{n\times n}$. Hodnota funkcie f je definovaná pre maticu $\boldsymbol{A}$, píšeme tiež $\boldsymbol{A}\in\operatorname{Dom}_n f$, ak pre každé vlastné číslo $\lambda\in\operatorname{Spec}\boldsymbol{A}$ s rádom r sú funkcia f a všetky derivácie $f',\dots,f^{(r-1)}$ definované v bode λ. (Ak $r>1$, tak to predpokladá, že funkcia f je definovaná v nejakom okolí bodu λ.) Samotnú definíciu hodnoty $f(\boldsymbol{A})$, za uvedených podmienok, rozložíme do troch krokov:

(1) Ak $\boldsymbol{A}=\boldsymbol{J}_n(\lambda)$ je Jordanova bunka, tak

$$f(\boldsymbol{A}) = f\big(\boldsymbol{J}_n(\lambda)\big) = \begin{pmatrix} f(\lambda) & \frac{1}{1!}f'(\lambda) & \dots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda)\\ 0 & f(\lambda) & \dots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda)\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & f(\lambda)\end{pmatrix}.$$

(2) Ak $\boldsymbol{A} = \boldsymbol{J} = \operatorname{diag}(\boldsymbol{J}_1, \dots, \boldsymbol{J}_k)$ je v JKT, kde $\boldsymbol{J}_i = \boldsymbol{J}_{n_i}(\lambda_i)$ pre $i \le k$ sú Jordanove bunky, tak
$$f(\boldsymbol{A}) = f(\boldsymbol{J}) = \operatorname{diag}\bigl(f(\boldsymbol{J}_1), \dots, f(\boldsymbol{J}_k)\bigr).$$

(3) Konečne, ak $\boldsymbol{A} = \boldsymbol{P} \cdot \boldsymbol{J} \cdot \boldsymbol{P}^{-1}$, kde $\boldsymbol{J}$ je JKT matice $\boldsymbol{A}$ a $\boldsymbol{P}$ je regulárna matica, tak
$$f(\boldsymbol{A}) = \boldsymbol{P} \cdot f(\boldsymbol{J}) \cdot \boldsymbol{P}^{-1}.$$

Pred chvíľou spomínané výsledky nám zaručujú, že pre funkcie dané polynómami či potenčnými radmi dáva táto definícia rovnaké výsledky ako pôvodné definície. V prípade všeobecných funkcií je však potrebné dokázať jej korektnosť. Táto úloha sa redukuje na dôkaz nezávislosti hodnoty $f(\boldsymbol{A})$ na matici prechodu $\boldsymbol{P}$ medzi maticou $\boldsymbol{A}$ a jej JKT $\boldsymbol{J}$, ako aj na poradí Jordanových buniek v matici $\boldsymbol{J}$.

Začneme jedným zovšeobecnením Lagrangeovho interpolačného polynómu (porovnaj s cvičením 10.16).

22.3.1. Lema. *Nech K je niektoré z polí $\mathbb{R}$ alebo $\mathbb{C}$, hodnoty $\lambda_1, \dots, \lambda_m \in K$ sú navzájom rôzne a $\boldsymbol{Y} = (y_{ik}) \in K^{m \times r}$ je matica (so stĺpcami číslovanými od 0 do $r-1$). Potom existuje jednoznačne určený polynóm $F(x) \in K[x]$ stupňa $\le mr - 1$ taký, že*
$$F^{(k)}(\lambda_i) = y_{ik}$$
pre všetky $1 \le i \le m$, $0 \le k < r$.

Dôkaz. Polynóm $F(x)$ budeme hľadať v tvare
$$F(x) = p_1(x)q_1(x) + \ldots + p_m(x)q_m(x),$$
kde
$$p_i(x) = \sum_{k=0}^{r-1} \frac{c_{ik}}{k!}(x - \lambda_i)^k$$
má tvar Taylorovho polynómu a
$$q_i(x) = \prod_{j \ne i} (x - \lambda_j)^r.$$
Takže zostáva určiť iba koeficienty c_{ik}.

Zvoľme ľubovoľné $1 \le i \le m$, $0 \le k < r$. Z tvaru polynómov $q_i(x)$ je zrejmé, že $q_i^{(k)}(\lambda_j) = 0$ pre $j \ne i$. V dôsledku toho s použitím Newtonovej binomickej formuly pre k-tu deriváciu súčinu dvoch funkcií dostávame
$$\begin{aligned} F^{(k)}(\lambda_i) &= \frac{\mathrm{d}^k}{\mathrm{d}x^k}\Bigl[p_i(x)q_i(x)\Bigr]_{x=\lambda_i} \\ &= \sum_{j=0}^{k} \binom{k}{j} p_i^{(j)}(\lambda_i) q_i^{(k-j)}(\lambda_i) = \sum_{j=0}^{k} \binom{k}{j} c_{ij} q_i^{(k-j)}(\lambda_i); \end{aligned}$$

pozri cvičnie 22.9. Pre pevné i vypočítame vektor koeficientov $c_{i0}, \dots, c_{i\,r-1}$ ako riešenie sústavy r lineárnych rovníc

$$\sum_{j=0}^{k} \binom{k}{j} q_i^{(k-j)}(\lambda_i)\, c_{ij} = y_{ik}$$

pre $0 \leq k < r$. Táto sústava má dolnú trojuholníkovú maticu, pričom všetky jej diagonálne prvky sú $q_i(\lambda_i)$. Keďže hodnoty $\lambda_1, \dots, \lambda_m$ sú navzájom rôzne, $q_i(\lambda_i) \neq 0$, preto matica sústavy je regulárna a sústava má jediné riešenie. Postupnou elimináciou dostávame

$$c_{i0} = \frac{y_{i0}}{q_i(\lambda_i)}, \qquad c_{i1} = \frac{y_{i1}}{q_i(\lambda_i)} - \frac{q_i'(\lambda_i)\, y_{i0}}{q_i(\lambda_i)^2},$$

$$c_{i2} = \frac{y_{i2}}{q_i(\lambda_i)} - \frac{2q_i'(\lambda_i)\, y_{i1}}{q_i(\lambda_i)^2} + \frac{\big(2q_i'(\lambda_i)^2 - q_i(\lambda_i)q_i''(\lambda_i)\big)\, y_{i0}}{q_i(\lambda_i)^3}, \qquad \text{atď.}$$

Konkrétny tvar riešenia nás v tejto chvíli príliš nezaujíma – dôležitá je len jeho existencia.

22.3.2. Veta. *Nech $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ a $\lambda_1, \dots, \lambda_m$ sú všetky navzájom rôzne vlastné hodnoty matice $\boldsymbol{A}$, pričom r_i označuje rád vlastnej hodnoty λ_i. Nech f je skalárna funkcia maticového argumentu taká, že $\boldsymbol{A} \in \mathrm{Dom}_n f$. Potom existuje polynóm $F(x) \in \mathbb{C}[x]$, ktorý závisí len na vlastných hodnotách λ_i matice $\boldsymbol{A}$, ich rádoch r_i a hodnotách funkcie f a jej derivácií až do rádu $r_i - 1$ v bodoch spektra λ_i, taký, že $f(\boldsymbol{A}) = F(\boldsymbol{A})$.*

Dôkaz. Označme $r = \max\{r_1, \dots, r_m\}$. Pre $1 \leq i \leq m$, $0 \leq k < r$ položme

$$y_{ik} = \begin{cases} f^{(k)}(\lambda_i), & \text{ak } k < r_i, \\ 0, & \text{ak } k \geq r_i. \end{cases}$$

Podľa práve dokázanej lemy existuje polynóm $F(x)$ taký, že $F^{(k)}(\lambda_i) = y_{ik}$ pre všetky i, k.

Nech $\boldsymbol{J} \approx \boldsymbol{A}$ je JKT matice $\boldsymbol{A}$ a $\boldsymbol{P}$ je regulárna matica taká, že $\boldsymbol{A} = \boldsymbol{P} \cdot \boldsymbol{J} \cdot \boldsymbol{P}^{-1}$. Potom maximálny rozmer Jordanových buniek v $\boldsymbol{J}$ je $r \times r$. Keďže hodnota $f(\boldsymbol{J})$ závisí len od hodnôt $f^{(k)}(\lambda_i)$ pre $1 \leq i \leq m$, $0 \leq k < r_i$, zrejme platí $f(\boldsymbol{J}) = F(\boldsymbol{J})$. Podľa definície výrazu $f(\boldsymbol{A})$ a tvrdenia 22.1.2 z toho vyplýva

$$f(\boldsymbol{A}) = \boldsymbol{P} \cdot f(\boldsymbol{J}) \cdot \boldsymbol{P}^{-1} = \boldsymbol{P} \cdot F(\boldsymbol{J}) \cdot \boldsymbol{P}^{-1} = F\big(\boldsymbol{P} \cdot \boldsymbol{J} \cdot \boldsymbol{P}^{-1}\big) = F(\boldsymbol{A}).$$

Z práve dokázanej vety okamžite vyplýva korektnosť našej definície výrazu $f(\boldsymbol{A})$: hodnota $f(\boldsymbol{A}) = F(\boldsymbol{A})$ totiž nezávisí na konkrétnej reprezentácii $\boldsymbol{A} = \boldsymbol{P} \cdot \boldsymbol{J} \cdot \boldsymbol{P}^{-1}$ matice $\boldsymbol{A}$ pomocou matíc $\boldsymbol{J}$ a $\boldsymbol{P}$.

V najjednoduchšom prípade, keď matica $\boldsymbol{A}$ je diagonalizovateľná, t. j. $\boldsymbol{A} = \boldsymbol{P} \cdot \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \cdot \boldsymbol{P}^{-1}$ pre nejakú regulárnu maticu $\boldsymbol{P}$ a $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, hodnota $f(\boldsymbol{A})$ je definovaná práve vtedy, keď funkcia f je definovaná vo všetkých bodoch λ_i spektra matice $\boldsymbol{A}$, pričom

$$f(\boldsymbol{A}) = \boldsymbol{P} \cdot \operatorname{diag}\bigl(f(\lambda_1), \ldots, f(\lambda_n)\bigr) \cdot \boldsymbol{P}^{-1}.$$

Napokon si ešte ujasnime, ako je to s možnosťou rozšírenia *reálnych funkcií reálnej premennej* na maticové argumenty. Z vykonaných úvah vyplýva, že otázka, či $\boldsymbol{A}$ je reálna alebo komplexná matica, pri tom nehrá podstatnú úlohu. Dôležité je, aby $\boldsymbol{A}$ mala *reálne spektrum.* Teda hodnotu $f(\boldsymbol{A})$ možno definovať podľa rovnakej schémy pre ľubovoľnú maticu $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, pre ktorú platí $\operatorname{Spec} \boldsymbol{A} \subseteq \mathbb{R}$ a pre každé $\lambda \in \operatorname{Spec} \boldsymbol{A}$ sú funkcia f a všetky jej derivácie až do rádu menšieho než rád λ definované (a konečné) v bode λ. Potom výsledná matica $f(\boldsymbol{A}) \in \mathbb{C}^{n\times n}$ má opäť reálne spektrum $\operatorname{Spec} f(\boldsymbol{A}) = \{f(\lambda);\, \lambda \in \operatorname{Spec} \boldsymbol{A}\}$. Ak $\boldsymbol{A}$ je navyše reálna matica s reálnym spektrom, tak aj výsledná matica $f(\boldsymbol{A})$ je reálna. (Rozmyslite si prečo – pozri cvičenie 22.11.)

22.4 Maticové a vektorové funkcie reálnej premennej

Funkciu $\boldsymbol{x}\colon S \to \mathbb{C}^n$, resp. $\boldsymbol{x}\colon S \to \mathbb{R}^n$, kde $S \subseteq \mathbb{R}$, budeme nazývať *komplexnou* resp. *reálnou vektorovou funkciou reálnej premennej.* Takáto funkcia sa prirodzene rozpadá na n zložiek, t. j. *skalárnych funkcií* $x_i\colon S \to \mathbb{C}$ resp. $x_i\colon S \to \mathbb{R}$ takých, že $\boldsymbol{x}(t) = (x_1(t), \ldots, x_n(t))^{\mathsf{T}}$ pre každé $t \in S$. Podobne funkciu $\boldsymbol{X}\colon S \to \mathbb{C}^{m\times n}$ resp. $\boldsymbol{X}\colon S \to \mathbb{R}^{m\times n}$ nazývame *komplexnou* resp. *reálnou maticovou funkciou reálnej premennej* na množine $S \subseteq \mathbb{R}$. I takúto funkciu možno rozložiť na mn skalárnych zložiek $x_{ij}\colon S \to \mathbb{C}$, resp. $x_{ij}\colon S \to \mathbb{R}$ takých, že pre každé $t \in S$ platí $\boldsymbol{X}(t) = \bigl(x_{ij}(t)\bigr)_{m\times n}$. Zrejme vektorové funkcie možno chápať ako špeciálny prípad maticových a reálne ako špeciálny prípad komplexných. [5]

Maticová funkcia $\boldsymbol{X} = (x_{ij})$ je *spojitá* v bode t_0 svojho definičného oboru S, prípadne na množine $N \subseteq S$, ak každá z jej zložiek x_{ij} má uvedenú vlastnosť.

[5] Pripomeňme, že komplexnú funkciu $g\colon S \to \mathbb{C}$ reálnej premennej $t \in S \subseteq \mathbb{R}$ možno jednoznačne rozložiť na tvar $g(t) = g_1(t) + \mathrm{i} g_2(t)$, kde $g_1, = \operatorname{Re} g$, $g_2 = \operatorname{Im} g$ sú funkcie $S \to \mathbb{R}$. Potom g je spojitá v bode $t_0 \in S$ práve vtedy, keď g_1 aj g_2 sú spojité v t_0. Podobne aj derivácia a integrál sú definované po zložkách: $g'(t_0)$ existuje práve vtedy, keď existujú $g_1'(t_0)$ a $g_2'(t_0)$, pričom $g'(t_0) = g_1'(t_0) + \mathrm{i} g_2'(t_0)$; pre spojitú $g\colon \langle a, b\rangle \to \mathbb{C}$ kladieme $\int_a^b g(t)\,\mathrm{d}t = \int_a^b g_1(t)\,\mathrm{d}t \,+\, \mathrm{i}\int_a^b g_2(t)\,\mathrm{d}t$.

Predpokladajme, že maticová funkcia $\boldsymbol{X}$ je definovaná v nejakom okolí N bodu t_0, t. j. všetky jej zložky x_{ij} sú definované na N. Hovoríme, že $\boldsymbol{X}$ má deriváciu v bode t_0, ak všetky zložky x_{ij} majú v bode t_0 (konečnú) deriváciu. Deriváciu funkcie $\boldsymbol{X}$ v bode t_0 značíme

$$\frac{\mathrm{d}\boldsymbol{X}(t_0)}{\mathrm{d}t} = \left(\frac{\mathrm{d}x_{ij}(t_0)}{\mathrm{d}t}\right)_{m\times n}, \qquad \text{prípadne} \qquad \boldsymbol{X}'(t_0) = \big(x'_{ij}(t_0)\big)_{m\times n}.$$

Niekedy, najmä pri typickej interpretácii premennej t ako času, sa tiež používa od Newtona pochádzajúce označenie $\dot{\boldsymbol{X}}(t_0) = \big(\dot{x}_{ij}(t_0)\big)_{m\times n}$. Vyššie derivácie (ak existujú) značíme $\mathrm{d}^k\boldsymbol{X}/\,\mathrm{d}t^k$, prípadne $\boldsymbol{X}^{(k)}$.

Nasledujúce maticové zovšeobecnenia pravidiel pre deriváciu lineárnej kombinácie a súčinu maticových funkcií, resp. pre kompozíciu skalárnej a maticovej funkcie možno overiť priamym výpočtom.

22.4.1. Tvrdenie. *Nech maticové funkcie $\boldsymbol{X}, \boldsymbol{Y}\colon S \to \mathbb{C}^{m\times n}$, $\boldsymbol{Z}\colon S \to \mathbb{C}^{n\times p}$ majú derivácie v bode $t_0 \in S$ a $a, b \in \mathbb{C}$. Potom aj funkcie $a\boldsymbol{X} + b\boldsymbol{Y} = (ax_{ij} + by_{ij})_{m\times n}$, $\boldsymbol{X}\cdot\boldsymbol{Z} = \left(\sum_j x_{ij}z_{jk}\right)_{m\times p}$ majú derivácie v bode t_0 a platí*

$$\begin{aligned}(a\boldsymbol{X} + b\boldsymbol{Y})'(t_0) &= a\boldsymbol{X}'(t_0) + b\boldsymbol{Y}'(t_0)\\ (\boldsymbol{X}\cdot\boldsymbol{Z})'(t_0) &= \boldsymbol{X}'(t_0)\cdot\boldsymbol{Z}(t_0) + \boldsymbol{X}(t_0)\cdot\boldsymbol{Z}'(t_0).\end{aligned}$$

Nech navyše reálna skalárna funkcia g je definovaná v okolí bodu s_0 a má v ňom (konečnú) deriváciu, pričom $g(s_0) = t_0$. Potom aj funkcia $\boldsymbol{X}\circ g = (x_{ij}\circ g)_{m\times n}$ má v bode s_0 deriváciu

$$(\boldsymbol{X}\circ g)'(s_0) = \boldsymbol{X}'(t_0)g'(s_0).$$

22.4.2. Dôsledok. *Nech maticová funkcia $\boldsymbol{X}\colon S \to \mathbb{R}^{n\times n}$ je definovaná v okolí bodu t_0 a má tam deriváciu. Potom pre každé $k \in \mathbb{N}$ aj maticová funkcia $\boldsymbol{X}^k$ má v bode t_0 deriváciu*

$$\frac{\mathrm{d}\boldsymbol{X}^k(t_0)}{\mathrm{d}t} = \sum_{j=1}^{k} \boldsymbol{X}^{j-1}(t_0)\cdot\boldsymbol{X}'(t_0)\cdot\boldsymbol{X}^{k-j}(t_0).$$

Ak navyše matice $\boldsymbol{X}(t_0)$ a $\boldsymbol{X}'(t_0)$ komutujú, tak

$$\frac{\mathrm{d}\boldsymbol{X}^k(t_0)}{\mathrm{d}t} = k\boldsymbol{X}^{k-1}(t_0)\cdot\boldsymbol{X}'(t_0) = k\boldsymbol{X}'(t_0)\cdot\boldsymbol{X}^{k-1}(t_0).$$

Dôkaz indukciou cez k.

22.4.3. Tvrdenie. *Nech potenčný rad $f(x) = \sum_{k=0}^{\infty} c_k x^k \in \mathbb{C}[[x]]$ má polomer konvergencie $r > 0$. Nech ďalej maticová funkcia $\boldsymbol{X}\colon S \to \mathbb{C}^{n\times n}$ je definovaná v okolí bodu $t_0 \in S$ a má v ňom deriváciu, pričom matice $\boldsymbol{X}(t_0)$ a $\boldsymbol{X}'(t_0)$ komutujú. Ak $\rho(\boldsymbol{X}(t_0)) < r$, tak funkcia $f(\boldsymbol{X}) = f \circ \boldsymbol{X}$ je definovaná v nejakom okolí bodu t_0 a má v ňom deriváciu*

$$f(\boldsymbol{X})'(t_0) = (f \circ \boldsymbol{X})'(t_0) = f'\big(\boldsymbol{X}(t_0)\big) \cdot \boldsymbol{X}'(t_0) = \boldsymbol{X}'(t_0) \cdot f'\big(\boldsymbol{X}(t_0)\big).$$

V dôkaze použijeme niekoľko argumentov z matematickej analýzy. Čitateľovi, ktorému veľa nehovoria, odporúčame prejsť priamo k výpočtu.

Z existencie derivácie $\boldsymbol{X}'(t_0)$ vyplýva spojitosť funkcie $\boldsymbol{X}$ v nejakom okolí $N \subseteq S$ bodu t_0. Zo spojitej závislosti koeficientov charakteristického polynómu na prvkoch matice ako aj koreňov polynómu na koeficientoch zasa vyplýva existencia okolia $M \subseteq N$ bodu t_0 a čísla $q > 0$ takých, že $\rho(\boldsymbol{X}(t)) \leq q < r$ pre $t \in M$. Označme $f_m(x) = \sum_{k=0}^{m} c_k x^k$. Vďaka rovnomernej konvergencii radu $f(\boldsymbol{X}(t)) = \sum c_k \boldsymbol{X}(t)^k$ na množine M a s použitím pravidiel z 22.4.1, 22.4.2 dostávame

$$\begin{aligned} f(\boldsymbol{X})'(t_0) &= \lim_{m\to\infty} f_m(\boldsymbol{X})'(t_0) = \lim_{m\to\infty} \sum_{k=0}^{m} c_k \frac{\mathrm{d}\boldsymbol{X}^k(t_0)}{\mathrm{d}t} \\ &= \lim_{m\to\infty} \sum_{k=1}^{m} k c_k \boldsymbol{X}^{k-1}(t_0) \cdot \boldsymbol{X}'(t_0) = f'\big(\boldsymbol{X}(t_0)\big) \cdot \boldsymbol{X}'(t_0). \end{aligned}$$

Keďže za uvedeného predpokladu komutujú aj $\boldsymbol{X}'(t_0)$ a $f'\big(\boldsymbol{X}(t_0)\big)$, platí i druhá rovnosť.

Z faktu, že matica $\boldsymbol{A}t$ vždy komutuje s maticou $(\boldsymbol{A}t)' = \boldsymbol{A}$, ako aj s maticami $\mathrm{e}^{\boldsymbol{A}t}$, $\cos \boldsymbol{A}t$, $\sin \boldsymbol{A}t$, na základe práve dokázanej vety okamžite vyplýva

22.4.4. Dôsledok. *Nech $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ je pevne zvolená matica. Potom maticové funkcie $\mathrm{e}^{\boldsymbol{A}t}$, $\cos \boldsymbol{A}t$, $\sin \boldsymbol{A}t$ sú definované pre každé $t \in \mathbb{R}$ a pre ich derivácie platí*

$$\big(\mathrm{e}^{\boldsymbol{A}t}\big)' = \boldsymbol{A} \cdot \mathrm{e}^{\boldsymbol{A}t}, \qquad \big(\cos \boldsymbol{A}t\big)' = -\boldsymbol{A} \cdot \sin \boldsymbol{A}t, \qquad \big(\sin \boldsymbol{A}t\big)' = \boldsymbol{A} \cdot \cos \boldsymbol{A}t.$$

Určitý integrál spojitej maticovej funkcie $\boldsymbol{X}\colon S \to \mathbb{C}^{m\times n}$ na intervale [6] $S \subseteq \mathbb{R}$ definujeme pre $s, t \in S$ opäť po zložkách:

$$\int_s^t \boldsymbol{X}(\tau)\,\mathrm{d}\tau = \left(\int_s^t x_{ij}(\tau)\,\mathrm{d}\tau\right)_{m\times n}$$

[6] Dohodneme sa, že pod *intervalom* budeme odteraz vždy rozumieť *netriviálny interval* – či už ohraničený alebo neohraničený, – t. j. ľubovoľnú podmnožinu $S \subseteq \mathbb{R}$, ktorá obsahuje aspoň dva body a spĺňa podmienku $(\forall\, a, b \in S)(\forall\, x \in \mathbb{R})(a \leq x \leq b \Rightarrow x \in S)$.

Prenecháme čitateľovi, aby si samostatne premyslel, že takto definovaný určitý integrál splýva s limitou maticových integrálnych súčtov

$$\int_s^t \boldsymbol{X}(\tau)\,\mathrm{d}\tau = \lim_{k\to\infty} \sum_{j=1}^{k} \boldsymbol{X}(s+jd_k)d_k,$$

kde $d_k = (t-s)/k$.

Nasledujúce tvrdenie o „maticovej linearite“ sprava i zľava a množinovej aditívnosti určitého integrálu je priamym zovšeobecnením analogických pravidiel pre integrál skalárnych funkcií.

22.4.5. Tvrdenie. *Nech $\boldsymbol{X}, \boldsymbol{Y} : S \to \mathbb{C}^{m\times n}$ sú spojité maticové funkcie na intervale $S \subseteq \mathbb{R}$, $s, t, u \in S$ a $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{k\times m}$, $\boldsymbol{C}, \boldsymbol{D} \in \mathbb{C}^{n\times p}$. Potom*

$$\int_s^t \big(\boldsymbol{A}\cdot\boldsymbol{X}(\tau)\cdot\boldsymbol{C} + \boldsymbol{B}\cdot\boldsymbol{Y}(\tau)\cdot\boldsymbol{D}\big)\,\mathrm{d}\tau = \boldsymbol{A}\cdot\left(\int_s^t \boldsymbol{X}(\tau)\,\mathrm{d}\tau\right)\cdot\boldsymbol{C} + \boldsymbol{B}\cdot\left(\int_s^t \boldsymbol{Y}(\tau)\,\mathrm{d}\tau\right)\cdot\boldsymbol{D},$$

$$\int_s^t \boldsymbol{X}(\tau)\,\mathrm{d}\tau + \int_t^u \boldsymbol{X}(\tau)\,\mathrm{d}\tau = \int_s^u \boldsymbol{X}(\tau)\,\mathrm{d}\tau.$$

Špeciálne pre skaláry $a, b \in \mathbb{C}$ a $s, t \in S$ platí

$$\int_s^t \big(a\boldsymbol{X}(\tau) + b\boldsymbol{Y}(\tau)\big)\,\mathrm{d}\tau = a\int_s^t \boldsymbol{X}(\tau)\,\mathrm{d}\tau + b\int_s^t \boldsymbol{Y}(\tau)\,\mathrm{d}\tau,$$

$$\int_t^t \boldsymbol{X}(\tau)\,\mathrm{d}\tau = \boldsymbol{0}, \qquad \int_s^t \boldsymbol{X}(\tau)\,\mathrm{d}\tau = -\int_t^s \boldsymbol{X}(\tau)\,\mathrm{d}\tau.$$

Derivácia a určitý integrál ako funkcia hornej medze sú opäť zviazané maticovým zovšeobecnením Newtonovej-Leibnizovej formuly, ako aj vyjadrením určitého integrálu rozdielom primitívnej funkcie hornej a dolnej medze. Inak povedané, ide o navzájom inverzné operácie v rovnakom zmysle ako v jednorozmernom prípade.

22.4.6. Tvrdenie. *Nech $\boldsymbol{X} : S \to \mathbb{C}^{m\times n}$ je spojitá maticová funkcia na intervale $S \subseteq \mathbb{R}$ a $s, t \in S$. Potom*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_s^t \boldsymbol{X}(\tau)\,\mathrm{d}\tau = \boldsymbol{X}(t).$$

Ak $\boldsymbol{X}$ navyše má na S spojitú deriváciu $\boldsymbol{X}'$, tak

$$\int_s^t \boldsymbol{X}'(\tau)\,\mathrm{d}\tau = \boldsymbol{X}(t) - \boldsymbol{X}(s).$$

22.5 Sústavy lineárnych diferenciálnych rovníc

Paragraf 22.2 sme začali vyzdvihnutím exponenciály e^x ako „vari najdôležitejšej funkcie v matematickej analýze“. Jeden z dôvodov takéhoto hodnotenia sme už naznačili v príklade 18.5.1 – každé riešenie diferenciálnej rovnice

$$\frac{\mathrm{d}x}{\mathrm{d}t} = ax,$$

kde $a \in \mathbb{C}$, má tvar funkcie $x(t) = q\,\mathrm{e}^{at}$ reálnej premennej t, kde $q \in \mathbb{C}$ je ľubovoľná konštanta. Túto možno určiť, ak máme navyše predpísanú i počiatočnú podmienku, t. j. hodnotu $x(t_0) = c \in \mathbb{C}$ v nejakom bode $t_0 \in \mathbb{R}$, – vtedy riešenie nadobúda tvar

$$x(t) = c\,\mathrm{e}^{a(t-t_0)} = q\,\mathrm{e}^{at},$$

kde $q = c\,\mathrm{e}^{-at_0}$. Riešenia zložitejších diferenciálnych rovníc sa preto často zostavujú vhodnými kombináciami funkcií e^{at}, prípadne funkcií $\cos at$, $\sin at$, pre rôzne a, ktoré však s exponenciálnymi funkciami tesne súvisia prostredníctvom Eulerovho vzťahu.

Uvedený tvar riešenia možno priamo zovšeobecniť aj na prípad, keď $a\colon S \to \mathbb{C}$ je ľubovoľná spojitá funkcia, definovaná na nejakom intervale $S \subseteq \mathbb{R}$, a nielen konštanta. Vtedy

$$x(t) = x(t_0)\exp\left(\int_{t_0}^{t} a(\tau)\,\mathrm{d}\tau\right),$$

pre $t_0, t \in S$. V tomto paragrafe preskúmame viacrozmerný analóg podobnej situácie.

Sústavou lineárnych diferenciálnych rovníc, prípadne *vektorovou lineárnou diferenciálnou rovicou* nazývame formulu tvaru

$$\boldsymbol{x}' = \boldsymbol{A}\cdot\boldsymbol{x} + \boldsymbol{b},$$

kde $\boldsymbol{A}\colon S \to \mathbb{C}^{n\times n}$ resp. $\boldsymbol{b}\colon S \to \mathbb{C}^n$ je spojitá maticová resp. vektorová funkcia na nejakom intervale $S \subseteq \mathbb{R}$. *Riešením* takejto *sústavy* rozumieme ľubovoľnú funkciu $\boldsymbol{x}\colon S \to \mathbb{C}^n$, ktorá má na S spojitú deriváciu a pre každé $t \in S$ platí

$$\boldsymbol{x}'(t) = \boldsymbol{A}(t)\cdot\boldsymbol{x}(t) + \boldsymbol{b}(t).$$

(V prípadných krajných bodoch intervalu sa pod deriváciou myslí derivácia sprava resp. zľava.) Takáto sústava býva často doplnená o *počiatočnú podmienku*

$$\boldsymbol{x}(t_0) = \boldsymbol{c},$$

ktorá predpisuje hodnotu $\boldsymbol{c} \in \mathbb{C}^n$ funkcie $\boldsymbol{x}$ v niektorom pevnom bode $t_0 \in S$. V takom prípade hovoríme o *počiatočnej úlohe*. V kurze obyčajných diferenciálnych rovníc sa zvykne dokazovať nasledujúca veta o existencii a jednoznačnosti jej riešenia.

22.5.1. Veta. *Nech $S \subseteq \mathbb{R}$ je interval, $\boldsymbol{A}\colon S \to \mathbb{C}^{n\times n}$ resp. $\boldsymbol{b}\colon S \to \mathbb{C}^n$ je spojitá maticová resp. vektorová funkcia na S, $t_0 \in S$ a $\boldsymbol{c} \in \mathbb{C}^n$. Potom existuje práve jedno riešenie počiatočnej úlohy*

$$\boldsymbol{x}' = \boldsymbol{A}\cdot\boldsymbol{x} + \boldsymbol{b}, \qquad\qquad \boldsymbol{x}(t_0) = \boldsymbol{c}.$$

Poznámka. (a) Keďže $\mathbb{C}^{n\times n}$ je vektorový priestor nad $\mathbb{C}$ a násobenie pevnou maticou $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ (či už zľava alebo sprava) je lineárne zobrazenie $\mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$, do rámca vektorových lineárnych diferenciálnych rovníc zapadajú aj *maticové lineárne diferenciálne rovnice* tvaru $\boldsymbol{X}' = \boldsymbol{A}_1\cdot\boldsymbol{X}\cdot\boldsymbol{A}_2 + \boldsymbol{B}$ či $\boldsymbol{X}' = \boldsymbol{A}_1\cdot\boldsymbol{X} + \boldsymbol{X}\cdot\boldsymbol{A}_2 + \boldsymbol{B}$ a pod., kde $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{B}\colon S \to \mathbb{C}^{n\times n}$ sú spojité maticové funkcie na nejakom intervale $S \subseteq \mathbb{R}$, prípadne doplnené počiatočnou podmienkou $\boldsymbol{X}(t_0) = \boldsymbol{C}$, pre $t_0 \in S$, $\boldsymbol{C} \in \mathbb{C}^{n\times n}$. Čitateľ si iste ľahko sformuluje pojem riešenia takýchto počiatočných úloh. Taktiež veta 22.5.1 zostáva s minimálnymi typografickými úpravami v platnosti aj pre takéto rovnice.

(b) Lineárnu (jednorozmernú) diferenciálnu rovnicu n-tého rádu

$$x^{(n)} - a_1x^{(n-1)} - \ldots - a_{n-1}x' - a_nx = b,$$

kde $a_1, \ldots, a_n, b\colon S \to \mathbb{C}$ sú spojité funkcie, možno substitúciou $x = x_n$ previesť na sústavu n lineárnych diferenciálnych rovníc $x_1' = a_1x_1 + \ldots a_nx_n + b$, $x_2' = x_1$, …, $x_n' = x_{n-1}$, t. j. $\boldsymbol{x}' = \boldsymbol{A}\cdot\boldsymbol{x} + \boldsymbol{b}$, kde

$$\boldsymbol{A}(t) = \begin{pmatrix} a_1(t) & \ldots & a_{n-1}(t) & a_n(t) \\ 1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 1 & 0 \end{pmatrix}, \qquad \boldsymbol{b}(t) = b(t)\boldsymbol{e}_1 = \begin{pmatrix} b(t) \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Pri uvedenej substitúcii totiž pre $1 \le i \le n$ platí $x_i = x^{(n-i)}$. Všimnite si, že pri pevnom t je $\boldsymbol{A}(t) = \boldsymbol{M}_{g_t}^{\mathsf{T}}$ transponovaná matica k pridruženej matici normovaného polynómu $g_t(x) = x^n - a_1(t)x^{n-1} - \ldots - a_{n-1}(t)x - a_n(t)$ (pozri paragraf 21.2).

Vetu 22.5.1 možno tiež interpretovať nasledujúcim spôsobom. Podobne ako u sústav lineárnych (algebraických) rovníc, aj tento raz ľahko nahliadneme, že všetky riešenia *homogénnej sústavy lineárnych diferenciálnych rovníc*

$$\boldsymbol{x}' = \boldsymbol{A}\cdot\boldsymbol{x}$$

tvoria lineárny podpriestor vektorového priestoru $\mathcal{C}^{(1)}(S, \mathbb{C}^n)$ všetkých spojito diferencovateľných funkcií $S \to \mathbb{C}^n$. Riešenia pôvodnej nehomogénnej sústavy potom tvoria afinný podpriestor v $\mathcal{C}^{(1)}(S, \mathbb{C}^n)$, ktorého zameraním je podpriestor riešení homogénnej sústavy. Veta 22.5.1 okrem iného hovorí, že dimenzia oboch priestorov riešení je n. V prípade homogénnej sústavy je totiž pre pevné $t_0 \in S$ priradením $\boldsymbol{x} \mapsto \boldsymbol{x}(t_0)$ definovaný lineárny izomorfizmus medzi priestorom jej riešení a $\mathbb{C}^n$. Na popis priestoru všetkých riešení homogénnej sústavy teda stačí nájsť nejakú jeho bázu – v tomto prípade sa jej hovorí *fundamentálny systém riešení*. Takúto bázu tvorí ľubovoľných n lineárne nezávislých riešení $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, t. j. takých, že vektory $\boldsymbol{x}_1(t_0), \ldots, \boldsymbol{x}_n(t_0) \in \mathbb{C}^n$ sú lineárne nezávislé pre nejaké $t_0 \in S$; potom už vektory $\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_n(t)$ sú lineárne nezávislé pre každé $t \in S$ (pozri cvičenie 22.17 (c)). Priestor všetkých riešení nehomogénnej sústavy je tak plne určený (ľubovoľným) jej jediným riešením a fundamentálnym systémom riešení homogénnej sústavy.

Maticová funkcia $\boldsymbol{F}\colon S \to \mathbb{C}^{n\times n}$, ktorej stĺpce tvoria fundamentálny systém riešení sústavy $\boldsymbol{x}' = \boldsymbol{A} \cdot \boldsymbol{x}$, sa nazýva *fundamentálna matica* tejto *sústavy*. Z pred chvíľou vykonaných úvah vyplýva, že spojitá maticová funkcia $\boldsymbol{F}\colon S \to \mathbb{C}^{n\times n}$ je fundamentálnou maticou sústavy $\boldsymbol{x}' = \boldsymbol{A} \cdot \boldsymbol{x}$ práve vtedy, keď matica $\boldsymbol{F}(t)$ je regulárna pre každé $t \in S$ (na čo stačí overiť jej regularitu v jedinom ľubovoľnom bode $t_0 \in S$) a vyhovuje maticovej diferenciálnej rovnici $\boldsymbol{X}' = \boldsymbol{A} \cdot \boldsymbol{X}$, t. j. pre každé $t \in S$ platí

$$\boldsymbol{F}'(t) = \boldsymbol{A}(t) \cdot \boldsymbol{F}(t);$$

z toho už vyplýva aj spojitosť maticovej funkcie $\boldsymbol{F}'\colon S \to \mathbb{C}^{n\times n}$.

Za predpokladu, že poznáme nejakú fundamentálnu maticu sústavy, už nie je ťažké napísať explicitný tvar riešenia príslušnej počiatočnej úlohy.

22.5.2. Veta. *Nech $\boldsymbol{F}$ je fundamentálna matica sústavy $\boldsymbol{x}' = \boldsymbol{A}\cdot\boldsymbol{x}$. Potom*
(a) riešenie homogénnej počiatočnej úlohy $\boldsymbol{x}' = \boldsymbol{A} \cdot \boldsymbol{x}$, $\boldsymbol{x}(t_0) = \boldsymbol{c}$ má tvar

$$\boldsymbol{x}(t) = \boldsymbol{F}(t) \cdot \boldsymbol{F}(t_0)^{-1} \cdot \boldsymbol{c}\,;$$

(b) Riešenie nehomogénnej počiatočnej úlohy $\boldsymbol{x}'(t) = \boldsymbol{A}\cdot\boldsymbol{x} + \boldsymbol{b}$, $\boldsymbol{x}(t_0) = \boldsymbol{c}$ má tvar

$$\begin{aligned}\boldsymbol{x}(t) &= \boldsymbol{F}(t) \cdot \boldsymbol{F}(t_0)^{-1} \cdot \boldsymbol{c} + \int_{t_0}^{t} \boldsymbol{F}(t) \cdot \boldsymbol{F}(s)^{-1} \cdot \boldsymbol{b}(s)\,\mathrm{d}s \\ &= \boldsymbol{F}(t) \cdot \left(\boldsymbol{F}(t_0)^{-1} \cdot \boldsymbol{c} + \int_{t_0}^{t} \boldsymbol{F}(s)^{-1} \cdot \boldsymbol{b}(s)\,\mathrm{d}s\right).\end{aligned}$$

Dôkaz. Zrejme obe postulované riešenia sú spojité a vyhovujú počiatočnej podmienke.

(a) V homogénnom prípade platí

$$\boldsymbol{x}'(t) = \boldsymbol{F}'(t) \cdot \boldsymbol{F}(t_0)^{-1} \cdot \boldsymbol{c} = \boldsymbol{A}(t) \cdot \boldsymbol{F}(t) \cdot \boldsymbol{F}(t_0)^{-1} \cdot \boldsymbol{c} = \boldsymbol{A}(t) \cdot \boldsymbol{x}(t),$$

z čoho vyplýva aj spojitosť derivácie $\boldsymbol{x}'(t)$.

(b) V nehomogénnom prípade stačí overiť, že druhý sčítanec

$$\boldsymbol{y}(t) = \boldsymbol{F}(t) \cdot \int_{t_0}^{t} \boldsymbol{F}(s)^{-1} \cdot \boldsymbol{b}(s)\,\mathrm{d}s$$

výrazu pre $\boldsymbol{x}(t)$ je riešením sústavy. Ako už vieme, prvý sčítanec je totiž riešením homogénnej sústavy. S použitím pravidla pre deriváciu súčinu a Newtonovej-Leibnizovej formuly z tvrdení 22.4.1 resp. 22.4.6 vypočítame

$$\begin{aligned}
\boldsymbol{y}'(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{F}(t) \cdot \int_{t_0}^{t} \boldsymbol{F}(s)^{-1} \cdot \boldsymbol{b}(s)\,\mathrm{d}s\right) \\
&= \boldsymbol{F}'(t) \cdot \int_{t_0}^{t} \boldsymbol{F}(s)^{-1} \cdot \boldsymbol{b}(s)\,\mathrm{d}s \;+\; \boldsymbol{F}(t) \cdot \boldsymbol{F}(t)^{-1} \cdot \boldsymbol{b}(t) \\
&= \boldsymbol{A}(t) \cdot \boldsymbol{F}(t) \cdot \int_{t_0}^{t} \boldsymbol{F}(s)^{-1} \cdot \boldsymbol{b}(s)\,\mathrm{d}s \;+\; \boldsymbol{b}(t) = \boldsymbol{A}(t) \cdot \boldsymbol{y}(t) \;+\; \boldsymbol{b}(t).
\end{aligned}$$

Vidíme, že i funkcia $\boldsymbol{y}'$ je spojitá a $\boldsymbol{y}$ vyhovuje príslušnej rovnici.

Poznámka. Vzorec pre riešenie nehomogénnej sústavy sa odvodzuje tzv. *Lagrangeovou metódou variácie konštánt* tak, že vo vzorci pre riešenie homogénnej sústavy nahradíme konštatný vektor $\boldsymbol{c}$ vektorovou funkciou $\boldsymbol{q}\colon S \to \mathbb{C}^n$. Dosadením riešenia $\boldsymbol{x}(t) = \boldsymbol{F}(t) \cdot \boldsymbol{q}(t)$ do nehomogénnej počiatočnej úlohy nakoniec dospejeme k tvaru $\boldsymbol{q}(t) = \boldsymbol{F}(t_0)^{-1} \cdot \boldsymbol{c} \;+\; \int_{t_0}^{t} \boldsymbol{F}(s)^{-1} \cdot \boldsymbol{b}(s)\,\mathrm{d}s$.

Vo všeobecnosti nevieme vyjadriť fundamentálnu maticu sústavy $\boldsymbol{x}' = \boldsymbol{A} \cdot \boldsymbol{x}$ rozumným spôsobom pomocou maticovej funkcie $\boldsymbol{A}\colon S \to \mathbb{C}^{n\times n}$. Za istých dodatočných predpokladov si však dokážeme poradiť. Pre $s, t \in S$ označme

$$\boldsymbol{E}_{\boldsymbol{A}}(t, s) = \exp\left(\int_{s}^{t} \boldsymbol{A}(\tau)\,\mathrm{d}\tau\right).$$

Ľahko nahliadneme, že $\boldsymbol{E}_{\boldsymbol{A}}(t, t) = \boldsymbol{I}$ pre každé $t \in S$.

Hovoríme, že *maticová funkcia* $\boldsymbol{A}\colon S \to \mathbb{C}^{n\times n}$ je *komutujúca*, ak matice $\boldsymbol{A}(s)$, $\boldsymbol{A}(t)$ komutujú pre všetky $s, t \in S$.

22.5.3. Veta. *Nech $\boldsymbol{A}\colon S \to \mathbb{C}^{n\times n}$ je spojitá komutujúca maticová funkcia na intervale $S \subseteq \mathbb{R}$. Potom pre ľubovoľné pevné $t_0 \in S$ je*

$$\boldsymbol{F}(t) = \boldsymbol{E}_{\boldsymbol{A}}(t, t_0),$$

uvažovaná ako funkcia premennej $t \in S$, fundamentálna matica sústavy $\boldsymbol{x}' = \boldsymbol{A} \cdot \boldsymbol{x}$.

Dôkaz. Určitý integrál $\int_s^t \boldsymbol{A}(\tau)\,\mathrm{d}\tau$ splýva s limitou integrálnych súčtov

$$\lim_{k\to\infty}\sum_{j=1}^{k}\boldsymbol{A}(s+jd_k)d_k,$$

kde $d_k=(t-s)/k$. Keďže všetky uvažované sčítance komutujú s každou maticou $\boldsymbol{A}(u)$, $u\in S$, limitným prechodom dostaneme, že ľubovoľné dve z matíc $\boldsymbol{A}(t)$, $\int_s^t \boldsymbol{A}(\tau)\,\mathrm{d}\tau$, $\int_t^u \boldsymbol{A}(\tau)\,\mathrm{d}\tau$ tiež komutujú. Preto podľa tvrdenia 22.2.2 platí

$$\boldsymbol{E}_{\boldsymbol{A}}(u,t)\cdot\boldsymbol{E}_{\boldsymbol{A}}(t,s)=\exp\left(\int_t^u \boldsymbol{A}(\tau)\,\mathrm{d}\tau+\int_s^t \boldsymbol{A}(\tau)\,\mathrm{d}\tau\right)=\boldsymbol{E}_{\boldsymbol{A}}(u,s).$$

Z toho už vyplýva, že matica $\boldsymbol{E}_{\boldsymbol{A}}(t,s)$ je vždy regulárna s inverznou maticou

$$\boldsymbol{E}_{\boldsymbol{A}}(t,s)^{-1}=\boldsymbol{E}_{\boldsymbol{A}}(s,t).$$

Pomocou tvrdení 22.4.3 a 22.4.6 môžeme počítať

$$\begin{aligned}\boldsymbol{F}'(t)&=\frac{\mathrm{d}}{\mathrm{d}t}\exp\left(\int_{t_0}^t \boldsymbol{A}(\tau)\,\mathrm{d}\tau\right)=\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{t_0}^t \boldsymbol{A}(\tau)\,\mathrm{d}\tau\right)\cdot\exp\left(\int_{t_0}^t \boldsymbol{A}(\tau)\,\mathrm{d}\tau\right)\\&=\boldsymbol{A}(t)\cdot\boldsymbol{E}_{\boldsymbol{A}}(t,t_0)=\boldsymbol{A}(t)\cdot\boldsymbol{F}(t).\end{aligned}$$

22.5.4. Dôsledok. *Pre spojitú komutujúcu maticovú funkciu $\boldsymbol{A}\colon S\to\mathbb{C}^{n\times n}$ platí:*

(a) *riešenie homogénnej počiatočnej úlohy $\boldsymbol{x}'=\boldsymbol{A}\cdot\boldsymbol{x}$, $\boldsymbol{x}(t_0)=\boldsymbol{c}$ má tvar*

$$\boldsymbol{x}(t)=\boldsymbol{E}_{\boldsymbol{A}}(t,t_0)\cdot\boldsymbol{c}\,;$$

(b) *riešenie nehomogénnej počiatočnej úlohy $\boldsymbol{x}'(t)=\boldsymbol{A}\cdot\boldsymbol{x}+\boldsymbol{b}$, $\boldsymbol{x}(t_0)=\boldsymbol{c}$ má tvar*

$$\boldsymbol{x}(t)=\boldsymbol{E}_{\boldsymbol{A}}(t,t_0)\cdot\boldsymbol{c}+\int_{t_0}^t \boldsymbol{E}_{\boldsymbol{A}}(t,s)\cdot\boldsymbol{b}(s)\,\mathrm{d}s.$$

I v takom špeciálnom prípade sme v riešení sústav lineárnych diferenciálnych rovníc stále iba na pol ceste. Niečo iné je totiž napísať pekné všeobecné vzorce pre ich riešenie (dôsledne vzaté, zaviesť preň len isté vhodné označenie) a niečo iné nájsť riešenia konkrétnych sústav. To si vyžaduje vypočítat maticové funkcie $\boldsymbol{E}_{\boldsymbol{A}}(t,t_0)$ a $\int_{t_0}^t \boldsymbol{E}_{\boldsymbol{A}}(t,s)\cdot\boldsymbol{b}(s)\,\mathrm{d}s$ pre rôzne typy komutujúcich maticových funkcií $\boldsymbol{A}(t)$ a vektorových funkcií $\boldsymbol{b}(t)$. Také niečo však vieme len v určitých prípadoch. Tým najjednoduchším, no veľmi dôležitým, keď $\boldsymbol{A}$ je konštantná funkcia, sa budeme podrobnejšie zaoberať v nasledujúcom paragrafe, čím náš výlet do sveta diferenciálnych rovníc zakončíme.

22.6 Autonómne sústavy

Hovoríme, že *sústava lineárnych diferenciálnych rovníc* $\boldsymbol{x}' = \boldsymbol{A} \cdot \boldsymbol{x} + \boldsymbol{b}$ je *autonómna*, prípadne *sústava s konštantnými koeficientmi*, ak $\boldsymbol{A} \in \mathbb{C}^{n \times n}$, t. j. $\boldsymbol{A}$ je konštantná funkcia, a $\boldsymbol{b}\colon S \to \mathbb{C}^n$ je spojitá vektorová funkcia na intervale $S \subseteq \mathbb{R}$.

Keďže konštantná maticová funkcia $\boldsymbol{A}$ je automaticky spojitá a komutujúca, fundamentálna matica homogénnej autonómnej sústavy $\boldsymbol{x}' = \boldsymbol{A} \cdot \boldsymbol{x}$ je podľa vety 22.5.3 daná formulou

$$\boldsymbol{F}(t) = \boldsymbol{E}_{\boldsymbol{A}}(t, t_0) = \exp\left(\int_{t_0}^{t} \boldsymbol{A} \,\mathrm{d}\tau\right) = \mathrm{e}^{\boldsymbol{A}(t-t_0)},$$

pre ľubovoľné pevné $t_0 \in \mathbb{R}$ a premenné $t \in \mathbb{R}$. Špeciálne je fundamentálnou maticou takejto sústavy funkcia

$$\boldsymbol{F}(t) = \mathrm{e}^{\boldsymbol{A}t}.$$

K tomuto záveru možno dôjsť aj priamo na základe dôsledku 22.4.4. Z dôsledku 22.5.4 vyplýva

22.6.1. Veta. *Nech $\boldsymbol{A} \in \mathbb{C}^{n \times n}$, $\boldsymbol{b}\colon S \to \mathbb{C}^n$ je spojitá vektorová funkcia na intervale $S \subseteq \mathbb{R}$, $t_0 \in S$ a $\boldsymbol{c} \in \mathbb{C}^n$. Potom*

(a) riešenie homogénnej počiatočnej úlohy $\boldsymbol{x}' = \boldsymbol{A} \cdot \boldsymbol{x}$, $\boldsymbol{x}(t_0) = \boldsymbol{c}$ má tvar

$$\boldsymbol{x}(t) = \mathrm{e}^{\boldsymbol{A}(t-t_0)} \cdot \boldsymbol{c};$$

(b) riešenie nehomogénnej počiatočnej úlohy $\boldsymbol{x}'(t) = \boldsymbol{A} \cdot \boldsymbol{x} + \boldsymbol{b}$, $\boldsymbol{x}(t_0) = \boldsymbol{c}$ má tvar

$$\boldsymbol{x}(t) = \mathrm{e}^{\boldsymbol{A}(t-t_0)} \cdot \boldsymbol{c} + \int_{t_0}^{t} \mathrm{e}^{\boldsymbol{A}(t-s)} \cdot \boldsymbol{b}(s) \,\mathrm{d}s;$$

(c) ak aj $\boldsymbol{b}$ je konštantná funkcia, tak riešením nehomogénnej počiatočnej úlohy je

$$\boldsymbol{x}(t) = \mathrm{e}^{\boldsymbol{A}(t-t_0)} \cdot \boldsymbol{c} + \left(\int_{t_0}^{t} \mathrm{e}^{\boldsymbol{A}(t-s)} \,\mathrm{d}s\right) \cdot \boldsymbol{b}.$$

Na dovŕšenie riešenia autonómnej homogénnej úlohy ako aj popisu jej fundamentálnej matice stačí vedieť vypočítať maticu $\mathrm{e}^{\boldsymbol{A}t}$ pre ľubovoľné $t \in \mathbb{R}$.

Ak $\boldsymbol{J} = \operatorname{diag}\big(\boldsymbol{J}_{n_1}(\lambda_1), \ldots, \boldsymbol{J}_{n_k}(\lambda_k)\big)$ je JKT matice $\boldsymbol{A}$ a $\boldsymbol{P} \in \mathbb{C}^{n \times n}$ je regulárna matica, pre ktorú $\boldsymbol{A} = \boldsymbol{P} \cdot \boldsymbol{J} \cdot \boldsymbol{P}^{-1}$, tak

$$\mathrm{e}^{\boldsymbol{A}t} = \boldsymbol{P} \cdot \operatorname{diag}\Big(\exp\big(\boldsymbol{J}_{n_1}(\lambda_1)t\big), \ldots, \exp\big(\boldsymbol{J}_{n_k}(\lambda_k)t\big)\Big) \cdot \boldsymbol{P}^{-1}.$$

Úloha sa teda redukuje na výpočet matíc $\exp\bigl(\boldsymbol{J}_n(\lambda)t\bigr)$. Pri pevnom $t \in \mathbb{R}$ uvažujme mocninný rad

$$f_t(x) = \sum_{k=0}^{\infty} \frac{t^k}{k!} x^k = \mathrm{e}^{xt}$$

v premennej x. Potom $f_t^{(p)}(x) = t^p\,\mathrm{e}^{xt}$ pre každé $p \in \mathbb{N}$. Dosadením do vzorca z vety 22.1.5 tak dostávame nasledujúce zovšeobecnenie tvrdenia 22.2.5:

22.6.2. Veta. *Nech $n \geq 1$, $\lambda \in \mathbb{C}$. Potom pre ľubovoľné $t \in \mathbb{R}$ (dokonca $t \in \mathbb{C}$) platí*

$$\exp\bigl(\boldsymbol{J}_n(\lambda)t\bigr) = \begin{pmatrix} \mathrm{e}^{\lambda t} & \frac{t\,\mathrm{e}^{\lambda t}}{1!} & \dots & \frac{t^{n-1}\,\mathrm{e}^{\lambda t}}{(n-1)!} \\ 0 & \mathrm{e}^{\lambda t} & \dots & \frac{t^{n-2}\,\mathrm{e}^{\lambda t}}{(n-2)!} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathrm{e}^{\lambda t} \end{pmatrix} = \begin{pmatrix} 1 & \frac{t}{1!} & \dots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & \dots & \frac{t^{n-2}}{(n-2)!} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \mathrm{e}^{\lambda t}.$$

V prípade komplexných vlastných čísel reálnej matice možno dvojicu Jordanových buniek $\boldsymbol{J}_n(\lambda)$, $\boldsymbol{J}_n(\overline{\lambda})$ nahradiť zovšeobecnenou Jordanovou bunkou $\boldsymbol{J}_n\bigl(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\bigr)$, kde $a = \operatorname{Re}\lambda$, $b = \operatorname{Im}\lambda$ (pozri paragraf 20.3). Nasledujúca veta, ktorú uvádzame bez dôkazu, umožňuje nahradiť vo fundamentálnej matici komplexné funkcie $\mathrm{e}^{\lambda t}$ reálnymi funkciami $\mathrm{e}^{at}\cos bt$, $\mathrm{e}^{at}\sin bt$.

22.6.3. Veta. *Nech $n \geq 1$, $a, b \in \mathbb{R}$. Potom pre ľubovoľné $t \in \mathbb{R}$ platí*

$$\exp\begin{pmatrix} a & -b \\ b & a \end{pmatrix} t = \begin{pmatrix} \mathrm{e}^{at}\cos bt & -\mathrm{e}^{at}\sin bt \\ \mathrm{e}^{at}\sin bt & \mathrm{e}^{at}\cos bt \end{pmatrix} = \mathrm{e}^{at}\,\boldsymbol{R}_{bt},$$

$$\exp \boldsymbol{J}_n\bigl(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\bigr)t = \begin{pmatrix} \mathrm{e}^{at}\,\boldsymbol{R}_{bt} & \frac{t}{1!}\,\mathrm{e}^{at}\,\boldsymbol{R}_{bt} & \dots & \frac{t^{n-1}}{(n-1)!}\,\mathrm{e}^{at}\,\boldsymbol{R}_{bt} \\ \boldsymbol{0} & \mathrm{e}^{at}\,\boldsymbol{R}_{bt} & \dots & \frac{t^{n-2}}{(n-2)!}\,\mathrm{e}^{at}\,\boldsymbol{R}_{bt} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \dots & \mathrm{e}^{at}\,\boldsymbol{R}_{bt} \end{pmatrix} =$$

$$\begin{pmatrix} \mathrm{e}^{at}\cos bt & -\mathrm{e}^{at}\sin bt & \frac{t\,\mathrm{e}^{at}\cos bt}{1!} & \frac{-t\,\mathrm{e}^{at}\sin bt}{1!} & \dots & \frac{t^{n-1}\,\mathrm{e}^{at}\cos bt}{(n-1)!} & \frac{-t^{n-1}\,\mathrm{e}^{at}\sin bt}{(n-1)!} \\ \mathrm{e}^{at}\sin bt & \mathrm{e}^{at}\cos bt & \frac{t\,\mathrm{e}^{at}\sin bt}{1!} & \frac{t\,\mathrm{e}^{at}\cos bt}{1!} & \dots & \frac{t^{n-1}\,\mathrm{e}^{at}\sin bt}{(n-1)!} & \frac{t^{n-1}\,\mathrm{e}^{at}\cos bt}{(n-1)!} \\ 0 & 0 & \mathrm{e}^{at}\cos bt & -\mathrm{e}^{at}\sin bt & \dots & \frac{t^{n-2}\,\mathrm{e}^{at}\cos bt}{(n-2)!} & \frac{-t^{n-2}\,\mathrm{e}^{at}\sin bt}{(n-2)!} \\ 0 & 0 & \mathrm{e}^{at}\sin bt & \mathrm{e}^{at}\cos bt & \dots & \frac{t^{n-2}\,\mathrm{e}^{at}\sin bt}{(n-2)!} & \frac{t^{n-2}\,\mathrm{e}^{at}\cos bt}{(n-2)!} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \mathrm{e}^{at}\cos bt & -\mathrm{e}^{at}\sin bt \\ 0 & 0 & 0 & 0 & \dots & \mathrm{e}^{at}\sin bt & \mathrm{e}^{at}\cos bt \end{pmatrix}$$

Na riešenie nehomogénnej autonómnej úlohy s konštantným členom $\boldsymbol{b}$ treba ešte vedieť vypočítať integrál $\int_{t_0}^{t} \mathrm{e}^{\boldsymbol{A}(t-s)}\,\mathrm{d}s = \int_0^{t-t_0} \mathrm{e}^{\boldsymbol{A}s}\,\mathrm{d}s$. Stačí sa teda obmedziť na integrály tvaru $\int_0^t \mathrm{e}^{\boldsymbol{A}s}\,\mathrm{d}s$. A znovu stará známa pesnička: ak $\boldsymbol{A} = \boldsymbol{P}\cdot\boldsymbol{J}\cdot\boldsymbol{P}^{-1}$, kde $\boldsymbol{P}$ je regulárna matica a $\boldsymbol{J} = \operatorname{diag}\bigl(\boldsymbol{J}_{n_1}(\lambda_1),\dots,\boldsymbol{J}_{n_k}(\lambda)\bigr)$ je JKT matice $\boldsymbol{A}$, tak

$$\int_0^t \mathrm{e}^{\boldsymbol{A}s}\,\mathrm{d}s = \boldsymbol{P}\cdot\operatorname{diag}\Bigl(\int_0^t \exp\bigl(\boldsymbol{J}_{n_1}(\lambda_1)s\bigr)\mathrm{d}s,\dots,\int_0^t \exp\bigl(\boldsymbol{J}_{n_k}(\lambda_k)s\bigr)\mathrm{d}s\Bigr)\cdot\boldsymbol{P}^{-1},$$

takže opäť stačí poznať príslušné integrály pre Jordanove bunky.

22.6.4. Veta. *Nech $n \ge 1$, $0 \ne \lambda \in \mathbb{C}$. Potom pre ľubovoľné $t \in \mathbb{R}$ platí*

$$\int_0^t \exp\bigl(\boldsymbol{J}_n(\lambda)s\bigr)\,\mathrm{d}s =$$

$$\begin{pmatrix} \frac{\mathrm{e}^{\lambda t}}{\lambda} & \frac{t\,\mathrm{e}^{\lambda t}}{\lambda\,1!} - \frac{\mathrm{e}^{\lambda t}}{\lambda^2} & \dots & \sum_{j=0}^{n-1} \frac{(-1)^j t^{n-1-j}\,\mathrm{e}^{\lambda t}}{\lambda^{j+1}(n-1-j)!} \\ 0 & \frac{\mathrm{e}^{\lambda t}}{\lambda} & \dots & \sum_{j=0}^{n-2} \frac{(-1)^j t^{n-2-j}\,\mathrm{e}^{\lambda t}}{\lambda^{j+1}(n-2-j)!} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\mathrm{e}^{\lambda t}}{\lambda} \end{pmatrix} - \begin{pmatrix} \frac{1}{\lambda} & -\frac{1}{\lambda^2} & \dots & \frac{(-1)^{n-1}}{\lambda^n} \\ 0 & \frac{1}{\lambda} & \dots & \frac{(-1)^{n-2}}{\lambda^{n-1}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\lambda} \end{pmatrix};$$

pre $\lambda = 0$ máme

$$\int_0^t \exp\bigl(\boldsymbol{J}_n(0)s\bigr)\,\mathrm{d}s = \begin{pmatrix} t & \frac{t^2}{2!} & \dots & \frac{t^n}{n!} \\ 0 & t & \dots & \frac{t^{n-1}}{(n-1)!} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & t \end{pmatrix}.$$

Náčrt dôkazu. Oba vzorce dostaneme priamo z vety 22.6.2 integrovaním jednotlivých členov matice $\exp\bigl(\boldsymbol{J}_n(\lambda)s\bigr)$. Pre $\lambda \ne 0$, $0 \le p \le n-1$, pri výpočte integrálu

$$\int_0^t \frac{s^p\,\mathrm{e}^{\lambda s}}{p!}\,\mathrm{d}s = \sum_{j=0}^{p} \frac{(-1)^j t^{p-j}\,\mathrm{e}^{\lambda t}}{\lambda^{j+1}(p-j)!} - \frac{(-1)^p}{\lambda^{p+1}}$$

použijeme p-krát za sebou metódu *per partes*.

Ak $\boldsymbol{A}$ je regulárna, môžeme sa uvedeným vzorcom vyhnúť. Ľahko totiž nahliadneme, že $\frac{\mathrm{d}}{\mathrm{d}s}\bigl(\mathrm{e}^{\boldsymbol{A}s}\cdot\boldsymbol{A}^{-1}\bigr) = \mathrm{e}^{\boldsymbol{A}s}$, teda podľa tvrdenia 22.4.6

$$\int_0^t \mathrm{e}^{\boldsymbol{A}s}\,\mathrm{d}s = \bigl(\mathrm{e}^{\boldsymbol{A}t} - \boldsymbol{I}\bigr)\cdot\boldsymbol{A}^{-1}.$$

To znamená, že výpočet hľadaného integrálu možno priamo previesť na výpočet exponenciály $\mathrm{e}^{\boldsymbol{A}t}$ podľa vety 22.6.2 a úvahy, ktorá ju predchádza. Z časti (c) vety 22.6.1 teraz vyplýva

22.6.5. Dôsledok. *Nech $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ je regulárna matica, $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{C}^n$. Potom riešenie nehomogénnej autonómnej počiatočnej úlohy $\boldsymbol{x}' = \boldsymbol{A}\cdot\boldsymbol{x} + \boldsymbol{b}$, $\boldsymbol{x}(t_0) = \boldsymbol{c}$ má tvar*

$$\boldsymbol{x}(t) = \mathrm{e}^{\boldsymbol{A}(t-t_0)} \cdot \boldsymbol{c} + \left(\mathrm{e}^{\boldsymbol{A}(t-t_0)} - \boldsymbol{I}\right) \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{b}.$$

V poznámke (b) za vetou 22.5.1 sme sa stručne zmienili o možnosti previesť (jednorozmernú) lineárnu diferenciálnu rovnicu n-tého rádu na vektorovú diferenciálnu rovnicu. V špeciálnom prípade homogénnej rovnice s konštantnými koeficientmi a_i dostaneme homogénnu autonómnu sústavu

$$\boldsymbol{x}' = \boldsymbol{M}_g^{\mathsf{T}} \cdot \boldsymbol{x},$$

kde $g(x) = x^n - a_1 x^{n-1} - \ldots - a_{n-1}x - a_n$. Niekedy môže byť naopak užitočné „znížiť počet rovníc a neznámych funkcií“, presnejšie, pre $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ upraviť sústavu $\boldsymbol{x}' = \boldsymbol{A}\cdot\boldsymbol{x}$ na jednorozmernú lineárnu diferenciálnu rovnicu n-tého rádu. Rôzne ekvivalentné podoby podmienky, za ktorej je také niečo možné, sú sformulované v tvrdení 21.3.1. My si vyberieme len jednu z nich.

22.6.6. Tvrdenie. *Nech $\boldsymbol{A} \in \mathbb{C}^{n\times n}$. Ak $f(x) = x^n - \sum_{j=1}^n c_j x^{n-j} \in \mathbb{C}[x]$ je normovaný polynóm taký, že $\boldsymbol{A}^{\mathsf{T}} \approx \boldsymbol{M}_f$, tak autonómnu homogénnu sústavu $\boldsymbol{x}' = \boldsymbol{A}\cdot\boldsymbol{x}$ možno vhodnou substitúciou upraviť na homogénnu lineárnu diferenciálnu rovnicu n-tého rádu*

$$y^{(n)} - c_1 y^{(n-1)} - \ldots - c_{n-1}y' - c_n y = 0.$$

Dôkaz. Za uvedených predpokladov platí $\boldsymbol{A} \approx \boldsymbol{M}_f^{\mathsf{T}}$, preto existuje regulárna matica $\boldsymbol{P} \in \mathbb{C}^{n\times n}$ taká, že $\boldsymbol{A} = \boldsymbol{P}\cdot\boldsymbol{M}_f^{\mathsf{T}}\cdot\boldsymbol{P}^{-1}$. Položme $\boldsymbol{y} = \boldsymbol{P}^{-1}\cdot\boldsymbol{x}$. Potom $\boldsymbol{x}$ vyhovuje sústave $\boldsymbol{x}' = \boldsymbol{A}\cdot\boldsymbol{x}$ práve vtedy, keď $\boldsymbol{y}$ vyhovuje sústave $\boldsymbol{y}' = \boldsymbol{M}_f^{\mathsf{T}}\cdot\boldsymbol{y}$, t. j.

$$y_1' = c_1 y_1 + \ldots + c_n y_n, \qquad y_2' = y_1, \quad \ldots, \quad y_n' = y_{n-1}.$$

Teda pre $y = y_n$, $1 \le j \le n$, platí $y_j = y^{(n-j)}$ a $y^{(n)} - \sum_{j=1}^n c_j y^{(n-j)} = 0$.

22.6.7. Dôsledok. *Nech $\boldsymbol{A} \in \mathbb{C}^{n\times n}$. Ak $f_1(x), \ldots, f_k(x) \in \mathbb{C}[x]$ sú normované polynómy také, že $f_i(x) = x^{n_i} - \sum_{j=1}^{n_i} c_{ij}x^{n_i-j}$ a $\boldsymbol{A}^{\mathsf{T}} \approx \operatorname{diag}\left(\boldsymbol{M}_{f_1}, \ldots, \boldsymbol{M}_{f_k}\right)$, tak autonómnu homogénnu sústavu $\boldsymbol{x}' = \boldsymbol{A}\cdot\boldsymbol{x}$ n lineárnych diferenciálnych rovníc pre n neznámych funkcií $x_1, \ldots, x_n$ možno vhodnou substitúciou upraviť na sústavu k homogénnych lineárnych diferenciálnych rovníc*

$$\begin{aligned} y_1^{(n_1)} - c_{11}y_1^{(n_1-1)} - \cdots - c_{1\,n_1-1}y_1' - c_{1n_1}y_1 &= 0,\\ &\vdots\\ y_k^{(n_k)} - c_{k1}y_k^{(n_k-1)} - \cdots - c_{k\,n_k-1}y_k' - c_{k\,n_k}y_k &= 0, \end{aligned}$$

rádov $n_1, \ldots, n_k$ pre k neznámych funkcií $y_1, \ldots, y_k$.

Doplnenie konkrétneho tvaru takejto substitúcie už prenechávame na rozmyslenie čitateľovi. Ešte poznamenajme, že pokiaľ chceme dosiahnuť čo najvýraznejšie zníženie počtu rovníc a neznámych funkcií, najvhodnejším kandidátom na polynómy $f_1(x), \dots, f_k(x)$ je sústava invariantných faktorov matice $\boldsymbol{A}^{\mathsf{T}}$ (pozri paragraf 21.4, úsek medzi vetami 21.4.4, 21.4.5). Potom matica $\operatorname{diag}\big(\boldsymbol{M}_{f_1}, \dots, \boldsymbol{M}_{f_k}\big)$ je racionálny kanonický tvar matice $\boldsymbol{A}^{\mathsf{T}}$.

22.7 Komutátor

V predchádzajúcej i celej tejto kapitole sme mohli vidieť, akú významnú úlohu hrá pri štúdiu maticových funkcií vzťah komutovania. V tomto záverečnom paragrafe zavedieme pojem komutátora matíc. Pomocou neho, po sérii aplikácii lineárnej algebry v teórii diferenciálnych rovníc, pre zmenu predvedieme jednu aplikáciu vety 22.5.1 o jednoznačnosti riešenia sústavy lineárnych diferenciálnych rovníc v lineárnej algebre.

Komutátorom matíc $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$ (nad ľubovoľným poľom K) nazývame maticu

$$[\boldsymbol{A}, \boldsymbol{B}] = \boldsymbol{A} \cdot \boldsymbol{B} - \boldsymbol{B} \cdot \boldsymbol{A}.$$

Zrejme $\boldsymbol{A}$, $\boldsymbol{B}$ komutujú práve vtedy, keď $[\boldsymbol{A}, \boldsymbol{B}] = \boldsymbol{0}$.

Priamym výpočtom možno overiť, že komutátor je antisymetrické bilineárne zobrazenie $K^{n\times n} \times K^{n\times n} \to K^{n\times n}$ a lineárny operátor $\operatorname{ad}_{\boldsymbol{A}} = [\boldsymbol{A}, \cdot\,] \colon K^{n\times n} \to K^{n\times n}$, ktorý dostaneme zafixovaním prvej zložky, sa správa voči súčinu matíc podobne ako derivácia. Analogicky chová aj lineárny operátor $-\operatorname{ad}_{\boldsymbol{B}} = [\,\cdot, \boldsymbol{B}] \colon K^{n\times n} \to K^{n\times n}$.

22.7.1. Tvrdenie. *Nech K je pole a $n \in \mathbb{N}$. Potom pre ľubovoľné $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in K^{n\times n}$, $c, d \in K$ platí*

$$\begin{aligned} [\boldsymbol{A}, \boldsymbol{B}] &= -[\boldsymbol{B}, \boldsymbol{A}], \\ [\boldsymbol{A}, c\boldsymbol{B} + d\boldsymbol{C}] &= c[\boldsymbol{A}, \boldsymbol{B}] + d[\boldsymbol{A}, \boldsymbol{C}], \\ [\boldsymbol{A}, \boldsymbol{B} \cdot \boldsymbol{C}] &= [\boldsymbol{A}, \boldsymbol{B}] \cdot \boldsymbol{C} + \boldsymbol{B} \cdot [\boldsymbol{A}, \boldsymbol{C}]. \end{aligned}$$

22.7.2. Dôsledok. *Nech $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$, $f(x) \in K[x]$. Ak matice $\boldsymbol{A}$, $[\boldsymbol{A}, \boldsymbol{B}]$ komutujú, tak*

$$[f(\boldsymbol{A}), \boldsymbol{B}] = [\boldsymbol{A}, \boldsymbol{B}] \cdot f'(\boldsymbol{A}) = f'(\boldsymbol{A}) \cdot [\boldsymbol{A}, \boldsymbol{B}].$$

V prípade $K = \mathbb{C}$ platí uvedený vzťah nielen pre polynómy ale aj pre potenčné rady $f(x) \in \mathbb{C}[[x]]$ a matice $\boldsymbol{A} \in \operatorname{Dom}_n f$, $\boldsymbol{B} \in \mathbb{C}^{n\times n}$.

Dôkaz. S použitím antisymetrie komutátora a poslednej rovnosti z tvrdenia 22.7.1 možno za uvedeného predpokladu indukciou cez $k \in \mathbb{N}$ jednoducho dokázať

$$[\boldsymbol{A}^k, \boldsymbol{B}] = k\boldsymbol{A}^{k-1} \cdot [\boldsymbol{A}, \boldsymbol{B}].$$

Potrebný záver pre polynómy vyplýva z linearity zobrazenia $\mathrm{ad}_{\boldsymbol{B}} \colon \boldsymbol{A} \mapsto [\boldsymbol{A}, \boldsymbol{B}]$; na rady sa prenesie limitným prechodom.

Podľa tvrdenia 22.2.2 pre *komutujúce* matice $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$ platí $\mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}} = \mathrm{e}^{\boldsymbol{A}} \cdot \mathrm{e}^{\boldsymbol{B}}$. Pre nekomutujúce matice to však nie je pravda. I v prípade, že $\boldsymbol{A}$, $\boldsymbol{B}$ nekomutujú, za predpokladu, že komutujú aspoň so svojím komutátorom, možno vzorec pre exponenciálu súčtu modifikovať do nasledujúcej podoby.

22.7.3. Veta. *Nech matice $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$ komutujú so svojim komutátorom, t. j.*

$$[\boldsymbol{A}, [\boldsymbol{A}, \boldsymbol{B}]] = [[\boldsymbol{A}, \boldsymbol{B}], \boldsymbol{B}] = \boldsymbol{0}.$$

Potom

$$\mathrm{e}^{\boldsymbol{A}} \cdot \mathrm{e}^{\boldsymbol{B}} = \exp(\boldsymbol{A} + \boldsymbol{B}) \cdot \exp \frac{1}{2}[\boldsymbol{A}, \boldsymbol{B}].$$

Dôkaz. Uvažujme ešte raz mocninný rad $f_t(x) = \sum \frac{t^k}{k!} x^k = \mathrm{e}^{xt} \in \mathbb{C}[[x]]$ pri pevnom $t \in \mathbb{R}$. Potom $f_t'(x) = t\,\mathrm{e}^{xt}$. Ak $\boldsymbol{A}$ a $[\boldsymbol{A}, \boldsymbol{B}]$ komutujú, tak podľa dôsledku 22.7.2 platí

$$\left[\mathrm{e}^{\boldsymbol{A}t}, \boldsymbol{B}\right] = [f_t(\boldsymbol{A}), \boldsymbol{B}] = f_t'(\boldsymbol{A}) \cdot [\boldsymbol{A}, \boldsymbol{B}] = t\,\mathrm{e}^{\boldsymbol{A}t} \cdot [\boldsymbol{A}, \boldsymbol{B}],$$

t. j.

$$\mathrm{e}^{\boldsymbol{A}t} \cdot \boldsymbol{B} = \big(\boldsymbol{B} + [\boldsymbol{A}, \boldsymbol{B}]t\big) \cdot \mathrm{e}^{\boldsymbol{A}t}.$$

Uvažujme teraz spojité maticové funkcie

$$\boldsymbol{Y}(t) = \mathrm{e}^{\boldsymbol{A}t} \cdot \mathrm{e}^{\boldsymbol{B}t}, \qquad \boldsymbol{Z}(t) = \exp(\boldsymbol{A} + \boldsymbol{B})t \cdot \exp \frac{1}{2}[\boldsymbol{A}, \boldsymbol{B}]t^2,$$

pre $t \in \mathbb{R}$. Zrejme $\boldsymbol{Y}(0) = \boldsymbol{Z}(0) = \boldsymbol{I}_n$. Pozrime sa bližšie na ich derivácie. S využitím tvrdení 22.4.1, 22.4.3 a komutačného vzťahu pre $\mathrm{e}^{\boldsymbol{A}t}$ a $\boldsymbol{B}$ dostávame

$$\begin{aligned}
\boldsymbol{Y}'(t) &= \big(\mathrm{e}^{\boldsymbol{A}t}\big)' \cdot \mathrm{e}^{\boldsymbol{B}t} + \mathrm{e}^{\boldsymbol{A}t} \cdot \big(\mathrm{e}^{\boldsymbol{B}t}\big)' = \boldsymbol{A} \cdot \mathrm{e}^{\boldsymbol{A}t} \cdot \mathrm{e}^{\boldsymbol{B}t} + \mathrm{e}^{\boldsymbol{A}t} \cdot \boldsymbol{B} \cdot \mathrm{e}^{\boldsymbol{B}t} \\
&= \boldsymbol{A} \cdot \mathrm{e}^{\boldsymbol{A}t} \cdot \mathrm{e}^{\boldsymbol{B}t} + \big(\boldsymbol{B} + [\boldsymbol{A}, \boldsymbol{B}]t\big) \cdot \mathrm{e}^{\boldsymbol{A}t} \cdot \mathrm{e}^{\boldsymbol{B}t} = \big(\boldsymbol{A} + \boldsymbol{B} + [\boldsymbol{A}, \boldsymbol{B}]t\big) \cdot \boldsymbol{Y}(t).
\end{aligned}$$

Ak si uvedomíme, že za daných predpokladov komutujú $\boldsymbol{A}+\boldsymbol{B}$ a $[\boldsymbol{A}, \boldsymbol{B}]$, teda aj $(\boldsymbol{A}+\boldsymbol{B})t$ a $\frac{1}{2}[\boldsymbol{A}, \boldsymbol{B}]t^2$, resp. $(\boldsymbol{A}+\boldsymbol{B})t + \frac{1}{2}[\boldsymbol{A}, \boldsymbol{B}]t^2$ a $\frac{\mathrm{d}}{\mathrm{d}t}\big((\boldsymbol{A}+\boldsymbol{B})t + \frac{1}{2}[\boldsymbol{A}, \boldsymbol{B}]t^2\big) = \boldsymbol{A} + \boldsymbol{B} + [\boldsymbol{A}, \boldsymbol{B}]t$, s použitím tvrdení 22.2.2, 22.4.3 nám postupne vyjde

$$\begin{aligned}
\boldsymbol{Z}(t) &= \exp\big((\boldsymbol{A} + \boldsymbol{B})t + \frac{1}{2}[\boldsymbol{A}, \boldsymbol{B}]t^2\big), \\
\boldsymbol{Z}'(t) &= \big(\boldsymbol{A} + \boldsymbol{B} + [\boldsymbol{A}, \boldsymbol{B}]t\big) \cdot \boldsymbol{Z}(t).
\end{aligned}$$

Vidíme, že $\boldsymbol{Y}$, $\boldsymbol{Z}$ vyhovujú rovnakej počiatočnej úlohe $\boldsymbol{X}' = \big(\boldsymbol{A}+\boldsymbol{B}+[\boldsymbol{A},\boldsymbol{B}]t\big)\cdot \boldsymbol{X}$, $\boldsymbol{X}(0) = \boldsymbol{I}$. Z jednoznačnosti riešenia takejto úlohy (pozri poznámku (a) za vetou 22.5.1) vyplýva, že $\boldsymbol{Y}(t) = \boldsymbol{Z}(t)$ pre každé $t \in \mathbb{R}$. Voľbou $t = 1$ dostávame požadovanú rovnosť.

Nekonečnorozmerný analóg práve dokázaného vzťahu hrá významnú úlohu v kvantovej mechanike. Ako uvidíme v kapitole 26, pre lineárne operátory X a P, reprezentujúce fyzikálne veličiny polohy a hybnosti, je totiž $[X,P] = \mathrm{i}\hbar\cdot\mathrm{id}$, teda rýdzo imaginárny násobok identického operátora. Keďže skalárne násobky identity komutujú s každým operátorom, (po vhodnom zakamuflovaní fyzikálnych jednotiek) máme

$$\mathrm{e}^X\cdot\mathrm{e}^P = \exp(X+P)\cdot\exp\left(\frac{\mathrm{i}\hbar}{2}\,\mathrm{id}\right) = \mathrm{e}^{\mathrm{i}\hbar/2}\exp(X+P).$$

To znamená, že lineárne operátory $\mathrm{e}^X\cdot\mathrm{e}^P$ a $\exp(X+P)$ sa líšia len násobkom komplexnej jednotky $\mathrm{e}^{\mathrm{i}\hbar/2} = \cos(\hbar/2) + \mathrm{i}\sin(\hbar/2)$, t. j. fázovým posunom o „uhol“ $\hbar/2$, kde $\hbar = h/2\pi \approx 1{,}055\cdot 10^{-34}\,\mathrm{J\,s}$ je tzv. *redukovaná Planckova konštanta.* [7]

Cvičenia

22.1. Dokážte, že všetky formálne potenčné rady tvoria vektorový priestor $K[[x]]$ nad poľom K, všetky polynómy nad K tvoria lineárny podpriestor $K[x]$ v $K[[x]]$, pričom platí $\dim K[x] = \dim K[[x]] = \infty$, ale $K[x] \neq K[[x]]$.

22.2. Odvoďte konvolučnú formulu $c_k = \sum_{i=0}^k a_i b_{k-i}$ pre koeficienty súčinu $f(x)\,g(x) = \sum c_k x^k$ potenčných radov $f(x) = \sum a_k x^k$, $g(x) = \sum b_k x^k$. Dokážte, že pre polomer r konvergencie radu $f(x)\,g(x)$ platí $r \geq \min\{p,q\}$, kde p, q sú polomery konvergencie radov $f(x)$ resp. $g(x)$.

22.3. Dokážte postupne nasledujúce tvrdenia:

(a) Rad $\sum_{k=0}^\infty x^k$, kde $x \in \mathbb{C}$, konverguje práve vtedy, keď $|x| < 1$, teda diverguje na celej hraničnej kružnici svojho oboru konvergencie.

(b) Rad $\sum_{k=1}^\infty \frac{x^k}{k}$ konverguje pre každé $|x| \leq 1$ okrem $x = 1$, teda na celej hraničnej kružnici svojho oboru konvergencie s výnimkou jedného bodu (využite fakt, že tzv. *harmonický rad* $\sum_{k=1}^\infty \frac{1}{k}$ diverguje).

(c) Rad $\sum_{k=1} \frac{x^k}{k^2}$ má polomer konvergencie 1 a konverguje pre každé $|x| \leq 1$, teda na celej hraničnej kružnici svojho oboru konvergencie.

22.4. Nech $\boldsymbol{A} \in \mathbb{C}^{n\times n}$. Dokážte nasledujúce tvrdenia:

(a) Rad $\sum_{k=0}^\infty \boldsymbol{A}^k$ konverguje práve vtedy, keď $\rho(\boldsymbol{A}) < 1$.

(b) Z podmienky $\rho(\boldsymbol{A}) < 1$ vyplýva regularita matice $\boldsymbol{I} - \boldsymbol{A}$; za tohto predpokladu

[7] Redukovaná Planckova konštanta $\hbar$ sa často nazýva len stručne Planckova konštanta, občas tiež *Diracova konštanta.*

platí $(\boldsymbol{I}-\boldsymbol{A})\cdot\sum_{k=0}^{\infty}\boldsymbol{A}^k=\boldsymbol{I}$, čiže $\sum_{k=0}^{\infty}\boldsymbol{A}^k=(\boldsymbol{I}-\boldsymbol{A})^{-1}$.
Riešte obdobnú úlohu pre druhý rad z príkladu 22.1.7.

22.5. Nech $\boldsymbol{A}\in\mathbb{C}^{n\times n}$. Dokážte nasledujúce tvrdenia:
(a) Ak $\rho(\boldsymbol{A})<1$, tak rad $\sum_{k=1}^{\infty}\frac{1}{k}\boldsymbol{A}^k$ konverguje a platí $\exp\Bigl(-\sum\frac{1}{k}\boldsymbol{A}^k\Bigr)=\boldsymbol{I}-\boldsymbol{A}$.

(b) Ak $\operatorname{Re}\lambda<0$ pre každé $\lambda\in\operatorname{Spec}\boldsymbol{A}$, tak $\rho\bigl(\boldsymbol{I}-\mathrm{e}^{\boldsymbol{A}}\bigr)<1$ a $-\sum\frac{1}{k}\bigl(\boldsymbol{I}-\mathrm{e}^{\boldsymbol{A}}\bigr)^k=\boldsymbol{A}$.
Riešte obdobnú úlohu pre druhý rad z príkladu 22.1.8.

22.6. Dokážte, že pre $\boldsymbol{A},\boldsymbol{B}\in\mathbb{C}^{n\times n}$ platí $\mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}=\mathrm{e}^{\boldsymbol{A}}\cdot\mathrm{e}^{\boldsymbol{B}}\ \Leftrightarrow\ \boldsymbol{A}\cdot\boldsymbol{B}=\boldsymbol{B}\cdot\boldsymbol{A}$.

22.7. Dokážte tvrdenie 22.2.5 na základe rovnosti $\boldsymbol{J}_n(\lambda)=\boldsymbol{J}_n+\lambda\boldsymbol{I}_n$ a tvrdenia 22.2.2.

22.8. Vypočítajte hodnoty $\mathrm{e}^{\boldsymbol{A}}$, $\cos\boldsymbol{A}$, $\sin\boldsymbol{A}$, $\cosh\boldsymbol{A}=\frac{1}{2}\bigl(\mathrm{e}^{\boldsymbol{A}}+\mathrm{e}^{-\boldsymbol{A}}\bigr)$, $\sinh\boldsymbol{A}=\frac{1}{2}\bigl(\mathrm{e}^{\boldsymbol{A}}-\mathrm{e}^{-\boldsymbol{A}}\bigr)$, pre matice z príkladov 20.2.1–5, 20.3.1, 20.6.1, 21.5.3 a cvičení 20.4, 20.5 (c).

22.9. Overte tzv. *Newtonovu binomickú formulu pre k-tu deriváciu súčinu* funkcií $\bigl[p(x)q(x)\bigr]^{(k)}=\sum_{j=0}^{k}\binom{k}{j}p^{(j)}(x)q^{(k-j)}(x)$.

22.10. Dokážte, že lema 22.3.1 platí nad ľubovoľným poľom charakteristiky ∞. V čom je problém v prípade polí konečnej charakteristiky?
V nasledujúcich dvoch cvičeniach vychádzame z definície rozšírenia skalárnych funkcií na maticové argumenty z paragrafu 22.3.

22.11. (a) Nech f je komplexná funkcia komplexnej premennej a $\boldsymbol{A}\in\operatorname{Dom}_n(f)$. Ako vyzerá spektrum matice $f(\boldsymbol{A})$?

(b) Ak f je reálna funkcia reálnej premennej a $\boldsymbol{A}\in\mathbb{R}^{n\times n}\cap\operatorname{Dom}_n f$ má reálne spektrum, tak aj $f(\boldsymbol{A})$ je reálna matica. Dokážte.

(c) Nech f je komplexná funkcia komplexnej premennej taká, že $f(\overline{z})=\overline{f(z)}$ pre každé $z\in\operatorname{Dom}f$. Dokážte, že pre každú reálnu maticu $\boldsymbol{A}\in\operatorname{Dom}_n f$ aj $f(\boldsymbol{A})$ je reálna matica.

22.12. (a) Uvažujme reálne funkcie reálnej premennej $f(x)=\sqrt{x}$, $g(x)=\ln x$, $h(x)=x^{-1}$. Pre každú z funkcií f, g, h určte podmienky, za ktorých je definovaná jej hodnota pre Jordanovu bunku $\boldsymbol{J}_n(\lambda)$, kde $\lambda\in\mathbb{R}$.

(b) Dokážte, že za príslušných podmienok pre Jordanovu bunku $\boldsymbol{J}=\boldsymbol{J}_n(\lambda)$ ľubovoľného rozmeru n platí $f(\boldsymbol{J})^2=\boldsymbol{J}$, $\exp g(\boldsymbol{J})=\boldsymbol{J}$ a $\boldsymbol{J}\cdot h(\boldsymbol{J})=\boldsymbol{I}$, t. j. $h(\boldsymbol{J})=\boldsymbol{J}^{-1}$ je „obyčajná“ inverzná matica k matici $\boldsymbol{J}$.

(c) Dokážte, že pre ľubovoľnú Jordanovu bunku $\boldsymbol{J}=\boldsymbol{J}_n(\lambda)$ platí: ak $n=1$ alebo $\lambda\neq0$, tak $\boldsymbol{J}^2\in\operatorname{Dom}_n f$ a $f(\boldsymbol{J}^2)=\boldsymbol{J}$; $\mathrm{e}^{\boldsymbol{J}}\in\operatorname{Dom}g$ a $g(\mathrm{e}^{\boldsymbol{J}})=\boldsymbol{J}$.

(d) Za príslušných podmienok vypočítajte hodnoty $f(\boldsymbol{J})$, $g(\boldsymbol{J})$, $h(\boldsymbol{J})$ pre Jordanove bunky $\boldsymbol{J}=\boldsymbol{J}_n(\lambda)$, kde $2\le n\le4$.

(e) Vypočítajte $\sqrt{\boldsymbol{A}}$, $\ln\boldsymbol{A}$, $\boldsymbol{A}^{-1}$ (ak sú definované) pre každú z matíc z príkladov 20.2.1–5, 20.3.1, 20.6.1, 21.5.3 a cvičení 20.4, 20.5 (c).

22.13. Deriváciu maticovej (vektorovej) funkcie $\boldsymbol{F}:M\to K^{n\times n}$ (kde K je $\mathbb{R}$ alebo $\mathbb{C}$ a $M\subseteq\mathbb{R}$ je otvorený interval) v bode $t_0\in M$ definujeme ako limitu $\boldsymbol{F}'(t_0)=\lim_{t\to t_0}\frac{\boldsymbol{F}(t)-\boldsymbol{F}(t_0)}{t-t_0}$. Odvoďte z toho: ak $\boldsymbol{F}(t)$ je z lineárneho podpriestoru $S\subseteq K^{n\times n}$ pre každé t z nejakého okolia bodu t_0 a $\boldsymbol{F}'(t_0)$ existuje, tak aj $\boldsymbol{F}'(t_0)\in S$.

22.14. Nech $\boldsymbol{X}:M\to\mathbb{C}^{n\times n}$ je maticová funkcia reálneho argumentu, $M\subseteq\mathbb{R}$ je otvorený interval a $t_0\in M$. Ak $\boldsymbol{X}(t)$ je regulárna pre t z nejakého okolia bodu t_0 a $\boldsymbol{X}$ má deriváciu v bode t_0, tak $\mathrm{d}\boldsymbol{X}^{-1}(t_0)/\,\mathrm{d}t=-\boldsymbol{X}^{-1}(t_0)\cdot\boldsymbol{X}'(t_0)\cdot\boldsymbol{X}^{-1}(t_0)$. Dokážte.

22.15. Sformulujte a dokážte pravidlá o substitúcií a metóde *per partes* pre integrály maticových funkcií reálnej premennej.

22.16. (a) Pre ľubovoľnú bilineárnu funkciu $F\colon \mathbb{C}^m \times \mathbb{C}^n \to \mathbb{C}^k$ a vektorové funkcie $\boldsymbol{x}\colon S \to \mathbb{C}^m$, $\boldsymbol{y}\colon S \to \mathbb{C}^n$ reálneho argumentu platí $\frac{\mathrm{d}}{\mathrm{d}t} F(\boldsymbol{x}(t), \boldsymbol{y}(t)) = F(\boldsymbol{x}'(t), \boldsymbol{y}(t)) + F(\boldsymbol{x}(t), \boldsymbol{y}'(t))$. Sformulujte vynechané predpoklady a dokážte. Zovšeobecnite na prípad multilineárnej funkcie F.

(b) Nech $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\colon S \to \mathbb{C}^n$ sú diferencovateľné vektorové funkcie. Dokážte, že potom $\frac{\mathrm{d}}{\mathrm{d}t} \det(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \sum_{j=1}^{n} \det(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j', \ldots, \boldsymbol{x}_n)$.

22.17. (a) Nech $S \subseteq \mathbb{R}$ je otvorený interval a $\boldsymbol{A}\colon S \to \mathbb{C}^{n\times n}$ je spojitá maticová funkcia. Uvažujme sústavu lineárnych diferenciálnych rovníc $\boldsymbol{X}'(t) = \boldsymbol{A}(t) \cdot \boldsymbol{X}(t)$ pre maticovú funkciu $\boldsymbol{X}\colon S \to \mathbb{C}^{n\times n}$. S použitím Liouvilleovej formuly (veta 22.2.6) dokážte, že funkcia $f(t) = \det \boldsymbol{X}(t)$ vyhovuje lineárnej diferenciálnej rovnici $f'(t) = f(t) \operatorname{tr} \boldsymbol{A}(t)$.

(b) Na základe (a) dokážte, že z podmienky $\det \boldsymbol{X}(t_0) \neq 0$ pre nejaké $t_0 \in S$ vyplýva $\det \boldsymbol{X}(t) \neq 0$ pre každé $t \in S$.

(c) Nech $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\colon S \to \mathbb{C}^n$ sú riešenia homogénnej sústavy lineárnych diferenciálnych rovníc $\boldsymbol{x}' = \boldsymbol{A} \cdot \boldsymbol{x}$ také, že vektory $\boldsymbol{x}_1(t_0), \ldots, \boldsymbol{x}_n(t_0)$ sú lineárne nezávislé pre nejaké $t_0 \in S$. Dokážte, že potom vektory $\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_k(t)$ sú lineárne nezávislé pre ľubovoľné $t \in S$. (*Návod*: Položte $f(t) = \det(\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_n(t))$ a využite (b).)

22.18. Nech $F\colon S \to \mathbb{C}^n$ je fundamentálna matica sústavy $\boldsymbol{x}' = \boldsymbol{A} \cdot \boldsymbol{x}$. Dokážte, funkcia $\boldsymbol{F}(t)^{-1}$ je spojitá na S. Zdôvodnite tým integrovateľnosť funkcie $\boldsymbol{F}(s)^{-1} \cdot \boldsymbol{b}(s)$ z vety 22.5.2 (b).

22.19. Za predpokladu, že všetky matice $\boldsymbol{A}(t)$ navzájom komutujú, overte vzorce pre riešenie počiatočných úloh z dôsledku 22.5.4.

22.20. Dokážte postupne nasledujúce tvrdenia:

(a) Nech $\varepsilon > 0$ a $f\colon (-\varepsilon, \varepsilon) \to \mathbb{R}$ je ľubovoľná diferencovateľná funkcia taká, že $f(0) = 1$ a $f'(0) = a$. Označme $f_k(x) = f(x/k)^k$ pre $x \in (-k\varepsilon, k\varepsilon)$ a $h(x) = \lim_{|x|/\varepsilon < k \to \infty} f_k(x)$. Potom $h(x) = \mathrm{e}^{ax}$. Odvoďte z toho, že $\mathrm{e}^a = \lim_{k\to\infty}(1 + a/k)^k$ pre ľubovoľné $a \in \mathbb{C}$.

(b) Nech $\varepsilon > 0$ a $\boldsymbol{F}\colon (-\varepsilon, \varepsilon) \to \mathbb{C}^{n\times n}$ je ľubovoľná diferencovateľná maticová funkcia taká, že $\boldsymbol{F}(0) = \boldsymbol{I}$ a $\boldsymbol{F}'(0) = \boldsymbol{A}$. Označme $\boldsymbol{F}_k(x) = \boldsymbol{F}(x/k)^k$ pre $x \in (-k\varepsilon, k\varepsilon)$ a $\boldsymbol{H}(x) = \lim_{|x|/\varepsilon < k \to \infty} \boldsymbol{H}_k(x)$. Potom $\boldsymbol{H}(x) = \mathrm{e}^{\boldsymbol{A}x}$. Odvoďte z toho, že $\mathrm{e}^{\boldsymbol{A}} = \lim_{k\to\infty} \left(\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\right)^k$ pre ľubovoľnú maticu $\boldsymbol{A} \in \mathbb{C}^{n\times n}$.
(*Návod*: V oboch prípadoch sformulujte vhodnú počiatočnú úlohu, ktorej vyhovujú obe porovnávané funkcie.)

22.21. (a) Pre každú z matíc $\boldsymbol{A}$ z príkladov 20.2.1–5, 20.3.1, 20.6.1, 21.5.3 a cvičení 20.4, 20.5 (c). nájdite nejakú fundamentálnu maticu homogénnej autonómnej sústavy $\boldsymbol{x}' = \boldsymbol{A} \cdot \boldsymbol{x}$. Aj v prípade komplexných vlastných čísel, nájdite reálnu fundamentálnu maticu.

(b) S využitím racionálneho kanonického tvaru nahraďte každú zo sústav z časti (a) sústavou s menším počtom rovníc vyšieho rádu.

22.22. Dokážte tvrdenie 22.7.1.

22.23. Nech $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in K^{n\times n}$ sú matice nad ľubovoľným poľom K. Odvoďte nasledujúce identity pre komutátor matíc:

(a) $\boldsymbol{A}^2 + 2\boldsymbol{A} \cdot \boldsymbol{B} + \boldsymbol{B}^2 - (\boldsymbol{A} + \boldsymbol{B})^2 = [\boldsymbol{A}, \boldsymbol{B}]$;

(b) $[\boldsymbol{A}, \boldsymbol{B} \cdot \boldsymbol{C}] + [\boldsymbol{C}, \boldsymbol{A} \cdot \boldsymbol{B}] + [\boldsymbol{B}, \boldsymbol{C} \cdot \boldsymbol{A}] = \boldsymbol{0}$;

(c) $[\boldsymbol{A}, [\boldsymbol{B}, \boldsymbol{C}]] = [[\boldsymbol{A}, \boldsymbol{B}], \boldsymbol{C}] + [\boldsymbol{B}, [\boldsymbol{A}, \boldsymbol{C}]]$ (lineárny operátor $\mathrm{ad}_{\boldsymbol{A}} = [\boldsymbol{A}, \cdot\,]$ sa chová voči komutátoru ako derivácia voči súčinu);

(d) $[\boldsymbol{A}, [\boldsymbol{B}, \boldsymbol{C}]] + [\boldsymbol{C}, [\boldsymbol{A}, \boldsymbol{B}]] + [\boldsymbol{B}, [\boldsymbol{C}, \boldsymbol{A}]] = \boldsymbol{0}$ (*Jacobiho identita*).

22.24. Nech $\boldsymbol{A}, \boldsymbol{B} \in K^{n\times n}$. Dokážte, že pre každé $k \geq 1$ platí $[\boldsymbol{A}^k, \boldsymbol{B}] = \sum_{i=1}^{k} \boldsymbol{A}^{i-1} \cdot [\boldsymbol{A}, \boldsymbol{B}] \cdot \boldsymbol{A}^{k-i}$. Odvoďte z toho, že ak $\boldsymbol{A}$ komutuje s $[\boldsymbol{A}, \boldsymbol{B}]$, tak $[\boldsymbol{A}^k, \boldsymbol{B}] = k\boldsymbol{A}^{k-1} \cdot [\boldsymbol{A}, \boldsymbol{B}]$.

22.25. Nech $\boldsymbol{A} \in \mathbb{C}^{n\times n}$. Vypočítame exponenciálu $\mathrm{e}^{\mathrm{ad}_{\boldsymbol{A}}} = \sum_{k=0}^{\infty} \frac{1}{k!}\, \mathrm{ad}_{\boldsymbol{A}}^k$ lineárneho operátora $\mathrm{ad}_A \colon \mathbb{C}^{n\times \boldsymbol{n}} \to \mathbb{C}^{n\times n}$.

(a) Overte, že pre $\boldsymbol{B} \in \mathbb{C}^{n\times n}$ platí $\mathrm{e}^{\mathrm{ad}_{\boldsymbol{A}}}(\boldsymbol{B}) = \boldsymbol{B} + \frac{1}{1!}[\boldsymbol{A}, \boldsymbol{B}] + \frac{1}{2!}[\boldsymbol{A}, [\boldsymbol{A}, \boldsymbol{B}]] + \frac{1}{3!}[\boldsymbol{A}, [\boldsymbol{A}, [\boldsymbol{A}, \boldsymbol{B}]]] + \ldots)$.

(b) Dokážte, že obe maticové funkcie $\boldsymbol{F}(t) = \mathrm{e}^{\boldsymbol{A}t} \cdot \boldsymbol{B} \cdot \mathrm{e}^{-\boldsymbol{A}t}$, $\boldsymbol{G}(t) = \boldsymbol{B} + \frac{t}{1!}[\boldsymbol{A}, \boldsymbol{B}] + \frac{t^2}{2!}[\boldsymbol{A}, [\boldsymbol{A}, \boldsymbol{B}]] + \frac{t^3}{3!}[\boldsymbol{A}, [\boldsymbol{A}, [\boldsymbol{A}, \boldsymbol{B}]]] + \ldots$ vyhovujú počiatočnej úlohe $\boldsymbol{X}'(t) = [\boldsymbol{A}, \boldsymbol{X}(t)]$, $\boldsymbol{X}(0) = \boldsymbol{B}$. Odvoďte z toho rovnosť $\boldsymbol{F}(t) = \boldsymbol{G}(t)$ pre každé $t \in \mathbb{R}$.

(c) Ako špeciálny prípad pre $t = 1$ odvoďte tzv. *Hadamardovu formulu* $\mathrm{e}^{\mathrm{ad}_{\boldsymbol{A}}}(\boldsymbol{B}) = \boldsymbol{B} + \frac{1}{1!}[\boldsymbol{A}, \boldsymbol{B}] + \frac{1}{2!}[\boldsymbol{A}, [\boldsymbol{A}, \boldsymbol{B}]] + \frac{1}{3!}[\boldsymbol{A}, [\boldsymbol{A}, [\boldsymbol{A}, \boldsymbol{B}]]] + \ldots) = \mathrm{e}^{\boldsymbol{A}} \cdot \boldsymbol{B} \cdot \mathrm{e}^{-\boldsymbol{A}}$.

23. Združené lineárne operátory v unitárnych a euklidovských priestoroch

Prítomnosť skalárneho súčinu v unitárnom, prípadne euklidovskom priestore V umožňuje zaviesť medzi lineárnymi operátormi na V zaujímavý vzťah duality a hovoriť o dvojiciach navzájom *združených* alebo tiež *adjungovaných* lineárnych operátorov. Rôzne podoby tohto vzťahu vedú k definíciam dôležitých tried lineárnych operátorov, napospol diagonalizovateľných vzhľadom na vhodné ortonormálne bázy. Východiskom pre dôkaz ich diagonalizovateľnosti ako aj detailný popis ich štruktúry je práve tesná súvislosť medzi spektrálnymi vlastnosťami navzájom združených operátorov.

Podrobne preskúmame najmä dva typy takýchto lineárnych operátorov: *hermitovské* alebo tiež *samodjungované operátory* a *unitárne operátory*. Výsledky potom aplikujeme na ich reálne obdoby – *symetrické* a *ortogonálne* operátory. Na záver stručne pojednáme o tzv. *normálnych* operátoroch, ktoré tvoria istú prirodzene definovanú triedu zahŕňajúcu hermitovské aj unitárne operátory. Ako sa ukáže, je to práve trieda všetkých tých lineárnych operátorov, ktoré sú diagonalizovateľné vzhľadom na ortonormálne bázy.

Táto kapitola má skôr teoretický charakter; niekoľko ukážok z množstva aplikácií spektrálnej teórie združených operátorov predvedieme v nasledujúcich dvoch kapitolách.

V celej kapitole V označuje (väčšinou konečnorozmerný) unitárny prípadne euklidovský priestor so skalárnym súčinom $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$.

23.1 Združené lineárne operátory

Nech $\varphi\colon V \to V$ je lineárny operátor. *Združeným* alebo *adjungovaným*, prípadne tiež *duálnym lineárnym operátorom* k operátoru φ nazývame lineárny operátor $\psi\colon V \to V$ taký, že pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí

$$\langle \varphi\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \psi\boldsymbol{y} \rangle.$$

23.1.1. Tvrdenie. *Ku každému lineárnemu operátoru φ na konečnorozmernom unitárnom priestore V existuje práve jeden združený lineárny operátor $\psi\colon V \to V$. Ak $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ je nejaká ortonormálna báza vo V a $\boldsymbol{A} = (\varphi)_{\boldsymbol{\beta}}$ je matica operátora φ v tejto báze, tak maticou k nemu adjungovaného operátora ψ v báze $\boldsymbol{\beta}$ je hermitovsky združená matica $\boldsymbol{A}^*$ k matici $\boldsymbol{A}$.*

Dôkaz. Platnosť tvrdenia vyplýva z nasledujúcej úvahy. Nech $\boldsymbol{\beta}$ je nejaká ortonormálna báza vo V a $\varphi, \psi \colon V \to V$ sú lineárne operátory s maticami $\boldsymbol{A}$ resp. $\boldsymbol{B}$ vzhľadom na $\boldsymbol{\beta}$. Ľahko nahliadneme, že predpismi

$$F(\boldsymbol{x}, \boldsymbol{y}) = \langle \varphi \boldsymbol{x}, \boldsymbol{y} \rangle = \big(\boldsymbol{A} \cdot (\boldsymbol{x})_{\boldsymbol{\beta}}\big)^{\mathsf{T}} \cdot \overline{(\boldsymbol{y})}_{\boldsymbol{\beta}} = (\boldsymbol{x})_{\boldsymbol{\beta}}^{\mathsf{T}} \cdot \boldsymbol{A}^{\mathsf{T}} \cdot \overline{(\boldsymbol{y})}_{\boldsymbol{\beta}},$$
$$G(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \psi \boldsymbol{y} \rangle = (\boldsymbol{x})_{\boldsymbol{\beta}}^{\mathsf{T}} \cdot \overline{\boldsymbol{B} \cdot (\boldsymbol{y})_{\boldsymbol{\beta}}} = (\boldsymbol{x})_{\boldsymbol{\beta}}^{\mathsf{T}} \cdot \overline{\boldsymbol{B}} \cdot \overline{(\boldsymbol{y})}_{\boldsymbol{\beta}}$$

sú definované dve poldruhalineárne formy na V s maticami $\boldsymbol{A}^{\mathsf{T}}$ resp. $\overline{\boldsymbol{B}}$ vzhľadom na bázu $\boldsymbol{\beta}$. Potom ψ je adjungovaný k φ práve vtedy, keď $F = G$, t. j. práve vtedy, keď $\boldsymbol{A}^{\mathsf{T}} = \overline{\boldsymbol{B}}$. To je zrejme ekvivalentné s podmienkou $\boldsymbol{B} = \boldsymbol{A}^*$.

Poznamenajme, že v reálnom prípade je $\boldsymbol{A}^* = \boldsymbol{A}^{\mathsf{T}}$.

Metódami presahujúcimi rámec lineárnej algebry možno dokázať, že jednoznačne určený združený operátor existuje ku každému lineárnemu operátoru φ na ľubovoľnom *Hilbertovom priestore* V, t. j. na (nie nutne konečnorozmernom) unitárnom priestore spĺňajúcom topologickú podmienku *úplnosti.*

Práve dokázané tvrdenie a uvedená poznámka, nás oprávňujú zaviesť pre adjungovaný operátor ψ k operátoru φ označenie $\psi = \varphi^*$.

Priamočiare overenie nasledujúceho tvrdenia a jeho dôsledku prenechávame ako cvičenie čitateľovi.

23.1.2. Tvrdenie. *Nech $\varphi, \psi \colon V \to V$ sú lineárne operátory, $a, b \in \mathbb{C}$. Potom*

$$\varphi^{**} = \varphi,$$
$$(a\varphi + b\psi)^* = \overline{a}\varphi^* + \overline{b}\psi^*,$$
$$(\varphi \circ \psi)^* = \psi^* \circ \varphi^*.$$

Podľa prvej rovnosti φ je adjungovaným operátorom k svojmu adjungovanému operátoru φ^*; inak povedané, ψ je združený k φ práve vtedy, keď φ je združený k ψ. Symetria tohto vzťahu nám tak umožňuje hovoriť jednoducho o adjungovaných či združených operátoroch φ, ψ.

23.1.3. Dôsledok. *Nech φ je lineárny operátor na konečnorozmernom unitárnom priestore V a $f(x) = \sum_{k=1}^{\infty} a_k x^k \in \mathbb{C}[[x]]$ je potenčný rad nad $\mathbb{C}$. Označme $\overline{f}(x) = \sum_{k=1}^{\infty} \overline{a}_k x^k$ rad s komplexne združenými koeficientmi. Ak rad $f(\varphi)$ konverguje, špeciálne ak $f(x)$ je polynóm, tak aj rad $\overline{f}(\varphi^*)$ konverguje a platí*

$$f(\varphi)^* = \overline{f}(\varphi^*),$$

t. j. združeným operátorom k lineárnemu operátoru $f(\varphi)$ je operátor $\overline{f}(\varphi^*)$. Ak navyše $f(x) \in \mathbb{R}[[x]]$, tak

$$f(\varphi)^* = f(\varphi^*).$$

Nasledujúce dve tvrdenia hrajú pri štúdiu spektrálnych vlastností a štruktúry združených operátorov kľúčovú úlohu.

23.1.4. Tvrdenie. *Nech $\varphi\colon V \to V$ je lineárny operátor. Ak $\boldsymbol{u}$ je vlastný vektor operátora φ prislúchajúci k jeho vlastnej hodnote λ a $\boldsymbol{v}$ je vlastný vektor združeného operátora φ^* prislúchajúci k jeho vlastnej hodnote μ, tak $\boldsymbol{u} \perp \boldsymbol{v}$ alebo $\lambda = \overline{\mu}$.*

Dôkaz. Počítajme:

$$\lambda\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \langle \lambda\boldsymbol{u}, \boldsymbol{v}\rangle = \langle \varphi\boldsymbol{u}, \boldsymbol{v}\rangle = \langle \boldsymbol{u}, \varphi^*\boldsymbol{v}\rangle = \langle \boldsymbol{u}, \mu\boldsymbol{v}\rangle = \overline{\mu}\langle \boldsymbol{u}, \boldsymbol{v}\rangle.$$

Preto z $\langle \boldsymbol{u}, \boldsymbol{v}\rangle \neq 0$ vyplýva $\lambda = \overline{\mu}$.

23.1.5. Tvrdenie. *Nech $\varphi\colon V \to V$ je lineárny operátor. Ak $S \subseteq V$ je invariantný lineárny podpriestor operátora φ, tak jeho ortokomplement $S^\perp$ je invariantný podpriestor združeného operátora φ^*.*

Dôkaz. Nech $\boldsymbol{y} \in S^\perp$, t. j. $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = 0$ pre každé $\boldsymbol{x} \in S$. Dokážeme, že aj $\varphi^*(\boldsymbol{y}) \in S^\perp$. Keďže S je φ-invariantný, pre ľubovoľné $\boldsymbol{x} \in S$ je tiež $\varphi(\boldsymbol{x}) \in S$, teda

$$\langle \boldsymbol{x}, \varphi^*\boldsymbol{y}\rangle = \langle \varphi\boldsymbol{x}, \boldsymbol{y}\rangle = 0,$$

t. j. $\varphi^*(\boldsymbol{y}) \in S^\perp$, čiže $S^\perp$ je φ^*-invariantný.

23.2 Hermitovské operátory

Lineárny operátor φ na unitárnom priestore V sa nazýva *hermitovský* alebo *samoadjungovaný*, ak $\varphi^* = \varphi$, t. j. ak pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí

$$\langle \varphi\boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \varphi\boldsymbol{y}\rangle.$$

V reálnom prípade sa zvykne používať názov *symetrický lineárny operátor.*

Špecifikáciou výsledkov 23.1.1–5 na hermitovské operátory dostávame ako dôsledky im zodpovedajúce tvrdenia 23.2.1–5.

23.2.1. Tvrdenie. *Nech $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ je nejaká ortonormálna báza vo V a $\boldsymbol{A} = (\varphi)_{\boldsymbol{\beta}}$ je matica lineárneho operátora $\varphi\colon V \to V$ v tejto báze. Potom φ je hermitovský práve vtedy, keď platí $\boldsymbol{A}^* = \boldsymbol{A}$, t. j. keď $\boldsymbol{A}$ je hermitovská matica.*

Teda v reálnom prípade φ je symetrický práve vtedy, keď $\boldsymbol{A}^\mathsf{T} = \boldsymbol{A}$, t. j. keď jeho matica $\boldsymbol{A}$ v ortonormálnej báze $\boldsymbol{\beta}$ je symetrická.

23.2.2. Tvrdenie. *Nech $\varphi, \psi \colon V \to V$ sú hermitovské lineárne operátory a $a, b \in \mathbb{R}$. Potom operátor $a\varphi + b\psi$ je hermitovský. Ak navyše φ, ψ komutujú, tak aj operátor $\varphi \circ \psi = \psi \circ \varphi$ je hermitovský.*

V dôsledku prvej časti tohto tvrdenia tvoria hermitovské operátory $V \to V$ lineárny podpriestor vektorového priestoru $\mathcal{L}(V, V)$ všetkých lineárnych zobrazení $V \to V$, chápaného (aj v komplexnom prípade) ako vektorový priestor nad poľom $\mathbb{R}$.

23.2.3. Dôsledok. *Nech φ je hermitovský operátor na konečnorozmernom unitárnom priestore V a $f(x) \in \mathbb{R}[[x]]$ je potenčný rad nad $\mathbb{R}$. Ak rad $f(\varphi)$ konverguje, špeciálne ak $f(x) \in \mathbb{R}[x]$ je polynóm, tak aj operátor $f(\varphi)$ je hermitovský.*

23.2.4. Tvrdenie. *Nech $\varphi \colon V \to V$ je hermitovský lineárny operátor. Potom všetky vlastné hodnoty operátora φ sú reálne a vlastné vektory prislúchajúce k rôznym vlastným hodnotám operátora φ sú navzájom kolmé.*

23.2.5. Tvrdenie. *Nech $\varphi \colon V \to V$ je hermitovský lineárny operátor. Ak $S \subseteq V$ je invariantný lineárny podpriestor operátora φ, tak aj ortokomplement $S^\perp$ je jeho invariantný podpriestor.*

Nasledujúce jednoduché tvrdenie, ktoré je operátorovou analógiou vyjadrenia komplexného čísla v tvare $z = z_0 + \mathrm{i}z_1$, kde $z_0, z_1 \in \mathbb{R}$, napovedá, že samoadjungované operátory sa v istom zmysle správajú ako „viacrozmerné reálne čísla“. Neskôr na základe ich spektrálnych vlastností uvidíme, že táto analógia siaha oveľa ďalej.

23.2.6. Tvrdenie. *Nech V je unitárny priestor. Potom každý lineárny operátor $\varphi \colon V \to V$ má tvar*

$$\varphi = \varphi_0 + \mathrm{i}\varphi_1$$

pre jednoznačne určené samoadjungované operátory $\varphi_0, \varphi_1 \colon V \to V$.

Dôkaz. Ľahko možno overiť, že lineárne operátory

$$\varphi_0 = \frac{1}{2}(\varphi + \varphi^*), \qquad \varphi_1 = \frac{1}{2\mathrm{i}}(\varphi - \varphi^*)$$

sú samoadjungované a spĺňajú podmienky

$$\varphi = \varphi_0 + \mathrm{i}\varphi_1, \qquad \varphi^* = \varphi_0 - \mathrm{i}\varphi_1.$$

Naopak, zo samoadjungovanosti φ_0, φ_1 vyplýva, že tieto dve podmienky sú ekvivalentné, teda ktorákoľvek z nich má za následok jednoznačnosť uvedeného vyjadrenia φ_0, φ_1.

23.3 Spektrálny rozklad hermitovského operátora

Dôležitosť hermitovských operátorov je dôsledkom najmä ich spektrálnych vlastností.

23.3.1. Veta. *Nech $\varphi\colon V \to V$ je lineárny operátor na konečnorozmernom unitárnom priestore V. Potom φ je hermitovský práve vtedy, keď existuje ortonormálna báza $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ priestoru V tvorená vlastnými vektormi operátora φ, vzhľadom na ktorú má φ diagonálnu maticu $(\varphi)_{\boldsymbol{\beta}} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, s diagonálou tvorenou vlastnými číslami operátora φ, pričom $\lambda_1, \dots, \lambda_n \in \mathbb{R}$.*

Pre istotu ešte podotknime, že vlastné číslo $\lambda \in \operatorname{Spec}\varphi$ sa na diagonále matice $(\varphi)_{\boldsymbol{\beta}}$ vyskytuje toľkokrát, aká je jeho (jedno či algebraická alebo geometrická) násobnosť.

Dôkaz. Nech φ je hermitovský a $\lambda_1, \dots, \lambda_k$ sú všetky jeho navzájom rôzne vlastné čísla. Podľa tvrdenia 23.2.4 sú všetky reálne. Pre $j \leq k$ označme $S_j = \operatorname{Ker}(\varphi - \lambda_j)$ príslušný vlastný podpriestor a položme $S = S_1 \oplus \ldots \oplus S_k$. Zrejme S je invariantný podpriestor operátora φ. Dokážeme, že $S = V$. V opačnom prípade by podľa tvrdenia 23.2.5 bol aj $S^\perp$ netriviálny invariantný podpriestor. Potom zúženie $\varphi \restriction S^\perp$ by bolo samoadjungovaným lineárnym operátorom na $S^\perp$. Keďže pole $\mathbb{C}$ je algebraicky uzavreté, operátor $\varphi \restriction S^\perp$ by mal aspoň jedno vlastné číslo μ a k nemu prislúchajúci vlastný vektor $\mathbf{0} \neq \boldsymbol{w} \in S^\perp$. Potom však μ by bolo vlastným číslom samého φ, teda $\mu = \lambda_j$ pre niektoré $j \leq k$ a $\boldsymbol{w} \in S_j$. Keďže $S_j \cap S^\perp = \{\mathbf{0}\}$, máme spor.

Nech $n_j = \dim S_j$ označuje geometrickú násobnosť vlastného čísla λ_j a $\boldsymbol{\beta}_j = (\boldsymbol{v}_{j1}, \dots, \boldsymbol{v}_{jn_j})$ je ľubovoľná ortonormálna báza podpriestoru S_j. Podľa tvrdenia 23.2.4 sú vektory ležiace v rôznych podpriestoroch S_j navzájom kolmé, teda spojením čiastkových báz $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_k$ dostaneme ortonormálnu bázu

$$\boldsymbol{\beta} = (\boldsymbol{v}_{11}, \dots, \boldsymbol{v}_{1n_1}, \dots\dots, \boldsymbol{v}_{k1}, \dots, \boldsymbol{v}_{kn_k})$$

priestoru $S = V$. Zrejme φ má v tejto báze diagonálnu maticu

$$(\varphi)_{\boldsymbol{\beta}} = \operatorname{diag}(\lambda_1 \boldsymbol{I}_{n_1}, \dots, \lambda_k \boldsymbol{I}_{n_k}).$$

Keďže každá reálna diagonálna matica je hermitovská, obrátená implikácia vyplýva z tvrdenia 23.2.1.

V unitárnom priestore $\mathbb{C}^n$ so štandardným skalárnym súčinom je násobenie $\boldsymbol{x} \mapsto \boldsymbol{A} \cdot \boldsymbol{x}$ hermitovskou maticou $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ hermitovský operátor $\mathbb{C}^n \to \mathbb{C}^n$. Podľa práve dokázanej vety v $\mathbb{C}^n$ existuje ortonormálna báza $\boldsymbol{\beta}$, v ktorej má tento operátor diagonálnu maticu

$$\boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\varepsilon}} \cdot \boldsymbol{A} \cdot \boldsymbol{P}_{\boldsymbol{\varepsilon},\boldsymbol{\beta}} = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

Keďže matica prechodu $\boldsymbol{P}_{\boldsymbol{\varepsilon},\boldsymbol{\beta}}$ (ktorej stĺpce sú vektory bázy $\boldsymbol{\beta}$) je unitárna (pozri paragraf 14.4), okamžite tak dostávame maticovú formuláciu vety 23.3.1.

23.3.2. Veta. *(a) Matica $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ je hermitovská práve vtedy, keď existuje unitárna matica $\boldsymbol{P}$ taká, že matica $\boldsymbol{P}^{-1}\cdot\boldsymbol{A}\cdot\boldsymbol{P} = \boldsymbol{P}^*\cdot\boldsymbol{A}\cdot\boldsymbol{P}$ je reálna a diagonálna, s diagonálou tvorenou vlastnými číslami matice $\boldsymbol{A}$.*

(b) Matica $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ je symetrická práve vtedy, keď existuje ortogonálna matica $\boldsymbol{P}$ taká, že matica $\boldsymbol{P}^{-1}\cdot\boldsymbol{A}\cdot\boldsymbol{P} = \boldsymbol{P}^{\mathsf{T}}\cdot\boldsymbol{A}\cdot\boldsymbol{P}$ je reálna a diagonálna, s diagonálou tvorenou vlastnými číslami matice $\boldsymbol{A}$.

K časti (b) treba ešte poznamenať, že stĺpce matice $\boldsymbol{P}$, t. j. vlastné vektory reálnej matice $\boldsymbol{A}$ prislúchajúce k jej reálnym vlastným číslam sú tiež nevyhnutne reálne.

Najjednoduchšími príkladmi samoadjungovaných operátorov sú ortogonálne projekcie na lineárne podpriestory konečnorozmerného unitárneho priestoru V. Ak je totiž $S \subseteq V$ taký podpriestor a $\dim S = k \le n = \dim V$, tak jeho ľubovoľnú ortonormálnu bázu $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$ možno doplniť do ortonormálnej bázy $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ priestoru V. Potom $(\boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_n)$ je báza ortokomplementu $S^\perp$ a ortogonálna projekcia pr_S má v báze $\boldsymbol{\beta}$ maticu

$$(\mathrm{pr}_S)_{\boldsymbol{\beta}} = \mathrm{diag}(\boldsymbol{I}_k, 0, \ldots, 0)$$

s $n-k$ nulami na diagonále. Taktiež naopak, ak $\varphi\colon V \to V$ je samoadjungovaný lineárny operátor s nanajvýš dvojprvkovým spektrom $\mathrm{Spec}\,\varphi \subseteq \{0,1\}$, tak φ má v nejakej ortonormálnej báze priestoru V maticu uvedeného tvaru, z čoho už ľahko nahliadneme, že φ je ortogonálna projekcia na vlastný podpriestor $S = \mathrm{Ker}(\varphi - \mathrm{id}) = \mathrm{Im}\,\varphi$ prislúchajúci vlastnému číslu 1.

V súvislostí s tým nás pri pozornom čítaní dôkazu vety 23.3.1 napadne ešte jeden spôsob, ako ju možno interpretovať. Aby sme si uľahčili vyjadrovanie, budeme hovoriť, že *lineárne podpriestory* $S_1,\ldots, S_k$ unitárneho priestoru V sú *navájom ortogonálne*, ak pre $i < j \le k$ a ľubovoľné $\boldsymbol{x} \in S_i$, $\boldsymbol{y} \in S_j$ platí $\boldsymbol{x} \perp \boldsymbol{y}$.

23.3.3. Veta. *Nech V je konečnorozmerný unitárny priestor a $\varphi\colon V \to V$ je lineárny operátor. Potom φ je samoadjungovaný práve vtedy, keď ho možno vyjadriť v tvare lineárnej kombinácie*

$$\varphi = \lambda_1\,\mathrm{pr}_{S_1} + \ldots + \lambda_k\,\mathrm{pr}_{S_k}$$

ortogonálnych projekcií $\mathrm{pr}_{S_1}, \ldots, \mathrm{pr}_{S_k}$ na navzájom ortogonálne vlastné podpriestory prislúchajúce k jeho po dvoch rôznym vlastným číslam $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$, pričom $V = S_1 \oplus \ldots \oplus S_k$.

Uvedená lineárna kombinácia ortogonálnych projekcií do vlastných podpriestorov sa nazýva *spektrálny rozklad samoadjungovaného operátora* φ.

Vety 23.3.1–2 umožňujú rozšíriť obor funkcií, ktoré pracujú na hermitovských operátoroch resp. maticiach, z potenčných radov na ľubovoľné reálne funkcie reálnej premennej. Sústredíme sa len na názornejšiu maticovú formuláciu vychádzajúcu z vety 23.3.2. Preklad do reči operátorov prenechávame čitateľovi.

Nech $f\colon M \to \mathbb{R}$ je reálna funkcia reálnej premennej s definičným oborom $M \subseteq \mathbb{R}$. Ak $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ je hermitovská matica taká, že $\operatorname{Spec} \boldsymbol{A} \subseteq M$, tak hodnota $f(\boldsymbol{A})$ funkcie f na matici $\boldsymbol{A}$ je podľa paragrafu 22.3 definovaná rovnosťou

$$f(\boldsymbol{A}) = \boldsymbol{P} \cdot \operatorname{diag}\bigl(f(\lambda_1), \dots, f(\lambda_n)\bigr) \cdot \boldsymbol{P}^{-1},$$

kde $\boldsymbol{P}$ je ľubovoľná regulárna matica taká, že $\boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Navyše, podľa vety 22.3.2 hodnota $f(\boldsymbol{A})$ nezávisí na matici $\boldsymbol{P}$. No práve možnosť zvoliť za $\boldsymbol{P}$ *unitárnu* maticu zaručuje, že aj výsledná matica $f(\boldsymbol{A})$ bude hermitovská. Podobne, pre reálnu symetrickú maticu $\boldsymbol{A}$ je vďaka existencii *ortogonálnej* matice prechodu $\boldsymbol{P}$ aj $f(\boldsymbol{A})$ reálna symetrická matica. Tým sme dokázali nasledujúcu vetu.

23.3.4. Veta. *Nech $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ je hermitovská matica a $f\colon M \to \mathbb{R}$ je funkcia taká, že $\operatorname{Spec} \boldsymbol{A} \subseteq M \subseteq \mathbb{R}$. Potom aj $f(\boldsymbol{A}) \in \mathbb{C}^{n\times n}$ je hermitovská matica. Ak navyše $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ je symetrická, tak aj $f(\boldsymbol{A}) \in \mathbb{R}^{n\times n}$ je symetrická.*

23.3.5. Príklad. *Odmocnina hermitovskej matice.* Nech $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ je hermitovská matica taká, že $\lambda \geq 0$ pre všetky $\lambda \in \operatorname{Spec} \boldsymbol{A}$. Potom existuje druhá odmocnina matice $\boldsymbol{A}$, t. j. matica

$$\sqrt{\boldsymbol{A}} = \boldsymbol{P} \cdot \operatorname{diag}\bigl(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}\,\bigr) \cdot \boldsymbol{P}^{-1},$$

kde $\boldsymbol{P} \in \mathbb{C}^{n\times n}$ je ľubovoľná regulárna (možno no nie nutne unitárna) matica taká, že $\boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Hodnota výrazu $\sqrt{\boldsymbol{A}}$ je automaticky hermitovská matica, ktorá nezávisí na matici $\boldsymbol{P}$.

23.4 Unitárne operátory

Pri formulácii výsledkov predchádzajúceho paragrafu sme hojne využívali unitárne, prípadne ortogonálne matice, t. j. matice prechodu medzi ortonormálnymi bázami. V tomto paragrafe zavedieme im zodpovedajúce lineárne operátory a opäť preskúmame ich spektrálne vlastnosti.

Lineárny operátor φ na unitárnom vektorovom priestore V sa nazýva *unitárny*, ak je bijektívny a k nemu inverzný operátor φ^{-1} je zároveň k nemu adjungovaný, t. j. $\varphi^* = \varphi^{-1}$. Potom, samozrejme,

$$\varphi^* \circ \varphi = \mathrm{id}_V = \varphi \circ \varphi^*.$$

V prípade reálneho vektorového priestoru so skalárnym súčinom hovoríme o *ortogonálnom lineárnom operátore.*

Nasledujúce tvrdenie je triviálne.

23.4.1. Tvrdenie. *Nech $\varphi, \psi \colon V \to V$ sú unitárne operátory. Potom aj φ^{-1} a $\varphi \circ \psi$ sú unitárne operátory.*

Z tvrdenia 23.1.1 okamžite dostávame

23.4.2. Tvrdenie. *Nech $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ je nejaká ortonormálna báza unitárneho priestoru V a $\boldsymbol{A} = (\varphi)_{\boldsymbol{\beta}}$ je matica lineárneho operátora $\varphi \colon V \to V$ v tejto báze. Potom φ je unitárny práve vtedy, keď platí $\boldsymbol{A}^* = \boldsymbol{A}^{-1}$, t. j. keď $\boldsymbol{A}$ je unitárna matica.*

Teda v reálnom prípade φ je ortogonálny práve vtedy, keď $\boldsymbol{A}^{\mathsf{T}} = \boldsymbol{A}^{-1}$, t. j. keď $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ je ortogonálna.

Nasledujúca charakterizácia unitárnych operátorov v konečnorozmerných priestoroch je bezprostredným dôsledkom tvrdenia 23.4.2 a vety 17.4.2 resp. 13.5.2. Vo všeobecnom prípade ju možno dokázať úplne rovnako ako vety 13.5.2, 17.4.2; kvôli nekonečnorozmerným priestorom je v častiach (ii) a (iii) pridaná (v konečnorozmernom prípade nadbytočná) požiadavka surjektívnosti.

23.4.3. Veta. *Nech $\varphi \colon V \to V$ je lineárny operátor na unitárnom priestore V. Potom nasledujúce podmienky sú ekvivalentné:*
(i) φ je unitárny;
(ii) φ je surjektívny a pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí $\langle \varphi\boldsymbol{x}, \varphi\boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$;
(iii) φ je surjektívny a pre všetky $\boldsymbol{x} \in V$ platí $\|\varphi\boldsymbol{x}\| = \|\boldsymbol{x}\|$.

Inak povedané, unitárne operátory sú práve tie lineárne izomorfizmy $V \to V$, ktoré zachovávajú skalárny súčin alebo, ekvivalentne, normu vektorov. V dôsledku toho pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí

$$\|\varphi\boldsymbol{x} - \varphi\boldsymbol{y}\| = \|\boldsymbol{x} - \boldsymbol{y}\|,$$

čiže unitárne operátory zachovávajú vzdialenosť v priestore V. Teda z geometrického hľadiska sú unitárne operátory *zhodné zobrazenia* alebo tiež *izometrie* priestoru V. Nie sú to však zďaleka všetky izometrie priestoru V. Izometriou $V \to V$ je očividne i každé posunutie o pevný vektor $\boldsymbol{u} \in V$. A keďže kompozícia dvoch izometrií je opäť izometria, sú izometriami unitárneho priestoru V aj všetky afinné zobrazenia $V \to V$, ktorých lineárna časť je unitárna (pozri paragraf 8.5). Neskôr uvidíme, že tiež naopak, každá izometria konečnorozmerného unitárneho priestoru má takýto tvar.

Unitárne a hermitovské operátory vykazujú do značnej miery podobné spektrálne vlastnosti. Jediný rozdiel spočíva v tom, že kým všetky vlastné čísla samoadjungovaného operátora sú reálne, všetky vlastné čísla unitárneho operátora sú komplexné jednotky.

23.4.4. Tvrdenie. *Nech $\varphi\colon V \to V$ je unitárny operátor. Potom všetky vlastné hodnoty operátora φ sú komplexné jednotky a vlastné vektory prislúchajúce k rôznym vlastným hodnotám operátora φ sú navzájom kolmé.*

Dôkaz. Ak $\lambda \in \mathbb{C}$ je vlastná hodnota unitárneho operátora φ a $\boldsymbol{v} \neq \mathbf{0}$ je k nej prislúchajúci vlastný vektor, tak

$$\|\boldsymbol{v}\| = \|\varphi\boldsymbol{v}\| = \|\lambda\boldsymbol{v}\| = |\lambda|\,\|\boldsymbol{v}\|\,,$$

preto $|\lambda| = 1$. Ďalej si stačí uvedomiť, že $\lambda \neq 0$ je vlastná hodnota operátora φ prislúchajúca k vlastnému vektoru $\boldsymbol{v}$ práve vtedy, keď $\lambda^{-1} = \overline{\lambda}$ je vlastná hodnota operátora $\varphi^{-1} = \varphi^*$ prislúchajúca k tomu istému vektoru $\boldsymbol{v}$, a odvolať sa na tvrdenie 23.1.4.

23.4.5. Tvrdenie. *Nech $\varphi\colon V \to V$ je unitárny operátor. Ak $S \subseteq V$ je invariantný lineárny podpriestor operátora φ, tak aj ortokomplement $S^\perp$ je jeho invariantný podpriestor.*

Dôkaz. Tentokrát si stačí uvedomiť, že ak $\varphi\colon V \to V$ je lineárny izomorfizmus, tak lineárny podpriestor $S \subseteq V$ je φ-invariatný práve vtedy, keď je φ^{-1}-invariantný. Keďže $\varphi^* = \varphi^{-1}$, potrebný záver vyplýva z tvrdenia 23.1.5.

Uvedomme si, že kľúčovým momentom pri dôkaze vety 23.3.1 boli tvrdenia 23.2.4 a najmä 23.2.5. Pre unitárne operátory im zodpovedajú tvrdenia 23.4.4 a 23.4.5. Nasledujúcu vetu teda možno dokázať úplne rovnako ako vetu 23.3.1.

23.4.6. Veta. *Nech $\varphi\colon V \to V$ je lineárny operátor na konečnorozmernom unitárnom priestore V. Potom φ je unitárny práve vtedy, keď existuje ortonormálna báza $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ priestoru V tvorená vlastnými vektormi operátora φ, vzhľadom na ktorú má φ diagonálnu maticu $(\varphi)_{\boldsymbol{\beta}} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, s diagonálou tvorenou vlastnými číslami operátora φ, pričom $|\lambda_1| = \ldots = |\lambda_n| = 1$.*

Ďalšie dva výsledky sú „unitárnou obdobou“ viet 23.3.2 a 23.3.3. Pripomíname, že unitárne matice rozmeru 2×2 sme plne charakterizovali vo vete 17.4.3. Pre väčšie rozmery podáme len charakterizáciu až na podobnosť prostredníctvom unitárnych matíc prechodu.

23.4.7. Veta. *Matica $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ je unitárna práve vtedy, keď existuje unitárna matica $\boldsymbol{P}$ taká, že matica $\boldsymbol{P}^{-1}\cdot\boldsymbol{A}\cdot\boldsymbol{P} = \boldsymbol{P}^*\cdot\boldsymbol{A}\cdot\boldsymbol{P}$ je diagonálna, pričom všetky jej diagonálne prvky sú komplexné jednotky.*

Náš čitateľ má teraz najskôr pocit, že sa pri formulácii tejto vety točíme dokola, a asi príliš nechápe, čo sme vlastne získali, keď sme od unitárnej matice $\boldsymbol{A}$ dospeli k ďalšej unitárnej matici $\boldsymbol{P}$. Preto pripomíname, že táto veta je ekvivalentná so zďaleka nie očividnou vetou 23.4.6. Význam oboch viet budeme môcť plne doceniť až v nasledujúcom paragrafe, keď na ich základe už jednoducho popíšeme štruktúru ortogonálnych, t. j. zhodných, lineárnych zobrazení v euklidovských priestoroch.

23.4.8. Veta. *Nech V je konečnorozmerný unitárny priestor a $\varphi\colon V \to V$ je lineárny operátor. Potom φ je unitárny práve vtedy, keď ho možno vyjadriť v tvare lineárnej kombinácie*

$$\varphi = \lambda_1 \operatorname{pr}_{S_1} + \ldots + \lambda_k \operatorname{pr}_{S_k}$$

ortogonálnych projekcií $\operatorname{pr}_{S_1}, \ldots, \operatorname{pr}_{S_k}$ na navzájom ortogonálne vlastné podpriestory prislúchajúce k jeho po dvoch rôznym vlastným číslam $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$, pričom $|\lambda_1| = \ldots = |\lambda_n| = 1$ a $V = S_1 \oplus \ldots \oplus S_k$.

Uvedená lineárna kombinácia ortogonálnych projekcií do vlastných podpriestorov sa, samozrejme, nazýva *spektrálny rozklad unitárneho operátora φ*.

Kým hermitovské operátory a matice, ako sme už spomínali, v mnohom pripomínajú reálne čísla, unitárne operátory a matice sú akousi viacrozmernou analógiou komplexných jednotiek. A podobne ako reálne čísla a komplexné jednotky, aj hermitovské a unitárne operátory či matice sú tesne zviazané prostredníctvom exponenciálnej funkcie. Opäť sa obmedzíme len na maticovú formuláciu.

23.4.9. Veta. *Nech $\boldsymbol{H} \in \mathbb{C}^{n\times n}$ je hermitovská matica. Potom matica $\exp(\mathrm{i}\boldsymbol{H})$ je unitárna. Naopak, každá unitárna matica $\boldsymbol{U} \in \mathbb{C}^{n\times n}$ má tvar*

$$\boldsymbol{U} = \exp(\mathrm{i}\boldsymbol{H}) = \cos\boldsymbol{H} + \mathrm{i}\sin\boldsymbol{H}$$

pre vhodnú hermitovskú maticu $\boldsymbol{H}$. Ak $\det\boldsymbol{U} = 1$, tak $\boldsymbol{H}$ možno zvoliť tak, aby navyše platilo $\operatorname{tr}\boldsymbol{H} = 0$.

Dôkaz. Nech $\boldsymbol{H}$ je hermitovská. Potom podľa dôsledkov 22.2.4 a 23.1.3

$$\exp(\mathrm{i}\boldsymbol{H})^{-1} = \exp(-\mathrm{i}\boldsymbol{H}) = \exp(\mathrm{i}\boldsymbol{H})^*,$$

teda $\exp(\mathrm{i}\boldsymbol{H})$ je unitárna. Nech naopak $\boldsymbol{U} \in \mathbb{C}^{n\times n}$ je unitárna. Potom

$$\boldsymbol{U} = \boldsymbol{P}\cdot\operatorname{diag}(\lambda_1, \ldots, \lambda_n)\cdot\boldsymbol{P}^*$$

pre vhodnú unitárnu maticu $\boldsymbol{P}$. Keďže $|\lambda_j| = 1$ pre $1 \leq j \leq n$, existujú reálne čísla α_j (určené jednoznačne až na sčítance tvaru $2k\pi$, kde $k \in \mathbb{Z}$) také, že $\lambda_j = \mathrm{e}^{\mathrm{i}\alpha_j}$. Potom matica

$$\boldsymbol{H} = \boldsymbol{P} \cdot \operatorname{diag}(\alpha_1, \ldots, \alpha_n) \cdot \boldsymbol{P}^*$$

je hermitovská a podľa tvrdenia 22.1.2 spĺňa podmienku $\exp(\mathrm{i}\boldsymbol{H}) = \boldsymbol{U}$. Tvar $\boldsymbol{U} = \cos \boldsymbol{H} + \mathrm{i} \sin \boldsymbol{H}$ vyplýva z tvrdenia 22.2.1.

Ak $\det \boldsymbol{U} = 1$, tak podľa Liouvilleovej formuly (veta 22.2.6) $\mathrm{e}^{\operatorname{tr}(\mathrm{i}\boldsymbol{H})} = \det \boldsymbol{U} = 1$, preto $\operatorname{tr} \boldsymbol{H} = 2k\pi$ pre nejaké $k \in \mathbb{Z}$. V takom prípade stačí jedno z čísel α_j nahradiť číslom $\alpha_j - 2k\pi$.

23.5 Štruktúra ortogonálnych operátorov

Nasledujúce tvrdenie je reálnou obdobou tvrdenia 23.4.1.

23.5.1. Tvrdenie. *Nech $\varphi, \psi\colon V \to V$ sú ortogonálne operátory na reálnom vektorovom priestore V so skalárnym súčinom. Potom aj φ^{-1} a $\varphi \circ \psi$ sú ortogonálne operátory.*

Podobne, nasledujúca veta je zhrnutím tvrdenia 23.4.2 a vety 23.4.3 pre prípad *euklidovských* (t. j. zároveň konečnorozmerných) priestorov.

23.5.2. Veta. *Nech φ je lineárny operátor na euklidovskom priestore $\boldsymbol{V}$, $\boldsymbol{\beta}$ je nejaká ortonormálna báza vo V a $\boldsymbol{A} = (\varphi)_{\boldsymbol{\beta}}$. Potom nasledujúce podmienky sú ekvivalentné:*

(i) $\langle \varphi\boldsymbol{x}, \varphi\boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ *pre všetky* $\boldsymbol{x}, \boldsymbol{y} \in V$;

(ii) $\|\varphi\boldsymbol{x}\| = \|\boldsymbol{x}\|$ *pre každé* $\boldsymbol{x} \in V$;

(iii) $\boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathsf{T}}$.

Inak povedané, ortogonálne lineárne operátory na euklidovských priestoroch sú charakterizované každou z troch ekvivalentných podmienok: zachovávajú skalárny súčin, zachovávajú normu vektorov a majú v nejakej (a tým pádom v každej) ortonormálnej báze ortogonálnu maticu. Preto sú zároveň zhodnými zobrazeniami (izometriami) priestoru V. Z toho dôvodu sú i všetky afinné zobrazenia $V \to V$ s ortogonálnou lineárnou časťou izometriami priestoru V. I v tomto prípade neskôr uvidíme, že každá izometria euklidovského priestoru V má už nevyhnutne takýto tvar.

V dôsledku vety 13.5.3 sú ortogonálne operátory na dvojrozmernom euklidovskom priestore práve všetky rotácie okolo počiatku a osové súmernosti podľa priamok prechádzajúcich počiatkom. Otočenie o uhol α má v ľubovoľnej ortonormálnej báze priestoru V maticu $\boldsymbol{R}_\alpha$ alebo $\boldsymbol{R}_{-\alpha}$ (v závislosti od orientácie bázy a zmyslu otočenia). Podobne, osová súmernosť má v takejto báze maticu $\boldsymbol{S}_\beta$ prípadne $\boldsymbol{S}_{-\beta}$, kde β je uhol, ktorý zviera prvý vektor bázy s osou súmernosti. Ak za bázu zvolíme jednotkový smerový vektor osi súmernosti a jednotkový smerový vektor jej ortokomplementu, tak daná osová súmernosť bude mať v tejto ortonormálne báze maticu $\boldsymbol{S}_0 = \operatorname{diag}(1,-1)$ – pozri príklad 18.4.1.

S využitím výsledkov o unitárnych operátoroch z predchádzajúceho paragrafu ukážeme, že práve rotácie v navzájom ortogonálnych dvojrozmerných podpriestoroch sú základnými stavebnými kameňmi všetkých ortogonálnych operátorov. Presnejšie, každý ortogonálny lineárny operátor možno rozložiť na takéto rotácie, s prípadnou výnimkou „zvyškového“ podpriestoru dimenzie najviac 2, na ktorom φ môže fungovať ako $\pm\,\mathrm{id}$ alebo osová súmernosť.

23.5.3. Veta. *Nech φ je lineárny operátor na euklidovskom priestore V a $n = \dim V \geq 1$. Potom φ je ortogonálny práve vtedy, keď existuje ortonormálna báza $\boldsymbol{\beta}$ priestoru V taká, že matica $(\varphi)_{\boldsymbol{\beta}}$ má jeden z nasledujúcich blokovo diagonálnych tvarov*

$$\operatorname{diag}\big(\boldsymbol{R}_{\alpha_1},\ldots,\boldsymbol{R}_{\alpha_m},1\big) \qquad \textit{alebo} \qquad \operatorname{diag}\big(\boldsymbol{R}_{\alpha_1},\ldots,\boldsymbol{R}_{\alpha_m},-1\big),$$

ak $n = 2m+1$ je nepárne, prípadne

$$\operatorname{diag}\big(\boldsymbol{R}_{\alpha_1},\ldots,\boldsymbol{R}_{\alpha_m}\big) \qquad \textit{alebo} \qquad \operatorname{diag}\big(\boldsymbol{R}_{\alpha_1},\ldots,\boldsymbol{R}_{\alpha_{m-1}},1,-1\big),$$

ak $n = 2m$ je párne.

Dôkaz. Zrejme každý lineárny operátor $\varphi\colon V \to V$, ktorý má v nejakej ortonormálnej báze maticu v niektorom z uvedených tvarov, je ortogonálny. Sústreďme sa preto len na obrátenú implikáciu.

Komplexifikácia $\varphi^{\mathbb{C}}\colon V^{\mathbb{C}} \to V^{\mathbb{C}}$ ortogonálneho operátora $\varphi\colon V \to V$, je unitárny operátor na n-rozmernom unitárnom priestore $V^{\mathbb{C}}$ (pozri paragraf 19.4). Podľa vety 23.4.6 existuje ortonormálna báza $\boldsymbol{\gamma} = (\boldsymbol{w}_1,\ldots,\boldsymbol{w}_n)$ priestoru $V^{\mathbb{C}}$ a komplexné jednotky $\lambda_1,\ldots,\lambda_n$ také, že $\big(\varphi^{\mathbb{C}}\big)_{\boldsymbol{\gamma}} = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$. Vďaka tvrdeniu 19.5.1 môžeme bez ujmy na všeobecnosti predpokladať, že vlastné čísla, ktoré nie sú reálne, sú zoradené na začiatku v dvojiciach λ_{2j-1}, $\lambda_{2j} = \overline{\lambda}_{2j-1}$ $(1 \leq j \leq k \leq n/2)$ a pre im zodpovedajúce vlastné vektory bázy γ platí $\boldsymbol{w}_{2j} = \overline{\boldsymbol{w}}_{2j-1}$. Za nimi nasledujú reálne vlastné čísla v dvojiciach $\lambda_{2j-1} = \lambda_{2j} = \pm 1$ $(k < j \leq m \leq n/2)$. Pre n nepárne, zoznam nevyhnutne uzatvára „nespárené“ vlastné číslo ± 1; pre n párne môže (no nemusí) nakoniec zostať dvojica $1, -1$. Pre $2k < j \leq n$ navyše platí $\boldsymbol{w}_j \in V$, lebo reálnym vlastným číslam zodpovedajú reálne vlastné vektory.

Nech $\alpha_1,\ldots,\alpha_m$ sú reálne čísla z intervalu $\langle 0, 2\pi)$ také, že $\lambda_{2j-1} = \mathrm{e}^{\mathrm{i}\alpha_j}$, $\lambda_{2j} = \mathrm{e}^{-\mathrm{i}\alpha_j}$ pre $1 \leq j \leq k$, a $\alpha_j = 0$, ak $\lambda_{2j-1} = \lambda_{2j} = 1$, resp. $\alpha_j = \pi$, ak $\lambda_{2j-1} = \lambda_{2j} = -1$, pre $k < j \leq m$. Položme $\boldsymbol{v}_{2j-1} = \frac{1}{\sqrt{2}}(\boldsymbol{w}_{2j-1} + \boldsymbol{w}_{2j})$, $\boldsymbol{v}_{2j} =$

$\frac{1}{\mathrm{i}\sqrt{2}}(\boldsymbol{w}_{2j}-\boldsymbol{w}_{2j-1})$ pre $j \le k$ a $\boldsymbol{v}_j = \boldsymbol{w}_j$ pre $2k < j \le n$. Podľa cvičenia 19.11 (d) a vety 19.5.2 je $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ ortonormálna báza euklidovského priestoru V a matica $(\varphi)_{\boldsymbol{\beta}}$ má jeden z uvedených tvarov.

Ešte si všimnime, že obe matice vľavo majú determinant $+1$, takže zodpovedajú tzv. *priamym zhodnostiam*, kým obe matice vpravo majú determinant -1, čiže zodpovedajú tzv. *nepriamym zhodnostiam*.

Špeciálne, pre „náš“ trojrozmerný euklidovský priestor dostávame

23.5.4. Dôsledok. *(Euler) Nech φ je lineárny operátor na trojrozmernom euklidovskom priestore V. Potom φ je ortogonálny práve vtedy, keď existuje ortonormálna báza $\boldsymbol{\beta}$ priestoru V a reálne číslo α také, že φ má v báze $\boldsymbol{\beta}$ blokovo diagonálnu maticu tvaru*

$$\operatorname{diag}(\boldsymbol{R}_\alpha, 1) \qquad \textit{alebo} \qquad \operatorname{diag}(\boldsymbol{R}_\alpha, -1).$$

Inak povedané, každý ortogonálny lineárny operátor v trojrozmernom euklidovskom priestore je buď otočením okolo nejakej osi prechádzajúcej počiatkom (ak ide o priamu zhodnosť), alebo takýmto otočením zloženým so súmernosťou podľa roviny prechádzajúcej počiatkom a kolmej na os otočenia (ak ide o nepriamu zhodnosť).

Podobne ako v rovine, aj kompozícia dvoch otočení trojrozmerného priestoru okolo tej istej osi o uhly α, resp. β je zrejme otočením okolo tej istej osi o uhol $\alpha + \beta$. Avšak to, že aj zložením otočení okolo dvoch rôznobežných osí dostaneme otočenie okolo ďalšej osi prechádzajúcej ich priesečníkom, už na prvý pohľad zrejmé nie je, rovnako ako nie sú z názoru zrejmé os ani uhol výsledného otočenia. Nasledujúci dôsledok tvrdenia 23.5.1 a dôsledku 23.5.4 preto nie je nijako očividný.

23.5.5. Dôsledok. *Kompozícia dvoch otočení trojrozmerného euklidovského priestoru okolo osí prechádzajúcich počiatkom je opäť otočením okolo osi prechádzajúcej počiatkom.*

23.5.6. Príklad. Otočenie euklidovského priestoru $\mathbb{R}^3$ (so štandardným skalárnym súčinom) okolo priamky so smerovým vektorom $\boldsymbol{w} \neq 0$ o uhol α v kladnom zmysle (t. j. proti smeru hodinových ručičiek) pri pohľade z koncového bodu vektora $\boldsymbol{w}$ budeme značiť $\boldsymbol{R}_\alpha^{\boldsymbol{w}}$, rovnako ako jeho maticu vzhľadom na kanonickú ortonormálnu bázu $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$. Zrejme každé takéto otočenie o uhol $\alpha \neq k\pi$, kde $k \in \mathbb{Z}$, má tvar $\boldsymbol{R}_\alpha^{\boldsymbol{w}}$ pre jednoznačne určený *jednotkový* vektor $\boldsymbol{w} \in \mathbb{R}^3$ (ak $\alpha = k\pi$, tak pre párne k ide o identitu, takže na $\boldsymbol{w}$ nezáleží, pre nepárne k je $\boldsymbol{R}_{k\pi}^{\boldsymbol{w}} = \boldsymbol{R}_{k\pi}^{-\boldsymbol{w}} = \boldsymbol{R}_\pi^{\boldsymbol{w}}$ osová súmernosť podľa priamky $[\boldsymbol{w}]$, takže $\boldsymbol{w}$ je určený jednoznačne až na orientáciu).

Pre otočenia okolo súradných osí dostávame

$$\boldsymbol{R}_\alpha^{\boldsymbol{e}_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} = \operatorname{diag}(1, \boldsymbol{R}_\alpha),$$

$$\boldsymbol{R}_\alpha^{\boldsymbol{e}_2} = \begin{pmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix},$$

$$\boldsymbol{R}_\alpha^{\boldsymbol{e}_3} = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} = \operatorname{diag}(\boldsymbol{R}_\alpha, 1).$$

Rozmyslite si, prečo je v matici $\boldsymbol{R}_\alpha^{\boldsymbol{e}_2}$ znamienko mínus pri výraze $\sin\alpha$ umiestnené v pozícii $(3,1)$ a nie $(1,3)$.

Ľubovoľný jednotkový vektor $\boldsymbol{w} \in \mathbb{R}^3$ možno s použitím jeho sférických súradníc zapísať v tvare

$$\boldsymbol{w} = (\cos\theta\cos\omega, \sin\theta\cos\omega, \sin\omega)^{\mathsf{T}},$$

kde $\theta \in (-\pi, \pi\rangle$ je „zemepisná dĺžka“ a $\omega \in \langle -\pi/2, \pi/2\rangle$ „zemepisná šírka“ jeho koncového bodu na „jednotkovom glóbuse.“ Toto vyjadrenie je dokonca jednoznačné pre všetky $\boldsymbol{w} \neq (0,0,\pm 1)^{\mathsf{T}}$, t. j. také, že $\omega \neq \pm\pi/2$, – pozri paragraf 14.4.

Otočenie $\boldsymbol{R}_\alpha^{\boldsymbol{w}}$ získame ako kompozíciu vhodných otočení okolo súradných osí, ktoré orientujeme pomocou ich smerových vektorov $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$: najprv otočíme celý priestor okolo osi z o uhol $-\theta$ tak, aby sa os otočenia $[\boldsymbol{w}]$ dostala do roviny xz, pokračujeme otočením okolo osi y o uhol ω tak, aby táto splynula s osou x (uvedomte si, že tentokrát má kladný zmysel otáčanie od vektora $\boldsymbol{e}_3$ k vektoru $\boldsymbol{e}_1$), teraz vykonáme vlastné otočenie okolo osi x o uhol α, a postupne vrátime súradné osi do pôvodnej polohy otočením okolo osi y o uhol $-\omega$ a následným otočením okolo osi z o uhol θ. Pre maticu pôvodného otočenia tak dostávame

$$\boldsymbol{R}_\alpha^{\boldsymbol{w}} = \boldsymbol{R}_\theta^{\boldsymbol{e}_3} \cdot \boldsymbol{R}_{-\omega}^{\boldsymbol{e}_2} \cdot \boldsymbol{R}_\alpha^{\boldsymbol{e}_1} \cdot \boldsymbol{R}_\omega^{\boldsymbol{e}_2} \cdot \boldsymbol{R}_{-\theta}^{\boldsymbol{e}_3}.$$

Výsledná matica by už bola príliš neprehľadná a komplikovaná, takže ju najmä z typografických dôvodov neuvádzame. Pre konkrétne hodnoty však dostávame rozumné výsledky (pozri cvičenia).

Pri inej voľbe pomocných transformácií dostávame napr. vyjadrenie

$$\boldsymbol{R}_\alpha^{\boldsymbol{w}} = \boldsymbol{R}_{\theta-\pi/2}^{\boldsymbol{e}_3} \cdot \boldsymbol{R}_{\omega-\pi/2}^{\boldsymbol{e}_1} \cdot \boldsymbol{R}_\alpha^{\boldsymbol{e}_3} \cdot \boldsymbol{R}_{\pi/2-\omega}^{\boldsymbol{e}_1} \cdot \boldsymbol{R}_{\pi/2-\theta}^{\boldsymbol{e}_3}.$$

Rozmyslite si jeho geometrický význam a napíšte matice jednotlivých dielčích otočení.

Ešte si všimnime, že pre matice

$$\boldsymbol{P} = \boldsymbol{R}_{\theta}^{\boldsymbol{e}_3} \cdot \boldsymbol{R}_{-\omega}^{\boldsymbol{e}_2} = \begin{pmatrix} \cos\theta\cos\omega & -\sin\theta & -\cos\theta\sin\omega \\ \sin\theta\cos\omega & \cos\theta & -\sin\theta\sin\omega \\ \sin\omega & 0 & \cos\omega \end{pmatrix},$$

$$\boldsymbol{Q} = \boldsymbol{R}_{\theta-\pi/2}^{\boldsymbol{e}_3} \cdot \boldsymbol{R}_{\omega-\pi/2}^{\boldsymbol{e}_1} = \begin{pmatrix} \sin\theta & \cos\theta\sin\omega & \cos\theta\cos\omega \\ -\cos\theta & \sin\theta\sin\omega & \sin\theta\cos\omega \\ 0 & -\cos\omega & \sin\omega \end{pmatrix}$$

platí $\boldsymbol{P}^{\mathsf{T}} \cdot \boldsymbol{P} = \boldsymbol{Q}^{\mathsf{T}} \cdot \boldsymbol{Q} = \boldsymbol{I}_3$, $\det \boldsymbol{P} = \det \boldsymbol{Q} = 1$ a

$$\begin{aligned} \boldsymbol{R}_{\alpha}^{\boldsymbol{w}} &= \boldsymbol{P} \cdot \boldsymbol{R}_{\alpha}^{\boldsymbol{e}_1} \cdot \boldsymbol{P}^{-1} = \boldsymbol{P} \cdot \boldsymbol{R}_{\alpha}^{\boldsymbol{e}_1} \cdot \boldsymbol{P}^{\mathsf{T}} \\ &= \boldsymbol{Q} \cdot \boldsymbol{R}_{\alpha}^{\boldsymbol{e}_3} \cdot \boldsymbol{Q}^{-1} = \boldsymbol{Q} \cdot \boldsymbol{R}_{\alpha}^{\boldsymbol{e}_3} \cdot \boldsymbol{Q}^{\mathsf{T}}, \end{aligned}$$

t. j. ich stĺpce tvoria kladne orientované ortonormálne bázy v $\mathbb{R}^n$, vzhľadom na ktoré má otočenie $\boldsymbol{R}_{\alpha}^{\boldsymbol{w}}$ maticu $\boldsymbol{R}_{\alpha}^{\boldsymbol{e}_1} = \operatorname{diag}(1, \boldsymbol{R}_{\alpha})$ resp. $\boldsymbol{R}_{\alpha}^{\boldsymbol{e}_3} = \operatorname{diag}(\boldsymbol{R}_{\alpha}, 1)$.

Túto úvahu možno preniesť aj na situáciu, keď os otočenia je zadaná nenulovým vektorom $\boldsymbol{u} \in \mathbb{R}^3$ v „neuhlovom", t. j. v súradnicovom tvare. Veľmi jednoducho možno totiž nájsť naň kolmý nenulový vektor $\boldsymbol{v} \in \mathbb{R}^3$, ktorý má aspoň jednu zložku nulovú (rozmyslite si ako). Potom $\boldsymbol{\beta} = (\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{u} \times \boldsymbol{v})$ je ortogonálna kladne orientovaná báza v $\mathbb{R}^3$. Znormovaním všetkých jej členov na jednotkovú dĺžku dostávame ortonormálnu kladne orientovanú bázu $\boldsymbol{\beta}_0$, vzhľadom na ktorú má otočenie $\boldsymbol{R}_{\alpha}^{\boldsymbol{u}}$ maticu $\boldsymbol{R}_{\alpha}^{\boldsymbol{e}_1} = \operatorname{diag}(1, \boldsymbol{R}_{\alpha})$. Preto

$$\boldsymbol{R}_{\alpha}^{\boldsymbol{u}} = \boldsymbol{\beta}_0 \cdot \operatorname{diag}(1, \boldsymbol{R}_{\alpha}) \cdot \boldsymbol{\beta}_0^{\mathsf{T}}.$$

Podobne možno nájsť aj ortonormálnu kladne orientovanú bázu, vzhľadom na ktorú má otočenie $\boldsymbol{R}_{\alpha}^{\boldsymbol{u}}$ maticu $\boldsymbol{R}_{\alpha}^{\boldsymbol{e}_3}$, prípadne $\boldsymbol{R}_{\alpha}^{\boldsymbol{e}_2}$.

23.5.7. Príklad. Uvažujme maticu

$$\boldsymbol{A} = \frac{1}{4} \begin{pmatrix} 1 & 1 - 2\sqrt{2} & 2 + \sqrt{2} \\ 1 + 2\sqrt{2} & 1 & \sqrt{2} - 2 \\ \sqrt{2} - 2 & 2 + \sqrt{2} & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Výpočtom sa ľahko presvedčíme, že $\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A} = \boldsymbol{I}_3$, teda $\boldsymbol{A}$ je ortogonálna, a $\det \boldsymbol{A} = 1$, teda (násobenie maticou) $\boldsymbol{A}$ predstavuje priamu zhodnosť, čiže otočenie okolo osi prechádzajúcej počiatkom. Analýzou vlastných čísel a vlastných vektorov matice $\boldsymbol{A}$ zistíme os a uhol tohto otočenia.

Charakteristický polynóm $\det(\boldsymbol{A} - x\boldsymbol{I}) = (1 - x)(x^2 + 1)$ má jeden reálny koreň $x_1 = 1$ a dva komplexne združené korene $x_{2,3} = \pm\mathrm{i} = \mathrm{e}^{\pm\mathrm{i}\pi/2}$. Vlastnej hodnote 1 zodpovedá vlastný vektor $\boldsymbol{w}_0 = (1, 1, \sqrt{2})^{\mathsf{T}}$. Teda $\boldsymbol{A}$ je maticou otočenia okolo osi $[\boldsymbol{w}_0]$ o uhol $\pi/2$, prípadne (ak sa na celú vec pozrieme „z druhej strany“) o uhol $-\pi/2$. Vyjasníme si aj tento detail, t. j. otázku, či platí $\boldsymbol{A} = \boldsymbol{R}^{\boldsymbol{w}_0}_{\pi/2}$ alebo $\boldsymbol{A} = \boldsymbol{R}^{\boldsymbol{w}_0}_{-\pi/2} = \boldsymbol{R}^{-\boldsymbol{w}_0}_{\pi/2}$.

Vlastnej hodnote $e^{i\pi/2} = i$ zodpovedá vlastný vektor $(1/\sqrt{2} + i, 1/\sqrt{2} - i, -1)^{\mathsf{T}} = \boldsymbol{u}_0 - i\boldsymbol{v}_0$, kde $\boldsymbol{u}_0 = (1/\sqrt{2}, 1/\sqrt{2}, -1)^{\mathsf{T}}$, $\boldsymbol{v}_0 = (-1, 1, 0)^{\mathsf{T}}$. Ľahko možno overiť (a nie je to náhoda – pozri cvičenie 19.11), že vektory $\boldsymbol{u}_0$, $\boldsymbol{v}_0$, $\boldsymbol{w}_0$ tvoria ortogonálnu bázu priestoru $\mathbb{R}^3$. Výpočtom determinatu $\det(\boldsymbol{u}_0, \boldsymbol{v}_0, \boldsymbol{w}_0) = 4 > 0$ sa presvedčíme, že báza $(\boldsymbol{u}_0, \boldsymbol{v}_0, \boldsymbol{w}_0)$ je kladne orientovaná (ak by sme dostali $\det(\boldsymbol{u}_0, \boldsymbol{v}_0, \boldsymbol{w}_0) < 0$, nahradili by sme vektor $\boldsymbol{w}_0$ vektorom $-\boldsymbol{w}_0$). Znormovaním všetkých bázických vektorov na jednotkovú dlžku dostaneme kladne orientovanú ortonormálnu bázu $\boldsymbol{\beta} = (\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$ priestoru $\mathbb{R}^3$ tvorenú stĺpcami matice

$$\boldsymbol{P} = \begin{pmatrix} 1/2 & -1/\sqrt{2} & 1/2 \\ 1/2 & 1/\sqrt{2} & 1/2 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{pmatrix}.$$

Potom $\boldsymbol{A} = \boldsymbol{R}^{\boldsymbol{w}}_{\pi/2}$ a $\boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P} = \operatorname{diag}(\boldsymbol{R}_{\pi/2}, 1)$. Inak povedané, lineárna transformácia $\boldsymbol{x} \mapsto \boldsymbol{A} \cdot \boldsymbol{x}$ má v báze $\boldsymbol{\beta}$ maticu

$$\operatorname{diag}(\boldsymbol{R}_{\pi/2}, 1) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

23.5.8. Príklad. V euklidovskom priestore $\mathbb{R}^3$ nájdeme maticu otočenia $\boldsymbol{R}^{\boldsymbol{u}}_{\alpha}$ o uhol $\alpha = \pi/6$ (v kladnom zmysle) okolo osi $[\boldsymbol{u}]$ danej vektorom $\boldsymbol{u} = (1, 1, 1)^{\mathsf{T}}$. Zrejme vektor $\boldsymbol{v} = (1, -1, 0)$ je kolmý na $\boldsymbol{u}$. Položme

$$\boldsymbol{w} = \boldsymbol{u} \times \boldsymbol{v} = (1, 1, -2)^{\mathsf{T}}.$$

Potom $\|\boldsymbol{u}\| = \sqrt{3}$, $\|\boldsymbol{v}\| = \sqrt{2}$, $\|\boldsymbol{w}\| = \sqrt{6}$, teda stĺpce matice

$$\boldsymbol{P} = \left(\frac{1}{\sqrt{3}} \boldsymbol{u}, \frac{1}{\sqrt{2}} \boldsymbol{v}, \frac{1}{\sqrt{6}} \boldsymbol{w} \right) = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & \sqrt{3} & 1 \\ \sqrt{2} & -\sqrt{3} & 1 \\ \sqrt{2} & 0 & -2 \end{pmatrix}$$

tvoria ortonormálnu kladne orientovanú bázu v $\mathbb{R}^3$, vzhľadom na ktorú má naše otočenie maticu

$$\boldsymbol{R}^{\boldsymbol{e}_1}_{\alpha} = \operatorname{diag}(1, \boldsymbol{R}_{\pi/6}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{3}/2 & -1/2 \\ 0 & 1/2 & \sqrt{3}/2 \end{pmatrix}.$$

Matica tohto otočenia v kanonickej báze $\boldsymbol{\varepsilon}$ je preto daná súčinom

$$\boldsymbol{R}_\alpha^{\boldsymbol{u}} = \boldsymbol{P} \cdot \boldsymbol{R}_\alpha^{\boldsymbol{e}_1} \cdot \boldsymbol{P}^\mathsf{T} = \frac{1}{3} \begin{pmatrix} 1+\sqrt{3} & 1-\sqrt{3} & 1 \\ 1 & 1+\sqrt{3} & 1-\sqrt{3} \\ 1-\sqrt{3} & 1 & 1+\sqrt{3} \end{pmatrix}.$$

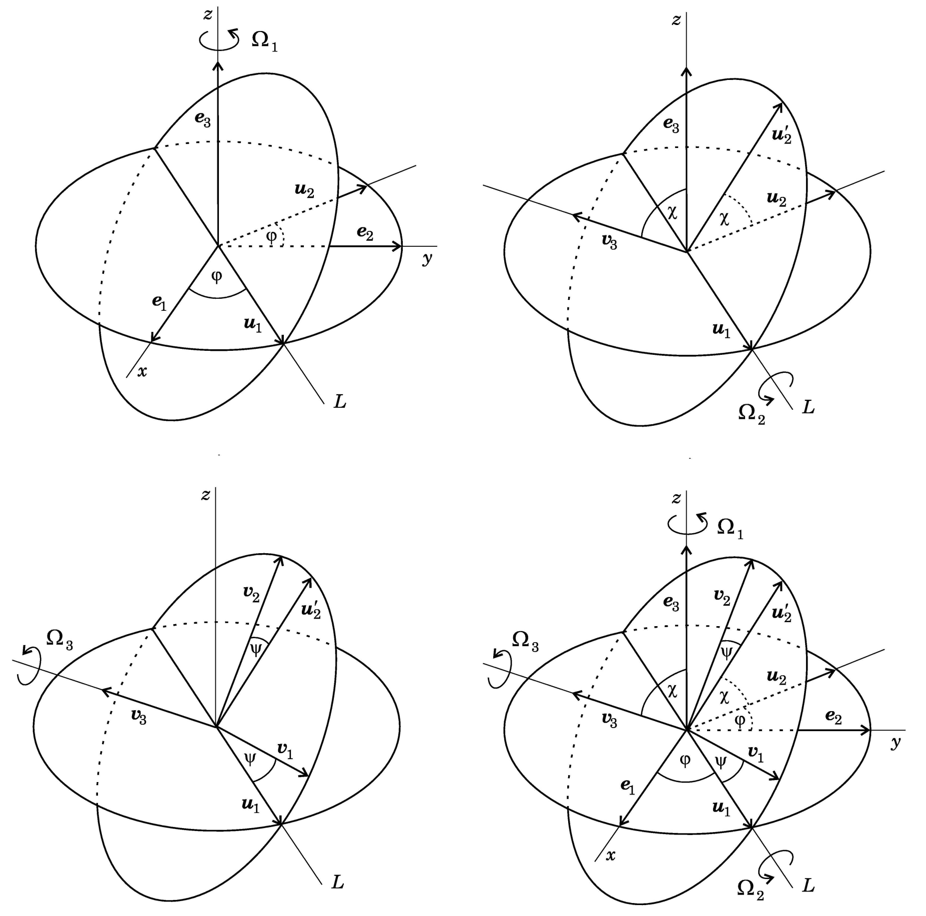
23.6 Eulerove uhly

Videli sme, že každé otočenie $\boldsymbol{R}_\alpha^{\boldsymbol{w}}$ v $\mathbb{R}^3$ možno rozložiť na kompozíciu piatich otočení okolo troch, prípadne iba dvoch súradných osí. Trochu náročnejšou geometrickou úvahou možno dokázať, že každé takéto otočenie možno rozložiť na kompozíciu dokonca len troch otočení okolo troch, prípadne iba dvoch súradných osí o tzv. *Eulerove uhly.* U objektov, ktoré majú svoje prirodzené súradné osi (napr. u lietadla je to pozdĺžna os, os krídel a os prechádzajúca ich priesečníkom kolmo na ich rovinu) zasa možno pomocou nich zadávať ich *priestorovú orientáciu*, t. j. natočenie osí objektu voči vopred zvoleným „nehybným“ súradným osiam. Tieto parametre sa (v rôznych modifikáciách) vyskytujú v mnohých aplikáciách, napr. v klasickej mechanike pri popise pohybu tuhých telies i v kvantovej mechanike pri popise uhlových momentov. Hojne sa využívajú napr. v aeronautike (kde sa im hovorí tiež *Taitove-Bryanove uhly*) pri riadení smeru letu a priestorovej orientácie lietadiel, kozmických lodí a satelitov.

Nevýhodou popisu rotácií v $\mathbb{R}^3$ pomocou Eulerových uhlov je to, že pri ňom priamo nevidíme os a nepoznáme uhol výsledného otočenia. Táto nevýhoda, ako aj trochu vyššie nároky na priestorovú predstavivosť sú však bohato vyvážené jednoduchosťou reprezentácie pomocou malého počtu rotácií okolo súradných osí.

Predvedieme, ako možno ľubovoľnú ortonormálnu bázu v trojrozmernom euklidovskom priestore pomocou troch vhodne zvolených otočení o *Eulerove uhly* φ, χ, ψ previesť na ľubovoľnú s ňou súhlasne orientovanú ortonormálnu bázu. Vzápätí metódou pohyblivej bázy (pozri paragraf 7.6) ukážeme, ako možno tieto otočenia nahradiť otočeniami o rovnaké uhly len v zmenenom poradí okolo súradných osí daných vektormi pôvodnej bázy. Pri tom sa môžeme bez ujmy na všeobecnosti obmedziť na euklidovský priestor $\mathbb{R}^3$ (so štandardným skalárnym súčinom) a za východziu bázu si zvoliť kanonickú bázu $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$. Tým zároveň dokážeme, že každé otočenie v $\mathbb{R}^3$ okolo nejakej osi prechádzajúcej počiatkom možno vyjadriť ako kompozíciu troch otočení okolo (niektorých dvoch zo) súradných osí x, y, z. Situácia je znázornená na štyroch obrázkoch: kým prvé tri znázorňujú jednotlivé dielčie otočenia, na štvrtom sú zhrnuté všetky tri fázy do jediného celku.

Nech $\boldsymbol{\beta} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ je ľubovoľná kladne orientovaná ortonormálna báza v $\mathbb{R}^3$. Predpokladajme najprv, že roviny $[\boldsymbol{e}_1, \boldsymbol{e}_2]$ a $[\boldsymbol{v}_1, \boldsymbol{v}_2]$ sú rôznobežné,



Obr. 23.1. Eulerove uhly (zxz-konvencia)

teda sa pretínajú v priamke, ktorú si označíme L. Nech $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ je kladne orientovaná ortonormálna báza roviny $[\boldsymbol{e}_1, \boldsymbol{e}_2]$ taká, že $\boldsymbol{u}_1$ je smerový vektor priamky L. Potom $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{e}_3)$ je kladne orientovaná ortonormálna báza priestoru $\mathbb{R}^3$ a tiež $\boldsymbol{u}_1 \perp \boldsymbol{v}_3$. Označme Ω_1 otočenie okolo osi $z = [\boldsymbol{e}_3]$ dané priradením $\boldsymbol{e}_1 \mapsto \boldsymbol{u}_1$, $\boldsymbol{e}_2 \mapsto \boldsymbol{u}_2$ a $\varphi \in \langle -\pi, \pi \rangle$ uhol tohto otočenia (obrázok vľavo hore). Nech ďalej Ω_2 je otočenie okolo osi $L = [\boldsymbol{u}_1]$, t. j. novej osi x, ktorým vektor $\boldsymbol{e}_3$ prejde do vektora $\boldsymbol{v}_3$ a vektor $\boldsymbol{u}_2$ do vektora $\boldsymbol{u}_2'$; označme $\chi \in \langle -\pi/2, \pi/2 \rangle$ uhol tohto otočenia (obrázok vpravo hore). Konečne Ω_3 je otočenie okolo osi $[\boldsymbol{v}_3]$, t. j. novej osi z, ktorým $\boldsymbol{u}_1$ prejde do $\boldsymbol{v}_1$ a $\boldsymbol{u}_2'$ do $\boldsymbol{v}_2$, a $\psi \in \langle -\pi, \pi \rangle$ je uhol tohto otočenia (obrázok vľavo dole).

Schématicky teda môžeme písať

$$\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3) \xmapsto{\Omega_1} \boldsymbol{\alpha} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{e}_3) \xmapsto{\Omega_2} \boldsymbol{\alpha}' = (\boldsymbol{u}_1, \boldsymbol{u}_2', \boldsymbol{v}_3) \xmapsto{\Omega_3} \boldsymbol{\beta} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3).$$

Potom $\Omega_1 = \boldsymbol{R}_\varphi^{\boldsymbol{e}_1}$, $\Omega_2 = \boldsymbol{R}_\chi^{\boldsymbol{u}_1}$ a $\Omega_3 = \boldsymbol{R}_\psi^{\boldsymbol{v}_3}$, teda

$$\Omega = \Omega_3 \circ \Omega_2 \circ \Omega_1 = \boldsymbol{R}_\psi^{\boldsymbol{v}_3} \circ \boldsymbol{R}_\chi^{\boldsymbol{u}_1} \circ \boldsymbol{R}_\varphi^{\boldsymbol{e}_3}.$$

Pritom uvedené vyjadrenie je jednoznačné, až na krajné hodnoty $\varphi = \pm\pi$, $\chi = \pm\pi/2$ a $\psi = \pm\pi$ (samozrejme, miesto intervalov $\langle -\pi, \pi\rangle$ resp. $\langle -\pi/2, \pi/2\rangle$ možno uvažovať intervaly $\langle 0, 2\pi\rangle$ resp. $\langle 0, \pi\rangle$ a pod.).

V prípade, že roviny $[\boldsymbol{e}_1, \boldsymbol{e}_2]$ a $[\boldsymbol{v}_1, \boldsymbol{v}_2]$ (teda aj vektory $\boldsymbol{e}_3$ a $\boldsymbol{v}_3$) splývajú, je Ω otočením okolo osi z o nejaký uhol γ. Uvedená rovnosť tak zostáva v platnosti s hodnotou $\chi = 0$; k spomínaným nejednoznačnostiam sa však pridružuje nová, podstatnejšia: φ a ψ sú teraz ľubovoľné hodnoty, pre ktoré platí $\varphi + \psi = \gamma$.

Vo všeobecnosti nevystupujú v uvedenom vyjadrení otočenia okolo pôvodných súradných osí x, y, z. Ortonormálne bázy $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}'$ a $\boldsymbol{\beta}$, ktoré dostaneme postupnými rotáciami pôvodnej bázy $\boldsymbol{\varepsilon}$ prostredníctvom otočení Ω_1, Ω_2 a Ω_3, však možno chápať ako jedinú kanonickú bázu spojenú s pohybujúcim sa objektom (napr. satelitom alebo lietadlom). Vyjadrenie pomocou otočení okolo pôvodných súradných osí možno dostať na základe zrejmých rovností

$$(\Omega_1)_{\boldsymbol{\varepsilon}} = \boldsymbol{R}_\varphi^{\boldsymbol{e}_3}, \qquad (\Omega_2)_{\boldsymbol{\alpha}} = \boldsymbol{R}_\chi^{\boldsymbol{e}_1}, \qquad (\Omega_3)_{\boldsymbol{\alpha}'} = \boldsymbol{R}_\psi^{\boldsymbol{e}_3},$$

z vety 7.6.1 (pri obvyklom stotožnení lineárnej transformácie na $\mathbb{R}^3$ s jej maticou v kanonickej báze $\boldsymbol{\varepsilon}$):

$$\begin{aligned} \Omega = (\Omega)_{\boldsymbol{\varepsilon}} = (\Omega_3 \circ \Omega_2 \circ \Omega_1)_{\boldsymbol{\varepsilon}} &= (\Omega_1)_{\boldsymbol{\varepsilon}} \cdot (\Omega_2)_{\boldsymbol{\alpha}} \cdot (\Omega_3)_{\boldsymbol{\alpha}'} \\ &= \boldsymbol{R}_\varphi^{\boldsymbol{e}_3} \cdot \boldsymbol{R}_\chi^{\boldsymbol{e}_1} \cdot \boldsymbol{R}_\psi^{\boldsymbol{e}_3} = \boldsymbol{R}_\varphi^{\boldsymbol{e}_3} \circ \boldsymbol{R}_\chi^{\boldsymbol{e}_1} \circ \boldsymbol{R}_\psi^{\boldsymbol{e}_3}. \end{aligned}$$

Vynásobením príslušných matíc dostávame pre maticu výsledného otočenia explicitné vyjadrenie pomocou Eulerových uhlov

$$\begin{aligned} (\Omega)_{\boldsymbol{\varepsilon}} &= \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\chi & -\sin\chi \\ 0 & \sin\chi & \cos\chi \end{pmatrix} \cdot \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\varphi\cos\psi - \sin\varphi\cos\chi\sin\psi & -\cos\varphi\sin\psi - \sin\varphi\cos\chi\cos\psi & \sin\varphi\sin\chi \\ \sin\varphi\cos\psi + \cos\varphi\cos\chi\sin\psi & -\sin\varphi\sin\psi + \cos\varphi\cos\chi\cos\psi & -\cos\varphi\sin\chi \\ \sin\chi\sin\psi & \sin\chi\cos\psi & \cos\chi \end{pmatrix}. \end{aligned}$$

Vyjadrenie pomocou otočení okolo osi z, potom x a zase z, t. j. tzv. *zxz-konvencia*, nie je jediné možné. Popri nej sa vyskytujú aj ďalšie konvencie využívajúce dve súradné osi, napr. zyz, xyx, xzx a pod. (dohromady je 6 možností). Taktiež sa používajú konvencie využívajúce všetky tri súradné osi, napr. xyz, xzy, zyx a pod. (spolu ďalších 6 možností). Preto je pri reprezentácií transformácie pomocou Eulerových uhlov vždy potrebné špecifikovať príslušnú konvenciu, ako aj fakt, či sa osi dielčích otočení vzťahujú na nehybný alebo pohyblivý súradnicový systém.

23.7 Normálne operátory

Čitateľovi asi neuniklo, že viaceré výsledky o hermitovských resp. unitárnych operátoroch, ktoré sme uviedli v paragrafoch 23.2–4, boli do značnej miery analogické, ba aj ich dôkazy využívali podobné myšlienky a postupy. Natíska sa preto otázka, či sa za touto podobnosťou neskrýva nejaký všeobecnejší pojem, ktorý by zahŕňal oba typy operátorov a umožnil jednotným spôsobom dokázať spomínané výsledky. Takýmto jednotiacim pojmom je pojem *normálneho operátora.*

Hovoríme, že *lineárny operátor* φ na unitárnom priestore V je *normálny*, ak komutuje so svojím združeným operátorom, t. j. ak platí

$$\varphi \circ \varphi^* = \varphi^* \circ \varphi.$$

Zrejme φ je normálny práve vtedy, keď je normálny jeho združený operátor φ^*.

Hermitovské aj unitárne operátory sú z triválnych dôvodov normálne. Ďalším príkladom normálnych operátorov sú *antihermitovské lineárne operátory*, definované podmienkou

$$\varphi^* = -\varphi,$$

ako aj ľubovoľné skalárne násobky unitárnych operátorov.

23.7.1. Tvrdenie. *Nech $\varphi\colon V \to V$ je normálny lineárny operátor. Potom združené operátory φ, φ^* majú spoločný vlastný vektor. Presnejšie, ak $\boldsymbol{v}$ je vlastný vektor operátora φ prislúchajúci k jeho vlastnej hodnote λ, tak $\boldsymbol{v}$ je zároveň vlastný vektor operátora φ^* prislúchajúci k jeho vlastnej hodnote $\overline{\lambda}$.*

Dôkaz. Keďže pole $\mathbb{C}$ je algebraicky uzavreté, φ má aspoň jednu vlastnú hodnotu a k nej prislúchajúci vlastný vektor. Stačí preto dokázať druhú časť tvrdenia.

Nech λ je vlastná hodnota operátora φ a $\boldsymbol{v}$ je k nej prislúchajúci vlastný vektor, t. j. $(\varphi - \lambda)(\boldsymbol{v}) = \mathbf{0}$. Podľa tvrdenia 23.1.2 platí $(\varphi - \lambda)^* = \varphi^* - \overline{\lambda}$, v dôsledku čoho aj združené operátory $\varphi - \lambda$, $(\varphi - \lambda)^*$ komutujú. Preto tiež

$$\begin{aligned}
\big\langle(\varphi - \lambda)^*(\boldsymbol{v}), (\varphi - \lambda)^*(\boldsymbol{v})\big\rangle &= \big\langle\big((\varphi - \lambda) \circ (\varphi - \lambda)^*\big)(\boldsymbol{v}), \boldsymbol{v}\big\rangle \\
&= \big\langle\big((\varphi - \lambda)^* \circ (\varphi - \lambda)\big)(\boldsymbol{v}), \boldsymbol{v}\big\rangle \\
&= \big\langle(\varphi - \lambda)(\boldsymbol{v}), (\varphi - \lambda)(\boldsymbol{v})\big\rangle \\
&= 0,
\end{aligned}$$

teda $\varphi^*(\boldsymbol{v}) - \overline{\lambda}\boldsymbol{v} = (\varphi - \lambda)^*(\boldsymbol{v}) = \mathbf{0}$. To znamená, že $\boldsymbol{v}$ je zároveň vlastným vektorom operátora φ^* prislúchajúcim k jeho vlastnej hodnote $\overline{\lambda}$.

Vzájomne združené *normálne* operátory možno súčasne diagonalizovať vzhľadom na *tú istú* ortonormálnu bázu.

23.7.2. Veta. *Nech φ je normálny operátor na n-rozmernom unitárnom priestore V. Potom existuje ortonormálna báza $\boldsymbol{\beta}$ priestoru V vzhľadom na ktorú majú združené operátory φ a φ^* diagonálne matice tvaru $(\varphi)_{\boldsymbol{\beta}} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ resp. $(\varphi^*)_{\boldsymbol{\beta}} = \operatorname{diag}(\overline{\lambda}_1, \dots, \overline{\lambda}_n)$.*

Dôkaz. Budeme postupovať indukciou podľa n. Prípad $n = 0$ je triviálny, pre $n = 1$ vyplýva potrebný záver z predchádzajúceho tvrdenia. Nech teda $n > 1$ a predpokladajme, že tvrdenie platí v unitárnych priestoroch dimenzie $< n$. Podľa tvrdenia 23.7.1 majú normálne operátory φ, φ^* spoločný vlastný vektor $\boldsymbol{v}$, ktorému zodpovedajú vlastné hodnoty λ, $\overline{\lambda}$ operátorov φ, resp. φ^*. Kedže $[\boldsymbol{v}]$ je invariantný podpriestor operátora φ, jeho ortokomplement $S = [\boldsymbol{v}]^\perp$ je podľa tvrdenia 23.1.5 invariantný podpriestor operátora φ^*. Podobne z φ^*-invariantnosti $[\boldsymbol{v}]$ vyplýva φ-invariantnosť S. Potom zúženia $\varphi \restriction S$, $\varphi^* \restriction S$ sú navzájom združené normálne operátory na $(n-1)$-rozmernom unitárnom priestore S. Podľa indukčného predpokladu v S možno nájsť ortonormálnu bázu $\boldsymbol{\beta}_0 = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_{n-1})$, vzhľadom na ktorú majú tieto operátory diagonálne matice $(\varphi \restriction S)_{\boldsymbol{\beta}_0} = (\lambda_1, \dots, \lambda_{n-1})$ resp. $(\varphi^* \restriction S)_{\boldsymbol{\beta}_0} = (\overline{\lambda}_1, \dots, \overline{\lambda}_{n-1})$. Potom $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_{n-1}, \boldsymbol{v})$ je ortonormálna báza priestoru V taká, že $(\varphi)_{\boldsymbol{\beta}} = \operatorname{diag}(\lambda_1, \dots, \lambda_{n-1}, \lambda)$ resp. $(\varphi^*)_{\boldsymbol{\beta}} = \operatorname{diag}(\overline{\lambda}_1, \dots, \overline{\lambda}_{n-1}, \overline{\lambda})$.

23.7.3. Dôsledok. *Nech φ je normálny operátor na n-rozmernom unitárnom priestore V. Potom existujú polynómy $f(x), g(x) \in \mathbb{C}[x]$ stupňa nanajvýš n také, že $\varphi^* = f(\varphi)$ a $\varphi = g(\varphi^*)$.*

Dôkaz. V označení predchádzajúcej vety za $f(x)$ resp. $g(x)$ stačí vziať Lagrangeov interpolačný polynóm taký, že $f(\lambda_1) = \overline{\lambda}_1, \dots, f(\lambda_n) = \overline{\lambda}_n$, resp. $g(\overline{\lambda}_1) = \lambda_1, \dots, g(\overline{\lambda}_n) = \lambda_n$ – pozri cvičenie 10.16, prípadne lemu 22.3.1.

Taktiež naopak, diagonalizovateľnosť vzhľadom na ortonormálnu bázu už má za následok normálnosť lineárneho operátora. Normálne operátory tak možno charakterizovať viacerými ekvivalentnými spôsobmi. Predošlime, že implikácia (i) ⇒ (v) je vlastne vetou o *spektrálnom rozklade normálneho operátora.*

23.7.4. Veta. *Pre ľubovoľný lineárny operátor $\varphi\colon V \to V$ nasledujúce podmienky sú ekvivalentné:*

(i) *φ je normálny, t. j. $\varphi \circ \varphi^* = \varphi^* \circ \varphi$;*

(ii) *pre hermitovské operátory $\varphi_0 = \operatorname{Re} \varphi = \frac{1}{2}(\varphi + \varphi^*)$, $\varphi_1 = \operatorname{Im} \varphi = \frac{1}{2\mathrm{i}}(\varphi - \varphi^*)$ platí $\varphi_0 \circ \varphi_1 = \varphi_1 \circ \varphi_0$;*

(iii) *pre každý invariantný podpriestor $S \subseteq V$ operátora φ aj jeho ortokomplement $S^\perp$ je invariantný;*

(iv) *existuje ortonormálna báza priestoru V, vzhľadom na ktorú má φ diagonálnu maticu s diagonálou tvorenou vlastnými číslami operátora φ;*

(v) *φ možno vyjadriť v tvare lineárnej kombinácie*

$$\varphi = \lambda_1 \operatorname{pr}_{S_1} + \ldots + \lambda_k \operatorname{pr}_{S_k}$$

ortogonálnych projekcií $\operatorname{pr}_{S_1}, \ldots, \operatorname{pr}_{S_k}$ na navzájom ortogonálne vlastné podpriestory prislúchajúce k jeho po dvoch rôznym vlastným číslam $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$, pričom $V = S_1 \oplus \ldots \oplus S_k$.

Dôkaz. (i) ⇔ (ii): Operátory φ_0, φ_1 sú hermitovské podľa tvrdenia 23.2.6. Na základe rovností $\varphi = \varphi_0 + \mathrm{i}\varphi_1$, $\varphi^* = \varphi_0 - \mathrm{i}\varphi_1$ a $\varphi_0 = \frac{1}{2}(\varphi+\varphi^*)$, $\varphi_1 = \frac{1}{2\mathrm{i}}(\varphi-\varphi^*)$ možno ľahko nahliadnuť (prípadne sa presvedčiť priamym výpočtom), že φ, φ^* komutujú práve vtedy, keď komutujú φ_0, φ_1.

(i) ⇒ (iii): Nech φ je normálny a $S \subseteq V$ je φ-invariantný lineárny podpriestor. Potom podľa tvrdenia 23.1.5 jeho ortokomplement $S^\perp$ je φ^*-invariantný podpriestor. Nech $g(x) \in \mathbb{C}[x]$ je polynóm taký, že $\varphi = g(\varphi^*)$, zaručený dôsledkom 23.7.3. Potom $S^\perp$ je zrejme invariatný podpriestor operátora $g(\varphi^*) = \varphi$.

(iii) ⇒ (iv) možno dokázať rovnako ako vetu 23.7.2, lebo podstatným momentom jej dôkazu bolo práve ustanovenie podmienky (iii), a (iv) ⇔ (v) je zrejmé, keďže podmienka (v) je len inou formuláciou podmienky (iv).

(iv) ⇒ (i): V ortonormálnej báze zaručenej podmienkou (iv), v ktorej má φ maticu $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, má združený operátor φ^* maticu $\operatorname{diag}(\overline{\lambda}_1, \ldots, \overline{\lambda}_n)$. Keďže tieto matice komutujú, aj φ, φ^* komutujú.

Zaznamenáme dva dôsledky práve dokázanej vety. Prvým z nich je istá extremálna vlastnosť vlastných čísel normálneho operátora.

23.7.5. Dôsledok. *Nech $\varphi\colon V \to V$ je normálny operátor. Označme*

$$\mu = \min\{|\lambda|;\ \lambda \in \operatorname{Spec}\varphi\}, \qquad M = \max\{|\lambda|;\ \lambda \in \operatorname{Spec}\varphi\}.$$

Potom pre každé $\boldsymbol{x} \in V$ platí

$$\mu\,\|\boldsymbol{x}\| \le \|\varphi\boldsymbol{x}\| \le M\,\|\varphi\boldsymbol{x}\|\,.$$

Poznamenajme, že M je *spektrálny polomer lineárneho operátora* φ (porovnaj so spektrálnym polomerom štvorcových matíc nad $\mathbb{C}$ v paragrafe 22.1).

Dôkaz. Nech $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ je ortonormálna báza priestoru V, vzhľadom na ktorú má φ diagonálnu maticu $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, t. j. $\varphi(\boldsymbol{v}_j) = \lambda_j\boldsymbol{v}_j$ pre $j \le n$. Zvoľme $\boldsymbol{x} = a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n \in V$. Potom $\varphi(\boldsymbol{x}) = a_1\lambda_1\boldsymbol{v}_1 + \ldots + a_n\lambda_n\boldsymbol{v}_n$, $\|\boldsymbol{x}\|^2 = a_1^2 + \ldots + a_n^2$ a $\|\varphi\boldsymbol{x}\|^2 = a_1^2\lambda_1^2 + \ldots + a_n^2\lambda_n^2$. Keďže $\mu \le |\lambda_j| \le M$ pre každé j,

$$\mu\,\|\boldsymbol{x}\| \le \|\varphi\boldsymbol{x}\| \le M\,\|\boldsymbol{x}\|\,.$$

Zrejme dolný (horný) odhad sa dosahuje práve vtedy, keď $\boldsymbol{x}$ je prvkom vlastného podpriestoru $\mathrm{Ker}(\varphi - \lambda)$ pre nejaké vlastné číslo λ operátora φ také, že $|\lambda| = \mu$ ($|\lambda| = M$).

Podľa vety 23.7.4 možno hermitovské resp. unitárne operátory vyčleniť z normálnych operátorov dodatočnou podmienkou pre prvky ich spektier.

23.7.6. Dôsledok. *Nech φ je normálny operátor. Potom*
(a) φ je hermitovský práve vtedy, keď má reálne spektrum;
(b) φ je unitárny práve vtedy, keď $|\lambda| = 1$ pre všetky $\lambda \in \mathrm{Spec}\,\varphi$.

Pojem normálnej matice možno definovať analogicky: $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ je *normálna matica*, ak

$$\boldsymbol{A} \cdot \boldsymbol{A}^* = \boldsymbol{A}^* \cdot \boldsymbol{A}.$$

Teda $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ je normálna práve vtedy, keď $\boldsymbol{A}\cdot\boldsymbol{A}^{\mathsf{T}} = \boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{A}$. Taktiež by malo byť zrejmé, že lineárny operátor na konečnorozmernom unitárnom priestore je normálny práve vtedy, keď jeho matica vzhľadom na nejakú (každú) ortonormálnu bázu je normálna.

Čitateľ si istotne aj sám bez ťažkostí preloží výsledky tohto paragrafu z jazyka operátorov do jazyka matíc. Len ako ukážku zaznamenáme nasledujúcu analógiu viet 23.3.2 a 23.4.6 a pridáme jednu ekvivalentnú podmienku navyše, ktorá znie prirodzenejšie v maticovej formulácii.

23.7.7. Veta. *Pre ľubovoľnú maticu $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ nasledujúce podmienky sú ekvivalentné:*
(i) A je normálna;
(ii) existuje unitárna matica $\boldsymbol{P}$ taká, že matica $\boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P} = \boldsymbol{P}^* \cdot \boldsymbol{A} \cdot \boldsymbol{P}$ je diagonálna, s diagonálou tvorenou vlastnými číslami matice $\boldsymbol{A}$;
(iii) ak $\lambda_1, \dots, \lambda_n$ sú všetky vlastné čísla matice $\boldsymbol{A}$ vrátane násobnosti, tak

$$\sum_{i,j=1}^{n} |a_{ij}|^2 = \sum_{i=1}^{n} |\lambda_i|^2.$$

Dôkaz. Ekvivalencia (i) $\Leftrightarrow$ (ii) je len prekladom ekvivalencie (i) $\Leftrightarrow$ (iv) z vety 23.7.4.

(ii) $\Rightarrow$ (iii): Zrejme $\sum_{i,j} |a_{ij}|^2$ je stopa matice $\boldsymbol{A} \cdot \boldsymbol{A}^*$ a $\sum_i |\lambda_i|^2$ je stopa matice $\mathrm{diag}\big(|\lambda_1|^2, \dots, |\lambda_n|^2\big)$. Ak $\boldsymbol{A} = \boldsymbol{P} \cdot \mathrm{diag}(\lambda_1, \dots, \lambda_n) \cdot \boldsymbol{P}^*$ pre nejakú unitárnu maticu $\boldsymbol{P}$, tak

$$\begin{aligned}\boldsymbol{A} \cdot \boldsymbol{A}^* &= \boldsymbol{P} \cdot \mathrm{diag}(\lambda_1, \dots, \lambda_n) \cdot \boldsymbol{P}^* \cdot \boldsymbol{P} \cdot \mathrm{diag}(\overline{\lambda}_1, \dots, \overline{\lambda}_n) \cdot \boldsymbol{P}^* \\ &= \boldsymbol{P} \cdot \mathrm{diag}\big(|\lambda_1|^2, \dots, |\lambda_n|^2\big) \cdot \boldsymbol{P}^{-1}.\end{aligned}$$

Podľa dôsledku 18.1.4 podobné matice $\boldsymbol{A} \cdot \boldsymbol{A}^*$ a $\mathrm{diag}\big(|\lambda_1|^2, \dots, |\lambda_n|^2\big)$ majú rovnaké stopy.

(iii) $\Rightarrow$ (ii): Podľa Schurovej vety o triangularizácii (veta 19.2.1, druhá časť), existuje unitárna matica $\boldsymbol{P}$ taká, že matica $\boldsymbol{C} = \boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P}$ je horná trojuholníková. Keďže podobné matice $\boldsymbol{A}$, $\boldsymbol{C}$ majú rovnaké spektrum, diagonála matice $\boldsymbol{C}$ je nevyhnutne tvorená prvkami $\lambda_1, \dots, \lambda_n$ (možno v inom poradí) spektra matice $\boldsymbol{A}$. Potom aj matice $\boldsymbol{A} \cdot \boldsymbol{A}^*$ a $\boldsymbol{C} \cdot \boldsymbol{C}^* = \boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{A}^* \cdot \boldsymbol{P}$ sú podobné, preto majú rovnaké stopy. Zrejme

$$\operatorname{tr}(\boldsymbol{C} \cdot \boldsymbol{C}^*) = \sum_i |\lambda_i|^2 + \sum_{i<j} |c_{ij}|^2,$$

a podľa predpokladu $\operatorname{tr}(\boldsymbol{A} \cdot \boldsymbol{A}^*) = \sum_i |\lambda_i|^2$. Z toho vyplýva, že v matici $\boldsymbol{C}$ sú všetky prvky mimo diagonály rovné 0. Teda $\boldsymbol{A}$ je podobná s diagonálnou maticou $\boldsymbol{C}$ prostredníctvom unitárnej matice $\boldsymbol{P}$.

23.7.8. Dôsledok. *Nech $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ je (či už horná alebo dolná) trojuholníková matica. Potom $\boldsymbol{A}$ je normálna práve vtedy, keď je diagonálna.*

Podobne ako v prípade reálnych funkcií reálnej premennej a hermitovských operátorov či matíc, výsledky tohto paragrafu umožňujú rozšíriť obor funkcií, ktoré pracujú na normálnych operátoroch či maticiach, z potenčných radov na ľubovoľné komplexné funkcie komplexnej premennej. Opäť sa sústredíme len na názornejšiu maticovú formuláciu.

Nech $f\colon M \to \mathbb{C}$ je komplexná funkcia komplexnej premennej s definičným oborom $M \subseteq \mathbb{C}$. Ak $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ je normálna matica taká, že $\operatorname{Spec} \boldsymbol{A} \subseteq M$, tak hodnota $f(\boldsymbol{A})$ funkcie f na matici $\boldsymbol{A}$ je podľa paragrafu 22.3 definovaná rovnosťou

$$f(\boldsymbol{A}) = \boldsymbol{P} \cdot \operatorname{diag}\bigl(f(\lambda_1), \dots, f(\lambda_n)\bigr) \cdot \boldsymbol{P}^{-1},$$

kde $\boldsymbol{P}$ je ľubovoľná regulárna matica taká, že $\boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Podľa vety 22.3.2 hodnota $f(\boldsymbol{A})$ nezávisí na matici $\boldsymbol{P}$. Možnosť zvoliť za $\boldsymbol{P}$ *unitárnu* maticu však zaručuje, že aj výsledná matica $f(\boldsymbol{A})$ bude normálna. Tým sme dokázali nasledujúcu vetu.

23.7.9. Veta. *Nech $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ je normálna matica a $f\colon M \to \mathbb{C}$ je funkcia taká, že $\operatorname{Spec} \boldsymbol{A} \subseteq M \subseteq \mathbb{C}$. Potom aj $f(\boldsymbol{A}) \in \mathbb{C}^{n\times n}$ je normálna matica.*

Na záver uvádzame ešte jeden výsledok o *súčasnej diagonalizácii* komutujúceho systému normálnych operátorov.

23.7.10. Veta. *Nech V je konečnorozmerný unitárny priestor a $\mathcal{F}$ je systém normálnych lineárnych operátorov na V. Potom nasledujúce podmienky sú ekvivalentné:*

(i) systém $\mathcal{F}$ je komutujúci, t. j. $\varphi \circ \psi = \psi \circ \varphi$ pre všetky $\varphi, \psi \in \mathcal{F}$;

(ii) existuje ortonormálna báza priestoru V, vzhľadom na ktorú má každý operátor $\varphi \in \mathcal{F}$ diagonálnu maticu;

(iii) existuje nejaká báza priestoru V, vzhľadom na ktorú má každý operátor $\varphi \in \mathcal{F}$ diagonálnu maticu.

Dôkaz. Keďže implikácie (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i) sú triviálne, sústredíme sa len na dôkaz zostávajúcej implikácie (i) $\Rightarrow$ (ii).

Nech $\mathcal{F}$ je komutujúci systém normálnych operátorov na V. Budeme postupovať indukciou podľa $n = \dim V$. Pre $n = 0$ resp. $n = 1$ niet čo dokazovať. Predpokladajme teda, že $n > 1$ a uvedená implikácia platí pre všetky $k < n$. Sú dve možnosti:

1. Každý operátor $\varphi \in \mathcal{F}$ má jednoprvkové spektrum. Podľa vety 23.7.2, resp. vety 23.7.4 z toho vyplýva, že φ má diagonálnu maticu $\lambda \boldsymbol{I}_n$, kde λ je jeho jediné vlastné číslo, vzhľadom na nejakú ortonormálnu bázu priestoru V. Potom nevyhnutne $\varphi = \lambda\, \mathrm{id}_V$ a φ má maticu $\lambda \boldsymbol{I}_n$ vzhľadom na akúkoľvek bázu priestoru V.

2. Existuje nejaký operátor $\varphi \in \mathcal{F}$ s aspoň dvojprvkovým spektrom. Nech λ je jedno z jeho vlastných čísel. Potom pre vlastný podpriestor $S = \mathrm{Ker}(\varphi - \lambda)$ platí $\{\mathbf{0}\} \neq S \neq V$, a takisto $\{\mathbf{0}\} \neq S^\perp \neq V$, teda $\dim S < n$ aj $\dim S^\perp < n$.

Ukážeme, že S je invariantný podpriestor každého operátora $\psi \in \mathcal{F}$. Zvoľme $\boldsymbol{x} \in S$, t. j. $\varphi(\boldsymbol{x}) = \lambda \boldsymbol{x}$. Potom tiež

$$\varphi(\psi \boldsymbol{x}) = \psi(\varphi \boldsymbol{x}) = \psi(\lambda \boldsymbol{x}) = \lambda \psi(\boldsymbol{x}),$$

teda $\psi(\boldsymbol{x}) \in S$. Podľa vety 23.7.4 z toho vyplýva, že aj $S^\perp$ je invariantný podpriestor každého operátora $\psi \in \mathcal{F}$.

Prechodom k zúženiam tak dostávame dva systémy komutujúcich normálnych operátorov $\mathcal{F}_0 = \{\psi \restriction S;\, \psi \in \mathcal{F}\}$ a $\mathcal{F}_1 = \{\psi \restriction S^\perp;\, \psi \in \mathcal{F}\}$ na podpriestoroch S resp. $S^\perp$. Podľa indukčného predpokladu existujú ortonormálne bázy $\boldsymbol{\beta}_0$ podpriestoru S a $\boldsymbol{\beta}_1$ podpriestoru $S^\perp$ také, že matice $(\psi \restriction S)_{\boldsymbol{\beta}_0}$, $(\psi \restriction S^\perp)_{\boldsymbol{\beta}_1}$ sú diagonálne pre všetky $\psi \in \mathcal{F}$. Spojením oboch báz dostávame hľadanú ortonormálnu bázu $\boldsymbol{\beta}$ priestoru V.

Cvičenia

23.1. Dokážte nasledujúce tvrdenia:

(a) Každý lineárny operátor na n-rozmernom vektorovom priestore nad poľom $\mathbb{C}$ má $(n-1)$-rozmerný invariantný podpriestor.

(b) Každý lineárny operátor na n-rozmernom vektorovom priestore nad poľom $\mathbb{R}$ má $(n-1)$- alebo $(n-2)$-rozmerný invariantný podpriestor; ak n je nepárne, tak má $(n-1)$-rozmerný invariantný podpriestor

(*Návod*: Uvedomte si, že na takom vektorovom priestore možno zaviesť skalárny súčin a využite dôsledok 19.5.3 a tvrdenia 23.1.4 a 23.1.5.)

23.2. Nech φ, ψ sú lineárne operátory na reálnom vektorovom priestore V so skalárnym súčinom. Potom φ, ψ sú združené operátory práve vtedy, keď ich komplexifikácie $\varphi^{\mathbb{C}}$, $\psi^{\mathbb{C}}$ sú združené operátory na unitárnom priestore $V^{\mathbb{C}}$ so skalárnym súčinom definovaným ako v cvičení 19.11. Dokážte. Odvoďte z toho, že $\varphi^{\mathbb{C}}$ je hermitovský (unitárny, normálny) práve vtedy, keď φ je symetrický (ortogonálny, normálny).

23.3. Nech φ, ψ sú hermitovské operátory na unitárnom priestore V.

(a) Na príklade ukážte, že už v prípade dvojrozmerného priestoru V operátor $\varphi \circ \psi$ nemusí byť hermitovský.

(b) Dokážte $\varphi \circ \psi + \psi \circ \varphi$ je hermitovský operátor.

(c) Na príklade ukážte, že už v prípade dvojrozmerného priestoru V komutátor $[\varphi, \psi] = \varphi \circ \psi - \psi \circ \varphi$ nemusí byť hermitovský.

(d) Dokážte, že operátor $\mathrm{i}\,[\varphi, \psi]$ je hermitovský.

23.4. Pripomeňme, že lineárny operátor na vektorovom priestore V sa nazýva *projekcia*, ak $p \circ p = p$ (pozri cvičenia 14.8, 14.9). Dokážte nasledujúce tvrdenia:
(a) Ak $p\colon V \to V$ je projekcia, tak $\operatorname{Spec} p \subseteq \{0, 1\}$.

(b) Nech V je unitárny priestor a $p\colon V \to V$ je projekcia. Potom p je samoadjungovaný operátor práve vtedy, keď p je ortogonálna projekcia, (t. j. $p(\boldsymbol{y}) \perp \boldsymbol{x} - p(\boldsymbol{x})$ pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$).

23.5. Uvažujme euklidovský priestor $\mathbb{R}^3$ so štandardným skalárnym súčinom. Dokážte, že pre ľubovoľný vektor $\boldsymbol{u} = (a, b, c)^{\mathsf{T}} \in \mathbb{R}^3$ taký, že $(a, b) \neq \mathbf{0}$, vektory $\boldsymbol{u}$, $\boldsymbol{v} = (-b, a, 0)^{\mathsf{T}}$, $\boldsymbol{w} = \boldsymbol{u} \times \boldsymbol{v}$ tvoria kladne orientovanú ortogonálnu bázu v $\mathbb{R}^3$. Nájdite explicitný tvar vektora $\boldsymbol{w}$. Ako možno čo najjednoduchšie doplniť do kladne orientovanej ortogonálnej bázy nenulový vektor $\boldsymbol{u} = (0, 0, c)^{\mathsf{T}}$?

23.6. Na základe cvičenia 23.5 napíšte maticu $\boldsymbol{R}^{\boldsymbol{u}}_{\alpha}$ otočenia okolo osi so smerovým vektorom $\boldsymbol{u}$ o uhol α (v kladnom zmysle z hľadiska vektora $\boldsymbol{u}$).

23.7. (a) Nájdite matice otočení $\boldsymbol{R}^{\boldsymbol{u}}_{\alpha}$ pre niekoľko vybraných dvojíc vektorov $\boldsymbol{u}$ a uhlov α, kde

$$\boldsymbol{u} \in \left\{ \begin{pmatrix} 1\\1\\0 \end{pmatrix}, \begin{pmatrix} 1\\2\\0 \end{pmatrix}, \begin{pmatrix} 1\\\sqrt{2}\\0 \end{pmatrix}, \begin{pmatrix} 1\\\sqrt{3}\\0 \end{pmatrix}, \begin{pmatrix} 3\\4\\0 \end{pmatrix}, \begin{pmatrix} 1\\1\\1 \end{pmatrix}, \begin{pmatrix} 1\\1\\2 \end{pmatrix}, \begin{pmatrix} 1\\2\\2 \end{pmatrix}, \begin{pmatrix} 1\\\sqrt{2}\\\sqrt{2} \end{pmatrix}, \begin{pmatrix} 1\\\sqrt{2}\\\sqrt{3} \end{pmatrix}, \begin{pmatrix} 3\\4\\5 \end{pmatrix} \right\},$$

$\alpha \in \{\pm\pi/6, \pm\pi/4, \pm\pi/3, \pm\pi/2, \pm 2\pi/3, \pi\}$.

(b) Nájdite matice uvedených otočení (priberte k nim aj otočenie o nulový uhol) zložených so zrkadlovou symetriou podľa roviny prechádzajúcej počiatkom a kolmej na os otočenia.

(c) Pre zopár dvojíc otočení z (a) okolo rôznych osí nájdite os a uhol otočenia, ktoré vznikne ich kompozíciou.

23.8. Zistite, aké transformácie trojrozmerného euklidovského priestoru $\mathbb{R}^3$ (otočenia resp. otočenia zložené so zrkadlovými súmernosťami) sú dané násobením nasledujúcimi ortogonálnymi maticami; zistite uhly a osi týchto otočení (vrátane orientácie):

(a) $\frac{1}{4}\begin{pmatrix} -2\sqrt{3} & \sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & 2+\sqrt{3} & 2-\sqrt{3} \\ \sqrt{2} & 2-\sqrt{3} & 2+\sqrt{3} \end{pmatrix}$, (b) $\frac{1}{4}\begin{pmatrix} 3 & 1 & -\sqrt{6} \\ 1 & 3 & -\sqrt{6} \\ \sqrt{6} & -\sqrt{6} & 2 \end{pmatrix}$, (c) $\frac{1}{6}\begin{pmatrix} 2\sqrt{3} & 2\sqrt{3} & -2\sqrt{3} \\ \sqrt{6} & \sqrt{6} & 2\sqrt{6} \\ 3\sqrt{2} & -3\sqrt{2} & 0 \end{pmatrix}$,

(d) $\frac{1}{4}\begin{pmatrix} 2+\sqrt{2} & 2-\sqrt{2} & -2 \\ 2-\sqrt{2} & 2+\sqrt{2} & 2 \\ -2 & 2 & -2\sqrt{2} \end{pmatrix}$, (e) $\frac{1}{2}\begin{pmatrix} -1 & -\sqrt{2} & 1 \\ -1 & \sqrt{2} & 1 \\ \sqrt{2} & 0 & \sqrt{2} \end{pmatrix}$, (f) $\frac{1}{5\sqrt{2}}\begin{pmatrix} 3 & 4 & 5 \\ -4\sqrt{2} & 3\sqrt{2} & 0 \\ 3 & 4 & -5 \end{pmatrix}$.

23.9. S použitím vety 23.4.9, tvrdenia 23.2.4 a Liouvilleovej formuly 22.2.6 dokážte, že pre unitárnu maticu $\boldsymbol{U}$ je $|\det \boldsymbol{U}| = 1$.

23.10. (Doplnok k vete 23.4.3.)

(a) Nech V je unitárny priestor $f\colon V \to V$ je ľubovoľné (teda nie nutne lineárne) zobrazenie. Ak pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí $\|f(\boldsymbol{x}) - f(\boldsymbol{y})\| = \|\boldsymbol{x} - \boldsymbol{y}\|$, t. j. f zachováva vzdialenosť, tak f je nevyhnutne injektívne. Dokážte.

(b) Označme $\ell_2(\mathbb{C})$ vektorový priestor všetkých postupností $\boldsymbol{x} = (x_k)_{k\in\mathbb{N}}$ komplexných čísel takých, že rad $\sum_{k=0}^{\infty} |x_k|^2$ konverguje. Dokážte, že $\ell_2(\mathbb{C})$ so skalárnym súčinom $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{k=0}^{\infty} x_k \overline{y}_k$ tvorí unitárny priestor (ide o komplexifikáciu vektorového priestoru $\ell_2 = \ell_2(\mathbb{R})$ z cvičenia 14.4).

(c) Nájdite príklad lineárneho operátora $\varphi\colon \ell_2(\mathbb{C}) \to \ell_2(\mathbb{C})$, ktorý spĺňa podmienku $\langle \varphi\boldsymbol{x}, \varphi\boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ pre všetky $\boldsymbol{x}, \boldsymbol{y} \in \ell_2$, no nie je surjektívny (teda ani unitárny).

23.11. Lineárny operátor na konečnorozmernom unitárnom priestore je anthihermitovský práve vtedy, keď má v nejakej ortonormálnej báze priestoru V diagonálnu maticu a všetky nenulové prvky diagonály sú rýdzo imaginárne čísla. Dokážte.

23.12. (a) Nájdite príklad normálnej matice $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$, ktorá nie je symetrická ani antisymetrická ani ortogonálna.

(b) Dokážte, že matica $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$ je normálna práve vtedy, keď je symetrická alebo skalárny násobok ortogonálnej. Ako je to v komplexnom prípade?

(c) Nájdite príklad normálnej matice $\boldsymbol{A} \in \mathbb{R}^{3\times 3}$, ktorá nie je symetrická ani antisymetrická ani skalárny násobok ortogonálnej.

(d) Nájdite príklad normálnej matice $\boldsymbol{A} \in \mathbb{C}^{2\times 2}$, ktorá nie je (komplexným) skalárnym násobkom hermitovskej ani unitárnej matice.

23.13. Sformulujte a dokážte maticovú analógiu vety 23.7.10.

23.14. Odvoďte výsledky o hermitovských a unitárnych operátoroch ako špeciálne prípady príslušných výsledkov o normálnych operátoroch.

24. Kvadriky

Kvadriky sú viacrozmernými zovšeobecneniami kužeľosečiek (t. j. kvadratických kriviek) a kvadratických plôch v euklidovských priestoroch ľubovoľnej dimenzie. Táto kapitola by teda tematicky skôr patrila do Časti II tejto knihy venovanej bilineárnym a kvadratickým formám a ich súvisu s geometriou. Ich zaradenie do tejto časti nám však umožní ich geometrickú klasifikáciu s využitím spektrálnej teórie symetrických lineárnych operátorov resp. symetrických matíc. Hlavným nástrojom tejto analýzy bude reformulácia viet 23.3.1 a 23.3.2 do reči hermitovských poldruhalineárných resp. symetrických bilineárnych foriem, známa ako *veta o hlavných osiach.* Všeobecné výsledky potom spätne aplikujeme na situáciu prístupnú nášmu geometrickému názoru, čiže na kužeľosečky v rovine a kvadratické plochy v trojrozmernom priestore.

24.1 Veta o hlavných osiach

Podľa vety 23.3.2 je každá hermitovská matica $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ podobná s diagonálnou maticou $\boldsymbol{D} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ prostredníctvom nejakej unitárnej matice $\boldsymbol{P} \in \mathbb{C}^{n\times n}$. Samozrejme $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ sú potom všetky vlastné čísla matice $\boldsymbol{A}$, vrátane násobnosti, a stĺpce matice $\boldsymbol{P}$ sú k nim prislúchajúce navzájom ortogonálne vlastné vektory jednotkovej dĺžky. Z rovností $\boldsymbol{P}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{P} = \boldsymbol{D}$ a $\boldsymbol{P}^{-1} = \boldsymbol{P}^*$ však vyplýva $\boldsymbol{P}^* \cdot \boldsymbol{A} \cdot \boldsymbol{P} = \boldsymbol{D}$, teda $\boldsymbol{A}$ je zároveň *hermitovsky kongruentná* s $\boldsymbol{D}$ prostredníctvom unitárnej matice $\boldsymbol{P}$. Každá symetrická matica $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ je zasa podobná aj kongruentná s diagonálnou maticou prostredníctvom ortogonálnej matice $\boldsymbol{P} \in \mathbb{R}^{n\times n}$. Tento výsledok možno sformulovať v jazyku kososymetrických poldruhalineárnych resp. symetrických bilineárnych foriem ako *vetu o hlavných osiach.*

24.1.1. Veta. *(a) Každá kososymetrická poldruhalineárna forma $F\colon V \times V \to \mathbb{C}$ na konečnorozmernom unitárnom priestore V má vo vhodnej ortonormálnej báze tohto priestoru reálnu diagonálnu maticu.*

(b) Každá symetrická bilineárna forma $F\colon V \times V \to \mathbb{R}$ na euklidovskom priestore V má vo vhodnej ortonormálnej báze tohto priestoru diagonálnu maticu.

Pre reálny prípad uvádzame aj formuláciu v jazyku kvadratických foriem.

24.1.2. Dôsledok. *Každá kvadratická forma $q\colon V \to \mathbb{R}$ na euklidovskom priestore V má vo vhodnej ortonormálnej báze tohto priestoru diagonálnu maticu.*

Upozorňujeme, že v oboch častiach vety aj v jej dôsledku je dôraz na slove *ortonormálna* – existenciu vôbec nejakej bázy, vzhľadom na ktorú je matica príslušnej formy diagonálna, sme totiž odôvodnili už v paragrafoch 17.2 resp. 11.3.

Hlavnými osami príslušnej formy rozumieme navzájom kolmé priamky, t. j. jednorozmerné lineárne podpriestory, generované vektormi spomínanej ortonormálnej bázy.

24.2 Kvadriky v euklidovských priestoroch

V tomto paragrafe preskúmame geometrický význam dôsledku 24.1.2 v kontexte euklidovských priestorov.

Kvadrikou v euklidovskom priestore V rozumieme množinu bodov tvaru

$$Q = \{\boldsymbol{x} \in V;\, q(\boldsymbol{x}) + \varphi(\boldsymbol{x}) + c = 0\},$$

kde $q\colon V \to \mathbb{R}$ je nenulová kvadratická forma, $\varphi\colon V \to \mathbb{R}$ je ľubovoľný lineárny funkcionál (lineárna forma) a $c \in \mathbb{R}$ je ľubovoľný skalár. Formulu

$$q(\boldsymbol{x}) + \varphi(\boldsymbol{x}) + c = 0$$

nazývame *rovnicou kvadriky Q*.

Kvadriky v dvojrozmernom euklidovskom priestore, t. j. v rovine, sa nazývajú *kužeľosečky.* Názov sa odvodzuje z pozorovania, že (s výnimkou prázdnej množiny a dvojice rovnobežných priamok) ich možno získať ako rezy rotačnej kužeľovej plochy vhodnými rovinami. Kvadriky v trojrozmernom euklidovskom priestore sa nazývajú tiež *kvadratické plochy.* Ide napospol o krivky a plochy dobre známe z elementárnej geometrie a hojne využívané v technickej praxi. Vo všeobecnosti sú kvadriky v n-rozmernom euklidovskom priestore (s výnimkou niektorých singulárnych prípadov) $(n-1)$-rozmernými nadplochami.

Veta o hlavných osiach umožňuje tzv. *metrickú klasifikáciu kvadrík.* Z jej hľadiska považujeme dve kvadriky za rovnaké, ak možno jednu na druhú zobraziť zhodným zobrazením (izometrickou transformáciou) euklidovského priestoru. Pri takejto klasifikácii určujeme nielen typ kvadriky (elipsa, hyperbola, parabola a pod.), ale tiež jej *metrické invarianty*, napr. dĺžky jej polosí, a *kanonický tvar jej rovnice.* Sú možné aj iné klasifikácie, založené na porovnávaní kvadrík pomocou iných typov zobrazení. Napr. pri *afinnej klasifikácii* by sme považovali dve kvadriky za rovnaké, ak by sme vedeli zobraziť jednu na druhú bijektívnou afinnou transformáciou príslušného priestoru. Pri takejto klasifikácii by sme napr. v euklidovskej rovine rozlišovali elipsy, hyperboly a paraboly, ale už by sme nedokázali rozlíšiť elipsu od kružnice ani dve elipsy či dve hyperboly s polosami rôznych dĺžok.

Hodnosťou resp. *signatúrou kvadriky* Q budeme rozumieť hodnosť resp. signatúru jej kvadratickej formy q; t. j. $h(Q) = h(q)$ a $\sigma(Q) = \sigma(q)$. [1]

Podľa dôsledku 24.1.2 existuje ortonormálna báza $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ priestoru V, vzhľadom na ktorú má kvadratická forma q diagonálnu maticu $\boldsymbol{D} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Diagonálne prvky λ_i matice $\boldsymbol{D}$ nájdeme ako vlastné čísla (symetrickej) matice formy q vzhľadom na akúkoľvek ortonormálnu bázu priestoru V (v typickom prípade euklidovského priestoru $V = \mathbb{R}^n$ so štandardným skalárnym súčinom to býva kanonická ortonormálna báza $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$). Báza $\boldsymbol{\beta}$ je potom tvorená navzájom ortogonálnymi vlastnými vektormi $\boldsymbol{v}_i$, prislúchajúcimi k vlastným číslam λ_i, znormovanými na jednotkovú dĺžku. Ak si označíme $(\boldsymbol{x})_{\boldsymbol{\beta}} = (x_1, \dots, x_n)^{\mathsf{T}}$ súradnice vektora $\boldsymbol{x}$ v báze $\boldsymbol{\beta}$ a položíme $b_i = \varphi(\boldsymbol{v}_i)$, môžeme rovnicu kvadriky Q vyjadriť v tvare

$$\sum_{i=1}^{n} \lambda_i x_i^2 + \sum_{i=1}^{n} b_i x_i + c = 0.$$

Ak si ďalej označíme $h = h(Q)$ hodnosť a $\sigma(Q) = (k, h-k, n-h)$ signatúru kvadriky Q, možno navyše bez ujmy na všeobecnosti predpokladať, že $\lambda_i > 0$ pre $i \le k$, $\lambda_i < 0$ pre $k < i \le h$ a $\lambda_i = 0$ pre $i > h$.

Lineárnych členov $b_i x_i$ pre $i \le h$ sa možno zbaviť doplnením na úplný štvorec a rovnicu kvadriky následne zapísať v tvare

$$\sum_{i=1}^{h} \lambda_i \left(x + \frac{b_i}{2\lambda_i} \right)^2 + \sum_{i=h+1}^{n} b_i x_i = d,$$

kde

$$d = \frac{1}{4} \sum_{i=1}^{h} \frac{b_i^2}{\lambda_i} - c.$$

V tejto chvíli rozlíšime dve možnosti:

(1) $h = n$ alebo $b_i = 0$ pre všetky $h < i \le n$, inak povedané, nezostali nám nijaké lineárne členy. *Kvadriky*, ktorých rovnicu možno upraviť na tvar bez lineárnych členov, nazývame *stredové*.

(2) $h < n$ a $b_i \ne 0$ pre aspoň jedno $h < i \le n$. *Kvadriky*, z ktorých rovnice nemožno eliminovať lineárne členy, nazývame *nestredové*.

(1) Pre stredovú kvadriku zavedieme nové súradnice substitúciou

$$y_i = \begin{cases} x_i + \frac{b_i}{2\lambda_i}, & \text{ak } 1 \le i \le h, \\ x_i, & \text{ak } h < i \le n, \end{cases}$$

[1] V tomto prípade, ako i miestami ďalej ide o nie celkom štandardnú terminológiu, ktorá sa však dobre hodí na účely nášho zámerne tak trochu kuchynského výkladu.

čo zodpovedá posunutiu počiatku súradníc o vektor, ktorého súradnice v báze $\boldsymbol{\beta}$ sú $(b_1/2\lambda_1, \ldots, b_h/2\lambda_h, 0, \ldots, 0)^\mathsf{T}$. Touto *izometrickou transformáciou* rovnica kvadriky nadobudne tvar

$$\sum_{i=1}^{h} \lambda_i y_i^2 = d.$$

Vidíme, že stredová kvadrika je skutočne symetrická podľa stredu, ktorý sme práve zvolili za nový počiatok súradnej sústavy (jej rovnica je invariantná voči substitúcii $y_i \mapsto -y_i$).

U stredových kvadrík je účelné ďalej rozlíšiť dva prípady:

(1a) $d = 0$, kedy ide o tzv. *homogénne stredové kvadriky.* Ich rovnicu možno upraviť na tvar

$$\sum_{i=1}^{k} \frac{y_i^2}{a_i^2} = \sum_{i=k+1}^{h} \frac{y_i^2}{a_i^2},$$

kde $a_i = 1/\sqrt{|\lambda_i|}$.

(1b) $d \neq 0$, čo sú tzv. *nehomogénne stredové kvadriky.* Ich rovnica prejde po prenásobení faktorom $1/d$ a substitúcii $a_i = \sqrt{|d/\lambda_i|}$ na tvar

$$\sum_{i=1}^{k} \frac{y_i^2}{a_i^2} - \sum_{i=k+1}^{h} \frac{y_i^2}{a_i^2} = 1,$$

ak $d > 0$, prípadne na tvar

$$-\sum_{i=1}^{k} \frac{y_i^2}{a_i^2} + \sum_{i=k+1}^{h} \frac{y_i^2}{a_i^2} = 1,$$

ak $d < 0$. Zrejme stačí uvažovať prvú možnosť, lebo druhú možno na ňu previesť nahradením pôvodnej rovnice kvadriky Q ekvivalentnou rovnicou $-q(\boldsymbol{x}) - \varphi(\boldsymbol{x}) - c = 0$.

Stredové kvadriky typu (1a), t. j. homogénne, a typu (1b), t. j. nehomogénne, pre ktoré navyše platí $h < n$, sa nazývajú *singulárne. Regulárne stredové kvadriky* sú potom nehomogénne, spĺňajúce podmienku $h = n$.

(2) V prípade nestredovej kvadriky sa na vektor $\boldsymbol{b} = (b_{h+1}, \ldots, b_n)$ pozrieme ako na prvok *riadkového* euklidovského priestoru $\mathbb{R}^{n-h}$ so štandardným skalárnym súčinom a položíme

$$p = \frac{1}{2}\,\|\boldsymbol{b}\| = \frac{1}{2}\sqrt{b_{h+1}^2 + \ldots + b_n^2}$$

(dôvody pre faktor $1/2$ sú viac-menej „tradično-kozmetické“ a majú pôvod v syntetickej geometrii – pozri cvičenie 24.5). Teraz si stačí uvedomiť, že jednotkový vektor $-(1/2p)\boldsymbol{b}$ možno doplniť do nejakej ortonormálnej bázy priestoru $\mathbb{R}^{n-h}$, ktorú možno zapísať ako ortogonálnu maticu $\boldsymbol{B} = (b_{ij})$, s *riadkami* tvorenými vektormi príslušnej bázy. (Ako hneď uvidíme, na konkrétnom tvare ďalších riadkov matice $\boldsymbol{B}$ nezáleží, teda ich vôbec nemusíme hľadať.) Nové súradnice zavedieme substitúciou

$$y_i = \begin{cases} x_i + \frac{b_i}{2\lambda_i}, & \text{ak } i \le h, \\ \frac{1}{2p}\Big(d - \sum_{j=h+1}^{n} b_j x_j\Big), & \text{ak } i = h+1, \\ \sum_{j=h+1}^{n} b_{ij} x_j, & \text{ak } h+1 < i \le n. \end{cases}$$

Vektor $(y_1, \dots, y_h, y_{h+1}, \dots, y_n)^{\mathsf{T}}$ dostaneme z vektora $(x_1, \dots, x_h, x_{h+1}, \dots, x_n)^{\mathsf{T}}$ vynásobením ortogonálnou maticou $\operatorname{diag}(\boldsymbol{I}_h, \boldsymbol{B})$ zľava a následným pričítaním vektora $\left(\frac{b_1}{2\lambda_1}, \dots, \frac{b_h}{2\lambda_h}, \frac{d}{2p}, 0, \dots, 0\right)^{\mathsf{T}}$. Teda substitúcia $x_i \mapsto y_i$ predstavuje *izometrickú afinnú transformáciu* priestoru $\mathbb{R}^n$ (násobenie ortogonálnou maticou zložené s posunutím).

Rovnica kvadriky v nových súradniciach prejde na tvar

$$\sum_{i=1}^{h} \lambda_i y_i^2 = 2p y_{h+1},$$

v ktorom nevystupujú premenné y_i pre $i > h+1$, teda nie sú v nich skryte prítomné ani prvky riadkov matice $\boldsymbol{B}$ okrem prvého. Konečne, ak položíme $a_i = 1/\sqrt{|\lambda_i|}$, môžeme zapísať rovnicu kvadriky v tvare

$$\sum_{i=1}^{k} \frac{y_i^2}{a_i^2} - \sum_{i=k+1}^{h} \frac{y_i^2}{a_i^2} = 2p y_{h+1}.$$

Nestredová kvadrika sa nazýva *regulárna*, ak $h = n-1$; v opačnom prípade, t. j. ak $h < n-1$, sa nazýva *singulárna*.

Skôr než zhrnieme naše úvahy do záverečnej podoby, pripomeňme, že *repérom* vo vektorovom priestore V nazývame usporiadanú dvojicu $\boldsymbol{\rho} = (\boldsymbol{r}, \boldsymbol{\beta})$, kde $\boldsymbol{r} \in V$ je ľubovoľný bod (vektor) a $\boldsymbol{\beta}$ je ľubovoľná báza vo V. *Afinné* alebo tiež *barycentrické súradnice bodu $\boldsymbol{x} \in V$ vzhľadom na repér $\boldsymbol{\rho}$* sú definované ako súradnice vektora $\boldsymbol{x} - \boldsymbol{r}$ vzhľadom na bázu $\boldsymbol{\beta}$, čiže

$$(\boldsymbol{x})_{\boldsymbol{\rho}} = (\boldsymbol{x} - \boldsymbol{r})_{\boldsymbol{\beta}}$$

(pozri cvičenia 8.7–10). *Repér $\boldsymbol{\rho} = (\boldsymbol{r}, \boldsymbol{\beta})$* euklidovského priestoru V sa nazýva *ortogonálny* resp. *ortonormálny*, ak báza $\boldsymbol{\beta}$ má príslušnú vlastnosť.

24.2.1. Veta. *Nech Q je ľubovoľná kvadrika s hodnosťou $h(Q) = h > 0$ a signatúrou $\sigma(Q) = (k, h-k, n-h)$ v n-rozmernom euklidovskom priestore V. Potom vo V existuje ortonormálny repér $\boldsymbol{\rho}$ a kladné čísla $a_1, \ldots, a_h$, prípadne tiež p, také, že Q má v afinných súradniciach $(\boldsymbol{x})_{\boldsymbol{\rho}} = (x_1, \ldots, x_n)^{\mathsf{T}}$ vzhľadom na $\boldsymbol{\rho}$ rovnicu v jednom z nasledujúcich* kanonických tvarov

$$\sum_{i=1}^{k} \frac{x_i^2}{a_i^2} = \sum_{i=k+1}^{h} \frac{x_i^2}{a_i^2},$$

ak Q je stredová homogénna,

$$\sum_{i=1}^{k} \frac{x_i^2}{a_i^2} - \sum_{i=k+1}^{h} \frac{x_i^2}{a_i^2} = 1,$$

ak Q je stredová nehomogéna, resp.

$$\sum_{i=1}^{k} \frac{x_i^2}{a_i^2} - \sum_{i=k+1}^{h} \frac{x_i^2}{a_i^2} = 2px_{h+1},$$

ak Q je nestredová. V prípade stredových homogénnych ako aj nestredových kvadrík možno navyše bez ujmy na všeobecnosti predpokladať, že $k \geq h-k$, t. j. $2k \geq h$.

Ešte poznamenajme, že *hlavnými osami kvadriky Q* nazývame navzájom kolmé priamky $\boldsymbol{r} + [\boldsymbol{v}_i]$, t. j. jednorozmerné afinné podpriestory prechádzajúce novým počiatkom súradnej sústavy $\boldsymbol{r}$ so zameraniami generovanými ortonormálnymi bázovými vektormi $\boldsymbol{v}_i$. Bod $\boldsymbol{r}$ nazývame v prípade stredových kvadrík *stredom kvadriky*; ľahko nahliadneme, že každá stredová kvadrika je naozaj symetrická podľa svojho stredu. V prípade nestredových kvadrík nazývame bod $\boldsymbol{r}$ *vrcholom kvadriky*; zrejme platí $\boldsymbol{r} \in Q$. Čísla a_i nazývame *dĺžkami polosí kvadriky Q*; číslo p nazývame *parametrom nestredovej kvadriky Q*. Tieto veličiny sú však určené jednoznačne iba pre nehomogénne stredové kvadriky. V prípade homogénnej stredovej kvadriky, ako aj v prípade nestredovej kvadriky sú určené len s presnosťou na ľubovoľnú kladnú multiplikatívnu konštantu. Zmysel tohto faktu, ako i niektoré ďalšie súvislosti by sa vyjasnili až v kontexte *projektívnej geometrie*, čo je však téma, ktorej sa v tomto kurze nebudeme venovať. Uvedenú nejednoznačnosť možno odstrániť napr. dodatočnou požiadavkou $a_1 = 1$, alebo $a_h = 1$, prípadne pre nestredové kvadriky $p = \frac{1}{2}$. Tieto možnosti občas využijeme pri nasledujúcej podrobnej klasifikácii kvadrík v rovine a v (trojrozmernom) priestore. Do euklidovských priestorov vyšších dimenzií už nezablúdime, lebo v dôsledku absencie priamej skúsenosti nám v nich zlyháva geometrická predstavivosť a chýbajú názvy pre rôzne typy kvadrík, s ktorými sa tam možno stretnúť.

24.3 Kužeľosečky

Pripomeňme, že *kužeľosečkami* nazývame kvadriky v dvojrozmernom euklidovskom priestore čiže v rovine. Na základe vety 24.2.1 ich možno klasifikovať podľa kanonického tvaru ich rovníc.

(1a) Stredové homogénne (t. j. singulárne) kužeľosečky:

signatúra	*rovnica*	*kužeľosečka*
$(1,0,1)$	$x^2 = 0$	priamka $x = 0$,
$(2,0,0)$	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 0$	bod $x = y = 0$,
$(1,1,0)$	$\dfrac{x^2}{a^2} = \dfrac{y^2}{b^2}$	dve rôznobežné priamky $y = \pm(b/a)x$.

(1b) Stredové nehomogénne kužeľosečky:

signatúra	*rovnica*	*kužeľosečka*
$(1,0,1)$	$\dfrac{x^2}{a^2} = 1$	dve rovnobežné priamky $x = \pm a$,
$(0,1,1)$	$-\dfrac{x^2}{a^2} = 1$	prázdna možina,
$(2,0,0)$	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$	elipsa,
$(1,1,0)$	$\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$	hyperbola,
$(0,2,0)$	$-\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$	prázdna množina.

Poznamenajme, že z uvedených kužeľosečiek považujeme za singulárny len prvý a druhý prípad. Teda za regulárne kužeľosečky považujeme popri elipse a hyperbole aj prázdnu množinu riešení poslednej rovnice. Toto rozlišovanie medzi dvoma prázdnymi množinami zodpovedajúcimi signatúram $(0,1,1)$ a $(0,2,0)$ najskôr čitateľovi pripadá umelé. Rozdiel medzi oboma typmi kužeľosečiek by sa prejavil, až keby sme vzali do úvahy imaginárne riešenia oboch rovníc.

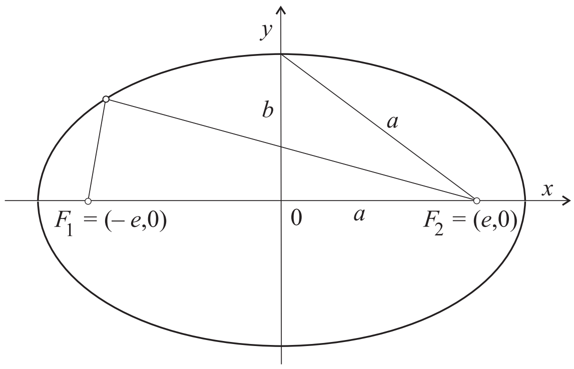
Z dĺžok a, b *polosí elipsy* za *hlavnú* považujeme väčšiu a *vedľajšiu* menšiu; ak $a = b$, ide samozrejme o *kružnicu* s polomerom a a rovnicou $x^2 + y^2 = a^2$. Pri *hyperbole* s rovnicou v uvedenom tvare je a vždy dĺžkou *hlavnej* a b *vedľajšej polosi*; ak $a = b$, hovoríme o *rovnoosej hyperbole*. Priamky $y = \pm(b/a)x$ sú *asymptoty hyperboly*.

(2) Nestredové kužeľosečky:

signatúra	*rovnica*	*kužeľosečka*
$(1, 0, 1)$	$x^2 = 2py$	parabola.

Na rozdiel od stredových kužeľosečiek, ktoré sú súmerné podľa oboch hlavných osí aj podľa stredu, parabola má len jednu os súmernosti – v uvedenom prípade je to os y.

Na obrázkoch sú znázornené len elipsa, hyperbola a parabola. Čitateľ by si však mal pri každom z uvedených typov samostatne rozmyslieť, ako treba viesť rez rotačnej kužeľovej plochy s rovnicou $x^2 + y^2 = z^2$ v trojrozmernom priestore rovinou tak, aby sme dostali tú ktorú kužeľosečku.



Obr. 24.1: Elipsa s ohniskami F_1, F_2, hlavnou polosou a, vedľajšou polosou b a excentricitou e (pozri cvičenie 24.3)

24.3.1. Príklad. Napíšeme rovnicu elipsy so stredom $\boldsymbol{s} = (-1, 2)^\mathsf{T}$, hlavnou polosou $a = \sqrt{5}$ v smere vektora $(3, -1)^\mathsf{T}$ a vedľajšou polosou $b = 1/2$ v smere vektora $(1, 3)^\mathsf{T}$.

Znormovaním uvedených vektorov dostaneme ortonormálnu bázu (tvorenú stĺpcami matice)

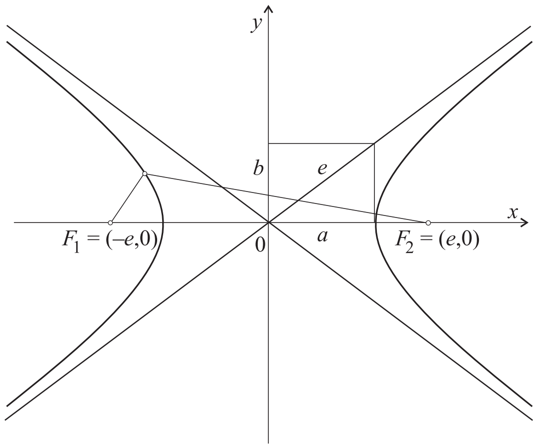
$$\boldsymbol{\gamma} = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}$$

V súradniciach $(u, v)^\mathsf{T}$ vzhľadom na repér $(\boldsymbol{s}, \boldsymbol{\gamma})$ má elipsa rovnicu v kanonickom tvare

$$\frac{u^2}{5} + 4v^2 = 1.$$

Vzťah medzi súradnicami $(u, v)^\mathsf{T}$ a súradnicami $(x, y)^\mathsf{T}$ vzhľadom na kanonický repér $(\boldsymbol{0}, \boldsymbol{\varepsilon})$ je daný rovnosťou

$$\begin{pmatrix} u \\ v \end{pmatrix} = \boldsymbol{P}_{\boldsymbol{\gamma},\boldsymbol{\varepsilon}} \cdot \left(\begin{pmatrix} x \\ y \end{pmatrix} - \boldsymbol{s} \right) = \boldsymbol{\gamma}^\mathsf{T} \cdot \begin{pmatrix} x + 1 \\ y - 2 \end{pmatrix} = \frac{1}{\sqrt{10}} \begin{pmatrix} 3x - y + 5 \\ x + 3y - 5 \end{pmatrix}$$



Obr. 24.2: Hyperbola s ohniskami F_1, F_2, hlavnou polosou a, vedľajšou polosou b a excentricitou e (pozri cvičenie 24.4)

(pozri cvičenie 8.10). Dosadením za u a v do pôvodnej rovnice dostávame

$$\frac{(3x-y+5)^2}{50} + \frac{2(x+3y-5)^2}{5} = 1,$$

čo po prenásobení faktorom 50 a ďalších úpravách dáva výslednú rovnicu

$$29x^2 + 181y^2 + 114xy - 170x - 610y + 475 = 0,$$

z ktorej by sme na prvý pohľad asi sotva uhádli, akú kužeľosečku popisuje.

24.3.2. Príklad. V euklidovskom priestore $\mathbb{R}^2$ je pre ľubovoľnú hodnotu parametra $t \in \mathbb{R}$ daná kužeľosečka Q_t s rovnicou

$$tx^2 + ty^2 + 2xy + \sqrt{2}\,x - \sqrt{2}\,y = 0.$$

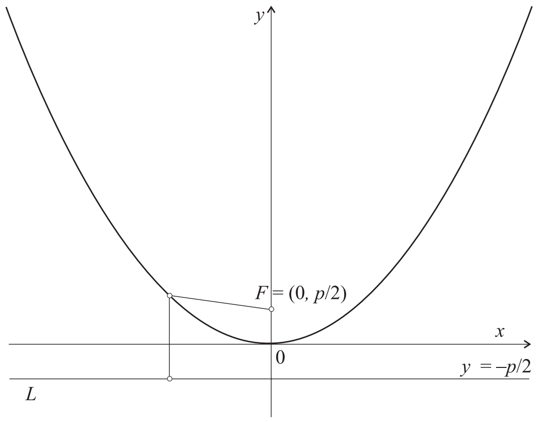
Zistíme o akú kužeľosečku ide v závislosti na t.

Matica príslušnej kvadratickej formy

$$\boldsymbol{A} = \begin{pmatrix} t & 1 \\ 1 & t \end{pmatrix}$$

má vlastné hodnoty $t+1$ a $t-1$, ktorým zodpovedajú vlastné vektory $(1,1)^{\mathsf{T}}$ resp. $(-1,1)^{\mathsf{T}}$. Príslušná orotonormálna báza je (tvorená stĺpcami matice)

$$\boldsymbol{\beta} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$



Obr. 24.3: Parabola s ohniskom F, riadiacou priamkou L a parametrom p pozri cvičenie 24.5)

V súradniciach $(u,v)^{\mathsf T} = \boldsymbol{\beta}^{\mathsf T} \cdot (x,y)^{\mathsf T}$ vzhľadom na túto bázu má rovnica kužeľosečky tvar $(t+1)u^2 + (t-1)v^2 + \bigl(\sqrt{2}, -\sqrt{2}\bigr) \cdot \boldsymbol{\beta} \cdot (u,v)^{\mathsf T} = 0$, t. j.

$$(t+1)u^2 + (t-1)v^2 - 2v = 0.$$

Odteraz budeme všetky údaje popisovať v súradniciach $(u,v)^{\mathsf T}$ vzhľadom na bázu $\boldsymbol{\beta}$. Ak teda budeme hovoriť napr. o bode $(3,-1)^{\mathsf T}$, budeme mať na mysli bod, ktorého súradnice vzhľadom na pôvodnú bázu $\boldsymbol{\varepsilon}$ sú $\boldsymbol{\beta} \cdot (3,-1)^{\mathsf T} = \bigl(2\sqrt{2}, \sqrt{2}\bigr)^{\mathsf T}$.

Rozlíšime tri základné prípady:

(a) $t = -1$. Rovnica $-2v^2 - 2v = 0$ po úprave dáva $v(v+1) = 0$, takže Q_{-1} je zjednotením dvojice rovnobežných priamok s rovnicami $v = 0$ a $v = -1$.

(b) $t = 1$. Rovnica $2u^2 - 2v = 0$ po úprave dáva $u^2 = v$, čo je rovnica paraboly s vrcholom v počiatku $\mathbf{0}$ a osou súmernosti v.

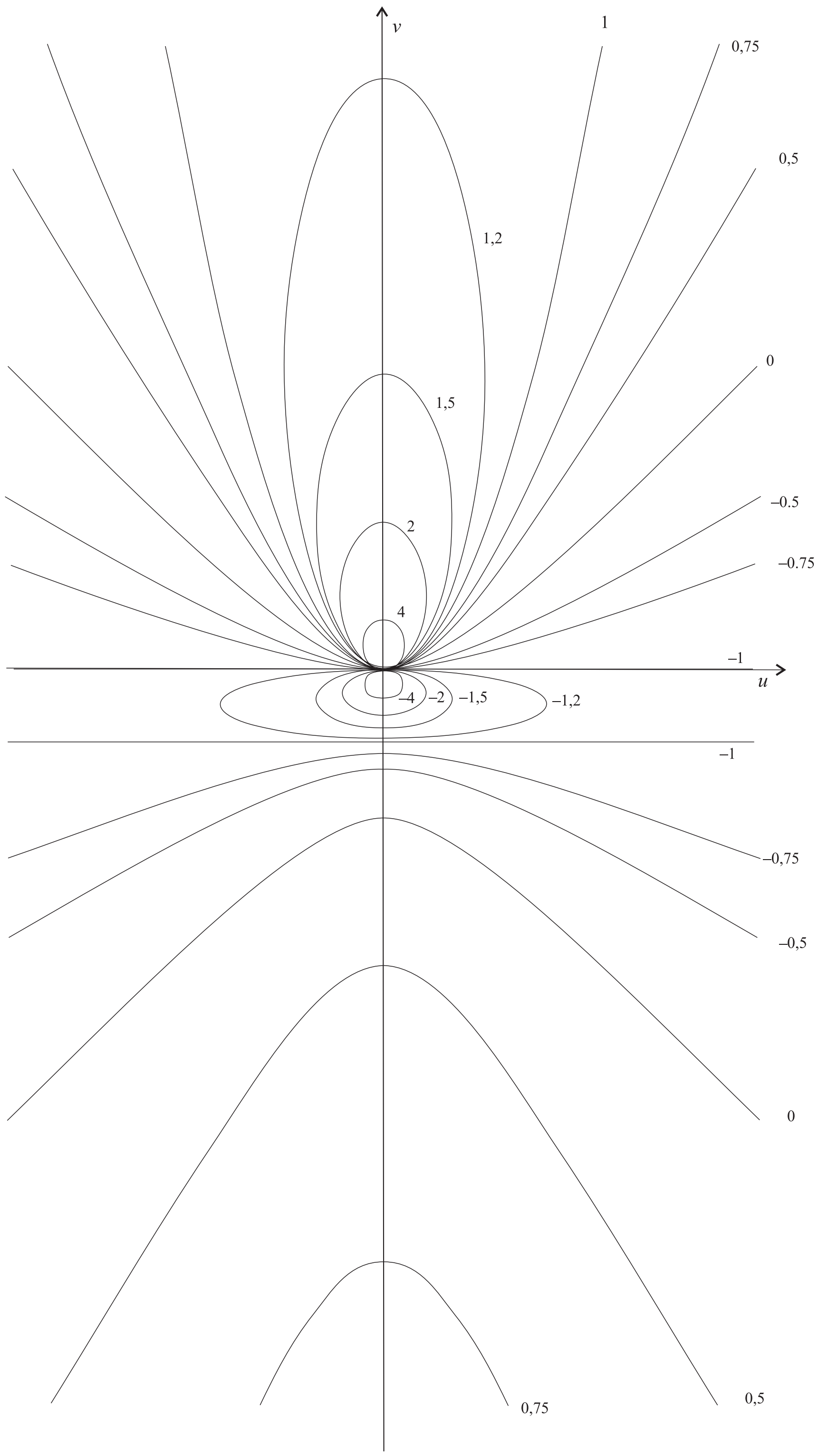
(c) $t \neq \pm 1$. Rovnicu kužeľosečky Q_t možno doplnením na štvorec upraviť na tvar

$$(t+1)u^2 + (t-1)\left(v - \frac{1}{t-1}\right)^2 - \frac{1}{t-1} = 0,$$

čiže

$$\bigl(t^2-1\bigr)u^2 + (t-1)^2\left(v - \frac{1}{t-1}\right)^2 = 1. \qquad (*)$$

Z toho je zrejmé, že Q_t je stredová kužeľosečka so stredom v bode $\bigl(0, \frac{1}{t-1}\bigr)^{\mathsf T}$ a jedným z vrcholov v počiatku. Jej hlavné osi ležia na priamkach s rovnicami $v = \frac{1}{t-1}$ (čo je rovnobežka s osou u) resp. $u = 0$ (čo je priamo os v). Pri ďalšej analýze opäť rozlíšime dva prípady:



Obr. 24.4. Príklad parametrického systému kužeľosečiek.

(c1) $|t| < 1$. $(*)$ je potom rovnicou hyperboly s hlavnou polosou dĺžky $\frac{1}{1-t}$ na priamke $u = 0$ a vedľajšou polosou dĺžky $\frac{1}{\sqrt{1-t^2}}$ na priamke $v = \frac{1}{t-1}$.

(c2) $|t| > 1$. $(*)$ je potom rovnicou elipsy s polosami dĺžok $\frac{1}{\sqrt{t^2-1}}$, $\frac{1}{|t-1|}$, ktoré ležia na priamkach $v = \frac{1}{t-1}$ resp. $u = 0$. Ak $t < -1$, tak prvá z nich je hlavná a druhá vedľajšia; pre $t > 1$ je tomu naopak.

Ak parameter t interpretujeme ako čas, tak vývoj krivky Q_t pripúšťa s trochou nadsádzky fabuláciu bezmála kozmologickú:

Pred úsvitom času (t. j. pre $t = -\infty$) je „všetko" sústredené v jedinom bode v počiatku súradného systému $\mathbf{0}uv$. Ten bod sa náhle začne rozpínať do elipsy, ktorej stred sa pohybuje po osi v nadol smerom k bodu $-\frac{1}{2}$, jej vedľajšia polos sa rozťahuje po osi v až k medznej dĺžke $\frac{1}{2}$, zatiaľ čo dĺžka jej hlavnej polosi rastie v smere rovnobežnom s osou u nad všetky medze. V okamihu $t = -1$ sa elipsa roztrhne na dve rovnobežné priamky $v = 0$ a $v = -1$. Tie sa vzápätí začnú prehýbať ako dve vetvy hyperboly – horná má vrchol v počiatku, kým vrchol dolnej sa po osi v vzdaľuje stále ďalej nadol, obe vetvy sa stále väčšmi prehýbajú a v čase $t = 0$ tvoria rovnoosú hyperbolu s rovnicou $(v+1)^2 - u^2 = 1$. Ohýbanie a vzďaľovanie dolnej vetvy stále pokračuje, až sa v okamihu $t = 1$ dolná vetva úplne stratí za obzorom a horná sa zmení na parabolu s rovnicou $u^2 = v$. Parabola sa vzápätí uzavrie do elipsy s jedným vrcholom v počiatku – jej hlavná polos v smere osi v ako aj vedľajšia polos v smere rovnobežnom s osou u sa postupne skracujú, až kým na súmraku času (t. j. pre $t = +\infty$) celá krivka neskolabuje do jediného bodu v počiatku, z ktorého kedysi povstala.

24.4 Kvadratické plochy

Opäť pripomíname, že kvadriky v trojrozmernom euklidovskom priestore nazývame aj *kvadratickými plochami*. Klasifikácia podľa kanonického tvaru ich rovníc na základe vety 24.2.1 je už pomerne rozsiahla.

(1a) Stredové homogénne (t. j. singulárne) kvadratické plochy:

signatúra	*rovnica*	*kvadrika*
$(1,0,2)$	$x^2 = 0$	rovina $x = 0$,
$(2,0,1)$	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 0$	priamka $x = y = 0$,
$(1,1,1)$	$\dfrac{x^2}{a^2} = \dfrac{y^2}{b^2}$	dve rôznobežné roviny $y = \pm(b/a)x$,
$(3,0,0)$	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 0$	bod $x = y = z = 0$,
$(2,1,0)$	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = z^2$	eliptická kužeľová plocha.

V poslednom prípade sú rezy eliptickej kužeľovej plochy vodorovnými rovinami $z = c$ elipsy s polosami dĺžok $a|c|$, $b|c|$ (pre $c = 0$ ide pochopiteľne o bod $x = y = 0$). Ak $a = b$, hovoríme o *rotačnej kužeľovej ploche.*

(1b) Stredové nehomogénne kvadratické plochy:

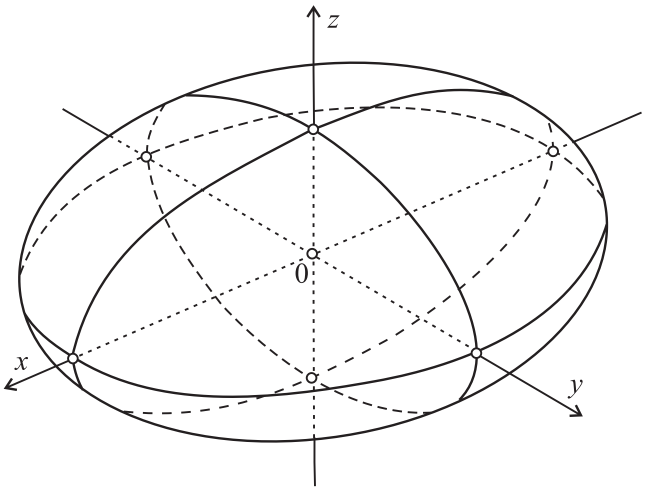
signatúra	*rovnica*	*kvadrika*
$(1,0,2)$	$\dfrac{x^2}{a^2} = 1$	dve rovnobežné roviny $x = \pm a$,
$(0,1,2)$	$-\dfrac{x^2}{a^2} = 1$	prázdna množina,
$(2,0,1)$	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$	eliptická valcová plocha,
$(1,1,1)$	$\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$	hyperbolická valcová plocha,
$(0,2,1)$	$-\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$	prázdna množina,
$(3,0,0)$	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$	trojosý elipsoid,
$(2,1,0)$	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$	jednodielny eliptický hyperboloid,
$(1,2,0)$	$\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$	dvojdielny eliptický hyperboloid,
$(0,3,0)$	$-\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$	prázdna množina.

Prvých päť z uvedených kvadratických plôch je singulárnych, ďalšie štyri sú regulárne. Opäť sa tak stretáme s javom tento raz dokonca „trojakej prázdnej množiny“, s možnosťou rozlíšenia až v komplexnom rozšírení.

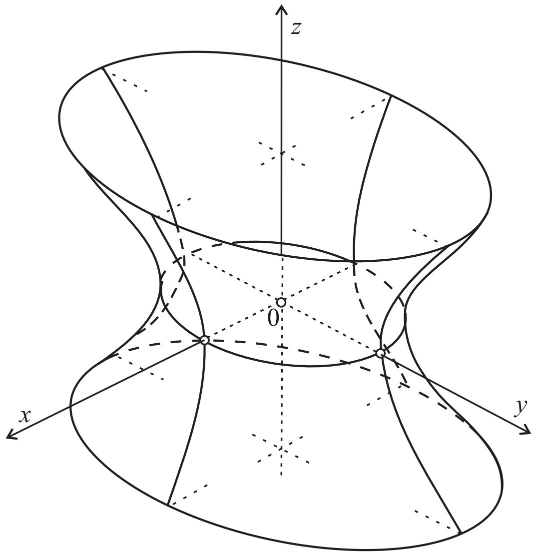
Všetky rezy uvedenej eliptickej valcovej plochy vodorovnými rovinami $z = c$ sú elipsy s polosami dĺžok a, b. Ak $a = b$, ide o *rotačnú valcovú plochu.* Podobne, rezy hyperbolickej valcovej plochy takýmito rovinami sú hyperboly s hlavnou polosou dĺžky a a vedľajšou polosou dĺžky b. Ak $a = b$, sú to rovnoosé hyperboly.

Trojosý elipsoid je niečo ako „guľová plocha s rôznymi polomermi“ v smeroch osí x, y, z. Tieto „rôzne polomery“ sú samozrejme dĺžky polosí a, b, c nášho elipsoidu. O naozajstnú *guľovú (sférickú) plochu* ide, len ak $a = b = c$. Ak sa dve z dĺžok polosí rovnajú a dĺžka zvyšnej polosi je väčšia, hovoríme o *vajcovitom* alebo *pretiahnutom rotačnom elipsoide*; ak je dĺžka zvyšnej polosi menšia, ide o *sploštený rotačný elipsoid.*

Podobná situácia je pri hyperboloidoch. V uvedenom prípade sú rezy jednodielneho eliptického hyperboloidu vodorovnými rovinami $z = d$ elipsy



Obr. 24.5. Trojosý elipsoid

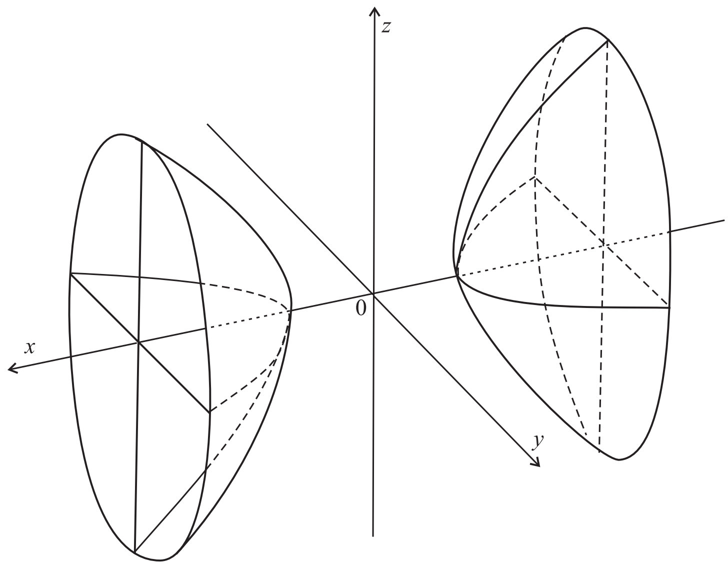


Obr. 24.6. Jednodielny eliptický hyperboloid

s polosami dĺžok $a\sqrt{1+(d/c)^2}$, $b\sqrt{1+(d/c)^2}$; ak $a=b$, sú tieto rezy kružnice a hovoríme o *rotačnom jednodielnom hyperboloide*. Rezy jednodielneho hyperboloidu zvislými rovinami $y=0$, resp. $x=0$ sú hyperboly s hlavnou polosou dĺžky a resp. b a vedľajšou polosou dĺžky c. „Vnútornou“ *asymptotickou plochou jednodielneho eliptického hyperboloidu* je v danom prípade eliptická kužeľová plocha s rovnicou $x^2/a^2+y^2/b^2=z^2/c^2$ (pozri tiež posledný obrázok k paragrafu 16.2).

Rezy uvedeného dvojdielneho eliptického hyperboloidu zvislými rovinami

$x = d$, kde $|d| > a$, sú elipsy s polosami dĺžok $b\sqrt{d^2/a^2 - 1}$, $c\sqrt{d^2/a^2 - 1}$ (ak $|d| = a$, je tento rez jediný bod, ak $|d| < a$, je rezom prázdna množina – preto sa plocha rozpadá na dve časti). Ak $b = c$, ide o *rotačný dvojdielny hyperboloid*. Rezy vodorovnou rovinou $z = 0$, resp. zvislou rovinou $y = 0$ sú hyperboly s hlavnou polosou dĺžky a a vedľajšou polosou dĺžky b resp. c. „Vonkajšou" *asymptotickou plochou dvojdielneho eliptického hyperboloidu* je tentokrát eliptická kužeľová plocha s rovnicou $x^2/a^2 = y^2/b^2 + z^2/c^2$ (pozri tiež posledný obrázok k paragrafu 16.2).



Obr. 24.7. Dvojdielny eliptický hyperboloid

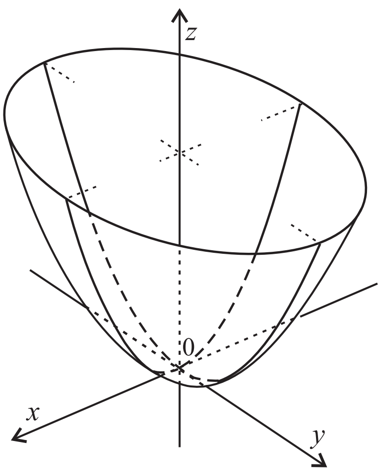
(2) Nestredové kvadratické plochy:

signatúra	*rovnica*	*kvadrika*
$(1, 0, 2)$	$x^2 = 2py$	parabolická valcová plocha,
$(2, 0, 1)$	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = z$	eliptický paraboloid,
$(1, 1, 1)$	$\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = z$	hyperbolický paraboloid.

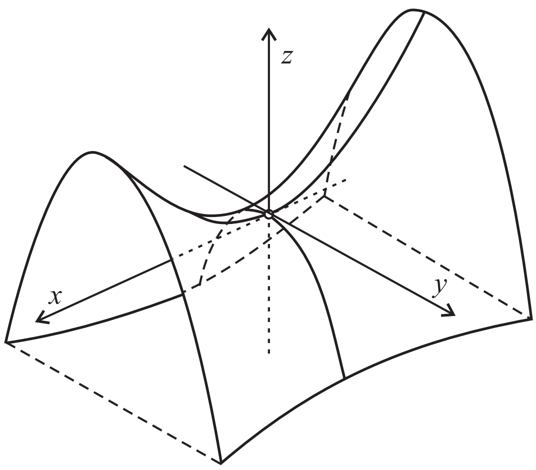
Prvá z uvedných plôch je singulárna, ďalšie dve sú regulárne.

Rezy parabolickej valcovej plochy vodorovnými rovinami $z = c$ sú paraboly zhodné s parabolou $x^2 = 2py$ v rovine $z = 0$.

Rezy eliptického paraboloidu vodorovnými rovinami $z = c$, kde $c > 0$, sú elipsy s polosami dĺžok $a\sqrt{c}$, $b\sqrt{c}$ (pre $c = 0$ je to bod, pre $c < 0$ prázdna množina). Ak $a = b$, ide o *rotačný paraboloid*. Rezy zvislými rovinami $x = 0$, resp. $y = 0$ sú paraboly s rovnicami $y^2 = b^2 z$, resp. $x^2 = a^2 z$.



Obr. 24.8. Eliptický paraboloid



Obr. 24.9. Hyperbolický paraboloid

Rezy hyperbolického paraboloidu vodorovnými rovinami $z = c$ sú pre $c > 0$ hyperboly s hlavnou polosou dĺžky $a\sqrt{c}$ v smere osi x a vedľajšou polosou dĺžky $b\sqrt{c}$ v smere osi y; pre $c < 0$ sú to hyperboly s hlavnou polosou dĺžky $b\sqrt{-c}$ v smere osi y a vedľajšou polosou dĺžky $a\sqrt{-c}$ v smere osi x. Ak $a = b$, tak všetky tieto rezy sú rovnoosé hyperboly. Pre $c = 0$ je rezom dvojica priamok $y = \pm(b/a)x$. Rezy zvislými rovinami $x = 0$, resp. $y = 0$ sú paraboly s rovnicami $y^2 = -b^2 z$, resp. $x^2 = a^2 z$. Hyperbolický paraboloid, ako graf funkcie $z = x^2/a^2 - y^2/b^2$, má sedlo v bode $(0, 0)$.

24.4.1. Príklad. V euklidovskom priestore $\mathbb{R}^3$ so štandardným skalárnym súčinom je rovnicou

$$x^2 + 2z^2 + 8\sqrt{6}\,xy - 2\sqrt{2}\,xz + 8\sqrt{3}\,yz = 12$$

daná kvadrika Q. Treba zistiť, o akú kvadriku ide, a určiť jej polohu, smery hlavných osí a metrické invarianty.

Príslušná kvadratická forma má maticu

$$\boldsymbol{A} = \begin{pmatrix} 1 & 4\sqrt{6} & -\sqrt{2} \\ 4\sqrt{6} & 0 & 4\sqrt{3} \\ -\sqrt{2} & 4\sqrt{3} & 2 \end{pmatrix}$$

s vlastnými číslami 3, 12 a -12, ktorým zodpovedajú vlastné vektory $\left(1, 0, -\sqrt{2}\right)^{\mathsf{T}}$, $\left(\sqrt{2}, \sqrt{3}, 1\right)^{\mathsf{T}}$ a $\left(\sqrt{2}, -\sqrt{3}, 1\right)^{\mathsf{T}}$. Hlavnými osami kvadriky Q sú tak priamky generované týmito vektormi. Po ich znormovaní dostávame ortonormálnu bázu

$$\boldsymbol{\beta} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

priestoru $\mathbb{R}^3$, tvorenú vlastnými vektormi matice $\boldsymbol{A}$ (ktorú, ako obvykle, stotožníme s uvedenou maticou).

Potom v súradniciach $(u, v, w)^{\mathsf{T}} = \boldsymbol{\beta}^{\mathsf{T}} \cdot (x, y, z)^{\mathsf{T}}$ vzhľadom na bázu $\boldsymbol{\beta}$ má rovnica kvadriky Q tvar $3u^2 + 12v^2 - 12w^2 = 12$. Z toho po úprave dostávame

$$\frac{u^2}{4} + v^2 - w^2 = 1,$$

čo je rovnica jednodielného eliptického hyperboloidu so stredom v počiatku $\mathbf{0}$. Rezom tejto kvadriky rovinou $w = 0$ je elipsa s hlavnou polosou $a = 2$ v smere osi u a vedľajšou polosou $b = 1$ v smere osi v. Jej rezom rovinou $v = 0$ je hyperbola s hlavnou polosou $a = 2$ v smere osi u a vedľajšou polosou $c = 1$ v smere osi w. Konečne jej rezom rovinou $u = 0$ je rovnoosá hyperbola s polosami $b = c = 1$ v smeroch osí v, w.

24.4.2. Príklad. V euklidovskom priestore $\mathbb{R}^3$ so štandardným skalárnym súčinom je rovnicou

$$41x^2 + 41y^2 + 8z^2 - 80xy - 4xz - 4yz - 48x - 48y - 51z + 180 = 0$$

daná kvadrika Q. Treba zistiť, o akú kvadriku ide, a určiť jej polohu, smery hlavných osí a metrické invarianty.

Príslušná kvadratická forma má maticu

$$\boldsymbol{A} = \begin{pmatrix} 41 & -40 & -2 \\ -40 & 41 & -2 \\ -2 & -2 & 8 \end{pmatrix}$$

s vlastnými číslami 81, 9 a 0 a im zodpovedajúcimi vlastnými vektormi $(1,-1,0)^{\mathsf{T}}$, $(1,1,-4)^{\mathsf{T}}$ a $(2,2,1)^{\mathsf{T}}$. Hlavnými osami kvadriky Q sú priamky generované týmito vektormi. Ich znormovaním dostávame ortonormálnu bázu

$$\boldsymbol{\beta} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ -\frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & -\frac{2\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix}$$

priestoru $\mathbb{R}^3$, tvorenú vlastnými vektormi matice $\boldsymbol{A}$.

V súradniciach $(\xi,\eta,\zeta)^{\mathsf{T}} = \boldsymbol{\beta}^{\mathsf{T}} \cdot (x,y,z)^{\mathsf{T}}$ vzhľadom na bázu $\boldsymbol{\beta}$ má rovnica kvadriky Q tvar $81\xi^2 + 9\eta^2 + (-48,-48,-51)\cdot\boldsymbol{\beta}\cdot(\xi,\eta,\zeta)^{\mathsf{T}} + 180 = 0$, t. j.

$$81\xi^2 + 9\eta^2 + 18\sqrt{2}\,\eta - 81\zeta + 180 = 0.$$

Z toho po niekoľkých úpravach a predelení oboch strán rovnice faktorom 81 dostaneme

$$\xi^2 + \frac{1}{9}(\eta+\sqrt{2})^2 - (\zeta-2) = 0.$$

Záverečnou substitúciou $u=\xi$, $v=\eta+\sqrt{2}$, $w=\zeta-2$ prejde rovnica na kanonický tvar

$$u^2 + \frac{v^2}{9} = w,$$

čo je rovnica eliptického paraboloidu s osou súmernosti w. Jeho rez rovinou $w=1$ je elipsa s hlavnou polosou $b=3$ v smere osi v a vedľajšou polosou $a=1$ v smere osi u.

Podrobnejšia analýza oboch substitúcií dáva

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \xi \\ \eta+\sqrt{2} \\ \zeta-2 \end{pmatrix} = \boldsymbol{\beta}^{\mathsf{T}} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 0 \\ -\sqrt{2} \\ 2 \end{pmatrix} = \boldsymbol{\beta}^{\mathsf{T}} \cdot \begin{pmatrix} x-1 \\ y-1 \\ z-2 \end{pmatrix}.$$

Vidíme tak, že $(u,v,w)^{\mathsf{T}}$ predstavujú súradnice vzhľadom na repér $(\boldsymbol{r},\boldsymbol{\beta})$, kde $\boldsymbol{r} = (1,1,2)^{\mathsf{T}} = \boldsymbol{\beta}\cdot\left(0,-\sqrt{2},2\right)^{\mathsf{T}}$ je vrchol kvadriky Q.

24.4.3. Príklad. Hyperbolický paraboloid $Q \subseteq \mathbb{R}^3$ má v súradniciach $(u,v,w)^{\mathsf{T}}$ vzhľadom na ortonormálny repér $(\boldsymbol{s},\boldsymbol{\alpha})$, kde $\boldsymbol{s} = (-2,0,1)^{\mathsf{T}}$ a

$$\boldsymbol{\alpha} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} & 0 \\ -\frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

rovnicu v kanonickom tvare

$$u^2 - \frac{v^2}{2} = w.$$

Napíšeme jeho rovnicu v súradniciach $(x, y, z)^\mathsf{T}$ vzhľadom na kanonický repér $(\mathbf{0}, \boldsymbol{\varepsilon})$.

Stačí si uvedomiť, že prechod od jedných súradníc k druhým je daný rovnicou

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\varepsilon}} \cdot \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \boldsymbol{s} \right) = \boldsymbol{\alpha}^\mathsf{T} \cdot \begin{pmatrix} x+2 \\ y \\ z-1 \end{pmatrix} = \begin{pmatrix} (3x-4y+6)/5 \\ (4x+3y+8)/5 \\ z-1 \end{pmatrix}.$$

Po dosadení posledného výrazu za $(u, v, w)^\mathsf{T}$ do kanonickej rovnice kvadriky Q dostávame

$$\frac{(3x-4y+6)^2}{25} - \frac{(4x+3y+8)^2}{50} = z - 1,$$

čo po prenásobení faktorom 50 a ďalších úpravách dáva

$$2x^2 + 23y^2 - 72xy + 68x + 26y - 50z + 53 = 0.$$

Cvičenia

24.1. Dokážte *vetu o súčasnej diagonalizácii dvoch kvadratických foriem*: Nech V je konečnorozmerný vektorový priestor nad poľom $\mathbb{R}$ a $p, q\colon V \to \mathbb{R}$ sú dve kvadratické formy na V, pričom p je kladne definitná. Potom existuje taká báza $\boldsymbol{\beta}$ priestoru V, že v súradniciach $(\boldsymbol{x})_{\boldsymbol{\beta}} = (x_1, \dots, x_n)^\mathsf{T}$ vzhľadom na túto bázu majú obe formy diagonálny tvar $p(\boldsymbol{x}) = x_1^2 + \ldots + x_n^2$, $q(x) = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2$.

24.2. Klasifikujte kvadriky na euklidovskej priamke podľa signatúry a načrtnite ich obrázky. Presvedčte sa, že neprázdne singulárne kužeľosečky, s výnimkou bodu a dvojice rôznobežných priamok, vzniknú vždy z vhodnej kvadriky na priamke tak, že každým jej bodom vedieme na túto priamku kolmicu. Načrtnite obrázok každej z nich.

Nasledujúce tri cvičenia sú venované vzťahu medzi tzv. syntetickým a analytickým popisom regulárnych kužeľosečiek.

24.3. Elipsa je syntetickým spôsobom definovaná ako množina bodov v euklidovskej rovine, ktoré majú rovnaký predpísaný súčet vzdialeností od daných dvoch bodov F_1, F_2 tejto roviny, tzv. *ohnísk elipsy.*

(a) Zvoľte si v rovine súradný systém tak, aby obe ohniská ležali na osi x súmerne voči počiatku, teda $F_1 = (-e, 0)$, $F_2 = (e, 0)$, kde $e \geq 0$ je tzv. *excentricita elipsy.* Predpísaný súčet vzdialeností bodov na elipse od oboch ohnísk označte $2a$ a položte

$b = \sqrt{a^2 - e^2}$. Dokážte, že bod roviny so súradnicami (x, y) leží na takto synteticky určenej elipse práve vtedy, keď $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. (Pozri obr. 24.3.)

(b) Akú podmienku musia spĺňať parametre a a e, aby šlo naozaj o elipsu? Aké množiny bodov v rovine môžeme dostať, ak táto podmienka nie je splnená?

(c) Ako sa nazýva elipsa s nulovou excentricitou?

24.4. Hyperbola je syntetickým spôsobom definovaná ako množina bodov v euklidovskej rovine, ktoré majú v absolútnej hodnote rovnaký predpísaný rozdiel vzdialeností od daných dvoch bodov F_1, F_2 tejto roviny, tzv. *ohnísk hyperboly.*

(a) Zvoľte si v rovine súradný systém tak, aby obe ohniská ležali na osi x súmerne voči počiatku, teda $F_1 = (-e, 0)$, $F_2 = (e, 0)$, kde $e \geq 0$ je tzv. *excentricita hyperboly.* Prepdpísanú absolútnu hodnotu rozdielu vzdialeností bodov na elipse od oboch ohnísk označte $2a$ a položte $b = \sqrt{e^2 - a^2}$. Dokážte, že bod roviny so súradnicami (x, y) leží na takto synteticky určenej hyperbole práve vtedy, keď $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$. (Pozri obr. 24.2.)

(b) Aké podmienky musia spĺňať parametre a a e, aby šlo naozaj o hyperbolu? Aké množiny bodov v rovine môžeme dostať, ak tieto podmienky nie sú splnené?

(c) Je rovnoosá hyperbola to isté ako hyperbola s nulovou excentricitou? Akým vzťahom medzi parametrami a, e zodpovedá prvý prípad? Akú množinu bodov dostaneme v druhom prípade?

24.5. Parabola je syntetickým spôsobom definovaná ako množina bodov v euklidovskej rovine, ktoré majú rovnakú vzdialenosť od danej *riadiacej priamky* L a daného bodu $F \notin L$, tzv. *ohniska paraboly*, ležiacich v tejto rovine.

(a) Zvoľte si v rovine súradný systém tak, aby os y prechádzala ohniskom F kolmo na riadiacu priamku a ich priesečník bol súmerne združený s ohniskom podľa počiatku. Teda $F = (0, p/2)$ a riadiaca priamka má rovnicu $y = -p/2$, kde $p > 0$ je tzv. *parameter paraboly.* Dokážte, že bod roviny so súradnicami (x, y) leží na takto synteticky určenej parabole práve vtedy, keď $x^2 = 2py$. (Pozri obr. 24.3.)

(b) Ako by vyzerala „parabola“ v prípade, že $F \in L$, t. j. $p = 0$?

24.6. Presvedčte sa, že neprázdne singulárne kvadratické plochy, s výnimkou bodu a eliptickej kužeľovej plochy, vzniknú vždy z vhodnej kužeľosečky tak, že každým jej bodom vedieme priamku kolmú na rovinu, v ktorej leží. Načrtnite obrázok každej z nich.

24.7. Pre každú z kvadratických plôch určte všetky jej osi a roviny súmernosti.

24.8. Pripomíname, že bod $\boldsymbol{s}$ euklidovského priestoru V sa nazýva *stred kvadriky* $Q \subseteq V$, ak kvadrika Q je symetrická podľa tohto stredu. Dokážte postupne tvrdenia (a) – (d):

(a) Bod $\boldsymbol{s} \in V$ je stredom kvadriky $Q \subseteq V$ práve vtedy, keď pre každý bod $\boldsymbol{x} \in V$ platí $\boldsymbol{x} \in Q \Rightarrow 2\boldsymbol{s} - \boldsymbol{x} \in Q$.

(b) Nech $(\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot \boldsymbol{A}(\boldsymbol{x})_{\boldsymbol{\alpha}} + \boldsymbol{b} \cdot (\boldsymbol{x})_{\boldsymbol{\alpha}} + c = 0$ je rovnica kvadriky Q v súradniciach vzhľadom na ľubovoľnú bázu $\boldsymbol{\alpha}$ euklidovského priestoru V, kde $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ je symetrická matica, $\boldsymbol{b} \in \mathbb{R}^n$ je riadkový vektor a $c \in \mathbb{R}$. Potom bod $\boldsymbol{s} \in V$ je stredom kvadriky Q práve vtedy, keď $2\boldsymbol{A} \cdot (\boldsymbol{s})_{\boldsymbol{\alpha}} + \boldsymbol{b}^{\mathsf{T}} = \boldsymbol{0}$.

(c) Množina všetkých stredov stredovej kvadriky tvorí afinný podpriestor vo V.

(d) Regulárna stredová kvadrika má práve jeden stred.

(e) Nájdite príklady singulárnych stredových kvadrík, ktoré majú práve jeden stred, ako aj singulárnych stredových kvadrík, ktorých stredy tvoria priamku resp. rovinu.

24.9. Určte druh, polohu a metrické invarianty kužeľosečiek v $\mathbb{R}^2$ daných rovnicami:

(a) $2x^2 + 2y^2 - 4y + 1 = 0$; (b) $x^2 + y^2 - 6x + 8y = 0$;

(c) $7x^2 + 2xy\sqrt{3} + 5y^2 = 8$; (d) $3x^2 + 2xy + 3y^2 - 8x - 8y + 4 = 0$;

(e) $x^2 - 4y^2 + 2x + 16y = 19$; (f) $11x^2 + 24xy + 4y^2 - 26x + 8y = 1$;

(g) $x^2 + 2x - 2y + 5 = 0$; (h) $x^2 - 2xy + y^2 - 2x\sqrt{2} - 2y\sqrt{2} = 0$;

(i) $9x^2 - 24xy + 16y^2 - 34x - 38y + 71 = 0$.

Zakaždým načrtnite obrázok.

24.10. V závislosti na parametri t určte druh, polohu a metrické invarianty kužeľosečiek v $\mathbb{R}^2$ daných rovnicami:

(a) $tx^2 + ty^2 - 4xy = 8$; (b) $x^2 - y^2 + 2txy = 1$;

(c) $(t+4)x^2 + (t-4)y^2 - 6xy = t - 5$; (d) $(t+3)x^2 + (t-3)y^2 + 8xy = (t-5)^2$;

(e) $(t+1)x^2 + (t-2)y^2 - 4xy = 1$; (f) $(t+2)x^2 + (t-1)y^2 + 4xy = (t+3)(t-2)$.

Podobne ako na obr. 24.4, načrtnite tieto kužeľosečky pre viacero hodnôt parametra t a pozorujte, ako sa transformujú s meniacim sa t.

24.11. Napíšte rovnice nasledujúcich kužeľosečiek v $\mathbb{R}^2$:

(a) kružnica so stredom $(-4, 3)$ a polomerom 5;

(b) elipsa so stredom v počiatku, hlavnou polosou $a = (1 + \sqrt{5})/2$ v smere vektora $(1, \sqrt{3})$ a vedľajšou polosou $b = 1$;

(c) rovnoosá hyperbola s ohniskami $F_1 = (-1, -1)$, $F_2 = (1, 1)$;

(d) parabola s ohniskom $F = (1, -5)$ a vrcholom $V = (0, -4)$.

24.12. Určte typ, polohu a metrické invarianty kvadrík v $\mathbb{R}^3$ daných rovnicami

(a) $x^2 + 3y^2 + z^2 + \sqrt{2}\,xy + 4xz + \sqrt{2}\,yz = 0$;

(b) $3x^2 + 6y^2 + 3z^2 + 2\sqrt{2}\,xy + 6xz + 2\sqrt{2}\,yz = 4$;

(c) $2x^2 + 3y^2 + 2z^2 + \sqrt{2}\,xy + 2xz + \sqrt{2}\,yz - \big(6+\sqrt{2}\big)x - 2\big(3+\sqrt{2}\big)y - \big(6+\sqrt{2}\big)z + 7 + \sqrt{2} = 0$;

(d) $x^2 + y^2 - 3z^2 - 6xy - 2xz - 2yz + 14x - 10y + 2z + 21 = 0$;

(e) $13x^2 + 13y^2 + 7z^2 - 22xy + 10xz + 10yz + 70x - 74y - 10z + 121 = 0$;

(f) $x^2 + y^2 + 7z^2 + 2xy + 10xz + 10yz + 12 = 0$;

(g) $13x^2 + 37y^2 + 30z^2 + 30xy + 2\sqrt{6}\,xz + 2\sqrt{6}\,yz - 32\sqrt{6}\,x + 32\sqrt{6}\,y - 64z = 0$;

(h) $7x^2 - 17y^2 - 6z^2 - 6xy - 10\sqrt{6}\,xz - 14\sqrt{6}\,yz - 4\sqrt{6}\,x + 4\sqrt{6}\,y - 8z = 0$;

(i) $2x^2 + 2y^2 + 2z^2 + 4xy - 4xz - 4yz - \big(9\sqrt{2} + 4\big)x + \big(9\sqrt{2} - 4\big)y + 4z + 2 = 0$.

24.13. Napíšte rovnice nasledujúcich kvadrík $\mathbb{R}^3$:

(a) rotačný elipsoid so stredom v počiatku, ktorý vznikne rotáciou elipsy s hlavnou polosou $a = 1$ a vedľajšou polosou $b = 1/\sqrt{2}$ okolo jej hlavnej osi so smerovým vektorom $(1, 1, 1)$;

(b) rotačný elipsoid so stredom v počiatku, ktorý vznikne rotáciou elipsy s hlavnou polosou $a = 1$ a vedľajšou polosu $b = 1/\sqrt{2}$ okolo jej vedľajšej osi so smerovým vektorom $(1, -2, 1)$;

(c) jednoplochý eliptický hyperboloid, ktorý vznikne otočením $\boldsymbol{R}_{\alpha}^{\boldsymbol{e}_3}$ základného eliptického hyperboloidu s rovnicou $x^2 + 2y^2 - z^2 = 1$ okolo osi z o uhol $\alpha = \pi/4$ a následným posunutím o vektor $(1, 2, 0)$;

(d) dvojplochý rotačný hyperboloid so stredom v počiatku, ktorý vznikne rotáciou rovnoosej hyperboly s excentricitou $e = \sqrt{2}$ okolo jej hlavnej osi so smerovým vektorom $(1, 1, 1)$;

(e) eliptická kužeľová plocha, ktorá vznikne otočením $\boldsymbol{R}_{\beta}^{\boldsymbol{e}_1}$ základnej eliptickej kužeľovej plochy s rovnicou $x^2 + 3y^2 - 3z^2 = 0$ okolo osi x o uhol $\beta = \pi/6$;

(f) eliptický paraboloid s vrcholom v bode $(0, 0, 1)$, ktorého os súmernosti je rovnobežná s osou z a rez rovinou $z = -1$ je elipsa s hlavnou polosou $a = 1$ v smere vektora $(1, \sqrt{3}, 0)$ a vedľajšou polosou $b = 1/\sqrt{3}$;

(g) hyperbolický paraboloid, ktorý vznikne otočením $\boldsymbol{R}_{\gamma}^{\boldsymbol{u}}$ základného hyperbolického paraboloidu s rovnicou $x^2 - y^2 = z$ okolo priamky prechádzajúcej počiatkom so smerovým vektorom $\boldsymbol{u} = (1, 1, 0)$ o uhol $\gamma = \pi/2$ a následným posunutím o vektor $(2, -1, -3)$.

25. Vybrané aplikácie združených operátorov

V tejto pomerne rôznorodej kapitole sa – po kužeľosečkách a kvadrikách – oboznámime s niekoľkými ďalšími aplikáciami združených (najmä hermitovských) operátorov.

Výsledky prvého a druhého paragrafu, venovaného *singulárnemu* a *polárnemu rozkladu* resp. *pseudoinverzným maticiam*, sa vo veľkej miere využívajú v numerickej matematike, v ktorej sa však hojne vyskytujú aj iné typy rozkladov. K nim patrí napr. *QR-rozklad*, s ktorým sme sa už stretli v cvičeniach 13.13 a 13.14 resp. 17.16 a 17.17, ďalej tzv. *LU-rozklad*, *LDU-rozklad* atď., no tým, ani ich aplikáciám sa už v tomto kurze venovať nebudeme. S aplikáciou iného druhu – tentoraz v geometrii – sa stretneme v treťom paragrafe: pomocou singulárnych čísel dokážeme elegantne vyjadriť odchýlku ľubovoľných lineárnych či afinných podpriestorov v euklidovskom priestore.

Dva záverečné paragrafy tvoria ucelenú časť. Riešením diferenciálnej rovnice *harmonického oscilátora* sa pokúsime motivovať stručnú analýzu vzťahu samoadjungovaných operátorov $\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}$ a $-\frac{\mathrm{d}^2}{\mathrm{d}t^2}$ (na vhodnom priestore funkcií) a *(trigonometrických) Fourierových radov*.

25.1 Singulárny a polárny rozklad

V tomto paragrafe budeme ďalej rozvíjať paralelu medzi reálnymi číslami a hermitovskými operátormi resp. maticami.

Hermitovský lineárny operátor φ na unitárnom priestore V sa nazýva

(a) *kladný* prípadne *pozitívny*, ak pre každý vektor $\mathbf{0} \neq \boldsymbol{x} \in V$ platí $\langle \varphi\boldsymbol{x}, \boldsymbol{x}\rangle > 0$;

(b) *nezáporný*, ak pre každý vektor $\boldsymbol{x} \in V$ platí $\langle \varphi\boldsymbol{x}, \boldsymbol{x}\rangle \geq 0$.

Vďaka samoadjungovanosti operátora φ je $F(\boldsymbol{x}, \boldsymbol{y}) = \langle \varphi\boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \varphi\boldsymbol{y}\rangle$ kososymetrická bilineárna forma na V. Preto φ je kladný práve vtedy, keď F je kladne definitná, a φ je nezáporný práve vtedy, keď F je kladne semidefinitná.

Formuláciu pojmov kladnej resp. nezápornej hermitovskej matice ako aj dôkaz nasledujúceho jednoduchého tvrdenia prenechávame ako cvičenie čitateľovi.

25.1.1. Tvrdenie. *Nech φ je hermitovský lineárny operátor na konečnorozmernom unitárnom priestore V.*

(a) φ je kladný práve vtedy, keď pre každé $\lambda \in \operatorname{Spec}\varphi$ platí $\lambda > 0$;

(b) φ je nezáporný práve vtedy, keď pre každé $\lambda \in \operatorname{Spec}\varphi$ platí $\lambda \geq 0$.

Ľahko nahliadneme, že pre ľubovoľný lineárny operátor $\varphi\colon V \to V$ sú oba operátory $\varphi^* \circ \varphi$, $\varphi \circ \varphi^*$ hermitovské. Navyše sú však aj nezáporné. Naozaj, pre každé $\boldsymbol{x} \in V$ platí

$$\big\langle \varphi^*(\varphi\boldsymbol{x}), \boldsymbol{x} \big\rangle = \langle \varphi\boldsymbol{x}, \varphi\boldsymbol{x} \rangle \geq 0,$$

a podobne pre operátor $\varphi \circ \varphi^*$.

Analogicky pre každú, dokonca nie nutne štvorcovú maticu $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ sú obe matice $\boldsymbol{A}^* \cdot \boldsymbol{A}$, $\boldsymbol{A} \cdot \boldsymbol{A}^*$ nezáporné hermitovské. Ich druhé odmocniny, t. j. matice $\sqrt{\boldsymbol{A}^* \cdot \boldsymbol{A}}$, $\sqrt{\boldsymbol{A} \cdot \boldsymbol{A}^*}$ sa nazývajú *pravý* resp. *ľavý modul matice* $\boldsymbol{A}$. Ako vyplýva z cvičenia 18.8, obe tieto matice majú rovnaké nenulové vlastné čísla. Nazývame ich *singulárne čísla matice* $\boldsymbol{A}$. Sú to vlastne (kladné) druhé odmocniny kladných vlastných čísel ktorejkoľvek z matíc $\boldsymbol{A}^* \cdot \boldsymbol{A}$, $\boldsymbol{A} \cdot \boldsymbol{A}^*$.

25.1.2. Veta. *Nech $\boldsymbol{A} \in \mathbb{C}^{m\times n}$. Potom existujú unitárne matice $\boldsymbol{U} \in \mathbb{C}^{m\times m}$ a $\boldsymbol{V} \in \mathbb{C}^{n\times n}$ také, že*

$$\boldsymbol{A} = \boldsymbol{U} \cdot \begin{pmatrix} \boldsymbol{D}_0 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \cdot \boldsymbol{V}^*,$$

kde $h = h(\boldsymbol{A}^* \cdot \boldsymbol{A})$ a $d_1, \dots, d_h$ sú singulárne čísla matice $\boldsymbol{A}$, pričom každé z nich sa v tomto zozname vyskytuje toľkokrát, aká je jeho násobnosť, $\boldsymbol{D}_0 = \operatorname{diag}(d_1, \dots, d_h)$ a $\boldsymbol{0}$-y sú nulové matice potrebných rozmerov. Ak $\boldsymbol{A}$ je reálna, tak matice $\boldsymbol{U}$, $\boldsymbol{V}$ možno zvoliť ortogonálne.

Dôkaz. Označme $\boldsymbol{B}$ uvedenú blokovú maticu. Pre určitosť predpokladajme, že $m \leq n$; druhý prípad možno zvládnuť analogicky. Matica $\boldsymbol{A}^* \cdot \boldsymbol{A} \in \mathbb{C}^{n\times n}$ je nezáporná hermitovská, preto existuje unitárna matica $\boldsymbol{P} \in \mathbb{C}^{n\times n}$ taká, že

$$\boldsymbol{A}^* \cdot \boldsymbol{A} = \boldsymbol{P} \cdot \boldsymbol{D}^2 \cdot \boldsymbol{P}^*.$$

kde $\boldsymbol{D} = \operatorname{diag}(d_1, \dots, d_n)$ je reálna diagonálna matica, pričom $d_1, \dots, d_h > 0$ sú singulárne čísla matice $\boldsymbol{A}$ a $d_j = 0$ pre $h < j \leq n$. Keďže pre každé $k \leq n$ je k-ty stĺpec matice $\boldsymbol{P}$ vlastný vektor matice $\boldsymbol{A}^* \cdot \boldsymbol{A}$ prislúchajúci k jej vlastnému číslu d_k^2, a $\boldsymbol{P}$ je unitárna, pre stĺpce matice $\boldsymbol{A} \cdot \boldsymbol{P}$ platí

$$\begin{aligned}\big\langle \boldsymbol{s}_j(\boldsymbol{A}\cdot\boldsymbol{P}), \boldsymbol{s}_k(\boldsymbol{A}\cdot\boldsymbol{P}) \big\rangle &= \big\langle \boldsymbol{A}\cdot\boldsymbol{s}_j(\boldsymbol{P}), \boldsymbol{A}\cdot\boldsymbol{s}_k(\boldsymbol{P}) \big\rangle = \big\langle \boldsymbol{s}_j(\boldsymbol{P}), \boldsymbol{A}^*\cdot\boldsymbol{A}\cdot\boldsymbol{s}_k(\boldsymbol{P}) \big\rangle \\ &= \big\langle \boldsymbol{s}_j(\boldsymbol{P}), d_k^2\boldsymbol{s}_k(\boldsymbol{P}) \big\rangle = d_k^2\big\langle \boldsymbol{s}_j(\boldsymbol{P}), \boldsymbol{s}_k(\boldsymbol{P}) \big\rangle = d_k^2\delta_{jk}\end{aligned}$$

Teda $\boldsymbol{s}_j(\boldsymbol{A}\cdot\boldsymbol{P}) = \boldsymbol{0}$ pre $j > h$, a všetky vektory tvaru $\boldsymbol{u}_k = d_k^{-1}\boldsymbol{s}_k(\boldsymbol{A}\cdot\boldsymbol{P})$, kde $k \leq h$, tvoria ortonormálny systém v $\mathbb{C}^m$. Možno ich teda doplniť do ortonormálnej bázy $\boldsymbol{u}_1, \dots, \boldsymbol{u}_m$ tohto priestoru. Potom $\boldsymbol{U} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ je unitárna matica, pre ktorej *všetky* stĺpce $\boldsymbol{u}_k$ platí $d_k\boldsymbol{u}_k = \boldsymbol{s}_k(\boldsymbol{A}\cdot\boldsymbol{P})$, t. j.

$$\boldsymbol{A}\cdot\boldsymbol{P} = (d_1\boldsymbol{u}_1, \dots, d_m\boldsymbol{u}_m, \boldsymbol{0}, \dots, \boldsymbol{0}) = \boldsymbol{U}\cdot\boldsymbol{B}.$$

Z toho vyplýva

$$A = U \cdot B \cdot P^{-1} = U \cdot B \cdot P^*.$$

Stačí teda položiť $\boldsymbol{V} = \boldsymbol{P}$.

Pre reálnu maticu $\boldsymbol{A}$ možno hneď na začiatku zvoliť ortogonálnu maticu $\boldsymbol{P} = \boldsymbol{V}$ a ortogonálnu maticu $\boldsymbol{U}$ získať doplnením reálnych vektorov $\boldsymbol{u}_k = d_k^{-1}\boldsymbol{s}_k(\boldsymbol{A} \cdot \boldsymbol{P})$, $k \leq h$, do ortonormálnej bázy v $\mathbb{R}^m$.

Práve dokázaný výsledok, nazývaný tiež veta o *singulárnom rozklade*, má názornú geometrickú interpretáciu. Lineárne zobrazenie $\boldsymbol{x} \mapsto \boldsymbol{A} \cdot \boldsymbol{x}$ zobrazuje jednotkovú sféru

$$S = \{\boldsymbol{x} \in \mathbb{C}^n;\ \|\boldsymbol{x}\| = 1\}$$

v unitárnom priestore $\mathbb{C}^n$ na (možno degenerovaný) elipsoid v $\mathbb{C}^m$. Singulárne čísla matice $\boldsymbol{A}$ sú dĺžkami jeho (nedegenerovaných) polosí. Dôkaz vety o singulárnom rozklade je vlastne návod, ako zostrojiť ortonormálnu bázu $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) = \boldsymbol{V}$ priestoru $\mathbb{C}^n$ a ortonormálnu bázu $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_m) = \boldsymbol{U}$ priestoru $\mathbb{C}^m$ tak, aby vektory $\boldsymbol{A} \cdot \boldsymbol{v}_1 = d_1\boldsymbol{u}_1, \dots, \boldsymbol{A} \cdot \boldsymbol{v}_h = d_h\boldsymbol{u}_h$ tvorili priamo nedegenerované polosi tohto elipsoidu a $\boldsymbol{A} \cdot \boldsymbol{v}_j = \boldsymbol{0}$ pre $j > h$.

Nasledujúca veta, nazývaná tiež vetou o *polárnom rozklade*, je maticovou analógiou goniometrického vyjadrenia komplexného čísla v goniometrickom tvare $z = ru = r\,\mathrm{e}^{\mathrm{i}\alpha}$, kde $0 \leq r \in \mathbb{R}$, $u \in \mathbb{C}$, $|u| = 1$, a $\alpha \in \mathbb{R}$.

25.1.3. Veta. *Nech $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ je ľubovoľná matica. Potom existujú nezáporné hermitovské matice $\boldsymbol{R}, \boldsymbol{R}' \in \mathbb{C}^{n\times n}$ a unitárne matice $\boldsymbol{U}, \boldsymbol{U}' \in \mathbb{C}^{n\times n}$ také, že*

$$\boldsymbol{A} = \boldsymbol{R} \cdot \boldsymbol{U} = \boldsymbol{U}' \cdot \boldsymbol{R}'.$$

Pritom matice $\boldsymbol{R}$, $\boldsymbol{R}'$ sú určené jednoznačne, a ak $\boldsymbol{A}$ je regulárna, tak i matice $\boldsymbol{U}, \boldsymbol{U}'$ sú určené jednoznačne. Navyše, ak $\boldsymbol{A}$ je reálna, tak aj matice $\boldsymbol{R}, \boldsymbol{U}$ resp. $\boldsymbol{R}'$, $\boldsymbol{U}'$ možno zvoliť reálne, t. j. $\boldsymbol{R}$, $\boldsymbol{R}'$ symetrické, kladne semidefinitné a $\boldsymbol{U}$, $\boldsymbol{U}'$ ortogonálne.

Dôkaz. Nech $\boldsymbol{A} = \boldsymbol{W} \cdot \boldsymbol{D} \cdot \boldsymbol{V}^*$ je singulárny rozklad matice $\boldsymbol{A}$. Keďže $\boldsymbol{D} \in \mathbb{C}^{n\times n}$ je nezáporná diagonálna (teda hermitovská) a $\boldsymbol{V}, \boldsymbol{W} \in \mathbb{C}^{n\times n}$ sú unitárne, tak aj $\boldsymbol{R} = \boldsymbol{W} \cdot \boldsymbol{D} \cdot \boldsymbol{W}^*$ je nezáporná hermitovská a $\boldsymbol{U} = \boldsymbol{W} \cdot \boldsymbol{V}^*$ je unitárna. Potom

$$\boldsymbol{A} = \boldsymbol{W} \cdot \boldsymbol{D} \cdot \boldsymbol{W}^* \cdot \boldsymbol{W} \cdot \boldsymbol{V}^* = \boldsymbol{R} \cdot \boldsymbol{U}.$$

Ak $\boldsymbol{A} = \boldsymbol{R} \cdot \boldsymbol{U}$ s nezápornou hermitovskou $\boldsymbol{R}$ a unitárnou $\boldsymbol{U}$, tak $\boldsymbol{A}^* = \boldsymbol{U}^* \cdot \boldsymbol{R}$ a $\boldsymbol{A} \cdot \boldsymbol{A}^* = \boldsymbol{R} \cdot \boldsymbol{U} \cdot \boldsymbol{U}^* \cdot \boldsymbol{R} = \boldsymbol{R}^2$, teda $\boldsymbol{R}$ je určená jednoznačne ako druhá odmocnina nezápornej hermitovskej matice $\boldsymbol{A} \cdot \boldsymbol{A}^*$. Navyše, ak $\boldsymbol{A}$ je regulárna, tak aj $\boldsymbol{A}^*$, $\boldsymbol{A} \cdot \boldsymbol{A}^*$ a $\boldsymbol{R} = \sqrt{\boldsymbol{A} \cdot \boldsymbol{A}^*}$ sú regulárne. Preto podmienkou $\boldsymbol{A} = \boldsymbol{R} \cdot \boldsymbol{U}$ je matica $\boldsymbol{U} = \boldsymbol{R}^{-1} \cdot \boldsymbol{A}$ určená jednoznačne.

Druhý polárny rozklad $\boldsymbol{A} = \boldsymbol{U}' \cdot \boldsymbol{R}'$ možno získať z prvého polárneho rozkladu $\boldsymbol{A}^* = \boldsymbol{Q} \cdot \boldsymbol{V}$ matice $\boldsymbol{A}^*$, s nezápornou hermitovskou $\boldsymbol{Q}$ a unitárnou $\boldsymbol{V}$. Potom $\boldsymbol{A} = \boldsymbol{V}^* \cdot \boldsymbol{Q}$; stačí teda položiť $\boldsymbol{R}' = \boldsymbol{Q}$, $\boldsymbol{U}' = \boldsymbol{V}^*$.

Podmienka reálnosti matíc $\boldsymbol{R}$, $\boldsymbol{U}$ resp. $\boldsymbol{R}'$, $\boldsymbol{U}'$ pre $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ vyplýva z analogickej podmienky pre singulárny rozklad.

Spojením práve dokázanej vety a vety 23.4.9 dostávame

25.1.4. Dôsledok. *Pre ľubovoľnú maticu $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ existujú hermitovské matice $\boldsymbol{H}, \boldsymbol{H}' \in \mathbb{C}^{n \times n}$ a jednoznačne určené nezáporné hermitovské matice $\boldsymbol{R}, \boldsymbol{R} \in \mathbb{C}^{n \times n}$ také, že*

$$\boldsymbol{A} = \boldsymbol{R} \cdot \exp(\mathrm{i}\boldsymbol{H}) = \exp(\mathrm{i}\boldsymbol{H}') \cdot \boldsymbol{R}'.$$

25.2 Pseudoinverzné lineárne zobrazenia a matice

Týmto paragrafom nadväzujeme na tému, ktorú sme začali rozvíjať v paragrafe 14.5, venovanom „riešeniu neriešiteľných sústav."

Nech X, Y sú ľubovoľné množiny. Hovoríme, že $g\colon Y \to X$ je *pseudoinverzné zobrazenie* k zobrazeniu $f\colon X \to Y$, ak platí

$$f \circ g \circ f = f, \qquad g \circ f \circ g = g.$$

Zrejme g je pseudoinverzné zobrazenie k f práve vtedy, keď f je pseudoinverzné zobrazenie ku g. Takisto ľahko nahliadneme, že ak $f\colon X \to Y$ je bijekcia, t. j. ak k nemu existuje *inverzné* zobrazenie $f^{-1}\colon Y \to X$, tak f^{-1} je zároveň jediným pseudoinverzným zobrazením k f. Ak však f nie je bijektívne a $\# X \geq 2$, tak k nemu existuje viacero pseudoinverzných zobrazení g. (Presvedčte sa o tom.)

Nás budú v tejto chvíli zaujímať najmä *pseudoinverzné lineárne zobrazenia* k lineárnym zobrazeniam medzi konečnorozmernými unitárnymi prípadne euklidovskými priestormi. Začneme však trochu všeobecnejším prípadom lineárnych zobrazení medzi konečnorozmernými vektorovými priestormi nad ľubovoľným poľom K. Z výpočtových dôvodov zavádzame aj pojem pseudoinverznej matice. $\boldsymbol{B} \in K^{n \times m}$ sa nazýva *pseudoinverzná matica* k matici $\boldsymbol{A} \in K^{m \times n}$, ak platí

$$\boldsymbol{A} \cdot \boldsymbol{B} \cdot \boldsymbol{A} = \boldsymbol{A}, \qquad \boldsymbol{B} \cdot \boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{B}.$$

Zrejme k regulárnej matici $\boldsymbol{A} \in K^{n \times n}$ existuje jediná pseudoinverzná matica, a to $\boldsymbol{A}^{-1}$. Ako čoskoro uvidíme, k obdĺžnikovým maticiam, ako i k singulárnym štvorcovým maticiam môže existovať viacero rôznych pseudoinverzných matíc.

Nech $\varphi\colon V \to U$ je lineárne zobrazenie medzi konečnorozmernými vektorovými priestormi nad poľom K. Označme $m = \dim U$, $n = \dim V$ a $h = h(\varphi)$. Podľa vety 7.5.4 existuje báza $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ priestoru V a báza $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$ priestoru U, vzhľadom na ktoré má φ maticu v blokovom tvare

$$\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \begin{pmatrix} \boldsymbol{I}_h & \mathbf{0}_{h,n-h} \\ \mathbf{0}_{m-h,h} & \mathbf{0}_{m-h,n-h} \end{pmatrix}.$$

Z tohto tvaru je zrejmé, že vektory $\boldsymbol{v}_{h+1}, \ldots, \boldsymbol{v}_n$ tvoria bázu lineárneho podpriestoru $\operatorname{Ker}\varphi \subseteq V$ a vektory $\boldsymbol{u}_1 = \varphi(\boldsymbol{v}_1), \ldots, \boldsymbol{u}_h = \varphi(\boldsymbol{v}_h)$ bázu lineárneho podpriestoru $\operatorname{Im}\varphi \subseteq U$. Navyše platí

$$V = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_h] \oplus \operatorname{Ker}\varphi, \qquad U = \operatorname{Im}\varphi \oplus [\boldsymbol{u}_{h+1}, \ldots, \boldsymbol{u}_m]$$

a φ zúžené na podpriestor $[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_h]$ je lineárny izomorfizmus tohto podpriestoru na podpriestor $\operatorname{Im}\varphi$. Nech teraz $\psi\colon U \to V$ je lineárne zobrazenie, ktoré má vzhľadom na bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ maticu

$$\boldsymbol{A}^{\mathsf{T}} = (\psi)_{\boldsymbol{\beta},\boldsymbol{\alpha}} = \begin{pmatrix} \boldsymbol{I}_h & \mathbf{0}_{h,m-h} \\ \mathbf{0}_{n-h,h} & \mathbf{0}_{n-h,m-h} \end{pmatrix}.$$

Potom $\operatorname{Ker}\psi = [\boldsymbol{u}_{h+1}, \ldots, \boldsymbol{u}_m]$ a $\psi \restriction \operatorname{Im}\varphi\colon \operatorname{Im}\varphi \to [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_h]$ je lineárny izomorfizmus, inverzný k zúženiu φ na $[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_h]$. Prenechávame čitateľovi, aby sa sám presvedčil, že lineárne zobrazenie ψ je pseudoinverzné k φ.

Taktiež naopak, ak $\psi\colon U \to V$ je pseudoinverzné k $\varphi\colon V \to U$, tak existujú bázy $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_h, \boldsymbol{v}_{h+1}, \ldots, \boldsymbol{v}_n)$ priestoru V a $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_h, \boldsymbol{u}_{h+1}, \ldots, \boldsymbol{u}_m)$ priestoru U také, že matice $(\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$, $(\psi)_{\boldsymbol{\beta},\boldsymbol{\alpha}}$ majú uvedený tvar.

Vo všeobecnosti pre dané lineárne zobrazenie $\varphi\colon V \to U$ medzi konečnorozmernými vektorovými priestormi ľubovoľná dvojica lineárnych podpriestorov $S \subseteq V$, $T \subseteq U$ takých, že

$$S \oplus \operatorname{Ker}\varphi = V, \qquad \operatorname{Im}\varphi \oplus T = U,$$

jednoznačne určuje lineárne zobrazenie $\psi\colon U \to V$ pseudoinverzné k φ, pre ktoré $\operatorname{Ker}\psi = T$ a $\operatorname{Im}\psi = S$. Stačí si uvedomiť, že $\varphi \restriction S\colon S \to \operatorname{Im}\varphi$ je lineárny izomorfizmus, takže k nemu existuje inverzné lineárne zobrazenie $(\varphi \restriction S)^{-1}\colon \operatorname{Im}\varphi \to S$. Potom jediná možnosť ako definovať ψ je

$$\psi(\boldsymbol{y}) = (\varphi \restriction S)^{-1}(\boldsymbol{y}_1),$$

kde $\boldsymbol{y} = \boldsymbol{y}_1 + \boldsymbol{y}_2$ je jednoznačný rozklad vektora $\boldsymbol{y} \in U$ na zložky $\boldsymbol{y}_1 \in \operatorname{Im}\varphi$, $\boldsymbol{y}_2 \in T$. Opäť možno jednoducho overiť, že lineárne zobrazenie ψ je pseudoinverzné k φ.

Z našej konštrukcie je taktiež zrejmé, že pokiaľ φ nie je bijektívne, tak aspoň jeden z podpriestorov S, T je netriviálny a možno ho voliť viacerými spôsobmi. Rôzne voľby S a T potom určujú rôzne pseudoinverzné lineárne zobrazenia k φ. V prípade konečnorozmerných vektorových priestorov nad všeobecným poľom nemáme k dispozícii prirodzený prostriedok, ako túto nejednoznačnosť odstrániť. K danému netriválnemu vlastnému lineárnemu podpriestoru $P \subseteq V$ totiž existuje mnoho lineárnych podpriestorov $P' \subseteq V$ takých, že $V = P \oplus P'$, a všetky možné voľby P' sú navzájom rovnocenné. Iná je však situácia v konečnorozmerných unitárnych resp. euklidovských priestoroch – tam je prirodzené položiť $P' = P^\perp$.

25.2.1. Tvrdenie. *Nech U, V sú konečnorozmerné unitárne, prípadne euklidovské priestory a $\varphi\colon V \to U$, $\psi\colon U \to V$ sú navzájom pseudoinverzné lineárne zobrazenia. Potom nasledujúce podmienky sú ekvivalentné:*

(i) lineárne podpriestory $\operatorname{Ker}\varphi$, $\operatorname{Im}\psi \subseteq V$ *ako aj* $\operatorname{Ker}\psi$, $\operatorname{Im}\varphi \subseteq U$ *sú navzájom ortogonálne;*

(ii) lineárne operátory $\psi \circ \varphi\colon V \to V$ a $\varphi \circ \psi\colon U \to U$ sú samoadjungované.

Dôkaz. Označme $S = \operatorname{Im}\psi$. Keďže φ, ψ sú navzájom pseudoinverzné, $V = S \oplus \operatorname{Ker}\varphi$ a $\psi\varphi(\boldsymbol{u}) = \boldsymbol{u}$ pre $\boldsymbol{u} \in S$.

(i) ⇒ (ii) Zvoľme $\boldsymbol{x}, \boldsymbol{y} \in V$. Potom $\boldsymbol{x}_S, \boldsymbol{y}_S \in S$, teda $\boldsymbol{x} - \boldsymbol{x}_S, \boldsymbol{y} - \boldsymbol{y}_S \in \operatorname{Ker}\varphi = S^\perp$, a platí $\varphi(\boldsymbol{x}) = \varphi(\boldsymbol{x}_S)$, $\langle \boldsymbol{x}_S, \boldsymbol{y} - \boldsymbol{y}_S\rangle = 0$. V dôsledku toho

$$\langle \psi\varphi(\boldsymbol{x}), \boldsymbol{y}\rangle = \langle \psi\varphi(\boldsymbol{x}_S), \boldsymbol{y}\rangle = \langle \boldsymbol{x}_S, \boldsymbol{y}\rangle = \langle \boldsymbol{x}_S, \boldsymbol{y}_S\rangle.$$

Z rovnakých dôvodov je $\langle \boldsymbol{x}, \psi\varphi(\boldsymbol{y})\rangle = \langle \boldsymbol{x}_S, \boldsymbol{y}_S\rangle$, teda $\langle \psi\varphi(\boldsymbol{x}), \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \psi\varphi(\boldsymbol{y})\rangle$, čiže operátor $\psi \circ \varphi$ je samoadjungovaný. Podobne možno dokázať samodjungovanosť operátora $\varphi \circ \psi$.

(ii) ⇒ (i) Ukážeme, že $\boldsymbol{x} \perp \boldsymbol{y}$ pre ľubovoľné $\boldsymbol{x} \in \operatorname{Im}\psi$, $\boldsymbol{y} \in \operatorname{Ker}\varphi$. Potom $\boldsymbol{x} = \psi(\boldsymbol{u})$ pre nejaké $\boldsymbol{u} \in U$ a $\varphi(\boldsymbol{y}) = \boldsymbol{0}$. Keďže $\psi \circ \varphi \circ \psi = \psi$ a $\psi \circ \varphi$ je samoadjungovaný,

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \langle \psi(\boldsymbol{u}), \boldsymbol{y}\rangle = \langle \psi\varphi\psi(\boldsymbol{u}), \boldsymbol{y}\rangle = \langle \psi(\boldsymbol{u}), \psi\varphi(\boldsymbol{y})\rangle = 0.$$

Podobne sa dokáže ortogonálnosť podpriestorov $\operatorname{Im}\varphi$ a $\operatorname{Ker}\psi$.

Nech $\varphi\colon V \to U$ je lineárne zobrazenie medzi konečnorozmernými unitárnymi prípadne euklidovskými priestormi. Pseudoinverzné lineárne zobrazenie $\psi\colon U \to V$ k zobrazeniu φ nazývame jeho *Mooreovou-Penroseovou pseudoinverziou*, ak je splnená jedna (teda nevyhnutne obe) z ekvivalentných podmienok (i), (ii) tvrdenia 25.2.1. Z našich predchádzajúcich úvah vyplýva, že Mooreova-Penroseova pseudoinverzia (na rozdiel od inverzného zobrazenia) existuje vždy a uvedenou dodatočnou podmienkou je určená jednoznačne – značíme ju $\psi = \varphi^\neg$.

Podobne, maticu $\boldsymbol{B} \in \mathbb{C}^{n\times m}$ nazývame *Mooreovou-Penroseovou pseudoinverziou* matice $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, ak $\boldsymbol{B}$ je pseudoinverzná k $\boldsymbol{A}$ a obe matice $\boldsymbol{A}\cdot\boldsymbol{B}$,

$\boldsymbol{B} \cdot \boldsymbol{A}$ sú hermitovské (v reálnom prípade symetrické). Takisto každá matica $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ má jednoznačne určenú Mooreovu-Penroseovu pseudoinverziu, ktorú opäť značíme $\boldsymbol{A}^{\neg}$.

Mooreovu-Penroseovu pseudoinverziu matice $\boldsymbol{A}$ možno tiež opísať s využitím singulárneho rozkladu.

25.2.2. Tvrdenie. *Nech $\boldsymbol{A} \in \mathbb{C}^{m\times n}$. Potom jej Mooreova-Penroseova pseudoinverzia má tvar*

$$\boldsymbol{A}^{\neg} = \boldsymbol{V} \cdot \operatorname{diag}\bigl(d_1^{-1}, \dots, d_h^{-1}, 0, \dots, 0\bigr) \cdot \boldsymbol{U}^*,$$

kde

$$\boldsymbol{A} = \boldsymbol{U} \cdot \operatorname{diag}(d_1, \dots, d_h, 0, \dots, 0) \cdot \boldsymbol{V}^*$$

je singulárny rozklad matice $\boldsymbol{A}$. Ak navyše $\boldsymbol{A}$ je reálna, tak aj $\boldsymbol{A}^{\neg}$ je reálna.

Dôkaz. Zrejme pre maticu $\boldsymbol{D} = \operatorname{diag}(d_1, \dots, d_h, 0, \dots, 0)$ platí

$$\boldsymbol{D}^{\neg} = \operatorname{diag}\bigl(d_1^{-1}, \dots, d_h^{-1}, 0, \dots, 0\bigr).$$

Označme $\boldsymbol{B} = \boldsymbol{V} \cdot \boldsymbol{D}^{\neg} \cdot \boldsymbol{U}^*$. Keďže $\boldsymbol{U}$, $\boldsymbol{V}$ sú unitárne,

$$\boldsymbol{A} \cdot \boldsymbol{B} \cdot \boldsymbol{A} = \boldsymbol{U} \cdot \boldsymbol{D} \cdot \boldsymbol{V}^* \cdot \boldsymbol{V} \cdot \boldsymbol{D}^{\neg} \cdot \boldsymbol{U}^* \cdot \boldsymbol{U} \cdot \boldsymbol{D} \cdot \boldsymbol{V}^* = \boldsymbol{U} \cdot \boldsymbol{D} \cdot \boldsymbol{V}^* = \boldsymbol{A}.$$

Podobne možno tiež dokázať $\boldsymbol{B} \cdot \boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{B}$. Rovnosti $(\boldsymbol{A} \cdot \boldsymbol{B})^* = \boldsymbol{A} \cdot \boldsymbol{B}$, $(\boldsymbol{B} \cdot \boldsymbol{A})^* = \boldsymbol{B} \cdot \boldsymbol{A}$ sú triviálne. Z jednoznačnosti Mooreovej-Penroseovej inverzie vyplýva $\boldsymbol{A}^{\neg} = \boldsymbol{B}$. Z vety 25.1.2 navyše vyplýva, že $\boldsymbol{A}^{\neg}$ je reálna pre reálnu $\boldsymbol{A}$.

Najdôležitejšie vlastnosti Mooreovej-Penroseovej pseudoinverzie sú zhrnuté v nasledujúcich dvoch tvrdeniach. Dôkaz prvého z nich sa redukuje na niekoľko priamych výpočtov, ktorých vzory možno nájsť v paragrafe 14.5.

25.2.3. Tvrdenie. *Nech $\boldsymbol{A} \in \mathbb{C}^{m\times n}$. Potom*
(a) ak $\boldsymbol{A} \cdot \boldsymbol{A}^*$ je regulárna, tak $\boldsymbol{A}^{\neg} = \boldsymbol{A}^* \cdot (\boldsymbol{A} \cdot \boldsymbol{A}^*)^{-1}$;
(b) ak $\boldsymbol{A}^* \cdot \boldsymbol{A}$ je regulárna, tak $\boldsymbol{A}^{\neg} = (\boldsymbol{A}^* \cdot \boldsymbol{A})^{-1} \cdot \boldsymbol{A}^*$.

Pre reálnu maticu $\boldsymbol{A}$ platí $\boldsymbol{A}^* = \boldsymbol{A}^{\mathsf{T}}$, teda z regularity matice $\boldsymbol{A} \cdot \boldsymbol{A}^{\mathsf{T}}$ resp. $\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A}$ vyplýva $\boldsymbol{A}^{\neg} = \boldsymbol{A}^{\mathsf{T}} \cdot (\boldsymbol{A} \cdot \boldsymbol{A}^{\mathsf{T}})^{-1}$ resp. $\boldsymbol{A}^{\neg} = (\boldsymbol{A}^* \cdot \boldsymbol{A})^{-1} \cdot \boldsymbol{A}^*$, čo sa zhoduje s výsledkami paragrafu 14.5.

Skôr než čitateľ prejde k druhému tvrdeniu, mal by si najprv samostatne premyslieť, že pojmy pseudoriešenia a minimálneho (pseudo)riešenia sústavy lineárnych rovníc z paragrafu 14.5 možno celkom mechanicky preniesť z euklidovského priestoru $\mathbb{R}^n$ do unitárneho priestoru $\mathbb{C}^n$. Taktiež výsledky o nich zostávajú v platnosti s jedinou zmenou: transponovanú maticu $\boldsymbol{A}^{\mathsf{T}}$ treba všade nahradiť hermitovsky združenou maticou $\boldsymbol{A}^*$.

25.2.4. Tvrdenie. *Nech* $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, $\boldsymbol{b} \in \mathbb{C}^m$ a $\boldsymbol{x}_0 = \boldsymbol{A}^{\neg} \cdot \boldsymbol{b} \in \mathbb{C}^n$. *Potom* $\boldsymbol{x}_0$ *je minimálne pseudoriešenie sústavy* $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$.

Dôkaz. Podľa tvrdení 14.5.1 a 14.5.3, modifikovaných spomínaným očividným spôsobom, je $\boldsymbol{x} \in \mathbb{C}^n$ minimálnym pseudoriešním sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ práve vtedy, keď $\boldsymbol{x} = \boldsymbol{A}^* \cdot \boldsymbol{A} \cdot \boldsymbol{z}$ pre nejaké $\boldsymbol{z} \in \mathbb{C}^n$ také, že $(\boldsymbol{A}^* \cdot \boldsymbol{A})^2 \cdot \boldsymbol{z} = \boldsymbol{A}^* \cdot \boldsymbol{b}$. Pre $\boldsymbol{x}_0 = \boldsymbol{A}^{\neg} \cdot \boldsymbol{b}$ stačí položiť

$$\boldsymbol{z} = \boldsymbol{A}^{\neg} \cdot \boldsymbol{A}^{\neg *} \cdot \boldsymbol{A}^{\neg} \cdot \boldsymbol{b}.$$

Skôr než pristúpime k overeniu požadovaných rovností, uvedomme si, že

$$\boldsymbol{A}^* \cdot \boldsymbol{A} \cdot \boldsymbol{A}^{\neg} = \boldsymbol{A}^* \cdot (\boldsymbol{A} \cdot \boldsymbol{A}^{\neg})^* = (\boldsymbol{A} \cdot \boldsymbol{A}^{\neg} \cdot \boldsymbol{A})^* = \boldsymbol{A}^*,$$

a z rovnakého dôvodu $\boldsymbol{A}^* \cdot \boldsymbol{A}^{\neg *} \cdot \boldsymbol{A}^{\neg} = \boldsymbol{A}^{\neg}$. Preto

$$\boldsymbol{A}^* \cdot \boldsymbol{A} \cdot \boldsymbol{z} = \boldsymbol{A}^* \cdot \boldsymbol{A} \cdot \boldsymbol{A}^{\neg} \cdot \boldsymbol{A}^{\neg *} \cdot \boldsymbol{A}^{\neg} \cdot \boldsymbol{b} = \boldsymbol{A}^* \cdot \boldsymbol{A}^{\neg *} \cdot \boldsymbol{A}^{\neg} \cdot \boldsymbol{b} = \boldsymbol{A}^{\neg} \cdot \boldsymbol{b} = \boldsymbol{x}_0,$$

z čoho ďalej vyplýva

$$(\boldsymbol{A}^* \cdot \boldsymbol{A})^2 \cdot \boldsymbol{z} = \boldsymbol{A}^* \cdot \boldsymbol{A} \cdot \boldsymbol{A}^{\neg} \cdot \boldsymbol{b} = \boldsymbol{A}^* \cdot \boldsymbol{b}.$$

Či už sústava $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ má alebo nemá riešenie, $\boldsymbol{x}_0 = \boldsymbol{A}^{\neg} \cdot \boldsymbol{b}$ je jej *pseudoriešením*, t. j. riešením sústavy $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}_S$, kde $\boldsymbol{b}_S$ je kolmý priemet vektora $\boldsymbol{b}$ do lineárneho podpriestoru $S \subseteq \mathbb{C}^m$ generovaného stĺpcami matice $\boldsymbol{A}$. To znamená, že pre všetky $\boldsymbol{y} \in \mathbb{C}^n$ platí

$$\|\boldsymbol{A} \cdot \boldsymbol{x}_0 - \boldsymbol{b}\| \leq \|\boldsymbol{A} \cdot \boldsymbol{y} - \boldsymbol{b}\|.$$

Navyše toto pseudoriešenie je minimálne, t. j. pre každé iné pseudoriešenie $\boldsymbol{x} \in \mathbb{C}^n$ platí

$$\|\boldsymbol{x}_0\| \leq \|\boldsymbol{x}\|.$$

Na druhej strane, ak sústava $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ má viac riešení, tak $\boldsymbol{x}_0 = \boldsymbol{A}^{\neg} \cdot \boldsymbol{b}$ je jej riešením s minimálnou normou. Mooreova-Penroseova pseudoinverzia matice $\boldsymbol{A}$ tak jednotným spôsobom rieši obe otázky diskutované v paragrafe 14.5 a zároveň ich zovšeobecňuje z reálnych matíc na komplexné.

25.3 Odchýlka dvoch lineárnych podpriestorov

Singulárne čísla matíc možno okrem iného využiť na určenie odchýlky afinných či lineárnych podpriestorov v euklidovskom priestore. Pripomeňme, že odchýlku dvoch netriviálnych afinných podpriestorov v euklidovskom priestore V sme v paragrafe 14.3 definovali ako odchýlku ich zameraní. Odchýlku netriviálnych lineárnych podpriestorov $S, T \subseteq V$ sme definovali ako odchýlku ich podpriestorov $S_1 = S \cap (S \cap T)^{\perp}$, $T_1 = T \cap (S \cap T)^{\perp}$, pre ktoré platí $S_1 \cap T_1 = \{\mathbf{0}\}$. Avšak odchýlku netriviálnych lineárnych podpriestorov S, T spĺňajúcich $S \cap T = \{\mathbf{0}\}$ sme vedeli určiť, len ak aspoň jeden z nich bol priamka alebo nadrovina. Túto medzeru zapĺňa nasledujúca veta.

25.3.1. Veta. *Nech S, T sú netriviálne lineárne podpriestory v euklidovskom priestore V, $S \cap T = \{\mathbf{0}\}$ a $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$, resp. $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l)$ sú ich ortonormálne bázy. Potom*

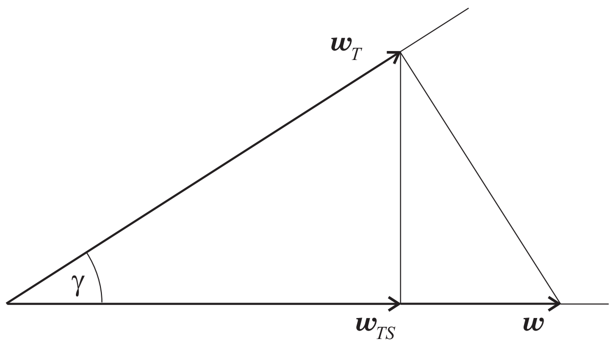
$$\cos \sphericalangle(S, T) = D = \sqrt{M},$$

kde D je najväčšie singulárne číslo matice

$$\boldsymbol{A} = \begin{pmatrix} \langle \boldsymbol{u}_1, \boldsymbol{v}_1 \rangle & \langle \boldsymbol{u}_2, \boldsymbol{v}_1 \rangle & \ldots & \langle \boldsymbol{u}_k, \boldsymbol{v}_1 \rangle \\ \langle \boldsymbol{u}_1, \boldsymbol{v}_2 \rangle & \langle \boldsymbol{u}_2, \boldsymbol{v}_2 \rangle & \ldots & \langle \boldsymbol{u}_k, \boldsymbol{v}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \boldsymbol{u}_1, \boldsymbol{v}_l \rangle & \langle \boldsymbol{u}_2, \boldsymbol{v}_l \rangle & \ldots & \langle \boldsymbol{u}_k, \boldsymbol{v}_l \rangle \end{pmatrix},$$

resp. M je najväčšie vlastné číslo matice $\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A}$.

Dôkaz. Označme $\varphi\colon S \to T$ zúženie ortogonálnej projekcie pr_T na podpriestor S a $\psi\colon T \to S$ zúženie ortogonálnej projekcie pr_S na podpriestor T. Potom lineárne zobrazenia φ, ψ majú vzhľadom na bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ matice $(\varphi)_{\boldsymbol{\beta},\boldsymbol{\alpha}} = \boldsymbol{A}$, $(\psi)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \boldsymbol{A}^{\mathsf{T}}$. Lineárny operátor $\psi \circ \varphi\colon S \to S$ má v báze $\boldsymbol{\alpha}$ maticu $\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A}$, teda je nezáporný hermitovský a všetky jeho vlastné čísla (t. j. vlastné čísla matice $\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A}$) sú reálne nezáporné.



Obr. 25.1. K odchýlke dvoch lineárnych podpriestorov

Podľa tvrdenia 14.3.1 a vlastností kolmého priemetu je

$$\cos \sphericalangle(S, T) = \sup\{\cos \sphericalangle(\boldsymbol{x}, T);\ \mathbf{0} \neq \boldsymbol{x} \in S\} = \sup\left\{\frac{\|\boldsymbol{x}_T\|}{\|\boldsymbol{x}\|};\ \mathbf{0} \neq \boldsymbol{x} \in S\right\}.$$

Nech $\boldsymbol{w} \in S$ je vlastný vektor operátora $\psi \circ \varphi$ prislúchajúci k jeho vlastnému číslu λ, t. j. $\boldsymbol{w}_{TS} = \psi\varphi(\boldsymbol{w}) = \lambda \boldsymbol{w}$. Označme $\gamma = \sphericalangle(\boldsymbol{w}, T)$. Potom

$$\cos\gamma = \frac{\|\boldsymbol{w}_T\|}{\|\boldsymbol{w}\|} = \frac{\|\varphi \boldsymbol{w}\|}{\|\boldsymbol{w}\|}$$

Z obrázku 25.1 je zrejmé, že taktiež

$$\|\boldsymbol{w}_{TS}\| = \|\psi\varphi(\boldsymbol{w})\| = \|\boldsymbol{w}_T\| \cos\gamma = \|\boldsymbol{w}\| \cos^2\gamma.$$

V dôsledku toho $\lambda \|\boldsymbol{w}\| = \|\boldsymbol{w}\| \cos^2\gamma$, čiže

$$\cos\gamma = \sqrt{\lambda}.$$

Zostáva dokázať, že ak $\lambda = M$ je *najväčšie* vlastné číslo matice $\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A}$, tak pre ľubovoľný nenulový vektor $\boldsymbol{x} \in S$ platí

$$\frac{\|\boldsymbol{x}_T\|}{\|\boldsymbol{x}\|} \le \sqrt{M} = D.$$

Keďže $\varphi(\boldsymbol{x}) = \boldsymbol{x}_T \perp \boldsymbol{x} - \boldsymbol{x}_T$, $\langle \varphi\boldsymbol{x}, \varphi\boldsymbol{x}\rangle = \langle \varphi\boldsymbol{x}, \boldsymbol{x}\rangle$. Ďalej z podmienky $\boldsymbol{x} \in S$ vyplýva $\psi(\boldsymbol{x}) = \boldsymbol{x}$. S využitím týchto faktov, Cauchyho-Schwartzovej nerovnosti a dôsledku 23.7.5 dostávame

$$\begin{aligned} \|\boldsymbol{x}_T\|^2 &= \langle \varphi\boldsymbol{x}, \varphi\boldsymbol{x}\rangle = \langle \varphi\boldsymbol{x}, \boldsymbol{x}\rangle = \langle \varphi\psi\boldsymbol{x}, \boldsymbol{x}\rangle \\ &\le \|\varphi\psi\boldsymbol{x}\| \, \|\boldsymbol{x}\| \le M \|\boldsymbol{x}\|^2 = D^2 \|\boldsymbol{x}\|^2, \end{aligned}$$

teda $\|\boldsymbol{x}_T\| \le D \|\boldsymbol{x}\|$.

25.4 Harmonický oscilátor

Harmonickým alebo aj *lineárnym harmonickým oscilátorom* budeme nazývať systém charakterizovaný jedinou číselnou veličinou $x = x(t)$, ktorej vývoj v čase t sa riadi diferenciálnou rovnicou

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -\omega^2 x,$$

kde $\omega \in \mathbb{R}$ je konštanta, nazývaná *kruhová frekvencia.*

Pomocou rovnice harmonického oscilátora možno – či už presne alebo približne – popísať mnoho rozmanitých jednoduchých fyzikálnych systémov, najmä v blízkosti rovnovážneho stavu. Tie zložitejšie možno zasa často „vyskladať“ z viacerých rôzne spriahnutých harmonických oscilátorov (a tie matematickou transformáciou analogickou diagonalizácii kvadratickej formy nahradiť súborom nespriahnutých oscilátorov). [1]

[1] V realistickejších matematických modeloch však treba brať do úvahy aj tlmenie kmitov napr. v dôsledku trenia, prípadne kmitanie vynútené premennou silou.

Okamžite vidíme, že ide o lineárnu diferenciálnu rovnicu druhého rádu s konštantnými koeficientmi (pozri paragrafy 22.5–6). Ak položíme $\boldsymbol{x} = (x_1, x_2)^{\mathsf{T}}$, kde $x_1 = \mathrm{d}x/\,\mathrm{d}t$, $x_2 = x$, môžeme ju zapísať v tvare homogénnej autonómnej sústavy

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \begin{pmatrix} 0 & -\omega^2 \\ 1 & 0 \end{pmatrix} \cdot \boldsymbol{x}.$$

Vlastnými hodnotami jej matice sú rýdzo imaginárne čísla $\lambda_{1,2} = \pm\mathrm{i}\omega$. Podmienky $-x_1 \pm \mathrm{i}\omega x_2 = 0$ pre zložky vlastných vektorov sú vlastne diferenciálne rovnice

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \pm\mathrm{i}\omega x.$$

Fundamentálny systém riešení pôvodnej rovnice tak tvoria funkcie $\mathrm{e}^{\pm\mathrm{i}\omega t}$. Ak nám ide len o reálne riešenia, môžeme ho nahradiť fundamentálnym systémom $(\cos\omega t, \sin\omega t)$. Všeobecné reálne riešenie má preto tvar

$$x(t) = C_1 \cos\omega t + C_2 \sin\omega t,$$

pričom koeficienty $C_1, C_2 \in \mathbb{R}$ možno určiť z prípadných počiatočných podmienok $C_1 = x(0)$, $C_2 = \frac{\mathrm{d}x}{\mathrm{d}t}(0)$. Alternatívne možno všeobecné riešenie zapísať v ľubovoľnom z tvarov

$$x(t) = A\cos(\omega t + \varphi) = A\sin(\omega t + \psi).$$

A potom predstavuje *amplitúdu* a φ, $\psi = \varphi + \pi/2$ *fázové posuny* kmitov harmonického oscilátora.

Ako príklad harmonického oscilátora môžu slúžiť kmity závažia zaveseného na pružine, uvedeného do pohybu vychýlením z rovnovážnej polohy. Pokiaľ je počiatočná výchylka malá a tlmenie v dôsledku trenia a odporu prostredia zanedbateľné, výchylka x z rovnovážnej polohy sa mení v čase podľa diferenciálnej rovnice

$$m\,\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -kx,$$

kde m je hmotnosť závažia a konštanta k, *tuhosť pružiny*, charakterizuje jej odpor proti pozdĺžnej deformácii. Rovnicu možno odvodiť z druhého Newtonovho pohybového zákona porovnaním dvoch výrazov pre zrýchlenie $a = \mathrm{d}^2x/\,\mathrm{d}t^2 = F/m$, kde sila $F = -kx$, ktorou sa pružina „bráni“ proti deformácii, je priamo úmerná veľkosti výchylky x. Ak položíme $\omega = \sqrt{k/m}$, dostaneme rovnicu v pôvodnom tvare.

Podobne možno rovnicou harmonického oscilátora pomerne verne opísať kmity ideálneho matematického kyvadla pri malých výchylkách z rovnovážnej polohy. Rovnomerný pohyb po kružnici s polomerom R a uhlovou rýchlosťou

ω možno zasa vyjadriť ako superpozíciu kmitov dvoch harmonických oscilátorov

$$x(t) = R\cos\omega t, \qquad y(t) = R\sin\omega t$$

v dvoch navzájom kolmých smeroch. V tomto prípade je však výhodnejšie položiť $z(t) = x(t) + \mathrm{i}y(t)$ a obe rovnice spojiť do jedinej rovnice v komplexnom tvare

$$z(t) = R\,\mathrm{e}^{\mathrm{i}\omega t}.$$

Príkladom *tlmeného lineárneho oscilátora* je elektrický obvod s kapacitou C, indukciou L a odporom (rezistenciou) R: prúd I (no taktiež náboj a napätie) v takomto obvode sa mení v čase podľa rovnice

$$\frac{\mathrm{d}^2 I}{\mathrm{d}t^2} + \frac{R}{L}\frac{\mathrm{d}I}{\mathrm{d}t} + \frac{I}{LC} = 0$$

Harmonický oscilátor dostaneme v ideálnom prípade obvodu s nulovou rezistenciou $R = 0$. Kruhová frekvencia vtedy je $\omega = 1/\sqrt{LC}$.

Vráťme sa však k harmonickému oscilátoru reprezentovanému kmitmi závažia na pružine. Jeho *hamiltonián* dostaneme ako súčet kinetickej a potenciálnej energie kmitavého pohybu

$$H(x,p) = \frac{1}{2}\,mv^2 + \int_0^x k\xi\,\mathrm{d}\xi = \frac{p^2}{2m} + \frac{1}{2}\,kx^2 = \frac{1}{2}\left(\frac{p^2}{m} + m\omega^2 x^2\right),$$

kde $v = \mathrm{d}x/\,\mathrm{d}t$ je okamžitá rýchlosť a $p = mv$ hybnosť sústavy (pozri paragraf 17.6). Diferenciálnu rovnicu pre harmonický oscilátor možno teraz odvodiť z Hamiltonových pohybových rovníc

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial H}{\partial p}, \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial x},$$

Kým prvá z nich je len vyjadrením známeho faktu $\mathrm{d}x/\,\mathrm{d}t = v = p/m = \partial H/\partial p$, druhá dáva

$$m\,\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = \frac{\mathrm{d}(mv)}{\mathrm{d}t} = \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial x} = -m\omega^2 x,$$

z čoho po vykrátení člena m v krajných výrazoch dostaneme pôvodnú rovnicu harmonického oscilátora.

25.5 Harmonický oscilátor a Fourierove rady

Rovnica pre harmonický oscilátor je ekvivalentná s úlohou nájsť všetky vlastné funkcie (t. j. vlastné vektory) lineárneho operátora $-\frac{\mathrm{d}^2}{\mathrm{d}t^2}$ na vhodnom priestore funkcií, prislúchajúce k danej vlastnej hodnote ω^2 (čím sa v prípade úspechu spätne overí, že ω^2 je naozaj jeho vlastná hodnota). K štúdiu podobných otázok je účelné pristupovať zo zorného uhla unitárnych priestorov a samoadjungovaných operátorov.

Všimnime si, že všetky riešenia $x(t)$ rovnice harmonického oscilátora sú *periodické funkcie* s periódou $T = 2\pi/\omega$, čiže $x(t+T) = x(t)$ pre každé $t \in \mathbb{R}$. Obráťme tak trochu úlohu: zvoľme pevné kladné číslo T a uvažujme množiny $V = \mathcal{C}_T^{(\infty)}(\mathbb{R},\mathbb{R})$, $W = \mathcal{C}_T^{(\infty)}(\mathbb{R},\mathbb{C})$ všetkých funkcií $\mathbb{R} \to \mathbb{R}$ resp. $\mathbb{R} \to \mathbb{C}$, ktoré majú v každom bode $t \in \mathbb{R}$ (konečné a spojité) derivácie všetkých rádov a sú periodické s periódou T. Zrejme V je vektorový priestor nad $\mathbb{R}$ a W je vektorový priestor nad $\mathbb{C}$. Rovnosťou

$$\langle f, g\rangle = \int_0^{\mathsf{T}} f(t)\,\overline{g}(t)\,\mathrm{d}t$$

je na nich definovaný skalárny súčin (vo V možno vynechať pruh nad g). Ak si uvedomíme, že každá funkcia $f \in W$ má tvar $f = f_0 + \mathrm{i} f_1$ pre funkcie $f_0 = \operatorname{Re} f$, $f_1 = \operatorname{Im} f$ z V, vidíme, že nekonečnorozmerný unitárny priestor $W = V^{\mathbb{C}}$ je komplexifikáciou nekonečnorozmerného reálneho priestoru so skalárnym súčinom V (pozri príklady 13.1.2, 18.5.1, cvičenia 13.7, 13.8, 17.12 a paragraf 19.4).

Uvažujme lineárny operátor $\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\colon W \to W$. S použitím integrácie *per partes* a periodicity funkcií f, g, vďaka ktorej sa anulujú „hranaté zátvorky“, dostávame

$$\begin{aligned}\left\langle \mathrm{i}\frac{\mathrm{d}f}{\mathrm{d}t}, g\right\rangle &= \mathrm{i}\int_0^{\mathsf{T}} f'(t)\,\overline{g}(t)\,\mathrm{d}t\\ &= \mathrm{i}\big[f(t)\,\overline{g}(t)\big]_0^{\mathsf{T}} - \mathrm{i}\int_0^{\mathsf{T}} f(t)\,\overline{g}'(t)\,\mathrm{d}t = \left\langle f, \mathrm{i}\frac{\mathrm{d}g}{\mathrm{d}t}\right\rangle\end{aligned}$$

pre ľubovoľné funkcie $f, g \in W$. To znamená, že $\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}$ je *samoadjungovaný* lineárny operátor na priestore W.

Keby sme uvažovali o $\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}$ ako o lineárnom operátore na priestore $\mathcal{C}^{(\infty)}(\mathbb{R},\mathbb{C})$ všetkých (a nielen periodických) nekonečne diferencovateľných funkcií $\mathbb{R} \to \mathbb{C}$, bolo by každé $\lambda \in \mathbb{C}$ jeho vlastným číslom a zodpovedala by mu vlastná funkcia

$$\mathrm{e}^{-\mathrm{i}\lambda t} = \mathrm{e}^{\operatorname{Im}\lambda t}\big(\cos(\operatorname{Re}\lambda t) - \mathrm{i}\sin(\operatorname{Re}\lambda t)\big).$$

Takáto funkcia je však periodická s periódou T (teda patrí do nášho priestoru W), len ak $\lambda \in \mathbb{R}$ a $\lambda T/2\pi$ je celé číslo. Vlastnými číslami operátora $\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}$ na priestore W sú teda všetky celočíselné násobky čísla $\omega = 2\pi/T$, pričom vlastnému číslu $\lambda_n = n\omega$ zodpovedá vlastná funkcia

$$\mathrm{e}^{-\mathrm{i}n\omega t} = \cos(n\omega t) - \mathrm{i}\sin(n\omega t),$$

to znamená, že všetky vlastné čísla λ_n operátora $\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}$ sú jednoduché.

V dôsledku samoadjungovanosti operátora $\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}$ je aj

$$-\frac{\mathrm{d}^2}{\mathrm{d}t^2} = \mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t} \circ \mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t} = \left(\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\right)^2$$

samoadjungovaný lineárny operátor na W, a keďže $V \subseteq W$ je jeho invariantný podpriestor, tak aj na V. Jeho vlastnými hodnotami sú všetky čísla tvaru $\lambda_n^2 = n^2\omega^2$, kde $n \in \mathbb{Z}$. Vlastné číslo $\lambda_0^2 = 0$ je jednoduché – zodpovedá mu vlastná funkcia identicky rovná 1 (ďalší kandidát, funkcia $x(t) = t$, síce vyhovuje podmienke $-\mathrm{d}^2x/\mathrm{d}t^2 = 0$, no nie je periodická). Vlastné čísla $\lambda_n^2 = \lambda_{-n}^2 = n^2\omega^2$ pre $n > 0$ sú dvojnásobné, s príslušnými vlastnými funkciami $\mathrm{e}^{\pm \mathrm{i}n\omega t} = \cos(n\omega t) \pm \mathrm{i}\sin(n\omega t)$. Pri obmedzení sa na priestor V ich možno nahradiť vlastnými funkciami $\cos(n\omega t)$, $\sin(n\omega t)$ (veta 22.6.3).

Ortogonálny systém funkcií $\{\cos(n\omega t);\, n \in \mathbb{N}\} \cup \{\sin(n\omega t);\, 1 \leq n \in \mathbb{N}\}$ funguje ako svojho druhu „nekonečná ortogonálna báza“ priestoru V (porovnaj s cvičením 13.8). Presnejšie, každú funkciu $f \in V$ možno vyjadriť v tvare *trigonometrického Fourierovho radu*[2]

$$f(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}\bigl(a_n \cos(n\omega t) + b_n \sin(n\omega t)\bigr)$$

s *Fourierovými koeficientmi*

$$a_n = \frac{2}{T}\bigl\langle f, \cos(n\omega t)\bigr\rangle = \frac{2}{T}\int_0^T f(t)\cos(n\omega t)\,\mathrm{d}t,$$
$$b_n = \frac{2}{T}\bigl\langle f, \sin(n\omega t)\bigr\rangle = \frac{2}{T}\int_0^T f(t)\sin(n\omega t)\,\mathrm{d}t.$$

Uvedený rad rovnomerne konverguje k funkcii f dokonca za omnoho všeobecnejších podmienok: stačí napr., aby $f\colon \mathbb{R} \to \mathbb{R}$ bola spojitá periodická funkcia s periódou $T = 2\pi/\omega$ a s po častiach spojitou prvou deriváciou. Pre komplexné funkcie $f\colon \mathbb{R} \to \mathbb{C}$ analogických vlastností máme k dispozícii vyjadrenie v tvare *exponenciálneho* alebo *komplexného Fourierovho radu*

$$f(t) = \sum_{n\in\mathbb{Z}} c_n\, \mathrm{e}^{\mathrm{i}n\omega t}$$

s *komplexnými Fourierovými koeficientmi*

$$c_n = \frac{1}{T}\bigl\langle f, \mathrm{e}^{\mathrm{i}n\omega t}\bigr\rangle = \frac{1}{T}\int_0^T f(t)\,\mathrm{e}^{-\mathrm{i}n\omega t}\,\mathrm{d}t.$$

[2]Názov je na počesť francúzskeho prírodovedca a matematika Josepha Fouriera, ktorý vo svojej práci z r. 1822 použil trigonometrické rady na analýzu a riešenie parciálnej diferenciálnej rovnice pre vedenie tepla.

Na druhej strane, Fourierov rad spojitej periodickej funkcie k nej vo všeobecnosti nemusí konvergovať ani bodovo.

Pre (reálnu či komplexnú) periodickú funkciu f s periódou T, lebesgueovsky integrovateľnú na intervale $\langle 0, T\rangle$ spolu so svojim štvorcom f^2 máme ešte stále zaručenú *konvergenciu* jej Fourierovho radu k funkcii f *podľa stredu*. To znamená, že postupnosť

$$\|f - s_n\|^2 = \int_0^T |f(t) - s_n(t)|^2 \,\mathrm{d}t,$$

kde

$$s_n(t) = \frac{1}{2}a_0 + \sum_{k=0}^{n} \big(a_k \cos(k\omega t) + b_k \sin(k\omega t)\big),$$

resp.

$$s_n(t) = \sum_{k=-n}^{n} c_k \,\mathrm{e}^{\mathrm{i}k\omega t}$$

označuje n-tý čiastočný súčet príslušného Fourierovho radu, konverguje k nule. V takom prípade však treba vo vzorci pre skalárny súčin funkcií uvažovať namiesto Riemannovho integrálu Lebesgueov.

Otázky, pre aké funkcie a ako ich Fourierov rad konverguje, prípadne kedy a ako konverguje k pôvodnej funkcii, sú vo všeobecnosti veľmi zložité a podnes nie sú úplne zodpovedané. Podobné otázky priviedli v druhej polovici 19. storočia Georga Cantora k rozpracovaniu základov teórie množín (presnejšie, v jeho prípade to bola otázka jednoznačnosti rozvoja funkcie do Fourierovho radu) a na dlhé obdobie sa stali zdrojom významných podnetov jej ďalšieho rozvoja, ako aj rozvoja nových na nej založených matematických disciplín: všeobecnej topológie, teórie miery a funkcionálnej analýzy.

Hlavný význam Fourierových radov však spočíva v ich aplikáciách – používajú sa v elektrotechnike, akustike, optike, pri prenose signálov, spracovaní obrazu, kompresii dát či v spektrálnej analýze v chémii aj astronómii. Navyše ich mnohostranné zovšeobecnenia, ako napr. spojitá či diskrétna Fourierova transformácia, hrajú kľúčovú úlohu v mnohých oblastiach modernej matematiky, fyziky a informatiky.

Cvičenia

25.1. Dokážte tvrdenie 25.1.1.

25.2. Nájdite singulárne (a s výnimkou prípadov (e), (f) a (h)) aj polárne rozklady nasledujúcich matíc:

(a) $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, (b) $\boldsymbol{B} = \begin{pmatrix} 1 & -2 \\ 1/3 & 7 \end{pmatrix}$,

(a) $\boldsymbol{C} = \begin{pmatrix} 4 & 3 \\ -1 & 1 \end{pmatrix}$, (d) $\boldsymbol{D} = \begin{pmatrix} 1+2\mathrm{i} & -1 \\ 1+3\mathrm{i} & \mathrm{i} \end{pmatrix}$,

(e) $\boldsymbol{E} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & -3 \end{pmatrix}$, (f) $\boldsymbol{F} = \begin{pmatrix} 3 & \mathrm{i} & \mathrm{i} \\ -2\mathrm{i} & 0 & 1+\mathrm{i} \end{pmatrix}$,

(g) $\boldsymbol{G} = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 3 & -2 \\ 1 & 1 & 0 \end{pmatrix}$, (h) $\boldsymbol{H} = \begin{pmatrix} 3 & 1 \\ -2 & 0 \\ -1 & 1 \end{pmatrix}$.

25.3. Nájdite Mooreovu-Penroseovu pseudoinverziu riadkového vektora $\boldsymbol{a} = (a_1, \ldots, a_n) \in \mathbb{C}^n$ a stĺpcového vektora $\boldsymbol{b} = (b_1, \ldots, b_n)^\mathsf{T} \in \mathbb{C}^n$.

25.4. Nájdite Mooreove-Penroseove pseudoinverzie matíc $\boldsymbol{E}$, $\boldsymbol{F}$, $\boldsymbol{G}$ a $\boldsymbol{H}$ z cvičenia 25.2. Ako je to s pseudoniverziami zvyšných matíc z tohto cvičenia?

25.5. Nech $\boldsymbol{A} = \boldsymbol{Q} \cdot \boldsymbol{R}$ je QR-rozklad matice $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, kde $m \geq n$ (pozri cvičenia 13.13 a 17.16) a stĺpce matice $\boldsymbol{A}$ sú lineárne nezávislé (t. j. $h(\boldsymbol{A}) = n$). Potom $\boldsymbol{A}^* \cdot \boldsymbol{A} = \boldsymbol{R}^* \cdot \boldsymbol{R}$ a $\boldsymbol{A}^\neg = \boldsymbol{R}^\neg \cdot \boldsymbol{Q}^*$. Dokážte. Vysvetlite, ako možno veľmi jednoducho vypočítať pseudoinverznú maticu $\boldsymbol{R}^\neg$ k hornej trojuholníkovej matici $\boldsymbol{R} = \binom{\boldsymbol{R}'}{\boldsymbol{0}}$, kde $\boldsymbol{R}' \in \mathbb{C}^{n\times n}$ je regulárna.

25.6. Nájdite singulárny aj polárny rozklad všeobecnej Jordanovej bunky $\boldsymbol{J}_n(\lambda)$ a Mooreovu-Penroseovu pseudoinverziu Jordanovej bunky $\boldsymbol{J}_n(0)$.

25.7. Na príklade dvoch matíc rozmeru 2×2 ukážte, že pre matice $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$ nemusí platiť $(\boldsymbol{A} \cdot \boldsymbol{B})^\neg = \boldsymbol{B}^\neg \cdot \boldsymbol{A}^\neg$.

25.8. (a) Nájdite pseudoriešenie neriešiteľnej sústavy $(1 + \mathrm{i})x - 2y = 1$, $3x - \mathrm{i}y = 0$, $\mathrm{i}x + (1 - \mathrm{i})y = 2 + \mathrm{i}$ a dokážte, že je jediné.

(b) Nájdite minimálne riešenie sústavy $\mathrm{i}x + 2y - \mathrm{i}z = 0$, $2x + (2 - \mathrm{i})z = 1$.

(c) Nájdite minimálne pseudoriešenie sústavy $x + \mathrm{i}y - \mathrm{i}z = 1$ $2x + 2\mathrm{i}y - 2\mathrm{i}z = 0$.

25.9. (a) S využitím cvičenia 19.14 odvoďte, že singulárny rozklad cirkulantnej matice $\boldsymbol{C} = \boldsymbol{C}(a_0, a_1, \ldots, a_{n-1})$ má tvar $\boldsymbol{C} = \boldsymbol{F}^* \cdot \boldsymbol{D} \cdot \boldsymbol{F}$, kde $\boldsymbol{F} = n^{-1/2}\big(\omega^{-jk}\big)$ je znormovaná (teda unitárna) matica diskrétnej Fourierovej transformácie $F\colon \mathbb{C}^n \to \mathbb{C}^n$, $\omega = \mathrm{e}^{2\pi\mathrm{i}/n} = \cos(2\pi/n) + \mathrm{i}\sin(2\pi/n)$ je primitívna n-tá odmocnina z jednotky, $g(x) = a_0 + a_1x + \ldots + a_{n-1}x^{n-1}$ a $\boldsymbol{D} = \operatorname{diag}\big(|g(1)|, |g(\omega)|, \ldots, |g(\omega^{n-1})|\big)$.

(b) Na základe (a) dokážte, že Mooreova-Penroseova pseudoinverzia cirkulantnej matice má tvar $\boldsymbol{C}^\neg = \boldsymbol{F}^* \cdot \boldsymbol{D}^\neg \cdot \boldsymbol{F}$.

25.10. S využitím predchádzajúceho cvičenia modifikujte postup riešenia sústavy $\boldsymbol{C} \cdot \boldsymbol{x}$ z cvičenia 19.15 tak, aby ho bolo možné použiť, aj keď Fourierov obraz $F(\boldsymbol{a}^\varrho)$ vektora $\boldsymbol{a}^\varrho = (a_0, a_{n-1}, \ldots, a_1)^\mathsf{T}$ má niektoré zložky nulové. Aký druh riešenia predstavuje vektor $\boldsymbol{x} = \boldsymbol{C}^\neg \cdot \boldsymbol{b}$?

25.11. Nech S, T sú lineárne podpriestory euklidovského priestoru V. Označme $\varphi = \mathrm{pr}_T \restriction S\colon S \to T$, $\psi = \mathrm{pr}_S \restriction T\colon T \to S$, rovnako ako v dôkaze vety 25.3.1. Dokážte, že všetky vlastné čísla lineárneho operátora $\psi \circ \varphi\colon S \to S$ ležia v intervale $\langle 0, 1\rangle$.

25.12. (a) Určte odchýlku rovín $[\boldsymbol{e}_1, \boldsymbol{e}_2]$ a $[\boldsymbol{e}_1 + \boldsymbol{e}_3, \boldsymbol{e}_2 + e_4]$ v euklidovskom priestore $\mathbb{R}^4$.

(b) Určte odchýlku roviny $[\boldsymbol{e}_1, \boldsymbol{e}_2]$ a trojrozmerného lineárneho podpriestoru $[\boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3, \boldsymbol{e}_2 + \boldsymbol{e}_3 + \boldsymbol{e}_4, \boldsymbol{e}_3 + \boldsymbol{e}_4 + \boldsymbol{e}_5]$ v euklidovskom priestore $\mathbb{R}^5$.

25.13. Overte vzťahy $c_0 = \frac{1}{2}a_0$, $c_n = \frac{1}{2}(a_n - \mathrm{i}b_n)$, $c_{-n} = \frac{1}{2}(a_n + \mathrm{i}b_n)$ pre $n \geq 1$ medzi koeficientmi exponenciálneho a trigonometrického Fourierovho radu periodickej funkcie f. Nájdite spätné vyjadrenie koeficientov a_n, b_n pomocou c_n a c_{-n}.

25.14. Označme $\mathcal{C}_T^\dagger(\mathbb{R}, \mathbb{R})$ resp. $\mathcal{C}_T^\dagger(\mathbb{R}, \mathbb{C})$ vektorové priestory všetkých po častiach spojitých periodických funkcií $\mathbb{R} \to \mathbb{R}$ resp. $\mathbb{R} \to \mathbb{C}$ s periódou $T > 0$ a $\omega = 2\pi/T$.

(a) Dokážte že rovnosťou $\langle f, g\rangle = \frac{1}{T}\int_0^{\mathsf{T}} f(t)\,\overline{g}(t)\,\mathrm{d}t$ je definovaný skalárny súčin na $\mathcal{C}_T^{\dagger}(\mathbb{R},\mathbb{C})$ (a bez pruhu nad g aj na $\mathcal{C}_T^{\dagger}(\mathbb{R},\mathbb{R})$). Ďalej overte, že $\langle f, g\rangle = \frac{1}{T}\int_a^{a+T} f(t)\,\overline{g}(t)\,\mathrm{d}t$ pre ľubovoľné $a \in \mathbb{R}$.

(b) *Konvolúciou funkcií* $f, g \in \mathcal{C}_T^{\dagger}(\mathbb{R},\mathbb{C})$ nazývame funkciu $f * g$ danú predpisom $(f * g)(x) = \frac{1}{T}\int_0^{\mathsf{T}} f(x-t)g(t)\,\mathrm{d}t$. Dokážte, že pre $f, g \in \mathcal{C}_T^{\dagger}(\mathbb{R},\mathbb{C})$ platí $f * g \in \mathcal{C}_T^{\dagger}(\mathbb{R},\mathbb{C})$ a odvoďte z toho, že konvolúcia je komutatívna a asociatívna bilineárna operácia na vektorovom priestore $\mathcal{C}_T^{\dagger}(\mathbb{R},\mathbb{C})$. (*Návod*: Najprv dokážte, že $f * g \in \mathcal{C}_T^{\dagger}(\mathbb{R},\mathbb{R})$ pre $f, g \in \mathcal{C}_T^{\dagger}(\mathbb{R},\mathbb{R})$.)

(c) Označme $\ell_1(\mathbb{C})$ vektorový priestor všetkých postupností komplexných čísel $\boldsymbol{a} = (a_n)_{n\in\mathbb{Z}}$ (indexovaných „na obe strany") takých, že rad $\|\boldsymbol{a}\|_1 = \sum_{n\in\mathbb{Z}} |a_n|$ konverguje. Dokážte, že $\ell_1(\mathbb{C})$ je lineárny podpriestor vektorového priestoru $\mathbb{C}^{\mathbb{Z}}$ všetkých postupností $\boldsymbol{a}\colon \mathbb{Z} \to \mathbb{C}$ a $\|\cdot\|_1$ je norma na $\ell_1(\mathbb{C})$.

(d) Dokážte, že pre súčin $\boldsymbol{ab} = (a_n b_n)_{n\in\mathbb{Z}}$ postupností $\boldsymbol{a}, \boldsymbol{b} \in \ell_1$ platí $\boldsymbol{ab} \in \ell_1$. Odvoďte z toho, že rovnosťou $\langle \boldsymbol{a}, \boldsymbol{b}\rangle = \sum_{n\in\mathbb{Z}} a_n \overline{b}_n$ je definovaný komplexný skalárny súčin na vektorovom priestore $\ell_1(\mathbb{C})$ (porovnajte s cvičením 23.10 (b)).

(e) *Konvolúciou postupností* $\boldsymbol{a}, \boldsymbol{b} \in \ell_1$ nazývame postupnosť $\boldsymbol{a} * \boldsymbol{b} = (c_n)_{n\in\mathbb{Z}}$, kde $c_n = \sum_{k\in\mathbb{Z}} a_{n-k} b_k$. Dokážte, že pre $\boldsymbol{a}, \boldsymbol{b} \in \ell_1$ platí $\|\boldsymbol{a} * \boldsymbol{b}\|_1 \le \|\boldsymbol{a}\|_1 \|\boldsymbol{b}\|_1$. Odvoďte z toho, že $\boldsymbol{a} * \boldsymbol{b} \in \ell_1(\mathbb{C})$ a následne, že konvolúcia je komutatívna a asociatívna bilineárna operácia na vektorovom priestore $\ell_1(\mathbb{C})$.

(f) Nech $a_n = \langle f, \mathrm{e}^{\mathrm{i}n\omega t}\rangle$, $b_n = \langle g, \mathrm{e}^{\mathrm{i}n\omega t}\rangle$ sú Fourierove koeficienty funkcií $f, g \in \mathcal{C}_T^{\dagger}(\mathbb{R},\mathbb{C})$. Dokážte *Parsevalovu rovnosť* $\langle f, g\rangle = \langle \boldsymbol{a}, \boldsymbol{b}\rangle$.

(g) Pri rovnakom označení ako v (f) dokážte, že pre Fourierove koeficienty $c_n = \langle f * g, \mathrm{e}^{\mathrm{i}n\omega x}\rangle$ konvolúcie $f * g$ platí $c_n = a_n b_n$.

(h) Predpokladajme navyše, že $\boldsymbol{a}, \boldsymbol{b} \in \ell_1(\mathbb{C})$. Dokážte, že postupnosť $\boldsymbol{d}$ Fourierových koeficientov $d_n = \langle f\,g, \mathrm{e}^{\mathrm{i}n\omega t}\rangle$ súčinu funkcií f, g je konvolúciou postupností Fourierových koeficienov funkcií f, g, t. j. $\boldsymbol{d} = \boldsymbol{a} * \boldsymbol{b}$.

25.15. Predpokladajme, že $f \in \mathcal{C}_T^{\dagger}(\mathbb{R},\mathbb{C})$ má po častiach spojitú deriváciu, čiže $f' \in \mathcal{C}_T^{\dagger}(\mathbb{R},\mathbb{C})$.

(a) Potom pre Fourierove koeficienty $c_n' = \langle f', \mathrm{e}^{\mathrm{i}n\omega t}\rangle$ funkcie f' platí $c_n' = -\mathrm{i}n\omega c_n$, kde $c_n = \langle f, \mathrm{e}^{\mathrm{i}n\omega t}\rangle$ sú Fourierove koeficienty pôvodnej funkcie f. Dokážte.

(b) Za predpokladu, že $f\colon \mathbb{R} \to \mathbb{R}$, vyjadrite koeficienty a_n', b_n' trigonometrického Fourierovho radu funkcie f' pomocou koeficientov a_n, b_n trigonometrického Fourierovho radu funkcie f. Využite (a) a cvičenie 25.13 a výsledok overte deriváciou trigonometrického Fourierovho radu funkcie f člen za členom.

25.16. (a) Sformulujte Parsevalovu rovnosť pre trigonometrické Fourierove rady.

(b) Dokážte, že $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$. (*Návod*: Vypočítajte trigonometrické Fourierove koeficienty funkcie $f(x) = x$ na intervale $\langle -\pi, \pi\rangle$ a použite Parsevalovu rovnosť.)

(c) Vypočítate súčet radu $\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2}$.

26. Hermitovské operátory v kvantovej mechanike

V tejto kapitole sa oboznámime s matematickým modelovaním merateľných fyzikálnych veličín (hovoríme im tiež *pozorovateľné*) v kvantovej mechanike pomocou hermitovských operátorov. To si však vynúti doposiaľ najrozsiahlejší výlet mimo územia lineárnej algebry do iných oblastí matematiky – najmä funkcionálnej analýzy.

V súlade s označením zaužívaným v kvantovej mechanike budeme stavové vektory značiť gréckymi písmenami φ, χ, ψ a pod., zatiaľ čo lineárne operátory na stavovom priestore budeme značiť najrôznejšími písmenami – typicky A, B, X, Y, Z, P, H –, len nie φ, ψ ako doteraz.

V prístupe ku klasickej mechanike, s ktorým sme sa zoznámili v paragrafe 17.6, je okamžitý stav nejakého systému matematicky zachytený ako vektor (bod) vo vhodnom stavovom priestore. V prípade sústavy n pohybujúcich sa hmotných bodov to je $6n$-rozmerný euklidovský priestor $\mathbb{R}^{n\times 6}$. Jednotlivé zložky x_{ij}, p_{ij} daného vektora $(\boldsymbol{X}, \boldsymbol{P})$, kde $\boldsymbol{X}, \boldsymbol{P} \in \mathbb{R}^{n\times 3}$, predstavujú súradnice polohy resp. hybnosti i-teho hmotného bodu v smeroch troch súradných osí.

Hodnota akejkoľvek fyzikálnej veličiny, ktorú možno priradiť takejto sústave, je jednoznačne určená jej stavom. Matematicky ju teda možno zachytiť ako funkciu zo stavového priestoru do reálnych čísel. No v dôsledku toho, že množinové univerzum obsahuje tiež obrovské množstvo „patologických“ funkcií, nie každá takáto funkcia predstavuje merateľnú fyzikálnu veličinu. Preto za matematické modely fyzikálnych veličín zvykneme považovať len v nejakom zmysle „slušné“ funkcie, napr. dostatočne hladké alebo dostatočne dobre integrovateľné. Príkladmi takých funkcií sú samotné jednotlivé zložky polôh a hybností, t. j. priradenia $(\boldsymbol{X}, \boldsymbol{P}) \mapsto x_{ij}$, resp. $(\boldsymbol{X}, \boldsymbol{P}) \mapsto p_{ij}$, alebo Hamiltonova funkcia

$$H(\boldsymbol{X}, \boldsymbol{P}) = \sum_{i=1}^{n} \frac{1}{2m_i} \|\boldsymbol{p}_i\|^2 + \sum_{i=1}^{n} U_i(\boldsymbol{x}_i) + \sum_{i<j} U_{ij}(\boldsymbol{x}_i - \boldsymbol{x}_j);$$

hmotnosti m_i jednotlivých hmotných bodov treba chápať ako vopred dané kladné konštanty, a potenciálne energie vonkajšieho silového poľa resp. vzájomného silového pôsobenia medzi jednotlivými hmotnými bodmi, ako vopred dané „slušné“ funkcie $U_i, U_{ij} \colon \mathbb{R}^n \to \mathbb{R}$.

Stav nejakého systému v kvantovej mechanike je matematicky zachytený ako jednorozmerný podpriestor vhodného Hilbertovho (t. j. úplného unitárneho) priestoru – pozri paragraf 17.8. Napr. kvantovomechanickému analógu sústavy n-hmotných bodov zodpovedá priestor $L_2(\mathbb{R}^{n\times 3})$ všetkých *kvadraticky integrovateľných funkcií* $\psi\colon \mathbb{R}^{n\times 3} \to \mathbb{C}$. [1] Uvidíme, že „hybnostnú" alebo naopak „polohovú polovicu" $\mathbb{R}^{n\times 3}$ klasického stavového priestoru $\mathbb{R}^{n\times 6} \cong \mathbb{R}^{n\times 3} \times \mathbb{R}^{n\times 3}$ je pri tom možné, ba dokonca nutné „ušetriť." Hoci typicky ide o nekonečnorozmerné priestory, kvôli jednoduchosti sa budeme snažiť – pokiaľ to pôjde – ilustrovať preberanú látku na konečnorozmerných (teda automaticky úplných) unitárnych priestoroch.

26.1 Pozorovateľné veličiny a hermitovské operátory

Jeden zo základných postulátov kvantovej mechaniky hovorí, že *pozorovateľným fyzikálnym veličinám zodpovedajú hermitovské operátory na stavovom priestore príslušného kvantovomechanického systému.* Podľa silnejšej, no čiastočne problematickej verzie tohto princípu, tiež naopak, každý hermitovský operátor na stavovom priestore reprezentuje nejakú pozorovateľnú veličinu. V tomto zmysle teda možno termíny *pozorovateľná* a *hermitovský (samoadjungovaný) operátor* používať ako synonymá.

Podrobnejšie, ak je nejaká pozorovateľná reprezentovaná samoadjungovaným operátorom A s konečným spektrom $\operatorname{Spec} A = \{\lambda_1, \dots, \lambda_k\} \subseteq \mathbb{R}$ na stavovom priestore V a

$$A = \lambda_1 \operatorname{pr}_{S_1} + \dots + \lambda_k \operatorname{pr}_{S_k}$$

je jeho spektrálny rozklad (veta 23.3.3), tak bez ohľadu na stav $[\varphi]$ systému pred meraním, nameraná hodnota pozorovateľnej veličiny bude niektorá z vlastných hodnôt λ_j a systém po meraní prejde do stavu $[\psi]$, kde stavový vektor ψ patrí do príslušného vlastného podpriestoru $S_j = \operatorname{Ker}(A - \lambda_j)$. Pravdepodobnosť namerania danej vlastnej hodnoty λ_j a prechodu do stavu z vlastného podpriestoru S_j sa rovná výrazu

$$\frac{\langle \varphi, \operatorname{pr}_{S_j}(\varphi)\rangle}{\|\varphi\|^2} = \frac{\|\operatorname{pr}_{S_j}(\varphi)\|^2}{\|\varphi\|^2},$$

ktorý očividne závisí len od stavu $[\varphi]$ a nie samotného vektora φ.

Špeciálne, ak bol systém pôvodne v stave reprezentovanom *vlastným vektorom* $\varphi \in S_j$, tak nameraná hodnota pozorovateľnej (predstavovanej operátorom) A je s istotou λ_j a jeho stav sa po meraní nezmení.

Keďže $V = S_1 \oplus \ldots \oplus S_k$ je priamym súčtom vlastných podpriestorov operátora A, pre ľubovoľný vektor $\varphi \in V$ platí $\varphi = \operatorname{pr}_{S_1}(\varphi) + \ldots + \operatorname{pr}_{S_k}(\varphi)$.

[1] T. j. funkcia $|\psi|^2$ je lebesgueovsky integrovateľná na $\mathbb{R}^{n\times 3}$.

Preto ak $\varphi \neq \mathbf{0}$, tak

$$\sum_{j=1} \frac{\langle \varphi, \mathrm{pr}_{S_j}(\varphi) \rangle}{\|\varphi\|^2} = 1.$$

To znamená, že uvedené sčítance možno zmysluplne intrepretovať ako pravdepodobnosti. Stojí za zmienku, že podiel $\|\mathrm{pr}_{S_j}(\varphi)\| / \|\varphi\|$ možno chápať aj ako kosínus odchýlky lineárnych podpriestorov $[\varphi]$ a S_j.

Rovnako vedie na pravdepodobnostnú interpretáciu obraz

$$A\varphi = \sum_{j=1}^{k} \lambda_j \, \mathrm{pr}_{S_j}(\varphi)$$

vektora $\mathbf{0} \neq \varphi \in V$. Presnejšie, výraz

$$\langle A \rangle_\varphi = \frac{\langle \varphi, A\varphi \rangle}{\|\varphi\|^2} = \sum_{j=1}^{k} \lambda_j \frac{\langle \varphi, \mathrm{pr}_{S_j}(\varphi) \rangle}{\|\varphi\|^2} = \frac{1}{\|\varphi\|^2} \sum_{j=1}^{k} \lambda_j \|\mathrm{pr}_{S_j}(\varphi)\|^2$$

predstavuje *strednú hodnotu pozorovateľnej* A v stave $[\varphi]$.

Ak je spektrum operátora A jednoduché, čiže jeho vlastné podpriestory sú jednorozmerné, generované navzájom kolmými vektormi ψ_j, tak spomínané pravdepodobnosti majú tvar

$$\frac{|\langle \varphi, \psi_j \rangle|^2}{\|\varphi\|^2 \, \|\psi_j\|^2},$$

čo sú druhé mocniny absolútnych hodnôt amplitúd pravdepodobnosti stavových prechodov $[\varphi] \to [\psi_j]$ – pozri paragraf 17.8. Pri reprezentácii príslušných stavov *normovanými* vektormi φ, ψ_j sa táto pravdepodobnosť zjednoduší na tvar $|\langle \varphi, \psi_j \rangle|^2$. Za takých okolností

$$\varphi = \sum_{j=1}^{n} \langle \varphi, \psi_j \rangle \psi_j,$$

$$A\varphi = \sum_{j=1}^{n} \lambda_j \langle \varphi, \psi_j \rangle \psi_j,$$

$$\langle A \rangle_\varphi = \sum_{j=1}^{n} \lambda_j |\langle \varphi, \psi_j \rangle|^2.$$

Fyzici v prípade viacnásobných vlastných hodnôt hovoria o tzv. *degenerovanom spektre*. Zvlášť veľký význam má tento jav pre hamiltonián, kde v dôsledku degenerácie môžu splynúť energetické hladiny systému, ktoré by inak boli rôzne. K ich rozštiepeniu dôjde, ak do hamiltoniánu zahrnieme dodatočné členy opisujúce interakciu vnútri systému (bežne znedbávanú pri prvom priblížení) alebo doň pridáme ďalší člen opisujúci pôsobenie vonkajšieho poľa.

26.2 Heisenbergov vzťah neurčitosti

V ďalších úvahách už budeme predpokladať – tak ako to bežne robia fyzici, – že $\|\varphi\| = 1$, čiže stav systému je reprezentovaný normovaným vektorom, a takéto vektory budeme priamo nazývať *stavmi systému*. Musíme si však byť vedomí, že normované vektory líšiace sa len fázovým činiteľom $\mathrm{e}^{\mathrm{i}\vartheta}$, $\vartheta \in \mathbb{R}$, predstavujú ten istý stav. Čitateľ by si mal občas samostatne premyslieť, ako by bolo treba upraviť ďalej uvedené formuly, aby platili aj pre reprezentácie stavov ľubovoľnými nenulovými vektormi.

Reálne číslo $\langle A\rangle_\varphi$ je vhodné stotožniť s hermitovským operátorom $\langle A\rangle_\varphi \,\mathrm{id}_V$. Stredná hodnota pozorovateľnej A v stave φ je tak sama pozorovateľnou, rovnako ako operátory $A - \langle A\rangle_\varphi$ a $(A - \langle A\rangle_\varphi)^2$. Stredná hodnota poslednej pozorovateľnej v stave φ, t. j. výraz

$$\left\langle (A - \langle A\rangle_\varphi)^2 \right\rangle_\varphi = \left\|(A - \langle A\rangle_\varphi)(\varphi)\right\|^2$$

sa nazýva *disperziou* alebo *rozptylom pozorovateľnej* A v stave φ, a konečne výraz

$$\Delta_\varphi A = \sqrt{\left\langle (A - \langle A\rangle_\varphi)^2 \right\rangle_\varphi} = \left\|(A - \langle A\rangle_\varphi)(\varphi)\right\|$$

predstavuje *strednú kvadratickú odchýlku pozorovateľnej* A od jej strednej hodnoty v stave φ (porovnaj s paragrafom 14.6).

Pre pozorovateľnú A s konečným spektrom $\{\lambda_1, \dots, \lambda_k\}$ a príslušnými vlastnými podpriestormi $S_j = \operatorname{Ker}(A - \lambda_j)$ dostávame

$$\Delta_\varphi A = \left(\sum_{j=1}^{k} \bigl(\lambda_j - \langle A\rangle_\varphi\bigr)^2 \left\|\mathrm{pr}_{S_j}(\varphi)\right\|^2 \right)^{1/2}.$$

Za povšimnutie stojí, že stredná hodnota pozorovateľnej A vzhľadom na stav reprezentovaný *vlastným vektorom* $\varphi \in S_j$ sa rovná príslušnej vlastnej hodnote λ_j a jej stredná kvadratická odchýlka vzhľadom na φ je 0. Tento fakt sa zvykne intrepretovať ako možnosť odmerať hodnotu jednotlivej fyzikálnej veličiny kvantovomechanického systému vo vlastnom stave v zásade s ľubovoľnou presnosťou.

Na presnosť súčasného merania hodnôt viacerých pozorovateľných veličín však matematický aparát kvantovej mechaniky kladie isté zásadné obmedzenie, ktoré je v dobrej zhode s experimentom, známe ako *Heisenbergov vzťah neurčitosti.*

Pred jeho vyslovením pripomeňme, že

$$[A, B] = AB - BA = A \circ B - B \circ A$$

označuje komutátor lineárnych operátorov $A, B\colon V \to V$ – pozri paragraf 22.7.

26.2.1. Veta. *Nech A, B sú samoadjungované lineárne operátory na unitárnom priestore V. Potom pre ľubovoľný normovaný vektor $\varphi \in V$ platí*

$$\varDelta_\varphi A \cdot \varDelta_\varphi B \geq \frac{1}{2}\left|\big\langle [A,B](\varphi), \varphi\big\rangle\right|.$$

Dôkaz. Označme $A_0 = A - \langle A\rangle_\varphi$, $B_0 = B - \langle B\rangle_\varphi$. Potom zrejme $\varDelta_\varphi A = \|A_0\varphi\|$, $\varDelta_\varphi B = \|B_0\varphi\|$, a keďže operátory $\langle A\rangle_\varphi$, $\langle B\rangle_\varphi$ komutujú s každým operátorom, taktiež

$$[A_0, B_0] = [A, B].$$

Keď ďalej využijeme samoadjungovanosť operátorov A_0, B_0 a Cauchyho-Schwartzovu nerovnosť (veta 17.3.1), postupne nám vyjde

$$\begin{aligned}\left|\big\langle [A,B](\varphi),\varphi\big\rangle\right| &= \left|\big\langle (A_0B_0 - B_0A_0)\varphi,\varphi\big\rangle\right| = \left|\langle B_0\varphi, A_0\varphi\rangle - \langle A_0\varphi, B_0\varphi\rangle\right| \\ &= 2\left|\mathrm{Im}\langle A_0\varphi, B_0\varphi\rangle\right| \leq 2\left|\langle A_0\varphi, B_0\varphi\rangle\right| \\ &\leq 2\,\|A_0\varphi\|\,\|B_0\varphi\| = 2\,\varDelta_\varphi A\cdot\varDelta_\varphi B.\end{aligned}$$

Teda rozptyly dvoch nekomutujúcich pozorovateľných A, B v stave φ takom, že $\langle [A,B](\varphi),\varphi\rangle \neq 0$, nemôžu byť oba súčasne ľubovoľne malé. Čoskoro si ukážeme príklady pozorovateľných s komutátorom

$$[A, B] = \mathrm{i}\hbar = \mathrm{i}\hbar \cdot \mathrm{id}_V,$$

kde $\hbar = h/2\pi \approx 1{,}055 \cdot 10^{-34}\,\mathrm{J\,s}$ je (redukovaná) Planckova konštanta. Pre ich stredné kvadratické odchýlky od stredných hodnôt *v ľubovoľnom stave* φ potom platí

$$\varDelta_\varphi A \cdot \varDelta_\varphi B \geq \frac{\hbar}{2}.$$

Ak teda v niektorom stave dosiahneme napríklad výrazné zmenšenie strednej odchýlky $\varDelta_\varphi A$, úmerne tomu sa zvýši odchýlka $\varDelta_\varphi B$. Pravda, treba si tiež uvedomiť, že hodnota $\hbar/2$ je z makroskopického hľadiska nesmierne malá – prakticky rovná 0. To je matematické zdôvodnenie faktu, že princíp neurčitosti a podobné kvantové efekty sa „v bežnom živote“ neprejavujú.

Ďalej si všimnime, že na konečnorozmerných unitárnych priestoroch neexistujú hermitovské operátory s takouto vlastnosťou. Ak totiž $\dim V = n$, tak $\mathrm{tr}\,\mathrm{id}_V = \mathrm{tr}\,\boldsymbol{I}_n = n$, zatiaľ čo $\mathrm{tr}[A,B] = \mathrm{tr}(AB) - \mathrm{tr}(BA) = 0$ – pozri tvrdenie 18.1.3. V dôsledku toho operátory $\mathrm{i}\hbar$ a $[A,B]$ majú nevyhnutne rôzne stopy – prvý $\mathrm{i}\hbar n$, druhý 0 –, preto sa nemôžu rovnať. Aj to je jeden z dôvodov prečo v kvantovej mechanike nevystačíme s konečnorozmernými priestormi.

26.3 De Broglieho vlnové funkcie

Skôr než si ukážeme konkrétne príklady pozorovateľných, čiže hermitovských operátorov priradených niektorým základným fyzikálnym veličinám, bude potrebné malé fyzikálno–matematické odbočenie, v rámci ktorého sa pokúsime aspoň trochu skonkretizovať naše zatiaľ značne všeobecné a nie príliš jasné reči z paragrafu 17.7 o *vlnovo-korpuskulárnom dualizme.*

Max Planck dokázal v r. 1900 zásadne novým spôsobom vysvetliť experimentálne pozorovanú závislosť intenzity a frekvencie (vlnovej dĺžky) tepelného žiarenia vo vnútri rovnomerne zahriatej kovovej nádoby. Za východisko mu pri tom poslúžil predpoklad, že steny nádoby sa skladajú z oscilátorov, ktorých energia môže nadobúdať len diskrétne hodnoty – celočíselné násobky energie základného stavu $E = h\nu$, kde ν je frekvencia elektromagnetickej vlny príslušného žiarenia a $h \approx 6,626 \cdot 10^{-34}\,\mathrm{J\,s}$ je to, čo dostalo neskôr na jeho počesť názov *Planckova konštanta.*

Planckovu kvantovú hypotézu úspešne využil v r. 1905 Albert Einstein na vysvetlenie *fotolektrického javu*, t. j. emisie elektrónov z atómov niektorých kovov pod vplyvom dopadajúceho svetla. Podľa nej sa energia monochromatických svetelných lúčov s frekvenciou ν a vlnovou dĺžkou $\lambda = c/\nu$ môže pohlcovať len v celočíselných násobkoch rovnakého elementárneho kvanta $E = h\nu$. Svetlo teda možno chápať jednak ako elektromagnetické vlnenie, jednak ako prúd takýchto energetických kvánt, t. j. akýchsi čiastočiek, nazývaných *fotóny.* Jednotlivému fotónu prislúcha zároveň relativistická hmotnosť

$$m = \frac{E}{c^2} = \frac{h\nu}{c^2}$$

a hybnosť (o veľkosti)

$$p = mc = \frac{h\nu}{c} = \frac{h}{\lambda}.$$

De Broglieho vlnová hypotéza je v istom zmysle obrátením Planckovej hypotézy: akejkoľvek pohybujúcej sa častici zodpovedá istá rovinná vlna, matematicky popísaná komplexnou vlnovou funkciou. Pri presnejšej formulácii sa kvôli jednoduchosti obmedzíme len na časticu s jedným stupňom voľnosti, pohybujúcu sa rovnomerne priamočiaro v smere osi x. Takejto častici s energiou E a hybnosťou p prislúcha rovinná vlna s frekvenciou ν a vlnovou dĺžkou λ, danými Planckovými-Einsteinovými vzťahmi

$$\nu = \frac{E}{h}, \qquad \lambda = \frac{h}{p}.$$

Ak nahradíme frekvenciu ν *uhlovou frekvenciou* $\omega = 2\pi\nu$ a Planckovu konštantu h redukovanou Planckovou konštantou $\hbar = h/2\pi$, dostaneme uvedené rovnice v obvyklejšom tvare

$$\omega = \frac{E}{\hbar}, \qquad \lambda = \frac{2\pi\hbar}{p}.$$

Navyše hybnosť a energia voľnej častice spolu súvisia rovnosťou

$$E = \frac{p^2}{2m},$$

z ktorej spolu s Planckovými-Einsteinovými vzťahmi vyplýva

$$\nu = \frac{h}{2\lambda^2 m}, \qquad \omega = \frac{2\pi^2\hbar}{\lambda^2 m},$$

teda (kruhová) frekvencia nie je nepriamo úmerná vlnovej dĺžke (ako sme zvyknutí), ale jej štvorcu. Závislosť takého vlnenia na polohe x a čase t popisuje komplexná *de Broglieho vlnová funkcia*, zvaná tiež *rovinná vlna*,

$$\psi(x,t) = C\exp\left[-\mathrm{i}\left(\omega t - \frac{2\pi x}{\lambda}\right)\right] = C\exp\left[\frac{\mathrm{i}}{\hbar}(px - Et)\right],$$

pričom amplitúda $0 < C \in \mathbb{R}$ sa spravidla volí s ohľadom na vhodnú normovaciu podmienku. Popri nej zavádzame aj *bezčasovú de Broglieho vlnovú funkciu*

$$\psi(x) = \psi(x,0) = C\,\mathrm{e}^{2\pi\mathrm{i}x/\lambda} = C\,\mathrm{e}^{\mathrm{i}px/\hbar}.$$

Vlny vernejšie zodpovedajúce reálnej situácii sú superpozíciou viacerých (typicky z matematického hľadiska nekonečného diskrétneho alebo spojite sa meniaceho súboru) rovinných de Broglieho vĺn rôznych amplitúd, vlnových dĺžok (hybností) a uhlových frekvencií (energií). Výslednú vlnovú funkciu potom môžeme zapísať v tvare

$$\Psi(x,t) = \sum_{n=0}^{\infty} C_n \exp\left[\mathrm{i}\left(\frac{2\pi x}{\lambda_n} - \omega_n t\right)\right] = \sum_{n=0}^{\infty} C_n \exp\left[\frac{\mathrm{i}}{\hbar}(p_n x - E_n t)\right],$$

resp.

$$\Psi(x,t) = \int_{\mathbb{R}\smallsetminus\{0\}} C(\lambda)\exp\left[\mathrm{i}\left(\frac{2\pi x}{\lambda} - \omega t\right)\right]\mathrm{d}\lambda = \int_{\mathbb{R}} C(p)\exp\left[\frac{\mathrm{i}}{\hbar}(px - Et)\right]\mathrm{d}p.$$

V zrejme najúspešnejšej *Bornovej štatistickej interpretácii* de Broglieho vĺn sa výraz $\Psi(x,t)$ interpretuje ako *amplitúda hustoty pravdepodobnosti*, že sa častica v čase t nachádza v bode x. Samotnú hustotu pravdepodobnosti predstavuje výraz $|\Psi(x,t)|^2$. Charakter vlnenia určujú práve koeficienty C_n resp. $C(\lambda)$ či $C(p)$.

26.4 Pozorovateľná polohy

Na začiatok si predstavme kvantovomechanický systém pozostávajúci z jedinej častice, ktorá môže nadobúdať iba n rôznych polôh v bodoch 1, 2, …, n na osi x. V kvantovej mechanike takejto situácii ako stavový priestor zodpovedá unitárny priestor $\mathbb{C}^n$ (so štandardným skalárnym súčinom). Vektor $\boldsymbol{e}_k$ kanonickej ortonormálnej bázy $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n)$ popisuje stav systému „častica sa s istotou (t. j. s pravdepodobnosťou 1) nachádza v bode k“ Hermitovský operátor X zodpovedajúci pozorovateľnej polohy častice má preto vlastné vektory $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, …, $\boldsymbol{e}_n$, pričom vlastnému vektoru $\boldsymbol{e}_k$ zodpovedá vlastné číslo k. Operátor X tak má v báze $\boldsymbol{\varepsilon}$ maticu

$$(X)_{\boldsymbol{\varepsilon}} = \operatorname{diag}(1, 2, \dots, n)$$

a funguje podľa predpisu

$$X\varphi = (1\varphi_1, 2\varphi_2, \dots, n\varphi_n)^{\mathsf{T}},$$

kde $\varphi = (\varphi_1, \dots, \varphi_n)^{\mathsf{T}} \in \mathbb{C}^n$.

Pripusťme teraz, že naša častica môže nadobúdať ľubovoľnú polohu na osi x. Takej situácii zodpovedá stavový priestor $L_2(\mathbb{R})$ všetkých kvadraticky integrovateľných funkcií $\mathbb{R} \to \mathbb{C}$ so skalárnym súčinom

$$\langle \varphi, \psi \rangle = \int_{\mathbb{R}} \varphi(x)\, \overline{\psi}(x)\, \mathrm{d}x.$$

Pozorovateľná polohy (súradnice x) by mala mať možnosť nadobúdať všetky hodnoty $x \in \mathbb{R}$. Poučení predošlým jednoduchým príkladom ju teda budeme reprezentovať lineárnym operátorom $X\colon L_2(\mathbb{R}) \to L_2(\mathbb{R})$ násobenia argumentom, čiže

$$(X\varphi)(x) = x\varphi(x)$$

pre $\varphi \in L_2(\mathbb{R})$, $x \in \mathbb{R}$. Jednoduché overenie, že ide naozaj o hermitovský operátor, prenechávame čitateľovi.

Očakávame, že každé reálne číslo bude vlastným číslom operátora X (porovnaj s príkladom 18.5.1). Poznamenajme, že v prípade hermitovských operátorov, ktorých spektrum obsahuje nejaký netriviálny interval $\langle a, b \rangle$, hovoríme o tzv. *spojitom spektre*. Skúsme teraz nájsť vlastnú funkciu δ_a prislúchajúcu k vlastnému číslu $a \in \mathbb{R}$. Takáto funkcia musí vyhovovať podmienke

$$a\delta_a(x) = x\delta_a(x)$$

pre každé $x \in \mathbb{R}$. Teda pre $x = a$ môže byť hodnota $\delta_a(x)$ akákoľvek, kým $\delta_a(x) = 0$ pre $x \neq a$. Hodnotu $\delta_a(a)$ je prirodzené voliť tak, aby vektor δ_a mal jednotkovú normu, t. j.

$$\|\delta_a\|^2 = \int_{\mathbb{R}} |\delta_a(x)|^2\, \mathrm{d}x = 1, \tag{$*$}$$

a spĺňal obdobu podmienky $\langle \varphi, \boldsymbol{e}_k \rangle = \varphi_k$ pre $\varphi = (\varphi_1, \ldots, \varphi_n)^\mathsf{T} \in \mathbb{C}^n$, čiže

$$\langle \varphi, \delta_a \rangle = \int_{\mathbb{R}} \varphi(x)\, \overline{\delta_a}(x)\, \mathrm{d}x = \varphi(a) \tag{$**$}$$

pre $\varphi \in L_2(\mathbb{R})$.

Nech by sme však za $\delta_a(a)$ zvolili akékoľvek reálne či komplexné číslo, funkcia δ_a bude mať nenulovú hodnotu len v jedinom bode, v dôsledku čoho budú oba uvedené integrály vždy rovné nule, takže δ_a bude v priestore $L_2(\mathbb{R})$ nerozlíšiteľná od konštantne nulovej funkcie $\mathbf{0}\colon \mathbb{R} \to \mathbb{C}$. Vidíme, že medzi funkciami $\mathbb{R} \to \mathbb{C}$ sa vlastné funkcie operátora polohy X nenachádzajú.

Prísne vzaté to znamená, že žiadne reálne (a tým skôr komplexné) číslo nie je vlastnou hodnotou operátora polohy X – aspoň nie podľa definície z paragrafu 18.2. Jednako, v istom dobre definovanom zmysle má operátor X spojité spektrum $\operatorname{Spec} X = \mathbb{R}$.

26.5 Vlastné funkcie operátora polohy – Diracova δ-funkcia

Gordický uzol okolo vlastných funkcií pozorovateľnej polohy rozťal Paul Dirac v r. 1930, keď *postuloval* existenciu tzv. *zovšeobecnenej funkcie* δ takej, že $\delta(x) = 0$ pre každé $0 \neq x \in \mathbb{R}$, kým bližšie neurčená hodnota $\delta(0)$ je kladná a „nekonečne veľká“ takým spôsobom, že

$$\int_{\mathbb{R}} f(x)\, \delta(x)\, \mathrm{d}x = f(0),$$

pre každú spojitú funkciu $f \in L_2(\mathbb{R})$; špeciálne

$$\int_{\mathbb{R}} \delta(x)\, \mathrm{d}x = 1,$$

a taktiež

$$\delta(x) = \delta(-x) = \overline{\delta(x)}$$

pre každé $x \in \mathbb{R}$.

Zovšeobecnená funkcia δ dostala názov *Diracova δ-funkcia* a rýchlo sa ujala. Matematikov však zo zovšeobecnených funkcií ešte dlho bolela hlava a trvalo im ďalších pätnásť rokov, kým ich dokázali rigorózne modelovať ako tzv. *distribúcie*, čiže spojité lineárne funkcionály na vhodnom podpriestore Hilbertovho priestoru $L_2(\mathbb{R})$.[2]

[2]Ide o podpriestor *hladkých fumkcií s kompaktným nosičom*, t. j. funkcií $\mathbb{R} \to \mathbb{C}$, ktoré sú nulové mimo nejakého ohraničeného intervalu a majú všetky derivácie v každom bode $x \in \mathbb{R}$. Diracovu δ-funkciu potom stotožňujeme s funkcionálom $f \mapsto \int f(x)\, \delta(x)\, \mathrm{d}x = f(0)$ na tomto priestore. Zásluhu na rozpracovaní teórie distribúcií mali najmä Sergej Sobolev, Laurent Schwartz a Israil Gelfand.

Vďaka vlastnostiam Diracovej δ-funkcie sa každá z distribúcií $\delta_a(x) = \delta(x-a)$ správa ako vlastný vektor operátora X prislúchajúci k jeho vlastnej hodnote a a taktiež spĺňa podmienku $(**)$. S podmienkou $(*)$ je to horšie – účelnejšie sa ukazuje na nej netrvať a miesto toho rozšíriť platnosť podmienky $(**)$ aj na (aspoň niektoré) zovšeobecnené funkcie f. Pre $f = \delta_b$ dostávame

$$\langle \delta_b, \delta_a \rangle = \int_{\mathbb{R}} \delta_b(x)\, \delta_a(x)\, \mathrm{d}x = \int_{\mathbb{R}} \delta(x-b)\, \delta(x-a)\, \mathrm{d}x = \delta(b-a),$$

čo pripomína podmienku ortonormálnosti pre bázy v konečnorozmených unitárnych priestoroch $\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \delta_{ij}$, kde δ_{ij} je Kroneckerova delta. Špeciálne hodnota

$$\|\delta_a\|^2 = \int_{\mathbb{R}} \delta_a(x)^2\, \mathrm{d}x = \delta(0)$$

je nekonečne veľká, preto žiaden skalárny násobok $c\delta_a$ pre $c \in \mathbb{R}$ ani pre $c \in \mathbb{C}$ nemôže mať normu 1.

Systém funkcií $(\delta_a)_{a \in \mathbb{R}}$ je potom niečo ako „spojitá ortonormálna báza“ Funkciu $f \in L_2(\mathbb{R})$ možno vyjadriť ako „spojitú lineárnu kombináciu“, t. j. integrál

$$f = \int_{\mathbb{R}} f(a)\, \delta_a\, \mathrm{d}a,$$

ktorej koeficienty, t. j. „súradnice“ vzhľadom na túto „bázu“, tvoria samotné funkčné hodnoty

$$\langle f, \delta_a \rangle = \int_{\mathbb{R}} f(x)\, \delta(x-a)\, \mathrm{d}x = f(a).$$

Aspoň trochu názorne si možno Diracovu funkciu predstaviť ako svojho druhu limitu postupnosti funkcií

$$\alpha_n(x) = \frac{1}{2\pi} \int_n^n \mathrm{e}^{\mathrm{i}px}\, \mathrm{d}p = \frac{\sin nx}{\pi x}$$

pre $n \to \infty$. (Rovnakú úlohu však dokážu zastať aj iné postupnosti, napr. $\frac{n}{\sqrt{\pi}}\, \mathrm{e}^{-n^2x^2}$.) V tomto zmysle treba tiež chápať symbolickú rovnosť

$$\delta(x) = \lim_{n\to\infty} \alpha_n = \lim_{n\to\infty} \frac{\sin nx}{\pi x} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}px}\, \mathrm{d}p.$$

Nekonečnej hodnote $\delta(0)$ zodpovedajúci výraz $\lim_{x\to 0} \sin nx/\pi x$ naozaj diverguje do nekonečna. Pokročilejšími metódami teórie funkcií reálnej resp.

komplexnej premennej možno potom dokázať analógy symbolických vzťahov $\int_{\mathbb{R}} f(x)\,\delta(x)\,\mathrm{d}x = f(0)$ a $\delta(x) = 0$ pre $x \neq 0$. Presnejšie, pre ľubovoľnú funkciu $f \in L_2(\mathbb{R})$, spojitú v bode 0 platí

$$\lim_{n\to\infty} \langle f, \alpha_n \rangle = \lim_{n\to\infty} \int_{\mathbb{R}} f(x)\,\frac{\sin nx}{\pi x}\,\mathrm{d}x = f(0),$$

a taktiež

$$\lim_{n\to\infty} \int_a^b \alpha_n(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_a^b \frac{\sin nx}{\pi x}\,\mathrm{d}x = 0,$$

akonáhle interval $\langle a, b \rangle$ neobsahuje bod 0. Posledná rovnosť je špeciálnym prípadom *Riemannovej-Lebesgueovej vety*;[3] intuitívne ju možno zdôvodniť stále rýchlejšou osciláciou funkcie $\sin nx$ pre rastúce n.

Realistickejšej častici s troma stupňami voľnosti pohybujúcej sa v trojrozmernom priestore $\mathbb{R}^3$ zodpovedá stavový priestor $L_2\big(\mathbb{R}^3\big)$ kvadraticky integrovateľných funkcií $\mathbb{R}^3 \to \mathbb{C}$ so skalárnym súčinom

$$\langle \varphi, \psi \rangle = \iiint_{\mathbb{R}^3} \varphi(x,y,z)\,\overline{\psi}(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z.$$

Pozorovateľné súradníc x, y, z sú reprezentované hermitovskými operátormi X, Y, Z násobenia príslušnou súradnicou, t. j.

$$\begin{aligned}(X\varphi)(x,y,z) &= x\varphi(x,y,z),\\ (Y\varphi)(x,y,z) &= y\varphi(x,y,z),\\ (Z\varphi)(x,y,z) &= z\varphi(x,y,z),\end{aligned}$$

pre $\varphi \in L_2(\mathbb{R}^3)$, $x, y, z \in \mathbb{R}$. Úlohu vlastných vektorov opäť preberajú distribúcie: k vlastnej hodnote $a \in \mathbb{R}$ operátora X, Y resp. Z prislúchajú postupne zovšeobecnené vlastné funkcie

$$\begin{aligned}\delta_a^{(1)}(x,y,z) &= \delta(x-a),\\ \delta_a^{(2)}(x,y,z) &= \delta(y-a),\\ \delta_a^{(3)}(x,y,z) &= \delta(z-a).\end{aligned}$$

Priamym výpočtom možno ľahko overiť, že pozorovateľné súradníc navzájom komutujú, čiže

$$[X,Y] = [X,Z] = [Y,Z] = \mathbf{0}.$$

Inak povedané, rôzne súradnice polohy sú súčasne merateľné.

[3] Podľa nej zostane uvedená rovnosť v platnosti, aj keď funkciu $1/\pi x$ nahradíme ľubovoľnou lebesgueovsky integrovateľnou funkciou $g\colon \langle a, b \rangle \to \mathbb{C}$.

Vektorová pozorovateľná (všetkých troch súradníc) *polohy* je reprezentovaná *vektorovým operátorom*

$$\boldsymbol{R} = (X, Y, Z) = \boldsymbol{i}X + \boldsymbol{j}Y + \boldsymbol{k}Z \colon L_2\big(\mathbb{R}^3\big) \to L_2\big(\mathbb{R}^3\big) \times L_2\big(\mathbb{R}^3\big) \times L_2\big(\mathbb{R}^3\big),$$

kde $\boldsymbol{i} = \boldsymbol{e}_1$, $\boldsymbol{j} = \boldsymbol{e}_2$, $\boldsymbol{k} = \boldsymbol{e}_3$ je kanonická ortonormálna báza v $\mathbb{C}^3$ a

$$\boldsymbol{R}\varphi = (X\varphi, Y\varphi, Z\varphi) = \boldsymbol{i}X\varphi + \boldsymbol{j}Y\varphi + \boldsymbol{k}Z\varphi,$$

pre $\varphi \in L_2(\mathbb{R}^3)$. Polohovému vektoru $\boldsymbol{r} = (a, b, c) = a\boldsymbol{i} + b\boldsymbol{j} + c\boldsymbol{k} \in \mathbb{R}^3$, zloženému z trojice vlastných hodnôt a, b, c operátorov X, Y resp. Z, zodpovedá zovšeobecnená vlastná funkcia

$$\delta_{\boldsymbol{r}}(x, y, z) = \delta_a(x)\,\delta_b(y)\,\delta_c(z) = \delta(x - a)\,\delta(y - b)\,\delta(z - c).$$

26.6 Pozorovateľná hybnosti

Rovnako ako poloha, aj hybnosť častice s jedným stupňom voľnosti môže (aspoň v princípe) nadobúdať akúkoľvek hodnotu $p \in \mathbb{R}$. Teda pozorovateľná hybnosti P by opäť mala mať spojité spektrum $\operatorname{Spec} P = \mathbb{R}$. Ak ju máme definovať, tak prvé, čo nás asi napadne, bude napodobniť spôsob, ktorým sme definovali pozorovateľnú polohy. Priestor $L_2(\mathbb{R})$ (funkcií definovaných na $\mathbb{R}$ ako klasickom priestore polohy) však nie je uspôsobený na to, aby sme operátor hybnosti mohli definovať pomocou násobenia vhodným argumentom stavovej funkcie φ – klasická hybnosť totiž nie je argumentom takýchto funkcií. Dokonca sa zdá, že už pri definícii operátora polohy sme mali pracovať v priestore $L_2\big(\mathbb{R}^2\big)$ funkcií definovaných na plnom klasickom stavovom priestore systému s jedným stupňom voľnosti. Funkcia $\varphi \in L_2\big(\mathbb{R}^2\big)$ má potom dve premenné x a p, reprezentujúce klasickú polohu a hybnosť, takže nič nebráni tomu, aby sme kvantovomechanické pozorovateľné polohy X a hybnosti P definovali symetrickým spôsobom

$$(X\varphi)(x, p) = x\varphi(x, p), \qquad\qquad (P\varphi)(x, p) = p\,\varphi(x, p).$$

V takom prípade však veľmi ľahko overíme rovnosť $[X, P] = \mathbf{0}$, t. j. operátory X a P komutujú. Také niečo však v kvantovej mechanike interpretujeme ako *súčasnú merateľnosť polohy a hybnosti*, čo (ako sme už spomínali v paragrafe 17.7) odporuje experimentom. Dospievame tak k poznaniu, že zúženie klasického stavového priestoru $L_2\big(\mathbb{R}^2\big)$ na „polohovú polovicu“ $L_2(\mathbb{R})$ nebolo chybou lež nevyhnutnoťou. Zároveň však musíme hľadať iný spôsob, ako definovať pozorovateľnú hybnosti. Pri tom nám prídu vhod práve de Broglieho vlnové funkcie.

Pripomeňme si, že de Broglieho vlnová funkcia zodpovedá častici pohybujúcej sa *rovnomerne priamočiaro*. Pre pevné $p, t \in \mathbb{R}$ tak vlnová funkcia

$$\psi_{p,t}(x) = \exp\left[\frac{\mathrm{i}}{\hbar}(px - Et)\right] = \mathrm{e}^{\mathrm{i}px/\hbar} \cdot \mathrm{e}^{-\mathrm{i}Et/\hbar}$$

predstavuje stav, v ktorom má hybnosť častice s istotou hodnotu p. Činiteľ $\mathrm{e}^{-\mathrm{i}Et/\hbar}$ navyše nezávisí od x a vyjadruje len fázový posun, preto funkcia $\psi_{p,t}$ predstavuje rovnaký stav ako bezčasová vlnová funkcia

$$\psi_{p,0}(x) = \mathrm{e}^{\mathrm{i}px/\hbar}.$$

Teda $\psi_{p,0}$ by mala byť vlastnou funkciou pozorovateľnej P zodpovedajúcou vlastnému číslu $p \in \mathbb{R}$, t. j.

$$(P\psi_{p,0})(x) = p\,\psi_{p,0}(x) = p\,\mathrm{e}^{\mathrm{i}px/\hbar}$$

pre $x \in \mathbb{R}$. Teraz už nie je ťažké nahliadnuť, že operátor *pozorovateľnej hybnosti* P, ktorý vyhovuje tejto podmienke, musí mať tvar

$$P = \frac{\hbar}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}x},$$

čiže

$$P\varphi = \frac{\hbar}{\mathrm{i}}\frac{\mathrm{d}\varphi}{\mathrm{d}x} = \frac{\hbar}{\mathrm{i}}\varphi'$$

pre $\varphi \in L_2(\mathbb{R})$.

Sotva sme to však napísali, mala by nás zaraziť nepríjemná skutočnosť, že nie všetky funkcie z $L_2(\mathbb{R})$ sú diferencovateľné v každom bode $x \in \mathbb{R}$. Takrečeno proti vlastnej vôli sme tak nútení pripustiť, že hermitovské operátory zodpovedajúce merateľným fyzikálnym veličinám nemusia byť definované na celom stavovom priestore. Stačí, ak sú definované na nejakom jeho *hustom* lineárnom podpriestore,[4] ako je tomu aj v tomto prípade, keďže podpriestor funkcií z $L_2(\mathbb{R})$, ktoré majú konečné derivácie všetkých rádov v každom bode $x \in \mathbb{R}$, je naozaj hustý v $L_2(\mathbb{R})$. Pri spätnom pohľade si nezaškodí uvedomiť, že aj s operátorom polohy je to podobne. Hoci so samotnou definíciou funkcie $x\varphi(x)$ nie je problém, nie pre všetky funkcie $\varphi \in L_2(\mathbb{R})$ nutne platí $X\varphi \in L_2(\mathbb{R})$; také niečo možno zaručiť len pre funkcie z istého hustého podpriestoru.

Samodjungovanosť operátora P sa overí rovnako ako samoadjungovanosť operátora $\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}$ v paragrafe 25.5. Treba si len uvedomiť, že „hranaté zátvorky“ $[f(x)\overline{g}(x)]_{-\infty}^{\infty}$, ktoré vzniknú pri integrácii *per partes*, sa tentokrát anulujú

[4]Topologická podmienka *hustoty* podmnožiny S unitárneho priestoru V znamená, že pre všetky $\boldsymbol{x} \in V$ a $\varepsilon > 0$ existuje $\boldsymbol{y} \in S$ také, že $\|\boldsymbol{x} - \boldsymbol{y}\| < \varepsilon$.

v dôsledku kvadratickej integrovateľnosti funkcií f, g (ich limity pre $x \to \pm\infty$ sú totiž rovné 0).

Avizovaný Heisenbergov vzťah neurčitosti pre pozorovateľné polohy a hybnosti možno dokázať veľmi jednoducho.

26.6.1. Veta. *Komutátor pozorovateľných polohy X a hybnosti P pre kvantovo-mechanický systém s jedným stupňom voľnosti je*

$$[X, P] = \mathrm{i}\hbar.$$

V dôsledku toho pre ich stredné kvadratické odchýlky v ľubovoľnom stave φ platí

$$\Delta_\varphi X \cdot \Delta_\varphi P \geq \frac{\hbar}{2}.$$

Dôkaz. Overíme len uvedenú rovnosť pre komutátor. Heisenbergova nerovnosť pre ich kvadratické odchýlky je potom dôsledkom vety 26.2.1. Zvoľme dostatočne slušnú (stačí diferencovateľnú) funkciu $\varphi \in L_2(\mathbb{R})$. Potom pre každé $x \in \mathbb{R}$ platí

$$\begin{aligned}\big([X,P]\varphi\big)(x) &= (XP\varphi)(x) - (PX\varphi)(x) = x\,\frac{\hbar}{\mathrm{i}}\,\varphi'(x) - \frac{\hbar}{\mathrm{i}}\,(x\varphi)'(x) \\ &= \frac{\hbar}{\mathrm{i}}\big(x\varphi'(x) - \varphi(x) - x\varphi'(x)\big) = \mathrm{i}\hbar\,\varphi(x).\end{aligned}$$

Pri pokuse znormovať vlastné funkcie $\psi_{p,0}$ narazíme na ďalší problém: integrál

$$\|\psi_{p,0}\|^2 = \int_{\mathbb{R}} |\psi_{p,0}(x)|^2\,\mathrm{d}x = \int_{-\infty}^{\infty} 1\,\mathrm{d}x = +\infty$$

diverguje. Vidíme, že hoci sú vlnové funkcie $\psi_{p,0}$ – na rozdiel od distribúcií δ_a – „normálne“ funkcie $\mathbb{R} \to \mathbb{C}$, jedno majú s nimi spoločné: ani jedny ani druhé nepatria do priestoru $L_2(\mathbb{R})$. Rozšírenie priestoru funkcií o nové objekty, s cieľom zabezpečiť vlastné funkcie pre niektoré hermitovské operátory, vzdialene pripomína rozšírenie poľa reálnych čísel pomocou komplexných, ktorým sme dosiahli, aby všetky lineárne operátory na konečnorozmerných vektorových priestoroch nad $\mathbb{R}$ mali spektrum plnej algebraickej váhy.

Z nekonečnej veľkosti výrazu $\|\psi_{p,0}\|^2$ navyše vyplýva, že žiadnou voľbou amplitúdy $0 < C \in \mathbb{R}$ nemožno dosiahnuť, aby pre funkciu $\psi_p = C\psi_{p,0}$ platilo $\|\psi_p\| = 1$. Normovaciu konštantu C preto zvolíme tak, aby sme dosiahli splnenie podmienky

$$\langle \psi_p, \psi_q \rangle = \delta(p - q)$$

pre $p, q \in \mathbb{R}$. Rovnako ako rovnosť $\langle \delta_a, \delta_b \rangle = \delta(a - b)$ pre Diracove funkcie, je to zrejme najvhodnejšia obdoba podmienky $\langle \boldsymbol{v}_j, \boldsymbol{v}_k \rangle = \delta_{jk}$, charakterizujúcej ortonormálne bázy v konečnorozmerných unitárnych priestoroch. S využitím symbolickej rovnosti

$$\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}px}\,\mathrm{d}x = 2\pi\delta(p)$$

dostávame

$$\langle\psi_p,\psi_q\rangle = \int_{\mathbb{R}} C\,\mathrm{e}^{\mathrm{i}px/\hbar}\cdot C\,\mathrm{e}^{-\mathrm{i}qx/\hbar}\,\mathrm{d}x = C^2\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}(p-q)x/\hbar}\,\mathrm{d}x = 2\pi\hbar\,C^2\delta(p-q).$$

Z toho dôvodu volíme $C = 1/\sqrt{2\pi\hbar}$, takže *normované de Broglieho vlnové funkcie* majú tvar

$$\psi_p(x,t) = \frac{1}{\sqrt{2\pi\hbar}}\,\exp\left[\frac{\mathrm{i}}{\hbar}(px-Et)\right]$$

pre časovo závislú rovinnú vlnu, resp.

$$\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{\mathrm{i}px/\hbar}$$

pre bezčasovú rovinnú vlnu.

Pre časticu s troma stupňami voľnosti pohybujúcu sa v trojrozmernom priestore sú pozorovateľné hybnosti P_x, P_y, P_z v smeroch súradných osí x, y resp. z reprezentované imaginárnymi násobkami operátorov parciálnych derivácií podľa príslušnej premennej na priestore $L_2(\mathbb{R}^3)$ (presnejšie na jeho vhodnom hustom podpriestore)

$$P_x = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}, \qquad P_y = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial y}, \qquad P_z = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial z}.$$

K nim prislúchajúce normované vlastné funkcie, zopovedajúce vlastným hodnotám p_x, p_y resp. p_z majú tvar

$$\begin{aligned}
\psi^{(1)}_{p_x}(x,y,z) &= \psi_{p_x}(x) = \frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{\mathrm{i}xp_x/\hbar},\\
\psi^{(2)}_{p_y}(x,y,z) &= \psi_{p_y}(y) = \frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{\mathrm{i}yp_y/\hbar},\\
\psi^{(3)}_{p_z}(x,y,z) &= \psi_{p_z}(z) = \frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{\mathrm{i}zp_z/\hbar}.
\end{aligned}$$

Priamym výpočtom možno overiť *Heisenbergove komutačné vzťahy*

$$\begin{aligned}
[X,Y] = [X,Z] = [Y,Z] = [P_x,P_y] = [P_x,P_z] = [P_y,P_z] = \mathbf{0},\\
[X,P_y] = [X,P_z] = [Y,P_x] = [Y,P_z] = [Z,P_x] = [Z,P_y] = \mathbf{0},
\end{aligned}$$

$$[X, P_x] = [Y, P_y] = [Z, P_z] = \mathrm{i}\hbar.$$

Inak povedané, rovnako ako operátory súradníc X, Y, Z, aj operátory hybnosti P_x, P_y, P_z v smeroch jednotlivých osí navzájom komutujú. Takisto komutujú operátory súradníc a hybností v smeroch rôznych osí. Avšak operátory polohy a hybnosti v smere tej istej osi majú nenulový komutátor $\mathrm{i}\hbar$, teda presnosť ich súčasnej merateľnosti je obmedzená Heisenbergovým vzťahom neurčitosti.

Vektorová pozorovateľná (všetkých troch zložiek) *hybnosti* je reprezentovaná *vektorovým operátorom*

$$\boldsymbol{P} = (P_x, P_y, P_z) = \boldsymbol{i}P_x + \boldsymbol{j}P_y + \boldsymbol{k}P_z \colon L_2\big(\mathbb{R}^3\big) \to L_2\big(\mathbb{R}^3\big) \times L_2\big(\mathbb{R}^3\big) \times L_2\big(\mathbb{R}^3\big),$$

ktorý má tvar imaginárneho násobku operátora gradientu

$$\boldsymbol{P} = \frac{\hbar}{\mathrm{i}}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) = \frac{\hbar}{\mathrm{i}}\left(\boldsymbol{i}\,\frac{\partial}{\partial x} + \boldsymbol{j}\,\frac{\partial}{\partial y} + \boldsymbol{k}\,\frac{\partial}{\partial z}\right) = \frac{\hbar}{\mathrm{i}}\,\mathrm{grad}.$$

Vektoru hybnosti $\boldsymbol{p} = (p_x, p_y, p_z) = p_x\,\boldsymbol{i} + p_y\,\boldsymbol{j} + p_z\,\boldsymbol{k} \in \mathbb{R}^3$, zloženému z trojice vlastných hodnôt p_x, p_y, p_z operátorov P_x, P_y resp. P_z, zodpovedá vlastná funkcia

$$\begin{aligned}\psi_{\boldsymbol{p}}(x,y,z) &= \psi_{p_x}(x)\,\psi_{p_y}(y)\,\psi_{p_z}(z)\\ &= \frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{\mathrm{i}xp_x/\hbar}\,\frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{\mathrm{i}yp_y/\hbar}\,\frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{\mathrm{i}zp_z/\hbar} = \frac{1}{(2\pi\hbar)^{3/2}}\,\mathrm{e}^{\mathrm{i}\boldsymbol{rp}/\hbar},\end{aligned}$$

kde $\boldsymbol{r} = (x, y, z)$ je polohový vektor a $\boldsymbol{rp} = xp_x + yp_y + zp_z$ je štandardný skalárny súčin vektorov $\boldsymbol{r}$ a $\boldsymbol{p}$.

26.7 Polohová a hybnostná reprezentácia – Fourierova transformácia

Napriek diametrálne odlišnému tvaru operátorov polohy a hybnosti vládne v ich vzťahu zvláštna, hoci na prvý pohľad nie celkom viditeľná symetria. Aj na systém de Broglieho vlnových funkcií $(\psi_p)_{p\in\mathbb{R}}$ sa možno dívať ako na „spojitú ortonormálnu bázu“ priestoru $L_2(\mathbb{R})$, podobne ako na systém Diracových δ-funkcií $(\delta_a)_{a\in\mathbb{R}}$. Pre funkciu $f \in L_2(\mathbb{R})$ označme

$$\widehat{f}(p) = \langle f, \psi_p\rangle = \frac{1}{\sqrt{2\pi\hbar}}\int_{\mathbb{R}} f(x)\,\mathrm{e}^{-\mathrm{i}px/\hbar}\,\mathrm{d}x$$

koeficient pri funkcii ψ_p vo vyjadrení pôvodnej funkcie f v tvare „spojitej lineárnej kombinácie“

$$f = \int_{\mathbb{R}} \widehat{f}(p)\,\psi_p\,\mathrm{d}p.$$

Ak si totiž uvedomíme symetriu premenných x a p v de Broglieho vlnovej funkcii

$$\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}\, \mathrm{e}^{\mathrm{i}px/\hbar} = \psi_x(p),$$

a podmienku ortonormality $\langle \psi_x, \psi_y \rangle = \delta(x-y)$, ľahko overíme, že pre každé $x \in \mathbb{R}$ skutočne platí

$$\begin{aligned}\int_{\mathbb{R}} \widehat{f}(p)\, \psi_p(x)\, \mathrm{d}p &= \int_{\mathbb{R}} \langle f, \psi_p \rangle\, \psi_p(x)\, \mathrm{d}p = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(y)\, \overline{\psi}_p(y)\, \mathrm{d}y \right) \psi_p(x)\, \mathrm{d}p \\ &= \int_{\mathbb{R}} f(y) \left(\int_{\mathbb{R}} \psi_x(p)\, \overline{\psi}_y(p)\, \mathrm{d}p \right) \mathrm{d}y = \int_{\mathbb{R}} f(y)\, \langle \psi_x, \psi_y \rangle\, \mathrm{d}y \\ &= \int_{\mathbb{R}} f(y)\, \delta(x-y)\, \mathrm{d}y = f(x).\end{aligned}$$

Funkciu $\widehat{f} \colon \mathbb{R} \to \mathbb{C}$ tak možno chápať ako „súradnice“ funkcie f vzhľadom na „bázu“ $(\psi_p)_{p\in\mathbb{R}}$.

Poznamenajme, že formula $\widehat{f}(p) = \langle f, \psi_p \rangle$ ako aj uvedený výpočet nefungujú pre všetky funkcie $f \in L_2(\mathbb{R})$, ale opäť iba pre funkcie z istého hustého podpriestoru. Nad „prízemné“ otázky typu, ktoré funkcie to sú a čo s tými ostatnými, sa však veľkoryso povznesieme.

Priradením $f \mapsto \widehat{f}$ je definované zobrazenie $L_2(\mathbb{R}) \to L_2(\mathbb{R})$, nazývané *spojitá Fourierova transformácia* na $\mathbb{R}$ alebo len krátko *Fourierova transformácia*. Samozrejme, z matematického hľadiska je výskyt Planckovej konštanty vo Fourieovej transformácii nepodstatný – inak povedané, Fourierovu transformáciu možno definovať pre ľubovoľnú hodnotu parametra $\hbar$. [5] To je jeden z dôvodov, prečo nielen v matematike no i vo fyzike často kladieme $\hbar = 1$, čo zodpovedá voľbe takej sústavy jednotiek, v ktorej je Planckova konštanta $\hbar$ jednotkou účinku. My však budeme práve z fyzikálnych dôvodov konštantu $\hbar$ predsa len explicitne uvádzať.

26.7.1. Veta. *Fourierova transformácia $f \mapsto \widehat{f}$ je unitárny lineárny operátor na Hilbertovom priestore $L_2(\mathbb{R})$, ktorý prevádza vektory „ortonormálnej bázy“ $(\psi_p)_{p\in\mathbb{R}}$ na vektory „ortonormálnej bázy“ $(\delta_p)_{p\in\mathbb{R}}$. Navyše, pre $\varphi \in L_2(\mathbb{R})$ platí*

$$(\widehat{P\varphi})(p) = p\, \widehat{\varphi}(p) \qquad \text{a} \qquad (\widehat{X\varphi})(p) = \mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}p}\, \widehat{\varphi}(p).$$

[5] K tomu treba poznamenať, že aplikáciami v kvanotvej mechanike sa význam Fourierovej transformácie zďaleka nevyčerpáva. Napr. fakt, že Fourierova transformácia prevádza operátor derivácie $\mathrm{d}/\mathrm{d}x$ na operátor násobenia funkcie násobkom argumentu $(\mathrm{i}/\hbar)p$ (pozri nasledujúcu vetu 26.7.1), z nej robí účinný prostriedok používaný pri riešení diferenciálnych rovníc.

Dôkaz. Keďže linearita Fourierovej transformácie je zrejmá, stačí overiť, že zachováva skalárny súčin. Zvoľme „slušné" funkcie $\varphi, \chi \in L_2(\mathbb{R})$ a (len pre účely tohto výpočtu) položme $\hbar = 1$. Potom

$$\begin{aligned}
\langle \widehat{\varphi}, \widehat{\chi} \rangle &= \int_{\mathbb{R}} \widehat{\varphi}(p)\, \overline{\widehat{\chi}}(p)\, \mathrm{d}p = \int_{\mathbb{R}} \langle \varphi, \psi_p \rangle \langle \psi_p, \chi \rangle\, \mathrm{d}p \\
&= \int_{\mathbb{R}} \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \varphi(x)\, \mathrm{e}^{-\mathrm{i}px}\, \mathrm{d}x \;\cdot\; \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}py}\, \overline{\chi}(y)\, \mathrm{d}y \right) \mathrm{d}p \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(x) \left(\int_{\mathbb{R}} \overline{\chi}(y) \left(\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}p(y-x)}\, \mathrm{d}p \right) \mathrm{d}y \right) \mathrm{d}x \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(x) \left(\int_{\mathbb{R}} \overline{\chi}(y) \cdot 2\pi\, \delta(y-x)\, \mathrm{d}y \right) \mathrm{d}x \\
&= \int_{\mathbb{R}} \varphi(x)\, \overline{\chi}(x)\, \mathrm{d}x = \langle \varphi, \chi \rangle.
\end{aligned}$$

Ďalej pre $p, q \in \mathbb{R}$ platí

$$\widehat{\psi_p}(q) = \langle \psi_p, \psi_q \rangle = \delta(p-q) = \delta_p(q),$$

teda $\widehat{\psi_p} = \delta_p$. Posledné dve rovnosti overíme jednoduchými výpočtami

$$\begin{aligned}
(\widehat{P\varphi})(p) &= \langle P\varphi, \psi_p \rangle = \langle \varphi, P\psi_p \rangle = \langle \varphi, p\, \psi_p \rangle = p\, \langle \varphi, \psi_p \rangle = p\, \widehat{\varphi}(p), \\
(\widehat{X\varphi})(p) &= \langle X\varphi, \psi_p \rangle = \langle \varphi, X\psi_p \rangle = \left\langle \varphi, \frac{\hbar}{\mathrm{i}} \frac{\mathrm{d}}{\mathrm{d}p}\, \psi_p \right\rangle = -\frac{\hbar}{\mathrm{i}} \frac{\mathrm{d}}{\mathrm{d}p}\, \langle \varphi, \psi_p \rangle = \mathrm{i}\hbar\, \frac{\mathrm{d}}{\mathrm{d}p}\, \widehat{\varphi}(p).
\end{aligned}$$

Klasický stavový priestor $\mathbb{R}$ môžeme rovnako dobre chápať aj ako priestor hybností častice s jedným stupňom voľnosti – a nie jej polôh. Následne sa môžeme na kvantovomechanický stavový priestor $L_2(\mathbb{R})$ pozerať ako na priestor kvadraticky integrovateľných funkcií $\varphi\colon \mathbb{R} \to \mathbb{C}$ premennej hybnosti p. Náš pôvodný spôsob opisu kvantovomechanických stavov sa nazýva *súradnicová* prípadne *polohová reprezentácia* – v druhom prípade hovoríme o *hybnostnej reprezentácii.* Prepojenie medzi oboma reprezentáciami je analogické zámene jednej ortonormálnej bázy druhou v konečnorozmernom unitárnom priestore. Inak povedané, funkcie z $L_2(\mathbb{R})$ môžeme vyjadrovať jednak vzhľadom na „ortonormálnu bázu" tvorenú vlastnými vektormi δ_x operátora polohy X (t. j. obvyklým spôsobom), jednak vzhľadom na „ortonormálnu bázu" tvorenú vlastnými vektormi ψ_p operátora hybnosti P. Vzhľadom na túto „bázu" má pozorovateľná hybnosti P (v polohovej reprezentácii daná operátorom $\frac{\hbar}{\mathrm{i}} \frac{\mathrm{d}}{\mathrm{d}x}$) prirodzene tvar operátora násobenia argumentom p so zovšeobecnenými vlastnými funkciami δ_p. Pozorovateľná polohy X (v polohovej reprezentácii daná násobením súradnicou x) má zasa v hybnostnej reprezentácii tvar imaginárneho násobku operátora derivácie $\mathrm{i}\hbar\,\frac{\mathrm{d}}{\mathrm{d}p}$. Pritom vlastnej hodnote x zodpovedá ako vlastná funkcia normovaná de Broglieho bezčasová vlna

$$\overline{\psi}_x(p) = \psi_x(-p) = \frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{-\mathrm{i}px/\hbar},$$

v ktorej sme antisymetrickým spôsobom zamenili úlohy premenných x a p.

Pre úplnosť už len poznamenajme, že v trojrozmernom prípade má Fourierova transformácia „slušnej“ funkcie $f \in L_2\big(\mathbb{R}^3\big)$ tvar

$$\widehat{f}(\boldsymbol{p}) = \langle f, \psi_{\boldsymbol{p}}\rangle = \iiint_{\mathbb{R}^3} f(\boldsymbol{r})\,\overline{\psi}_{\boldsymbol{p}}(\boldsymbol{r})\,\mathrm{d}^3\boldsymbol{r} = \frac{1}{(2\pi\hbar)^{3/2}}\iiint_{\mathbb{R}^3} f(x,y,z)\,\mathrm{e}^{-\mathrm{i}\boldsymbol{r}\boldsymbol{p}/\hbar}\,\mathrm{d}^3\boldsymbol{r},$$

kde $\boldsymbol{r} = (x,y,z)$, $\boldsymbol{p} = (p_x,p_y,p_z)$ a $\mathrm{d}^3\boldsymbol{r} = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$. Vektorovým pozorovateľným hybnosti $\boldsymbol{P}$ resp. polohy $\boldsymbol{R}$ zodpovedajú v hybnostnej reprezentácii operátory dané predpismi

$$\begin{aligned}(\boldsymbol{P}\varphi)(\boldsymbol{p}) &= \boldsymbol{p}\,\varphi(\boldsymbol{p}) = (p_x,p_y,p_z)\,\varphi(\boldsymbol{p}),\\ (\boldsymbol{R}\varphi)(\boldsymbol{p}) &= \mathrm{i}\hbar\,\mathrm{grad}\varphi(\boldsymbol{p}) = \mathrm{i}\hbar\bigg(\frac{\partial\varphi(\boldsymbol{p})}{\partial p_x},\frac{\partial\varphi(\boldsymbol{p})}{\partial p_x},\frac{\partial\varphi(\boldsymbol{p})}{\partial p_x}\bigg).\end{aligned}$$

26.8 Ďalšie pozorovateľné – momenty hybnosti

V klasickej mechanike sú fyzikálne veličiny matematicky modelované ako („slušné“) funkcie na stavovom priestore, t. j. ako funkcie $F(x_1,\dots,x_n,p_1,\dots,p_n) = F(\boldsymbol{x},\boldsymbol{p})$ premenných polôh x_i a hybností p_j, kde n je počet stupňov voľnosti sústavy. V kvantovej mechanike zodpovedá takejto situácii stavový priestor $L_2(\mathbb{R}^n)$ a pozorovateľným polohy resp. hybnosti hermitovské operátory X_i, P_j, ktorých konkrétny matematický tvar závisí na zvolenej reprezentácii. Potom zrejme najprirodzenejší spôsob, ako zachytiť kvantovomechanický analóg klasickej fyzikálnej veličiny danej funkciou $F\colon \mathbb{R}^{2n} \to \mathbb{R}$, spočíva v jej reprezentácii operátorom

$$A = F(X_1,\dots,X_n,P_1,\dots,P_n).$$

Takýto prístup hladko funguje pre fyzikálne veličiny, ktoré závisia len na jedinej z premenných x_i, p_j. Ak je to napr. premenná x_i, t. j. $F(\boldsymbol{x},\boldsymbol{p}) = f(x_i)$, kde $f\colon \mathbb{R} \to \mathbb{R}$, tak operátor $A = f(X_i)$ funguje v polohovej reprezentácii ako násobenie funkčnou hodnotou argumentu x_i:

$$(A\varphi)(\boldsymbol{x}) = f(x_i)\,\varphi(\boldsymbol{x})$$

pre $\varphi \in L_2(\mathbb{R}^n)$, $\boldsymbol{x} \in \mathbb{R}^n$. Ak ide o funkciu závislú na jedinej premennej p_j, t. j. $F(\boldsymbol{x},\boldsymbol{p}) = g(p_j)$, kde $g\colon \mathbb{R} \to \mathbb{R}$, tak operátor $B = g(P_j)$ je výhodnejšie opisovať v hybnostnej reprezentácii:

$$(B\chi)(\boldsymbol{p}) = g(p_j)\,\chi(\boldsymbol{p})$$

pre $\chi \in L_2(\mathbb{R}^n)$, $\boldsymbol{p} \in \mathbb{R}^n$.

Možnosť voľby výhodnejšej z oboch reprezentácií sa však stráca, ak funkcia F závisí zároveň na niektorej zo súradníc polohy X_i aj hybnosti P_j. V takom prípade si vieme jednoducho poradiť, ak funkcia F je polynóm v premenných $x_1, \dots, x_n$, $p_1, \dots, p_n$, ktorý neobsahuje žiaden súčin $x_j^m p_j^k$ mocnín premenných zodpovedajúcich nekomutujúcim pozorovateľným X_j, P_j.

Mocninám p_j^k premennej hybnosti p_j (v polohovej reprezentácii) zodpovedajú mocniny operátora P_j:

$$P_j^k = \left(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x_j}\right)^k = \left(\frac{\hbar}{\mathrm{i}}\right)^k \frac{\partial^k}{\partial x_j^k}.$$

Ak F je polynóm pozostávajúci zo sčítancov tvaru $a x_1^{m_1} \dots x_n^{m_n} p_1^{k_1} \dots p_n^{k_n}$, pričom $m_j k_j = 0$ pre všetky $j \le n$, tak operátor $F(X_1, \dots, X_n, P_1, \dots, P_n)$ pozostáva z tomu zodpovedajúceho súčtu operátorov

$$a X_1^{m_1} \dots X_n^{m_n} P_1^{k_1} \dots P_n^{k_n}.$$

To znamená, že aplikáciou takého operátora na (dostatočne mnohokrát diferencovateľnú) funkciu $\varphi \in L_2(\mathbb{R}^n)$ dostaneme funkciu, ktorá je súčtom funkcií tvaru

$$a\left(\frac{\hbar}{\mathrm{i}}\right)^k x_1^{m_1} \dots x_n^{m_n} \frac{\partial^k \varphi}{\partial x_1^{k_1} \dots \partial x_n^{k_n}},$$

kde $k = k_1 + \ldots + k_n$.

Pre zložitejšie funkcie F, a taktiež ak F obsahuje súčiny mocnín niektorých dvojíc premenných x_j, p_j, výraz $F(X_1, \dots, X_n, P_1, \dots, P_n)$ zostavený podľa uvedeného receptu nemusí dávať hermitovský operátor. Vo všeobecnosti neexistuje nijaký jednoduchý univerzálny návod, ako priradiť funkcii F príslušnú pozorovateľnú. Problém sa zvykne riešiť pre jednotlivé prípady *ad hoc*. Napr. funkcii $F(x,p) = xp$ obyčajne priraďujeme pozorovateľnú $\frac{1}{2}(XP + PX)$.

Príkladom, kde naznačný prístup funguje, sú *pozorovateľné momentov hybnosti* vzhľadom na jednotlivé súradné osi. Klasický *moment hybnosti* je vektorová veličina definovaná ako vektorový súčin

$$\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p} = \begin{vmatrix} x & y & z \\ p_x & p_y & p_z \\ \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \end{vmatrix}$$

vektorov polohy $\boldsymbol{r} = (x, y, z)$ a hybnosti $\boldsymbol{p} = (p_x, p_y, p_z)$. Jeho jednotlivé zložky, t. j. momenty hybnosti vzhľadom na osi x, y resp. z, sú

$$l_x = yp_z - zp_y, \qquad l_y = zp_x - xp_z, \qquad l_z = xp_y - yp_x.$$

Pozorovateľným momentov hybnosti vzhľadom na jednotlivé osi preto zodpovedajú hermitovské operátory

$$\begin{aligned}
L_x &= YP_z - ZP_y = \frac{\hbar}{\mathrm{i}}\left(y\,\frac{\partial}{\partial z} - z\,\frac{\partial}{\partial y}\right),\\
L_y &= ZP_x - XP_z = \frac{\hbar}{\mathrm{i}}\left(z\,\frac{\partial}{\partial x} - x\,\frac{\partial}{\partial z}\right),\\
L_z &= XP_y - YP_x = \frac{\hbar}{\mathrm{i}}\left(x\,\frac{\partial}{\partial y} - y\,\frac{\partial}{\partial x}\right).
\end{aligned}$$

Vektorová pozorovateľná momentu hybnosti má potom tvar

$$\boldsymbol{L} = \boldsymbol{R} \times \boldsymbol{P} = \begin{vmatrix} X & Y & Z \\ P_x & P_y & P_z \\ \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \end{vmatrix} = \frac{\hbar}{\mathrm{i}} \begin{vmatrix} x & y & z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \end{vmatrix}.$$

Ešte poznamenjame, že nie všetky pozorovateľné majú klasickú analógiu, t. j. vznikajú z vhodných funkcií $F\colon \mathbb{R}^{2n} \to \mathbb{R}$. Napr. komutátor $[X, P] = XP - PX$ nie je hermitovský ale antihermitovský operátor – hermitovský je až operátor $(1/\mathrm{i})[X, P]$. Táto pozorovateľná však nemá analógiu v klasickej teórii.

Podstatne dôležitejším príkladom pozorovateľnej bez klasickej analógie je *spin*, čiže vnútorný moment hybnosti. Rovnako nazývame spinom aj najväčšiu hodnotu priemetu vnútorného momentu hybnosti v jednotkách $\hbar$. Táto veličina nadobúda polocelé hodnoty $1/2$, $3/2$, $5/2, \dots$ pre fermióny (elektróny, protóny, neutróny,…) a celé hodnoty 0, 1, $2, \dots$ pre bozóny (fotóny, W- a Z-bozóny, pióny, …). Ak chceme brať do úvahy aj spin častíc, treba stavový priestor vhodne modifikovať. Napr. pre jednu časticu so spinom $1/2$ musíme pôvodný unitárny priestor V nahradiť tenzorovým súčinom $V \otimes \mathbb{C}^2$. Pre n identických častíc so spinom $1/2$ poslúži tenzorový súčin n exemplárov priestoru $V \otimes \mathbb{C}$ – v ňom sa však musíme obmedziť na podpriestor antisymetrických tenzorov (pozri paragraf 33.6).

26.9 Hamiltonián a Schrödingerova rovnica

26.9.1. Hamiltonián. Vari najdôležitejšou pozorovateľnou popisujúcou vlastnosti kvantovomechanickej sústavy je – podobne ako v klasickej mechanike

– *pozorovateľná energie* alebo *hamiltonián*. Kvôli jednoduchosti sa pre začiatok obmedzíme len na jedinú časticu o hmotnosti m a s troma stupňami voľnosti, pohybujúcu sa v $\mathbb{R}^3$ v *konzervatívnom poli* danom potenciálnou energiou, ktorá sa nemení v čase.

Klasický výraz pre Hamiltonovu funkciu je v takom prípade súčtom dvoch častí: kinetickej energie

$$\frac{1}{2m}\|\boldsymbol{p}\|^2 = \frac{1}{2m}\big(p_x^2 + p_y^2 + p_z^2\big)$$

a potenciálnej energie $V(\boldsymbol{r})$, ktorá je výlučne funkciou (vektora) polohy $\boldsymbol{r} = (x, y, z)$, a plne určuje charakter poľa. V dôsledku toho je *Hamiltonov operátor* alebo *hamiltonián* (v polohovej reprezentácii) súčtom

$$\begin{aligned} H &= \frac{1}{2m}\big(P_x^2 + P_y^2 + P_z^2\big) + V(\boldsymbol{r}) \\ &= -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + V(\boldsymbol{r}) = -\frac{\hbar^2}{2m}\Delta + V(\boldsymbol{r}), \end{aligned}$$

kde $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ je *Laplaceov operátor*. Prvý člen $-\frac{\hbar^2}{2m}\Delta$ predstavuje *pozorovateľnú kinetickej* a druhý $V(\boldsymbol{r})$ *pozorovateľnú potenciálnej energie*. Pre istotu ešte pripomíname, že H je lineárny operátor, ktorý (dvakrát diferencovateľnej) funkcii $\psi \in L_2(\mathbb{R})$ priradí funkciu

$$H\psi = -\frac{\hbar^2}{2m}\Delta\psi + V(\boldsymbol{r})\psi = -\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}\right) + V(\boldsymbol{r})\psi.$$

Hamiltonián zodpovedajúci systému n častíc v $\mathbb{R}^3$ je hermitovský operátor na priestore $L_2\big(\mathbb{R}^{3n}\big)$ kvadraticky integrovateľných funkcií tvaru $\psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n)$, kde $\boldsymbol{r}_j = (x_j, y_j, z_j)$ je polohový vektor j-tej častice. Dostaneme ho sčítaním príspevkov kinetických a potenciálnych energií jednotlivých častíc, ku ktorým navyše pristupuje súčet potenciálnych energií vzájomných interakcií

$$H = -\frac{\hbar^2}{2}\sum_{j=1}^{n}\frac{1}{m_j}\left(\frac{\partial^2}{\partial x_j^2} + \frac{\partial^2}{\partial y_j^2} + \frac{\partial^2}{\partial z_j^2}\right) + \sum_{j=1}^{n} V_j(\boldsymbol{r}_j) + \sum_{j<k} V_{jk}(\boldsymbol{r}_j - \boldsymbol{r}_k).$$

26.9.2. Schrödingerova rovnica. Až doposiaľ sme pri našich úvahách o kvantovej mechanike nebrali do úvahy čas. Presnejšie, stavy našich kvantovomechanických systémov sa nevyvíjali. Ba čo viac, keď sa čas objavil v de Broglieho rovinných vlnách, ponáhľali sme sa ho z nich „vyhnať“ a rýchlo prejsť k bezčasovým vlnovým funkciám. Je už najvyšší čas niečo s tým robiť.

Znakmi φ, ψ začneme odteraz označovať nielen v čase nemenné ale aj časovo premenné stavy, reprezentované (normovanými) vektormi z priestoru

$L_2(\mathbb{R}^n)$ (kde n je počet stupňov voľnosti sústavy), čiže vlastne (vzhľadom na t dostatočne hladké) funkcie $\mathbb{R} \to L_2(\mathbb{R}^n)$ časovej premennej t. Miesto $\psi(t)(\boldsymbol{x})$ budeme písať $\psi(\boldsymbol{x}, t)$, teda v čase premenný vektor $\psi\colon \mathbb{R} \to L_2(\mathbb{R}^n)$ budeme chápať ako funkciu $n+1$ premenných $\psi\colon \mathbb{R}^{n+1} \to \mathbb{C}$.

Časový vývoj stavu ψ kvantovomechanického systému s hamiltoniánom H sa riadi parciálnou diferenciálnou rovnicou

$$\mathrm{i}\hbar \frac{\partial \psi}{\partial t} = H\psi,$$

nazývanou *Schrödingerova rovnica.*

Schrödingerova rovnica je jedným zo základných postulátov kvantovej mechaniky. Možno ju síce dobre zdôvodniť rôznymi heuristickými úvahami a najmä zhodou na jej základe získaných predpovedí s experimentmi, no – podobne ako Newtonove pohybové zákony či Maxwellove rovnice – ju nemožno striktne matematicky odvodiť z nejakých jednoduchších princípov.

Kedže operátory H aj $(1/\hbar)Ht$ sú hermitovské, podľa (nekonečnorozmernej operátorovej verzie) vety 23.4.8 sú operátory

$$U = \exp\left(\frac{1}{\mathrm{i}\hbar} H\right) = \mathrm{e}^{-(\mathrm{i}/\hbar)H} \qquad \text{a} \qquad U^t = \exp\left(\frac{1}{\mathrm{i}\hbar} Ht\right) = \mathrm{e}^{-(\mathrm{i}/\hbar)Ht}$$

unitárne pre každé $t \in \mathbb{R}$, teda

$$\left\| U^t \varphi \right\| = \|\varphi\|$$

pre každé $\varphi \in L_2(\mathbb{R}^n)$. Preto aplikáciou operátora U^t na normovaný stavový vektor φ dostaneme opäť normovaný stavový vektor $U_t \varphi$. Navyše pre všetky $s, t \in \mathbb{R}$ očividne platí

$$U^0 = \mathrm{id}, \qquad U^{s+t} = U^s\, U^t, \qquad U^{-t} = \left(U^t\right)^{-1}.$$

S vedomím týchto súvislostí možno *Schrödingerovu rovnicu* vysloviť v ekvivalentnom *integrálnom tvare*: ak je stav kvantovomechanického systému s hamiltoniánom H v čase t_0 daný normovaným vektorom φ, tak jeho stav čase t je daný normovaným vektorom

$$\psi(t) = U^{t-t_0} \varphi = \mathrm{e}^{-(\mathrm{i}/\hbar)H(t-t_0)}\, \varphi.$$

Pre také ψ skutočne platí $\psi(t_0) = \varphi$ a

$$\frac{\partial \psi}{\partial t} = \frac{\partial}{\partial t}\, \mathrm{e}^{-(\mathrm{i}/\hbar)H(t-t_0)}\, \varphi = -\frac{\mathrm{i}}{\hbar} H\, \mathrm{e}^{-(\mathrm{i}/\hbar)H(t-t_0)}\, \varphi = \frac{1}{\mathrm{i}\hbar} H\psi.$$

Časový vývoj kvantovomechanického systému je tak plne popísaný *jednoparametrickou grupou unitárnych operátorov* $\left(U^t\right)_{t\in\mathbb{R}}$ (pozri príklady 27.3.8 a 28.2.5).

V závislosti na hamiltoniáne H majú riešenia Schrödingerovej rovnice rôzne vlastnosti. Pri ich hľadaní hrá kľúčovú úlohu znalosť spektra hamiltoniánu. K danej vlastnej hodnote $E \in \operatorname{Spec} H$ treba najprv nájsť vlastnú funkciu (v prípade degenerovaného spektra ich môže byť aj viac) riešením tzv. *bezčasovej* alebo *stacionárnej Schrödingerovej rovnice*

$$H\psi = E\psi.$$

Cieľom je získať vlastný vektor ψ_E alebo – v prípade degenerovanej vlastnej hodnoty E – nejakú (či už diskrétnu alebo spojitú) ortonormálnu bázu $(\psi_{E,s})$ vlastného podpriestoru $\operatorname{Ker}(H - E)$, indexovanú vhodným parametrom s. Časovo závislé riešenie pôvodnej rovnice prislúchajúce k E a k ľubovoľnému bázickému vektoru $\psi_{E,s}$ sa potom nájde ako riešenie počiatočnej úlohy

$$\mathrm{i}\hbar\,\frac{\partial\psi}{\partial t} = E\psi, \qquad\qquad \psi(\boldsymbol{x}, 0) = \psi_{E,s}(\boldsymbol{x}),$$

čiže

$$\psi(\boldsymbol{x}, t) = \mathrm{e}^{-\mathrm{i}Et/\hbar}\,\psi_{E,s}(\boldsymbol{x}).$$

Časovo závislé riešenie zodpovedajúce všeobecnej počiatočnej podmienke

$$\psi(\boldsymbol{x}, 0) = \varphi(\boldsymbol{x}),$$

kde $\varphi \in L_2(\mathbb{R}^n)$, sú potom (či už diskrétnymi alebo spojitými) lineárnymi kombináciami riešení zodpovedajúcich jednotlivým dvojiciam E, $\psi_{E,s}$. Koeficient pri funkcii $\mathrm{e}^{-\mathrm{i}Et/\hbar}\,\psi_{E,s}(\boldsymbol{x})$ v takejto kombinácii sa nájde ako skalárny súčin

$$\langle\varphi, \psi_{E,s}\rangle = \int_{\mathbb{R}^n} \varphi(\boldsymbol{x})\,\overline{\psi}_{E,s}(\boldsymbol{x})\,\mathrm{d}^n\boldsymbol{x}\,.$$

26.9.3. Voľná častica na priamke. Najjednoduchší kvantovomechanický systém tvorí jediná *voľná častica* na priamke. V tomto prípade má hamiltonián tvar

$$H = -\frac{\hbar^2}{2m}\,\frac{\partial^2}{\partial x^2},$$

keďže „voľnosť“ práve znamená všade nulovú potenciálnu energiu. Príslušná bezčasová Schrödingerova rovnica

$$-\frac{\hbar^2}{2m}\,\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = E\psi$$

je vlastne rovnicou lineárneho oscilátora v zmenenom označení (pozri paragraf 25.4)

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = -\frac{p^2}{\hbar^2}\,\psi,$$

kde $p = \sqrt{2Em}$ je nezáporná hodnota hybnosti prislúchajúca k energii E. To znamená, že hamiltonián H má spojité spektrum a tvoria ho všetky reálne čísla $E \geq 0$. Každému z nich zodpovedajú dve lineárne nezávislé vlastné funkcie

$$\psi_{E,p}(x) = \psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}\, \mathrm{e}^{\mathrm{i}px/\hbar}, \qquad \overline{\psi}_{E,p}(x) = \psi_{-p}(x) = \frac{1}{\sqrt{2\pi\hbar}}\, \mathrm{e}^{-\mathrm{i}px/\hbar}$$

(s výnimkou $E = 0$, kedy sú obe konštantne rovné 1), v ktorých náš čitateľ isto spozná (normované) stacionárne de Broglieho vlnové funkcie. Zodpovedajúce riešenia časovej Schrödingerovej rovnice

$$\mathrm{i}\hbar\, \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\, \frac{\partial^2 \psi}{\partial x^2},$$

sú tak práve de Broglieho vlnové funkcie

$$\psi_p(x,t) = \mathrm{e}^{-\mathrm{i}Et/\hbar}\, \psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}\, \exp\left[\frac{\mathrm{i}}{\hbar}(px - Et)\right],$$

v ktorých sú energia E a hybnosť p zviazané vzťahom $E = p^2/2m$ (pričom hybnosť môže nadobúdať ľubovoľnú hodnotu $p \in \mathbb{R}$).

Všeobecnej počiatočnej podmienke $\psi(x,0) = \varphi(x)$ zodpovedá riešenie

$$\psi(x,t) = \int_{\mathbb{R}} \widehat{\varphi}(p)\, \psi_p(x,t)\, \mathrm{d}p = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \widehat{\varphi}(p) \exp\left[\frac{\mathrm{i}p}{\hbar}\left(x - \frac{pt}{2m}\right)\right] \mathrm{d}p,$$

v ktorom koeficienty pri de Broglieho funkciách $\psi_p(x,t)$ sú práve hodnoty $\widehat{\varphi}(p)$ Fourierovej transformácie funkcie φ. Všetko do seba krásne zapadá.

26.10 Kvantový harmonický oscilátor

Kvantový harmonický alebo tiež *kvantový lineárny harmonický oscilátor* je kvantovomechanickou obdobou harmonického oscilátora z klasickej mechaniky. Jeho význam je dvojaký. Na jednej strane popisuje so značnou presnosťou správanie sa veľkého množstvo rôznych kvantovomechanických sústav, ako napr. kmity dvojatómových molekúl. Zložitejšie systémy, hlavne v blízkosti rovnovážneho stavu, možno zasa často modelovať pomocou sústav viacerých rôznym spôsobom spriahnutých harmonických oscilátorov. Príkladom sú tzv. *fonóny*, t. j. kvantované vibračné módy (kvantá zvukových vĺn) v kryštalických mriežkach tuhých látok. Na druhej strane, vďaka matematickej jednoduchosti, je kvantový harmonický oscilátor jedným z mála kvantovomechanických systémov, ktorého Schrödingerovu rovnicu vieme presne vyriešiť, t. j. určiť spektrum a explicitne vyjadriť príslušné vlastné funkcie jej hamiltoniánu. V tom spočíva jeho metodický a pedagogický význam. Pre nás to navyše bude prvý (a zároveň jediný) príklad hamiltoniánu s *diskrétnym spektrom*, čo je inak pre kvantovú mechaniku charakteristické.

Pripomeňme, že Hamiltonova funkcia klasického harmonického oscilátora má tvar

$$H(x,p) = \frac{1}{2}\left(\frac{p^2}{m} + m\omega^2 x^2\right),$$

pričom členy $p^2/2m$ a $m\omega^2x^2/2$ zodpovedajú kinetickej resp. potenciálnej energii – pozri paragraf 25.4.

Tomu zodpovedajúci *hamiltonián kvantového harmonického oscilátora* (v súradnicovej reprezentácii) je hermitovský operátor na priestore $L_2(\mathbb{R})$

$$H = \frac{1}{2}\left(\frac{P^2}{m} + m\omega^2 X^2\right) = \frac{1}{2}\left(-\frac{\hbar^2}{m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + m\omega^2 x^2\right),$$

kde operátory $P^2/2m$ a $m\omega^2X^2/2$ predstavujú pozorovateľné kinetickej resp. potenciálnej energie. Príslušná stacionárna Schrödingerova rovnica $H\psi = E\psi$ má preto tvar

$$\frac{1}{2}\left(\frac{P^2}{m} + m\omega^2 X^2\right)\psi = E\psi,$$

prípadne, rozmenená na drobné,

$$\frac{1}{2}\left(-\frac{\hbar^2}{m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + m\omega^2 x^2\psi\right) = E\psi.$$

Veľkú časť „spektrálnej analýzy“ tejto rovnice možno vykonať len na základe komutačných vzťahov medzi šikovne zvolenými operátormi. Zaveďme si teda *anihilačný operátor* A a k nemu združený *kreačný operátor* A^*

$$\begin{aligned}
A &= \frac{1}{\sqrt{2\hbar}}\left(\sqrt{m\omega}\,X + \frac{\mathrm{i}}{\sqrt{m\omega}}\,P\right) = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{m\omega}{\hbar}}\,x + \sqrt{\frac{\hbar}{m\omega}}\,\frac{\mathrm{d}}{\mathrm{d}x}\right),\\
A^* &= \frac{1}{\sqrt{2\hbar}}\left(\sqrt{m\omega}\,X - \frac{\mathrm{i}}{\sqrt{m\omega}}\,P\right) = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{m\omega}{\hbar}}\,x - \sqrt{\frac{\hbar}{m\omega}}\,\frac{\mathrm{d}}{\mathrm{d}x}\right).
\end{aligned}$$

Na základe bilinearity a antisymetrie komutátora a Heisenbergovho komutačného vzťahu $[X,P] = \mathrm{i}\hbar$ dostávame

$$\begin{aligned}
[A,A^*] &= \frac{1}{2\hbar}\left(m\omega\,[X,X] + \frac{1}{m\omega}\,[P,P] - \mathrm{i}\,[X,P] + \mathrm{i}\,[P,X]\right)\\
&= \frac{1}{2\hbar}(-2\mathrm{i}[X,P]) = -\frac{\mathrm{i}}{\hbar}\,\mathrm{i}\hbar = 1,
\end{aligned}$$

t. j. $[A, A^*]$ je identický operátor. Označme

$$M = A^*A = \frac{1}{2\hbar}\left(m\omega X^2 + \frac{1}{m\omega}\,P^2 + \mathrm{i}\,[X,P]\right) = \frac{1}{\hbar\omega}\,H - \frac{1}{2}.$$

Z paragrafu 25.1 vieme, že M je nezáporný hermitovský operátor, teda $0 \le \lambda \in \mathbb{R}$ pre každé $\lambda \in \operatorname{Spec} M$. Jeho tesný súvis s pôvodným hamiltoniánom H nás oprávňuje nazvať ho *modifikovaným hamiltoniánom* a navyše zaručuje, že

$$\operatorname{Spec} H = \left\{\left(\lambda + \frac{1}{2}\right)\hbar\omega;\ \lambda \in \operatorname{Spec} M\right\},$$

pričom vlastným číslam λ a $E = (\lambda + 1/2)\hbar\omega$ operátorov M resp. H zodpovedajú rovnaké vlastné funkcie. Stačí sa teda zaoberať modifikovaným hamiltoniánom M. Začneme výpočtom komutátorov

$$[M, A] = [A^*A,\, A] = \big(A^*[A, A] + [A^*, A]\,A\big) = -A,$$
$$[M, A^*] = [A^*A,\, A^*] = \big(A^*[A, A^*] + [A^*,\, A^*]\,A\big) = A^*.$$

Ďalej si uvedemme, že

$$AA^* = A^*A + [A, A^*] = M + 1.$$

Predpokladajme teraz, že ψ je vlastná funkcia modifikovaného hamiltoniánu M prislúchajúca k jeho vlastnému číslu $0 \le \lambda \in \mathbb{R}$. Potom

$$M(A\psi) = [M, A]\psi + A(M\psi) = -A\psi + A(\lambda\psi) = (\lambda - 1)(A\psi),$$
$$M(A^*\psi) = [M, A^*]\psi + A^*(M\psi) = A^*\psi + A^*(\lambda\psi) = (\lambda + 1)(A^*\psi),$$

a taktiež

$$\|A\psi\|^2 = \langle A\psi, A\psi\rangle = \langle A^*A\psi, \psi\rangle = \langle M\psi, \psi\rangle = \langle\lambda\psi, \psi\rangle = \lambda\,\|\psi\|^2,$$
$$\|A^*\psi\|^2 = \langle A^*\psi, A^*\psi\rangle = \langle AA^*\psi, \psi\rangle = \langle (M+1)\psi, \psi\rangle = (\lambda + 1)\,\|\psi\|^2.$$

Teda okrem prípadu $\lambda = 0$ je aj $\lambda - 1$ vlastným číslom operátora M a zodpovedá mu vlastná funkcia $A\psi$. Podobne je aj $\lambda + 1$ je vlastným číslom hamiltoniánu M a zodpovedá mu vlastná funkcia $A^*\psi$.

Náš prvý čiastočný záver teda znie: Ak M má čo len jednu vlastnú hodnotu λ, tak je to nezáporné celé číslo, a následne platí

$$\operatorname{Spec} M = \{0, 1, 2, \dots\} = \mathbb{N}.$$

Navyše, ak ψ je vlastná funkcia operátora M prislúchajúca k vlastnému číslu 0, tak $(A^*)^n\psi$ je jeho vlastná funkcia prislúchajúca k vlastnému číslu n a platí

$$\big\|(A^*)^n\psi\big\| = \sqrt{n!}\;\|\psi\|.$$

Ukážeme, že 0 je naozaj vlastným číslom operátora M a nájdeme príslušný vlastný podpriestor $\operatorname{Ker} M$. Hamiltonián M si vyjadríme v explicitnom tvare

$$\begin{aligned} M &= \frac{1}{\hbar\omega}\left(\frac{1}{2m}\,P^2 + \frac{m\omega^2}{2}\,X^2\right) - \frac{1}{2} \\ &= \frac{1}{2}\left(-\frac{\hbar}{m\omega}\,\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{m\omega}{\hbar}\,x^2 - 1\right). \end{aligned}$$

Diferenciálna rovnica $M\psi = 0$ prejde po miernej úprave a substitúcii

$$\xi = \sqrt{\frac{m\omega}{\hbar}}\,x, \qquad\qquad \psi(x) = \varphi(\xi),$$

na tvar

$$\frac{\mathrm{d}^2\varphi}{\mathrm{d}\xi^2} = (\xi^2 - 1)\varphi$$

(uvedomte si, že potom $\mathrm{d}\xi = \sqrt{m\omega/\hbar}\,\mathrm{d}x$). To je lineárna diferenciálna rovnica druhého rádu s nekonštantými koeficientmi, na ktorú metódy z paragrafu 22.5 nezaberajú – (pozrite si poznámku (b) za vetou 22.5.1 a vetu 22.5.3 a rozmyslite si prečo). Jedno riešenie sa však dá uhádnuť a druhé čitateľovi prezradíme: fundamentálny systém jej riešení tvoria funkcie

$$\varphi_0(\xi) = \mathrm{e}^{-\xi^2/2} \qquad \text{a} \qquad \varphi_0^+(\xi) = \mathrm{e}^{-\xi^2/2}\int_0^\xi \mathrm{e}^{u^2}\,\mathrm{d}u$$

– dodatočné overenie je už priamočiare.

Norma riešenia φ_0 vedie na známy *Laplaceov integrál* $\int_{-\infty}^{\infty} \mathrm{e}^{-\xi^2}\,\mathrm{d}\xi = \sqrt{\pi}$ (pozri cvičenie 14.20), preto

$$\|\varphi_0\|^2 = \int_{\mathbb{R}} |\varphi(\xi)|^2\,\mathrm{d}x = \sqrt{\frac{\hbar}{m\omega}}\int_{-\infty}^{\infty} \mathrm{e}^{-\xi^2}\,\mathrm{d}\xi = \sqrt{\frac{\pi\hbar}{m\omega}}.$$

Z toho vyplýva, že $\varphi_0 \in L_2(\mathbb{R})$ a normovaná vlastná funkcia prislúchajúca k vlastnej hodnote 0 je

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\!\left(-\frac{m\omega}{2\hbar}x^2\right).$$

Druhé riešenie φ_0^+ nepatrí do priestoru $L_2(\mathbb{R})$. To by totiž obe limity $\varphi_0^+(\xi)$ pre $\xi \to \pm\infty$ museli byť rovné 0. Táto funkcia sa však pre malé hodnoty $|\xi|$ správa približne ako $\xi\,\mathrm{e}^{\xi^2}$ a pre veľké hodnoty $|\xi|$ má asymptotické chovanie ako $\mathrm{e}^{\xi^2}/2\xi$, teda $\lim_{\xi\to\pm\infty} \varphi_0^+(\xi) = \pm\infty$. To sú matematické dôvody, pre ktoré funkcia φ_0^+ nepopisuje žiaden stav harmonického oscilátora.

Kritického čitateľa, ktorý si spomenie, že ako vlastné stavy operátora hybnosti P sme pripustili de Broglieho vlnové funkcie $\mathrm{e}^{ipx/\hbar}$, ktoré rovnako nepatrili do priestoru $L_2(\mathbb{R})$ (hoci až tak prudko neubiehali do $\pm\infty$ ako funkcia φ_0^+), a ako vlastné funkcie operátora polohy X sme pripustili dokonca Diracove δ-funkcie, ktoré ani nie sú funkcie ale distribúcie, však takéto vysvetlenie sotva uspokojí. Dodajme ešte, že v prípade spojitého spektra vládne „väčšia benevolencia“ – jednoducho preto, lebo nájsť k takýmto vlastným hodnotám vlastné funkcie v priestore $L_2(\mathbb{R}^n)$ nie je vždy možné. Naopak, pre vlastné hodnoty z diskrétneho spektra to vždy možné je, preto si voči vlastným funkciám ležiacim mimo tohto priestoru môžeme dovoliť uplatňovať „diskriminačnú politiku“.

Rozhodujúce dôvody sú však fyzikálne. Degenerovanosť spektra hamiltoniánu a následne dvojnásobný počet stupňov voľnosti systémov popisovaných pomocou kvantového harmonického oscilátora by sa museli prejaviť pozorovateľnými fyzikálnymi efektmi – napr. v odlišných hodnotách merného tepla rôznych látok, ako boli skutočne namerané. Taktiež matematicky formulovaná požiadavka $\lim_{\xi\to\pm\infty}\varphi(\xi)=0$ pre riešenia rovnice $\mathrm{d}^2\varphi/\,\mathrm{d}\xi^2=(\xi^2-1)\xi$ má fyzikálnu interpretáciu: Podľa klasickej teórie by častica v základnom stave mala byť lokalizovaná v blízkosti rovnovážnej polohy (počiatku súradného systému). V kvantovej mechanike vychádza istá nenulová pravdepodobnosť, že ju nájdeme ľubovoľne ďaleko od rovnovážnej polohy. No keďže klasická mechanika je svojho druhu limitnou aproximáciou kvantovej mechaniky, musí táto pravdepodobnosť s rastúcou vzdialenosťou od počiatku klesať k nule. V závere paragrafu uvidíme, že podobnú podmienku spĺňajú aj vlastné funkcie zodpovedajúce ostatným vlastným hodnotám.

Zmierme sa teda s tým, že $\operatorname{Spec} M=\mathbb{N}$ a vlastné funkcie prislúchajúce k vlastným hodnotám $n\in\mathbb{N}$ tvoria postupnosť $\varphi_n=(A^*)^n\varphi_0$. Explicitný tvar kreačného operátora A^* v premennej ξ je

$$A^*=\frac{1}{\sqrt{2}}\left(\xi-\frac{\mathrm{d}}{\mathrm{d}\xi}\right).$$

Vzhľadom na „odolnosť“ exponenciálnej funkcie voči derivácii je rozumné hľadať vlastné funkcie v tvare

$$\varphi_n(\xi)=\varphi_0(\xi)\,F_n(\xi)=\mathrm{e}^{-\xi^2/2}\,F_n(\xi)$$

pre vhodné funkcie F_n. Potom

$$(A^*\varphi_n)(\xi)=\varphi_{n+1}(\xi)=\mathrm{e}^{-\xi^2/2}\,F_{n+1}(\xi)$$

a zároveň

$$
\begin{aligned}
(A^*\varphi_n)(\xi) &= \frac{1}{\sqrt{2}}\left(\xi\,\varphi_n(\xi) - \frac{\mathrm{d}\varphi_n(\xi)}{\mathrm{d}\xi}\right) \\
&= \frac{1}{\sqrt{2}}\left(\xi\,\mathrm{e}^{-\xi^2/2}\,F_n(\xi) - \frac{\mathrm{d}}{\mathrm{d}\xi}\left[\mathrm{e}^{-\xi^2/2}\,F_n(\xi)\right]\right) \\
&= \frac{1}{\sqrt{2}}\,\mathrm{e}^{-\xi^2/2}\big(2\,\xi F_n(\xi) - F_n'(\xi)\big).
\end{aligned}
$$

Pre postupnosť funkcií F_n tak dostávame rekurentný vzťah

$$
F_0(\xi) = 1, \qquad\qquad F_{n+1}(\xi) = \frac{1}{\sqrt{2}}\big(2\,\xi F_n(\xi) - F_n'(\xi)\big),
$$

z ktorého je zrejmé, že $F_n(\xi)$ je polynóm stupňa n v premennej ξ. Polynómy $F_n(\xi)$, až na multiplikatívny faktor $\sqrt{2^n}$, splývajú s nám už známymi *Hermitovými polynómami*

$$
H_n(\xi) = (-1)^n\,\mathrm{e}^{\xi^2}\,\frac{\mathrm{d}^n\,\mathrm{e}^{-\xi^2}}{\mathrm{d}\xi^n} = \sqrt{2^n}F_n(\xi)
$$

– pozri **cvičenie 13.16**. Pre $n \le 5$ máme

$$
\begin{aligned}
&H_0(\xi) = 1, && H_1(\xi) = 2\xi,\\
&H_2(\xi) = 4\xi^2 - 2, && H_3(\xi) = 8\xi^3 - 12\xi,\\
&H_4(\xi) = 16\xi^4 - 48\xi^2 + 12, && H_5(\xi) = 32\xi^5 - 160\xi^3 + 120\xi.
\end{aligned}
$$

Explicitný tvar vlastných funkcií φ_n teda je

$$
\varphi_n(\xi) = \frac{1}{\sqrt{2^n}}\,\mathrm{e}^{-\xi^2/2}\,H_n(\xi) = \frac{(-1)^n}{\sqrt{2^n}}\,\mathrm{e}^{\xi^2/2}\,\frac{\mathrm{d}^n\,\mathrm{e}^{-\xi^2}}{\mathrm{d}\xi^n}.
$$

Ak sa vrátime k pôvodnej premennej

$$
x = \sqrt{\frac{\hbar}{m\omega}}\;\xi
$$

a funkcie $\psi_n(x) = C_n\varphi_n(\xi)$ sa pokúsime znormovať tak, aby platilo

$$
\|\psi_n\|^2 = \int_{\mathbb{R}} |\psi_n(x)|^2\,\mathrm{d}x = 1,
$$

zistíme, že treba voliť

$$
C_n = \frac{1}{\sqrt{n!}}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}.
$$

Tak dostaneme normované vlastné funkcie

$$\begin{aligned}\psi_n(x) &= \frac{1}{\sqrt{n!}}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\varphi_n\left(\sqrt{\frac{m\omega}{\hbar}}\,x\right)\\ &= \frac{1}{\sqrt{2^n n!}}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\exp\left(-\frac{m\omega}{2\hbar}\,x^2\right)H_n\left(\sqrt{\frac{m\omega}{\hbar}}\,x\right)\end{aligned}$$

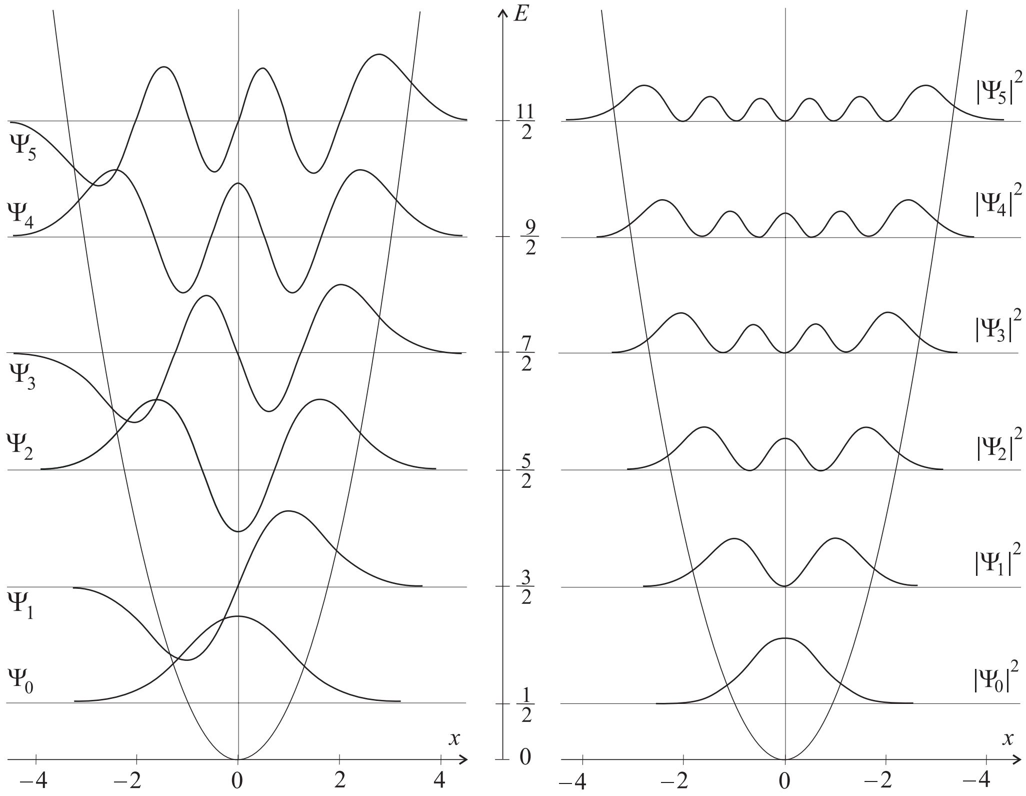
prislúchajúce k vlastným hodnotám n modifikovaného hamiltoniánu M a zároveň k vlastným hodnotám

$$E_n = \left(n+\frac{1}{2}\right)\hbar\omega$$

pôvodného hamiltoniánu H.

Energia kvantového harmonického oscilátora tak môže nadobúdať len diskrétne hodnoty E_n a jeho stav sa môže meniť iba pohltením alebo vyžiarením celočíselného násobku energetického kvanta $\hbar\omega$. Ak sa nachádza na energetickej hladine E_n, t. j. v stave ψ_n, tak pohltením energie $k\hbar\omega$ prejde na energetickú hladinu E_{n+k} do stavu ψ_{n+k}; vyžiarením energie $k\hbar\omega$, kde $k \leq n$, prejde na energetickú hladinu E_{n-k} do stavu ψ_{n-k}. Najnižšia možná energetická hladina zodpovedajúca *základnému stavu* ψ_0 je $E_0 = \frac{1}{2}\hbar\omega > 0$, a túto energiu už oscilátor nemôže vyžiariť. *Stavy* ψ_n, kde $n \geq 1$, zodpovedajúce vyšším energetickým hladinám nazývame *vzbudené* alebo *excitované*.

Obrázky vľavo znázorňujú základný stav ψ_0 a vzbudené stavy ψ_n pre $1 \leq n \leq 5$. Vpravo sú im zodpovedajúce hustoty pravdepodobnosti $|\psi_n|^2$. Všimnite si, že pravdepodobnosť základného stavu je sústredená v blízkosti rovnovážnej polohy $x = 0$, čo je prirodzené pre stav s malou energiou. Vo všeobecnosti miesta lokálnych maxím pravdepodobnosti výskytu častice splývajú s miestami lokálnych extrémov (maxím aj miním) funkcie ψ_n, kde je derivácia $\psi_n'(x)$ teda aj hybnosť $(P\psi)(x) = \frac{\hbar}{\mathrm{i}}\psi_n'(x)$ nulová. Časticu v danom stave tak pravdepodobnejšie nájdeme tam, kde je jej hybnosť (t. j. aj rýchlosť) najmenšia. S rastúcou energiou sa počet lokálnych maxím pravdepodobnosti výskytu zvyšuje vždy o jednotku a jej globálne maximá sa posúvajú do blízkosti miest zodpovedajúcich klasickým bodom obratu (tie ležia na priesečníkoch paraboly $E = \frac{1}{2}x^2$ a vodorovných priamok vyznačujúcich príslušné energetické hladiny a zároveň osi x pre jednotlivé stavy ψ_n), kde celá energia stavu oscilátora má podobu potenciálnej energie (a kinetická energia je nulová). Podľa klasickej teórie sa častica môže nachádzať len „vovnútri paraboly“, t. j. na úsečke ohraničenej klasickými bodmi obratu. V kvantovej mechanike existuje nenulová pravdepodobnosť objaviť ju aj mimo tejto úsečky, teda tam, kde sa z klasického hľadiska nemôže vyskytovať. Táto pravdepodobnosť však s rastúcou hodnotou $|x|$ rýchle klesá k nule.



Obr. 26.1. Kvantový harmonický oscilátor:
vľavo vlastné stavy, vpravo im zodpovedajúce hustoty pravdepodobnosti.

V kvantovej teórii poľa sa elektromagnetické pole matematicky popisuje ako sústava nekonečne mnohých (až kontinua) harmonických oscilátorov s rôznymi kruhovými frekvenciami ω. Presnejšie, každý stav (oscilačný mód) elektromagnetického poľa (ako aj iných bozónových polí) s daným vektorom hybnosti a polarizáciou je reprezentovaný kvantovým harmonickým oscilátorom. Kvantové číslo n (vlastné číslo príslušného modifikovaného hamiltoniánu $M = A^*A$) udáva počet fotónov v danom stave. Základný stav, tzv. *vákuum*, t. j. „nula fotónov", je superpozíciou základných stavov všetkých jednotlivých oscilátorov. Matematický model tak predpisuje energii poľa v tomto stave nekonečnú hodnotu, hoci z klasického hľadiska by jeho energia mala byť nulová. S podobnými matematickými ťažkosťami sa súčasné fyzikálne teórie vyrovnávajú zďaleka nie uspokojivým spôsobom.

Cvičenia

26.1. Nech V je unitárny priestor a $\dim V \geq 2$.

(a) Nájdite príklady hermitovských operátorov $A, B\colon V \to V$ takých, že ich komutátor $[A, B]$ nie je hermitovský operátor.

(b) Dokážte, že komutátor $[A, B]$ hermitovských operátorov $A, B\colon V \to V$ je antihermitovský operátor, teda $\mathrm{i}[A, B]$ je hermitovský operátor.

26.2. Nech $L_2(\mathbb{R})$ označuje vektorový priestor všetkých funkcií $f\colon \mathbb{R} \to \mathbb{C}$, pre ktoré má nevlastný integrál $\int_{\mathbb{R}} |f(x)|^2 \,\mathrm{d}x$ konečnú hodnotu, so skalárnym súčinom $\langle f, g\rangle = \int_{\mathbb{R}} f(x)\overline{g(x)}\,\mathrm{d}x$.

(a) Dokážte, že pre $f, g \in L_2(\mathbb{R})$ má nevlastný integrál $\int_{\mathbb{R}} f(x)\overline{g(x)}\,\mathrm{d}x$ konečnú hodnotu.

(b) Uvedomte si, že $\langle\cdot\,,\cdot\rangle$ je síce kososymetrická poldruhalineárna forma na $L_2(\mathbb{R})$, nie je však kladne definitná ale iba kladne semidefinitná. Nájdite príklady funkcií $f \in L_2(\mathbb{R})$, ktoré nie sú identicky rovné 0, jednako $\langle f, f\rangle = 0$.

(c) Dokážte, že $N = \{f \in L_2(\mathbb{R});\ \|f\| = 0\}$ je lineárny podpriestor v $L_2(\mathbb{R})$. Vysvetlite, ako možno na faktorový priestor $L_2(\mathbb{R})/N$ (pozri cvičenie 9.8) preniesť skalárny súčin z $L_2(\mathbb{R})$, a dokážte, že $L_2(\mathbb{R})/N$ je unitárny priestor. ($L_2(\mathbb{R})/N$ je dokonca Hilbertov priestor; v matematickej literatúre sa pod Hilbertovým priestorom $L_2(\mathbb{R})$ väčšinou rozumie faktorový priestor $L_2(\mathbb{R})/N$.)

26.3. Dokážte, že nasledujúce predpisy definujú samoadjungované lineárne operátory na vhodných „dosť veľkých“ lineárnych podpriestoroch priestoru $L_2(\mathbb{R})$:

(a) $X(f)(x) = xf(x)$, t. j. operátor polohy;

(b) $P(f)(x) = \mathrm{i}\frac{\mathrm{d}f(x)}{\mathrm{d}x}$, t. j. (až na multiplikatívnu konštantu) operátor hybnosti;

(c) $H(f)(x) = -\frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2} + x^2 f(x)$, t. j. (až na multiplikatívnu konštantu) operátor energie čiže hamiltonián.

26.4. Overte komutačné vzťahy pre operátory polohy a hybnosti v smeroch osí x, y, z:

(a) $[X, Y] = [X, Z] = [Y, Z] = \mathbf{0}$;

(b) $[P_x, P_y] = [P_x, P_z] = [P_y, P_z] = \mathbf{0}$;

(c) $[X, P_y] = [X, P_z] = [Y, P_x] = [Y, P_z] = [Z, P_x] = [Z, P_y] = \mathbf{0}$;

(d) $[X, P_x] = [Y, P_y] = [Z, P_z] = \mathrm{i}\hbar$.

26.5. (a) Symbolickým integrovaním (prevedením trojného integrálu na trojnásobný) overte rovnosť $\iiint_{\mathbb{R}^3} f(x, y, z)\,\delta_{\boldsymbol{r}}(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = f(a, b, c)$ pre $\boldsymbol{r} = (a, b, c)$ a ľubovoľnú „slušnú“ funkciu $f\colon \mathbb{R}^3 \to \mathbb{C}$.

(b) Odvoďte z toho zovšeobecnenú podmienku pre vektor vlastných hodnôt a vlastnú funkciu vektorového lineárneho operátora polohy $\boldsymbol{R}\,\delta_{\boldsymbol{r}} = \boldsymbol{r}\delta_{\boldsymbol{r}} = (a\delta_{\boldsymbol{r}}, b\delta_{\boldsymbol{r}}, c\delta_{\boldsymbol{r}})$.

26.6. Overte formulu pre *inverznú Fourierovu transformáciu*: Ak funkcia $F(t) = \widehat{f}(t) = \langle f, \psi_t\rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)\,\mathrm{e}^{-\mathrm{i}tx}\,\mathrm{d}x$ je Fourierovou transformáciou funkcie f vzhľadom na systém funkcií $\psi_t(x) = \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{\mathrm{i}tx}$, tak pôvodná funkcia $f(x) = \widehat{F}(-x) = \langle F, \psi_{-x}\rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} F(t)\,\mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}t$ je Fourierovou transformáciou funkcie F vzhľadom na systém funkcií $\psi_{-x}(t) = \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\mathrm{i}tx}$. (Predpokladajte, že obe funkcie f a F sú „slušné“ natoľko, že umožňujú všetky obvyklé manipulácie s príslušnými nevlastnými integrálmi.)

26.7. *Normálne* alebo tiež *Gaussovo pravdepodobnostné rozdelenie* so strednou hodnotou 0 a strednou kvadratickou odchýlkou $\sigma > 0$ je dané funkciou $\phi(x) = \frac{1}{\sigma\sqrt{2\pi}}\,\mathrm{e}^{-x^2/2\sigma^2}$, nazývanou *hustota normálneho rozdelenia.*

(a) Odvoďte formulu pre Fourierovu transformáciu hustoty takého normálneho $\widehat{\phi}(t) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \phi(x)\,\mathrm{e}^{-\mathrm{i}tx}\,\mathrm{d}x = \frac{1}{2\pi\sigma}\int_{\mathbb{R}} \exp\bigl(-\frac{x^2}{2\sigma^2} - \mathrm{i}tx\bigr)\,\mathrm{d}x = \frac{\sigma}{\sqrt{2\pi}}\,\mathrm{e}^{-\sigma^2t^2/2}$. (*Návod:* Úlohu preveďte vhodnou substitúciou na výpočet Laplaceovho integrálu $\int_{\mathbb{R}} \mathrm{e}^{-y^2}\,\mathrm{d}y = \sqrt{\pi}$ – pozri cvičenie 14.20.) Funkciu $\widehat{\phi}(t)$ nazývame tiež *charakteristickou funkciou* normálneho rozdelenia.

(b) Odvoďte z (a), že charakteristická funkcia normálneho rozdelenia so strednou hodnotou 0 a strednou kvadratickou odchýlkou σ je sama hustotou normálneho pravdepodobnostného rozdelenia s rovnakou strednou hodnotou 0 a strednou kvadratickou odchýlkou $\widehat{\sigma} = 1/\sigma$. Pre stredné kvadratické odchýlky normálnych rozdelení $\phi(x)$ a $\widehat{\phi}(t)$ teda platí $\sigma\widehat{\sigma} = 1$ (čo možno tiež interpretovať ako istý prípad Heisenbergovho vzťahu neurčitosti).

26.8. (a) Overte komutačné vzťahy pre operátory momentov hybnosti vzhľadom na osi x, y, z: $[L_x, L_y] = \mathrm{i}\hbar L_z$, $[L_y, L_z] = \mathrm{i}\hbar L_x$, $[L_z, L_x] = \mathrm{i}\hbar L_y$.

(b) Označme $L^2 = L_x^2 + L_y^2 + L_z^2$ operátor štvorca celkovej hybnosti. Overte komutačné vzťahy $\bigl[L_x, L^2\bigr] = \bigl[L_y, L^2\bigr] = \bigl[L_z, L^2\bigr] = \mathbf{0}$.

26.9. Zaveďme v $\mathbb{R}^3$ sférické súradnice $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$ (všimnite si tradovanú malú redakčnú úpravu oproti sférickým súradniciam z paragrafu 14.4). Uhly θ, ϕ sa nazývajú *polárny uhol* resp. *azimut.*

(a) Overte vyjadrenia operátorov momentov hybnosti vzhľadom na sférické súradnice: $L_x = \mathrm{i}\hbar\bigl(\sin\phi\,\frac{\partial}{\partial\theta} + \operatorname{ctg}\theta\cos\phi\,\frac{\partial}{\partial\phi}\bigr)$, $L_y = -\mathrm{i}\hbar\bigl(\cos\phi\,\frac{\partial}{\partial\theta} - \operatorname{ctg}\theta\sin\phi\,\frac{\partial}{\partial\phi}\bigr)$, $L_z = -\mathrm{i}\hbar\,\frac{\partial}{\partial\phi}$.

(b) Overte vyjadrenie operátora štvorca celkového momentu hybnosti vo sférických súradniciach: $L^2 = -\frac{\hbar^2}{\sin\theta}\Bigl(\frac{1}{\sin\theta}\frac{\partial^2}{\partial\phi^2} + \frac{\partial}{\partial\theta}\bigl(\sin\theta\,\frac{\partial}{\partial\theta}\bigr)\Bigr)$.

(c) Pokúste sa názorne geometricky zdôvodniť, prečo uvedené operátory závisia len na uhlových súradniciach θ, ϕ a nezávisia na dĺžkovej súradnici r.

26.10. Zúženie kreačného operátora A^* na vlastný podpriestor $\operatorname{Ker}(M - n)$ modifikovaného hamiltoniánu M je lineárny izomorfizmus tohto podpriestoru na vlastný podpriestor $\operatorname{Ker}(M - (n+1))$ operátora M. Dokážte. Akú medzeru v úvahách paragrafu 26.10 o kvantovom harmonickom oscilátore zapĺňa toto pozorovanie?

26.11. (a) Overte, že funkcie $\varphi_0(x) = \mathrm{e}^{-x^2/2}$ a $\varphi_0^+(x) = \mathrm{e}^{-x^2/2}\int_0^x \mathrm{e}^{u^2}\,\mathrm{d}u$ sú lineárne nezávislé a vyhovujú diferenciálnej rovnici $\frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2} = (x^2 - 1)\varphi$, teda tvoria jej fundamentálny systém riešení.

(b) Riešte uvedenú diferenciálnu rovnicu rozvojom funkcie φ do mocninného radu $\varphi(x) = \sum_{n=0}^{\infty} \frac{a_n}{n!}\,x^n$. (Na základe diferenciálnej rovnice odvoďte najprv rekurentný a potom explicitný vzťah pre koeficienty a_n.)

(c) Pre aké hodnoty počiatočných koeficientov a_0, a_1 dostávame riešenia φ_0 resp. φ_0^+?

26.12. (a) Ukážte, že lokálne extrémy vlastnej stavovej funkcie ψ_n kvantového harmonického oscilátora sú v takých bodoch $x \in \mathbb{R}$, kde $P\psi_n(x) = 0$.

(b) Vysvetlite, prečo hustota pravdepodobnosti výskytu častice v danom stave ψ_n

nadobúda lokálne maximá v tých istých bodoch $x \in \mathbb{R}$ ako v (a).

(c) V akých bodoch $x \in \mathbb{R}$ nadobúda hustota pravdepodobnosti výskytu častice v danom stave ψ_n svoje lokálne minimá?

26.13. (a) Zmodifikujte definície *cyklického podpriestoru* a *cyklického generátora* lineárneho operátora z paragrafu 21.3 tak, aby dávali zmysel pre každý lineárny operátor $T\colon V \to V$ na (nie nutne konečnorozmernom) vektorovom priestore V, aj keď je celá nekonečná postupnosť $\boldsymbol{x}, T(\boldsymbol{x}), T^2(\boldsymbol{x}), \dots, T^n(\boldsymbol{x}), \dots$ iterovaných obrazov vektora $\boldsymbol{x} \in V$ lineárne nezávislá.

(b) Presvedčte sa, že vlastná funkcia $\psi(x) = \exp\bigl(-\frac{m\omega}{2\hbar}x^2\bigr)$ oboch hamiltoniánov M a H je cyklickým generátorom cyklického podpriestoru kreačného operátora A^*. (Pokročilejšími metódami možno ukázať, že tento podpriestor je hustý v $L_2(\mathbb{R})$.)

Časť IV

Grupy a algebry

27. Úvod do teórie grúp

Pojem grupy hrá natoľko kľúčovú úlohu nielen v algebre, ale v celej modernej matematike a jej početných aplikáciách, že ani v našom kurze by už nebolo ďalej únosné sa mu vyhýbať. Systematické štúdium teórie grúp však nie je predmetom lineárnej algebry, teda ani tohto kurzu. Preto sa sústredíme hlavne na jej základné pojmy a výsledky v miere, ktorá nám umožní ich využitie pri hlbšom objasnení algebraickej a geometrickej štruktúry vektorových priestorov, bilineárnych a kvadratických foriem a lineárnych transformácií. Potom sa trochu podrobnejšie pozrieme na niektoré maticové grupy. V záujme ucelenosti výkladu a na ilustráciu niektorých pojmov, výsledkov a metód však neraz neodoláme pokušeniu povedať toho o grupách trochu viac, než je z hľadiska uvedených zámerov naozaj nevyhnutné.

Pri analýze akéhokoľvek (nielen matematického) štruktúrovaného oboru objektov hrá dôležitú úlohu otázka jeho *symetrie*. Znalosť transformácií, ktoré zachovávajú príslušnú štruktúru (t. j. jej symetrií), nám totiž neraz umožňuje výrazne sprehľadniť a zjednodušiť jej popis. Čím je takýchto transformácií viac, tým symetrickejšiu štruktúru nesie spomínaný obor, malé množstvo takýchto transformácií naopak svedčí o nízkom stupni symetrie.

Ukazuje sa, že (bijektívne) transformácie zachovávajúce danú štruktúru tvoria vždy *grupu*, t. j. množinu transformácií uzavretú vzhľadom na kompozíciu, obsahujúcu identickú transformáciu a spolu s každou transformáciou aj transformáciu k nej inverznú. Typickými príkladmi takýchto grúp sú kryštalografické grupy, alebo grupy transformácií euklidovských priestorov zachovávajúcich rôzne invarianty, ako napr. dĺžku, objem či uhol. Grupa transformácií daného štruktúrovaného oboru však v sebe nesie podstatne viac informácií o jeho symetrii, než len to, či je ich „veľa“ alebo „málo“. I „rovnako veľké“ grupy sa totiž môžu výrazne líšiť svojou vlastnou vnútornou štruktúrou, a tým spätne mnoho vypovedať o symetrii a štruktúre pôvodných oborov.

V tejto kapitole sa stručne oboznámime len s celkom základnými pojmami a výsledkami teórie grúp. V duchu modernej algebry ich však budeme študovať v abstraktnom poňatí, t. j. bez toho, aby sme predpokladali, že ide nutne o grupy transformácií. Tým sa budeme podrobnejšie venovať až v nasledujúcej kapitole.

27.1 Abstraktný pojem grupy

Grupou nazývame množinu G vybavenú binárnou operáciou $\cdot : G \times G \to G$, ktorá spĺňa nasledujúce podmienky, nazývané tiež *axiómami teórie grúp*:

(a) $(\forall\, a,b,c \in G)\big(a(bc) = (ab)c\big)$,
t. j. operácia $\cdot$ je asociatívna;

(b) $(\exists\, e \in G)(\forall\, a \in G)\big(ae = ea = a \ \&\ (\exists\, b \in G)(ab = ba = e)\big)$,
t. j. existuje neutrálny prvok $e \in G$ operácie $\cdot$ a ku každému $a \in G$ existuje inverzný prvok $b \in G$ vzhľadom na operáciu $\cdot$.

Ako už vieme z paragrafu 0.4, neutrálny prvok $e \in G$ je prvou časťou podmienky (b) určený jednoznačne; podobne je jej druhou časťou jednoznačne určený inverzný prvok $b \in G$ k danému $a \in G$.

Hovoríme, že *grupa* G je *komutatívna*, alebo tiež *abelovská*, ak operácia $\cdot$ je komutatívna, t. j. platí $ab = ba$ pre všetky $a, b \in G$.

Uvedený spôsob zápisu, pri ktorom grupovú operáciu značíme $\cdot$ (a jej znak väčšinou vynechávame), prípadne $\circ$, nazývame *multiplikatívny zápis*. Grupovú operáciu vtedy nazývame *súčinom* alebo *násobením*, prípadne *skladaním* alebo *kompozíciou*. Neutrálny prvok nazývame tiež *jednotkovým prvkom* alebo *jednotkou*, prípadne *identitou* a značíme ho väčšinou e, ε alebo 1, prípadne $\boldsymbol{I}$, id, ι a pod. Inverzný prvok k prvku $a \in G$ značíme a^{-1}, občas tiež a' alebo $\bar{a}$. Zrejmým spôsobom (porovnaj s paragrafom 1.2) zavádzame výrazy a^n pre $a \in G$, $n \in \mathbb{Z}$.

Grupovú operáciu, neutrálny prvok resp. inverzný prvok k danému môžeme, samozrejme, označiť hocako. Popri multiplikatívnom zápise sa však bežne používa už len tzv. *aditívny zápis*, pri ktorom grupovú operáciu značíme $+$ a nazývame *sčítaním*, neutrálny prvok značíme 0 a nazývame *nulou* alebo *nulovým prvkom* a inverzný prvok značíme $-a$ a nazývame *opačným prvkom* k prvku $a \in G$. Čitateľ by si mal samostane premyslieť, ako sa zmení formulácia grupových axióm (a), (b) pri prechode k aditívnemu zápisu. Výrazy $a - b$, na, pre $a, b \in G$, $n \in \mathbb{Z}$, zavádzame obdobne ako v paragrafe 1.2.

Aditívny zápis je rezervovaný takmer výlučne pre abelovské grupy. To neznamená, že by sme sa s komutatívnou grupou nemohli stretnúť v multiplikatívnom zápise. No uvedením nejakej grupy v aditívnom zápise už vlastne dávame najavo (pokiaľ výslovne nezdôrazníme opak), že ide o abelovskú grupu.

Ako sme už naznačili, grupu väčšinou označujeme rovnakým znakom ako jej základnú množinu. Niekedy je však účelné zahrnúť do označenia grupy i príslušnú binárnu operáciu, prípadne aj jej neutrálny prvok, či tiež unárnu operáciu inverzného prvku; vtedy hovoríme napr. o grupe $(G, \cdot)$, grupe $(A, +, 0)$, grupe $\big(H, \cdot, e, {}^{-1}\big)$, a pod.

Hovoríme, že *grupa* G je *konečná* resp. *nekonečná*, ak jej základná množina má príslušnú vlastnosť. *Rádom konečnej grupy* nazývame počet jej prvkov.

S niektorými jednoduchými príkladmi grúp sme sa už v našom kurze stretli.

27.1.1. Príklad. Množina $\mathbb{Z}$ všetkých celých čísel tvorí grupu vzhľadom na operáciu sčítania. Podobne, pre $n \geq 1$, tvorí grupu množina $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ všetkých zvyškových tried s operáciou sčítania modulo n (pozri paragraf 1.3). Zrejme $(\mathbb{Z}, +)$ aj všetky $(\mathbb{Z}_n, +)$ sú napospol abelovské grupy.

27.1.2. Príklad. Každé pole K určuje hneď dve komutatívne grupy. Je to jednak aditívna grupa $(K, +, 0)$, jednak multiplikatívna grupa $(K \smallsetminus \{0\}, \cdot, 1)$ jeho nenulových prvkov, ktorú zvykneme tiež značiť $(K^\star, \cdot, 1)$ alebo len krátko $K^\star$. V prípade polí $\mathbb{Q}$ a $\mathbb{R}$ sa k nim pridružujú ešte multiplikatívne grupy $(\mathbb{Q}^+, \cdot, 1)$ resp. $(\mathbb{R}^+, \cdot, 1)$ kladných prvkov daného poľa.

Podobne určuje každý vektorový priestor V nad ľubovoľným poľom K abelovskú grupu $(V, +, \mathbf{0})$.

27.1.3. Príklad. Z úvah vykonaných v paragrafe 0.5 vyplýva, že množina $\mathcal{S}(X)$ všetkých permutácií ľubovoľnej množiny X tvorí grupu vzhľadom na operáciu $\circ$ skladania zobrazení, s jednotkou id_X. Táto grupa je pre $\# X \geq 3$ nekomutatívna. Pre konečnú množinu $X = \{1, \ldots, n\}$ nazývame grupu $\mathcal{S}(X) = \mathcal{S}_n$ *symetrickou grupou stupňa* n; jej rád je zrejme $n!$.

27.1.4. Príklad. Množinu všetkých regulárnych matíc rozmeru $n \times n$ nad poľom K budeme odteraz značiť $\mathrm{GL}(n, K)$. Z výsledkov paragrafu 7.2 vyplýva, že $\mathrm{GL}(n, K)$ tvorí grupu vzhľadom na operáciu násobenia matíc, s jednotkou $\boldsymbol{I}_n$; nazývame ju *všeobecná lineárna grupa (stupňa n nad poľom K)* (GL je skratka anglického *general linear*). Pre $n \geq 2$ je $\mathrm{GL}(n, K)$ nekomutatívna grupa.

Poznámka. Pri pohľade na pred chvíľou uvedenú definíciu a za ňou nasledujúce dôverne známe príklady hĺbavejšieho čitateľa asi nevdojak napadne otázka, prečo sme s definíciou grupy tak dlho otáľali. Pritom je to definícia – najmä v porovnaní s definíciami poľa a vektorového priestoru (pozri paragrafy 1.2 a 1.5) – veľmi jednoduchá. Vlastne už v paragrafe 0.4 sme mali pohromade všetky pojmy potrebné nato, aby sme ju mohli vysloviť. Navyše, keby sme vtedy boli tak učinili, mohli sme trochu neskôr definície poľa a vektorového priestoru sformulovať podstatne kratšie a jednoduchšie.

Napr. v definícii poľa (pozri paragraf 1.2) možno prvé štyri formuly ľavého stĺpca zhrnúť do podmienky, že množina K tvorí vzhľadom na sčítanie $+$ komutatívnu grupu s nulovým prvkom 0, a celý ľavý stĺpec zasa do podmienky, že množina $K^\star = K \smallsetminus \{0\}$ tvorí komutatívnu grupu vzhľadom na násobenie s jednotkovým prvkom 1 (potom nutne $1 \in K^\star$, teda $0 \neq 1$). Zostáva už len jediná formula – distributívny zákon –, ktorá dáva do súvisu obe operácie. S istou dávkou zjednodušenia možno povedať, že pole pozostáva z dvoch komutatívnych grúp spojených distributívnym zákonom.

Podobne možno prvé štyri formuly v definícii vektorového priestoru (pozri paragraf 1.5) nahradiť požiadavkou, že $(V, +, \mathbf{0})$ je abelovská grupa.

No takýto prístup by popri spomínanom zisku mohol zároveň u čitateľa vyvolať mylný dojem, že pojem grupy je len verbálnou skratkou, slúžiacou na zjednodušenie niektorých zložitejších definícií iných algebraických pojmov. Navyše, pokiaľ by sme nechceli neorganicky odbočovať od témy, prípadne na prítomnosť grúp v našom výklade umelo upozorňovať, boli by sme obmedzení v podstate na grupy uvedené v príkladoch 27.1.1–4, v ktorých prevládajú abelovské grupy. Takéto obmedzenie by však zastieralo viaceré podstatné znaky sveta grúp, v ktorom naopak prevládajú grupy neabelovské. Práve stručnosť a jednoduchosť definície grupy má totiž za následok, že jej vyhovuje obrovské množstvo nesmierne rozmanitých matematických objektov, a tým aj prekvapivú zložitosť možnej štruktúry grúp. Abelovské grupy, a obzvlášť vektorové priestory patria práve k tým štruktúrne najjednoduchším predstaviteľom grúp.

V nasledujúcom tvrdení, ktorého dôkaz prenechávame ako jednoduché cvičenie čitateľovi, je zhrnutých niekoľko najelementárnejších pravidiel pre počítanie v grupách.

27.1.5. Tvrdenie. *Nech $(G, \cdot, e)$ je grupa. Potom pre ľubovoľné prvky $a, b, c \in G$ a $m, n \in \mathbb{Z}$ platí*

$$\begin{aligned} e^{-1} &= e, & \left(a^{-1}\right)^{-1} &= a, & (ab)^{-1} &= b^{-1}a^{-1}, \\ a^0 &= e, & a^{m+n} &= a^m a^n, & a^{mn} &= (a^m)^n, \end{aligned}$$

v G sú splnené pravidlá o krátení zľava aj sprava, t. j.

$$ab = ac \ \Rightarrow\ b = c, \qquad ac = bc \ \Rightarrow\ a = b,$$

a každá z rovníc $ax = b$, resp. $ya = b$ má v G jediné riešenie $x = a^{-1}b$, resp. $y = ba^{-1}$.

27.2 Podgrupy, generujúce množiny, cyklické grupy

Nech $(G, \cdot, e)$ je grupa. Hovoríme, že podmnožina $S \subseteq G$ je *podgrupa* grupy G, ak $e \in S$ a pre ľubovoľné $a, b \in S$ platí $ab \in S$ aj $a^{-1} \in S$. Inak povedané, podgrupa grupy G je jej podmnožina, ktorá obsahuje neutrálny prvok a je uzavretá vzhľadom na operácie súčinu a inverzného prvku v G. Zrejme každá podgrupa grupy G je zároveň sama grupou vzhľadom na grupovú operáciu zdedenú z G.

Pri overovaní, či daná pomnožina grupy je jej podgrupou, býva niekedy užitočné nasledujúce tvrdenie.

27.2.1. Tvrdenie. *Nech $(G, \cdot, e)$ je grupa a $S \subseteq G$. Potom S je podgrupa grupy G práve vtedy, keď $S \neq \emptyset$ a pre každé $a, b \in S$ platí $ab^{-1} \in S$.*

Dôkaz. Zrejme každá podgrupa grupy G je neprázdna a uzavretá vzhľadom na operáciu $(a, b) \mapsto ab^{-1}$. Naopak, nech $\emptyset \neq S \subseteq G$ je uzavretá na uvedenú operáciu a $s \in S$ je ľubovoľný prvok. Potom $e = ss^{-1} \in S$. Ďalej pre $a, b \in S$ platí $b^{-1} = eb^{-1} \in S$ a taktiež $ab = a(b^{-1})^{-1} \in S$. Teda S je podgrupa grupy G.

Pojem podgrupy danej grupy má niektoré spoločné črty s pojmom lineárneho podpriestoru daného vektorového priestoru: v oboch prípadoch ide o neprázdnu podmnožinu uzavretú vzhľadom na príslušné operácie. Navyše, lineárny podpriestor S vektorového priestoru V je zároveň podgrupou grupy $(V, +, \mathbf{0})$. Naopak, podgrupa S (aditívnej grupy) vektorového priestoru V je jeho lineárnym podpriestorom práve vtedy, keď je uzavretá aj na skalárne násobky.

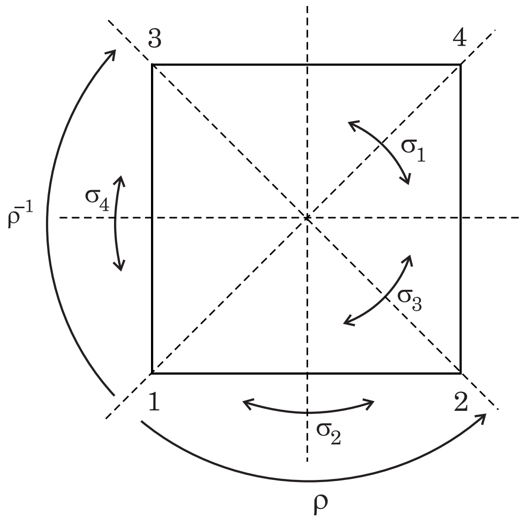
Každá grupa G obsahuje tzv. *triviálnu podgrupu* $\{e\}$ a *nevlastnú podgrupu* G. Pritom, okrem prípadu $G = \{e\}$, ide zrejme o dve rôzne podgrupy. V nasledujúcich príkladoch si ukážeme niekoľko dôležitých typov netriviálnych vlastných podgrúp, čím zároveň trochu rozšírime našu zatiaľ skromnú zbierku grúp.

27.2.2. Príklad. Nech $n \geq 3$ je prirodzené číslo. Označme ι identickú permutáciu množiny $\{1, \dots, n\}$ a ϱ cyklickú permutáciu $1 \mapsto 2 \mapsto \dots \mapsto n{-}1 \mapsto n \mapsto 1$. Potom permutácie $\iota = \varrho^0$, $\varrho = \varrho^1$, ..., ϱ^k, ..., ϱ^{n-1} predstavujú otočenia (proti smeru hodinových ručičiek) pravidelného n-uholníka s vrcholmi $1, \dots, n$ o uhly $2k\pi/n$ pre $0 \leq k \leq n-1$. Ak $n = 2m$ je párne, tak pre $1 \leq k \leq m$ označme σ_{2k-1} permutáciu zodpovedajúcu súmernosti n-uholníka podľa osi spájajúcej vrcholy k, $m{+}k$ a σ_{2k} permutáciu zodpovedajúcu súmernosti podľa osi strany spájajúcej vrcholy k, $k+1$ (a protiľahlej strany). Ak n je nepárne, tak σ_k pre $1 \leq k \leq n$ označuje permutáciu zodpovedajúcu osovej súmernosti podľa spojnice vrchola k so stredom protiľahlej strany n-uholníka.

Prípad nepárneho $n = 3$ je znázornený na obrázku v paragrafe 0.5 (kde zrejme $\varrho^{-1} = \varrho^2$). Obrázok na dalšej strane ukazuje situáciu pre párne $n = 4$ (tentokrát $\varrho^{-1} = \varrho^3$).

Pre každé $n \geq 3$ tvorí množina $\{\iota, \varrho, \dots, \varrho^{n-1}, \sigma_1, \dots, \sigma_n\}$ podgrupu symetrickej grupy $\mathcal{S}_n$. Označujeme ju Δ_n a nazývame *grupou symetrií pravidelného n-uholníka* alebo tiež *dihedrálnou grupou stupňa n*.[1] Dihedrálna grupa Δ_n má rád $2n$ a – okrem prípadu $n = 3$, kedy $\Delta_3 = \mathcal{S}_3$, – je to netriválna vlastná podgrupa symetrickej grupy $\mathcal{S}_n$.

[1]Používa sa tiež označenie D_n prípadne D_{2n}.



Obr. 27.1. Symetrie štvorca

27.2.3. Príklad. Z vety 0.5.1 vyplýva, že množina všetkých párnych permutácií množiny $\{1, \dots, n\}$ tvorí podgrupu symetrickej grupy $\mathcal{S}_n$. Hovoríme jej *alternujúca grupa stupňa* n a značíme ju $\mathcal{A}_n$. Zrejme $\mathcal{A}_0 = \mathcal{S}_0$, $\mathcal{A}_1 = \mathcal{S}_1$, no pre $n \geq 2$ má $\mathcal{A}_n$ rád $n!/2$ (rozmyslite si prečo), teda je to vlastná (a pre $n \geq 3$ tiež netriviálna) podgrupa symetrickej grupy $\mathcal{S}_n$.

27.2.4. Príklad. Ako sme už spomínali, množina $\mathbb{C}^\star = \mathbb{C} \smallsetminus \{0\}$ tvorí grupu vzhľadom na násobenie. Z vlastností komplexných čísel možno ľahko nahliadnuť, že množina $\{z \in \mathbb{C};\ |z| = 1\}$ je jej podgrupou. Hovoríme jej *grupa komplexných jednotiek* (nepliesť si s jednotkou ako neutrálnym prvkom grupy) a z dôvodov, ktoré vysvitnú neskôr, ju značíme $\mathrm{U}(1)$. Zrejme $\mathrm{U}(1)$ je nekonečná vlastná podgrupa grupy $(\mathbb{C}^\star, \cdot)$.

Podobne ako vo vektorovom priestore V generuje ľubovoľná množina $X \subseteq V$ lineárny podpriestor $[X]$, každá podmnožina X grupy G generuje istú podgrupu grupy G, ktorú teraz opíšeme. Pre $X \subseteq G$ označme

$$\langle X \rangle = \left\{ x_1^{k_1} \dots x_n^{k_n};\ n \in \mathbb{N}\ \&\ k_1, \dots, k_n \in \mathbb{Z}\ \&\ x_1, \dots, x_n \in X \right\}.$$

Ak $X = \{x_1, \dots, x_m\}$ je konečná množina, tak miesto $\langle \{x_1, \dots, x_m\} \rangle$ píšeme len $\langle x_1, \dots, x_m \rangle$. Množinu $\langle X \rangle$ nazývame *podgrupa generovaná množinou* X. Tento názov je oprávnený nasledujúcim tvrdením.

27.2.5. Tvrdenie. *Nech X je podmnožina grupy G. Potom množina $\langle X \rangle$ je najmenšia podgrupa grupy G taká, že $X \subseteq \langle X \rangle$.*

Náčrt dôkazu. Podobne ako v tvrdení 4.2.1 pre vektorové priestory, i teraz možno ľahko dokázať, že

(a) $\langle X\rangle$ je podgrupa grupy G;
(b) pre každú podgrupu $S \subseteq G$ platí $X \subseteq S \Rightarrow \langle X\rangle \subseteq S$.

Ak $\langle X\rangle = G$, hovoríme, že *množina X generuje grupu G*, alebo, že X je *množinou generátorov grupy G*. Prvky množiny X potom nazývame *generátory grupy G*. Grupa G sa nazýva *konečne generovaná*, ak G má nejakú konečnú množinu generátorov.

Štruktúrne najjednoduchšími grupami sú tzv. *cyklické grupy*, t. j. grupy generované jediným generátorom. Príkladom cyklickej grupy je grupa $(\mathbb{Z}, +)$ všetkých celých čísel, a taktiež grupy $(\mathbb{Z}_n, +)$ pre každé kladné celé číslo n. Každá z týchto grúp (okrem prípadu $n = 1$, kedy však $\mathbb{Z}_1 = \{0\} = \langle 0\rangle$) je totiž generovaná svojim prvkom 1.

Každý prvok x grupy G v nej generuje cyklickú podgrupu $\langle x\rangle$. Ak je konečná, tak jej rád $\#\langle x\rangle$ nazývame *rádom prvku x v G*; ak je nekonečná, hovoríme, že prvok $x \in G$ má *nekonečný rád*.

Nasledujúce tvrdenie by sme mohli dostať ako jednoduchý dôsledok našich neskorších úvah. Zámerne ho však dokážeme celkom elementárnymi, no o to názornejšími prostriedkami.

27.2.6. Tvrdenie. *Prvok x grupy G má konečný rád práve vtedy, keď existuje kladné celé číslo r také, že $x^r = e$. Potom rádom prvku x je najmenšie kladné celé číslo r s touto vlastnosťou a $\langle x\rangle = \{e, x, \dots, x^{r-1}\}$. Ak x má nekonečný rád, tak $\langle x\rangle = \{x^n;\, n \in \mathbb{Z}\}$, pričom $x^m \neq x^n$ pre všetky celé čísla $m \neq n$.*

Dôkaz. Vezmime $x \in G$ a uvažujme postupnosť mocnín $x^1 = x$, $x^2 = xx$, $x^3 = xxx$, atď. Všetky jej členy patria do podgrupy $\langle x\rangle$.

Pokiaľ x má konečný rád, t. j. $\langle x\rangle$ je konečná, musia sa v tejto postupnosti vyskytnúť aspoň dva rovnaké členy, napr. x^k a x^{k+r}, kde k a r sú kladné celé čísla. Potom však $x^k e = x^k = x^{k+r} = x^k x^r$, z čoho krátením zľava dostávame $e = x^r$. Ak r je najmenšie kladné celé číslo s touto vlastnosťou, tak všetky prvky $x^0 = e$, $x^1 = x$, …, x^{r-1} sú navzájom rôzne (rozmyslite si, prečo), a ďalšie mocniny sa už cyklicky opakujú: $x^r = e$, $x^{r+1} = x$, …, $x^{2r-1} = x^{r-1}$, atď. Ak si ešte uvedomíme, že pre $1 \leq k \leq r-1$ potom platí $\left(x^k\right)^{-1} = x^{r-k}$, je jasné, že $\langle x\rangle = \{x^k;\, 0 \leq k \leq r-1\}$.

Z prvej časti dôkazu je zrejmé, že ak $x \in G$ má nekonečný rád, tak všetky prvky uvažovanej nekonečnej postupnosti $x, x^2, \dots, x^n, \dots$ sú navzájom rôzne. Potom sa v nej nemôže vyskytnúť neutrálny prvok e. Navyše, žiadne dva z nich nemôžu byť navzájom inverzné. Teda $x^m \neq x^n$ pre navzájom rôzne $m, n \in \mathbb{Z}$ a $\langle x\rangle = \{x^n;\, n \in \mathbb{Z}\}$.

27.3 Homomorfizmy a izomorfizmy

Nech $(G,\cdot)$, $(H,\cdot)$ sú grupy. Zobrazenie $\varphi\colon G\to H$ sa nazýva *homomorfizmus grúp*, ak pre všetky $a,b\in G$ platí

$$\varphi(ab)=\varphi(a)\cdot\varphi(b).$$

Inak povedané, homomorfizmus je zobrazenie, ktoré zachováva operáciu súčinu v grupách. Nasledujúce tvrdenie ukazuje, že grupový homomorfizmus už nevyhnutne zachováva aj jednotku a operáciu inverzného prvku.

27.3.1. Tvrdenie. *Nech $\varphi\colon G\to H$ je homomorfizmus grúp. Potom $\varphi(e_G)=e_H$ a pre každé $a\in G$ platí $\varphi\big(a^{-1}\big)=\varphi(a)^{-1}$.*

Dôkaz. Čitateľ asi sám prišiel na to, že kvôli rozlíšeniu jednotiek v grupách G a H sme ich pre potreby tohto tvrdenia a jeho dôkazu označili e_G resp. e_H. S využitím vlastnosti homomorfizmu dostávame

$$\varphi(e_G)\cdot e_H=\varphi(e_G)=\varphi(e_G\cdot e_G)=\varphi(e_G)\cdot\varphi(e_G),$$

z čoho po krátení zľava vyplýva $e_H=\varphi(e_G)$.

Na dôkaz rovnosti $\varphi\big(a^{-1}\big)=\varphi(a)^{-1}$ stačí overiť, že $\varphi\big(a^{-1}\big)$ sa správa ako inverzný prvok k prvku $\varphi(a)$, t. j. platí $\varphi(a)\cdot\varphi\big(a^{-1}\big)=e_H$. Vďaka vlastnosti homomorfizmu a už dokázanej prvej rovnosti nám vyjde

$$\varphi(a)\cdot\varphi\big(a^{-1}\big)=\varphi\big(aa^{-1}\big)=\varphi(e_G)=e_H.$$

Pre ľubovoľné dve grupy G, H je konštantné zobrazenie $x\mapsto e_H$, ktoré každému prvku $x\in G$ priradí jednotkový prvok grupy H, homomorfizmus grúp; nazývame ho *triválny homomorfizmus*. Čoskoro sa budeme mať možnosť zoznámiť aj s netriválnymi homomorfizmami.

Opäť sa stretáme so zrejmou analógiou spájajúcou pojmy lineárneho zobrazenia medzi vektorovými priestormi a homomorfizmu grúp: sú to zobrazenia medzi ich základnými množinami, ktoré zachovávajú príslušné operácie. Navyše, lineárne zobrazenie $\varphi\colon V\to U$ je zároveň homomorfizmus grúp $\varphi\colon (V,+)\to(U,+)$. Naopak, homomorfizmus $\varphi\colon (V,+)\to(U,+)$ aditívnych grúp vektorových priestorov V a U je lineárnym zobrazením práve vtedy, keď φ zachováva aj skalárne násobky.

Z toho dôvodu nie je potrebné uvádzať dôkazy zostávajúcich tvrdení tohto paragrafu. Čitateľ by si však mal samostatne premyslieť, ako ich dostane malými obmenami dôkazov príslušných tvrdení o lineárnych zobrazeniach, ktorých čísla mu na uľahčenie zakaždým uvedieme.

27.3.2. Tvrdenie. *Nech* $\psi\colon F \to G$, $\varphi\colon G \to H$ *sú homomorfizmy grúp. Potom aj ich kompozícia* $\varphi \circ \psi\colon F \to H$ *je homomorfizmus grúp.*

Dôkaz. Pozri tvrdenie 6.1.2.

27.3.3. Tvrdenie. *Nech* $\varphi\colon G \to H$ *je homomorfizmus grúp,* $S \subseteq G$ *je podgrupa grupy* G *a* $T \subseteq H$ *je podgrupa grupy* H*. Potom aj* $\varphi(S) \subseteq H$ *je podgrupa grupy* H *a* $\varphi^{-1}(T) \subseteq G$ *je podgrupa grupy* G*.*

Dôkaz. Pozri tvrdenie 6.1.3.

Nech $\varphi\colon G \to H$ je homomorfizmus grúp. Jeho *jadrom* resp. *obrazom* nazývame množinu

$$\operatorname{Ker}\varphi = \varphi^{-1}\{e_H\} = \{x \in G;\, \varphi(x) = e_H\},$$

resp.

$$\operatorname{Im}\varphi = \varphi(G) = \{\varphi(x);\, x \in G\}.$$

Ako bezprostredný dôsledok tvrdenia 27.3.3 (pozri tiež tvrdenie 6.2.1) dostávame

27.3.4. Tvrdenie. *Nech* $\varphi\colon G \to H$ *je homomorfizmus grúp. Potom* $\operatorname{Ker}\varphi$ *je podgrupa grupy* G *a* $\operatorname{Im}\varphi$ *je podgrupa grupy* H*.*

Podobne ako vo vete 6.2.2 pre lineárne zobrazenia, možno pomocou jadra a obrazu charakterizovať aj injektívnosť resp. surjektívnosť grupových homomorfizmov.

27.3.5. Tvrdenie. *Nech* $\varphi\colon G \to H$ *je homomorfizmus grúp. Potom*
(a) φ *je injektívny práve vtedy, keď* $\operatorname{Ker}\varphi = \{e_G\}$*;*
(b) φ *je surjektívny práve vtedy, keď* $\operatorname{Im}\varphi = H$*.*

Bijektívny homomorfizmus grúp $\varphi\colon G \to H$ nazývame *izomorfizmus grúp*. Hovoríme, že *grupy* G, H sú *izomorfné*, označenie $G \cong H$, ak existuje nejaký izomorfizmus $\varphi\colon G \to H$.

Zrejme injektívny homomorfizmus $\varphi\colon G \to H$ je zároveň izomorfizmom grupy G na podgrupu $\operatorname{Im}\varphi$ grupy H.

Aj pre grupové izomorfizmy platí obdoba tvrdenia 6.3.1 pre lineárne izomorfizmy.

27.3.6. Tvrdenie. *Nech* F, G, H *sú grupy.*
(a) $\mathrm{id}_G\colon G \to G$ *je izomorfizmus grúp.*
(b) Ak $\varphi\colon G \to H$ *je izomorfizmus grúp, tak aj inverzné zobrazenie* $\varphi^{-1}\colon H \to G$ *je izomorfizmus grúp.*

(c) *Ak $\psi\colon F \to G$, $\varphi\colon G \to H$ sú izomorfizmy grúp, tak aj $\varphi \circ \psi\colon F \to H$ je izomorfizmus grúp.*

V dôsledku toho pre ľubovoľné grupy F, G, H platí

$$\begin{gathered} G \cong G, \\ G \cong H \Rightarrow H \cong G, \\ F \cong G \;\&\; G \cong H \Rightarrow F \cong H. \end{gathered}$$

Inak povedané, vzťah $\cong$ izomorfnosti grúp je *reflexívny*, *symetrický* a *tranzitívny*, teda je to vzťah *ekvivalencie.* Izomorfné grupy môžeme z hľadiska ich štruktúry považovať za totožné.

27.3.7. Príklad. Nech $(G, \cdot)$ je ľubovoľná grupa, $a \in G$. Keďže pre všetky $m, n \in \mathbb{Z}$ platí $a^{m+n} = a^m a^n$, znamená to, že predpisom $\varphi(n) = a^n$ je definovaný homomorfizmus $\varphi\colon (\mathbb{Z}, +) \to (G, \cdot)$. Zrejme $\operatorname{Im}\varphi = \langle a \rangle$ je cyklická podgrupa grupy G generovaná prvkom a. Ak a má konečný rád r, tak $\operatorname{Ker}\varphi = r\mathbb{Z} = \{rn;\, n \in \mathbb{Z}\}$; ak x má nekonečný rád, tak $\operatorname{Ker}\varphi = \{0\}$. Teda φ je surjektívne práve vtedy, keď $G = \langle a \rangle$, t. j. keď a generuje grupu G, a φ je injektívne práve vtedy, keď a má nekonečný rád.

27.3.8. Príklad. Nech $0 < a \in \mathbb{R}$. Keďže $a^x > 0$ a $a^{x+y} = a^x a^y$ pre všetky $x, y \in \mathbb{R}$, je predpisom $\varphi(x) = a^x$ definovaný homomorfizmus grúp $\varphi\colon (\mathbb{R}, +) \to (\mathbb{R}^+, \cdot)$. Ak $a = 1$, ide o triviálny homomorfizmus $\varphi(x) = 1$ pre každé $x \in \mathbb{R}$, teda $\operatorname{Ker}\varphi = \mathbb{R}$ a $\operatorname{Im}\varphi = \{1\}$. Pre $a \neq 1$ je to však bijektívny homomorfizmus, teda izomorfizmus. Inverzný izomorfizmus $\varphi^{-1}\colon (\mathbb{R}^+, \cdot) \to (\mathbb{R}, +)$ je daný predpisom $\varphi^{-1}(u) = \log_a u$ pre $u \in \mathbb{R}^+$. Známa formula $\log_a uv = \log_a u + \log_a v$ pre $u, v \in \mathbb{R}^+$ nie je vlastne nič iné než vlastnosť homomorfizmu zobrazenia $u \mapsto \log_a u$.

Okrem iného sme práve dokázali, že aditívna grupa $(\mathbb{R}, +)$ a multiplikatívna grupa $(\mathbb{R}^+, \cdot)$ sú izomorfné.

27.3.9. Príklad. Pripomeňme, že $\boldsymbol{R}_\alpha = \left(\begin{smallmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{smallmatrix}\right) \in \mathrm{GL}(2, \mathbb{R})$ je matica otočenia euklidovskej roviny $\mathbb{R}^2$ okolo počiatku o uhol α (pozri príklad 6.4.3). Keďže pre $\alpha, \beta \in \mathbb{R}$ platí známy vzťah $\boldsymbol{R}_\alpha \cdot \boldsymbol{R}_\beta = \boldsymbol{R}_{\alpha+\beta}$, znamená to, že priradením $\alpha \mapsto \boldsymbol{R}_\alpha$ je definovaný homomorfizmus grúp $\boldsymbol{R}\colon (\mathbb{R}, +) \to (\mathrm{GL}(2, \mathbb{R}), \cdot)$. Zrejme jadro $\operatorname{Ker}\boldsymbol{R} = 2\pi\mathbb{Z} = \{2k\pi;\, k \in \mathbb{Z}\} \subseteq \mathbb{R}$ tvorí podgrupa všetkých celočíselných násobkov čísla 2π a obraz $\operatorname{Im}\boldsymbol{R} = \{\boldsymbol{R}_\alpha;\, \alpha \in \mathbb{R}\} \subseteq \mathrm{GL}(2, \mathbb{R})$ je práve podgrupa všetkých matíc otočení $\boldsymbol{R}_\alpha$.

27.3.10. Príklad. Exponenciála rýdzo imaginárneho čísla $\mathrm{i}x$, kde $x \in \mathbb{R}$, je definovaná Eulerovým vzťahom $\mathrm{e}^{\mathrm{i}x} = \cos x + \mathrm{i}\sin x$. Opäť platí $\mathrm{e}^{\mathrm{i}(x+y)} =$

$e^{ix}e^{iy}$ pre všetky $x, y \in \mathbb{R}$. To znamená, že priradením $x \mapsto e^{ix}$ je definovaný homomorfizmus grúp $\varphi\colon (\mathbb{R}, +) \to (\mathbb{C}^\star, \cdot)$. Ľahko možno nahliadnuť, že $\operatorname{Ker}\varphi = 2\pi\mathbb{Z}$ a $\operatorname{Im}\varphi = \mathrm{U}(1)$.

27.3.11. Príklad. Predošlé tri príklady majú veľmi dôležité spoločné zovšeobecnenie. Ak $\boldsymbol{F}$ je spojitý homomorfizmus aditívnej grupy $\mathbb{R}$ do všeobecnej lineárnej grupy $\mathrm{GL}(n, \mathbb{R})$ resp. $\mathrm{GL}(n, \mathbb{C})$, tak jeho obraz $\operatorname{Im}\boldsymbol{F}$ sa nazýva *jednoparametrická maticová grupa* (pritom spojitosť maticovej funkcie $\boldsymbol{F} = (f_{ij})_{n\times n}$ chápeme v zmysle paragrafu 22.4, čiže ako spojitosť všetkých jej zložiek f_{ij}). Často striktne nerozlišujeme medzi homomorfizmom $\boldsymbol{F}$ a maticovou grupou $\operatorname{Im}\boldsymbol{F}$ a používame názov jednoparametrická (maticová) grupa aj pre homomorfizmus $\boldsymbol{F}$. Podobne možno definovať aj *jednoparametrické grupy lineárnych operátorov.* Jednoparametrické podgrupy sú akousi spojitou obdobou cyklických podgrúp grupy $\mathrm{GL}(n, \mathbb{R})$ resp. $\mathrm{GL}(n, \mathbb{C})$. Každá taká cyklická podgrupa má totiž tvar $\operatorname{Im}\boldsymbol{\Phi} = \{\boldsymbol{A}^k;\ k \in \mathbb{Z}\}$, kde $\boldsymbol{A}$ je jej generátor a $\boldsymbol{\Phi}(k) = \boldsymbol{A}^k$ je homomorfizmus aditívnej grupy celých čísel $\mathbb{Z}$ do príslušnej všeobecnej lineárnej grupy (porovnaj s cvičením 27.5(c)).

Jednoparametrické grupy úzko súvisia s autonómnymi sústavami lineárnych diferenciálnych rovníc (pozri paragraf 22.6). Presnejšie, pre $\boldsymbol{A}$ z $\mathbb{R}^{n\times n}$, prípadne $\mathbb{C}^{n\times n}$, je fundamentálna matica $\boldsymbol{F}(t) = e^{\boldsymbol{A}t}$ autonómnej homogénnej sústavy $\boldsymbol{x}' = \boldsymbol{A}\cdot\boldsymbol{x}$ zároveň jednoparametrickou grupou matíc. Ak položíme $\boldsymbol{U} = e^{\boldsymbol{A}}$, môžeme písať $\boldsymbol{F}(t) = e^{\boldsymbol{A}t} = \boldsymbol{U}^t$; príslušnú jednoparametrickú grupu potom obvykle zapisujeme v exponenciálnom tvare $\big(\boldsymbol{U}^t\big)_{t\in\mathbb{R}}$.

27.4 Rozklad grupy podľa podgrupy, normálne podgrupy

Skôr než čitateľ pristúpi k štúdiu tohto a nasledujúceho paragrafu, mal by si zopakovať základné fakty o ekvivalenciách a rozkladoch z paragrafu 0.6.

Pre ľubovoľné podmnožiny X, Y grupy $(G, \cdot)$ označíme

$$XY = \{xy;\ x \in X\ \&\ y \in Y\}.$$

Miesto $X\{y\}$ píšeme len krátko Xy a miesto $\{x\}Y$ iba xY. V aditívnom zápise používame označenie $X + Y = \{x + y;\ x \in X\ \&\ y \in Y\}$, $X + y$ a $x + Y$.

Nech S je podgrupa grupy $(G, \cdot, e)$. Potom zrejme $x = ex \in Sx$, lebo $e \in S$. Navyše pre ľubovoľné $x, y \in G$ platí

$$Sx \cap Sy \neq \emptyset\ \Rightarrow\ Sx = Sy;$$

inak povedané množiny, Sx, Sy sú disjunktné alebo sa rovnajú. Ak totiž $z \in Sx\cap Sy$, tak existujú $s_1, s_2 \in S$ také, že $z = s_1x = s_2y$. Potom $x = s_1^{-1}s_2y$, a každý prvok množiny Sx má pre vhodné $s \in S$ tvar $sx = ss_1^{-1}s_2y \in Sy$, lebo – keďže $S \subseteq G$ je podgrupa – platí $ss_1^{-1}s_2 \in S$. Teda $Sx \subseteq Sy$ a obrátenú inklúziu možno dostať rovnako, zámenou úloh x a y. Tým sme dokázali prvú časť nasledujúceho tvrdenia.

27.4.1. Tvrdenie. *Nech S je podgrupa grupy G. Potom systém množín $\{Sx;\, x \in G\}$ tvorí rozklad množiny G. Prvky $x, y \in G$ patria do tej istej triedy tohto rozkladu práve vtedy, keď $xy^{-1} \in S$.*

Dôkaz. Zostáva dokázať už len druhú časť. Keďže uvedený systém množín je rozklad a $y \in Sy$, tak x a y patria do tej istej triedy rozkladu práve vtedy, keď $x \in Sy$, t. j. $x = sy$ pre nejaké $s \in S$. To je zrejme ekvivalentné s podmienkou $xy^{-1} = s \in S$.

Množiny Sx, $x \in G$, sa nazývajú *ľavé triedy rozkladu grupy G podľa podgrupy S*. Reláciu ekvivalencie zodpovedajúcu tomuto rozkladu značíme $\equiv_S$ a môžeme ju vyjadriť nasledujúcimi piatimi ekvivalentnými spôsobmi:

$$\begin{aligned} x \equiv_S y \;&\Leftrightarrow\; xy^{-1} \in S \;\Leftrightarrow\; Sx = Sy \\ &\Leftrightarrow\; x \in Sy \;\Leftrightarrow\; y \in Sx \;\Leftrightarrow\; Sx \cap Sy \neq \emptyset. \end{aligned}$$

Príslušnú faktorovú množinu (t. j. vlastne rozklad) značíme

$$G/S = G/{\equiv_S} = \{Sx;\, x \in G\}.$$

Počet prvkov $\#(G/S)$ faktorovej množiny G/S (ak je konečná) nazývame *indexom podgrupy S* v grupe G a značíme ho tiež $[G : S]$; ak G/S je nekonečná, kladieme $[G : S] = \infty$ a hovoríme, že S má v G *nekonečný index.* Rád $\# G$ samotnej grupy G zrejme splýva s indexom $[G : e]$ jej triviálnej podgrupy $\{e\}$.

Analogicky možno zaviesť aj *pravé triedy rozkladu grupy G podľa podgrupy S*, t. j. množiny xS, $x \in G$, a dokázať pre ne obdobu tvrdenia 27.4.1.[2] Príslušnú reláciu ekvivalencie možno vyjadriť zodpovedajúcimi piatimi ekvivalentnými formulami:

$$\begin{aligned} x \,{}_S{\equiv}\, y \;&\Leftrightarrow\; x^{-1}y \in S \;\Leftrightarrow\; xS = yS \\ &\Leftrightarrow\; x \in yS \;\Leftrightarrow\; y \in xS \;\Leftrightarrow\; xS \cap yS \neq \emptyset. \end{aligned}$$

Taktiež rozklad grupy G na pravé triedy rozkladu podľa podgrupy S značíme rovnako

$$G/S = G/{}_S{\equiv} = \{xS;\, x \in G\};$$

[2] Používaná terminológia nie je v tomto smere jednotná. Niektorí autori nazývajú pravými triedami rozkladu grupy podľa podgrupy to, čo my nazývame ľavými triedami, a naopak.

ak je potrebné blžšie špecifikovať, či ide o rozklad na ľavé alebo pravé triedy, vyjadríme to väčšinou slovne.

Nech S je podgrupa grupy G a Sx, Sy sú ľubovoľné (nie nevyhnutne rôzne) ľavé triedy rozkladu G podľa S. Takpovediac na prstoch možno overiť, že predpisom $u \mapsto ux^{-1}y$ je definované bijektívne zobrazenie $Sx \to Sy$, ktoré prvok $u = sx \in Sx$ ($s \in S$) zobrazí na prvok $sy \in Sy$; k nemu inverzné zobrazenie $Sy \to Sx$ je dané predpisom $v \mapsto vy^{-1}x$, t. j. prvok $v = sy \in Sy$ ním prejde na prvok $sx \in Sx$. V prípade, že podgrupa S je konečná, to však znamená, že všetky ľavé (no rovnako aj pravé) triedy rozkladu G podľa S majú ten istý počet prvkov rovný rádu $\# S = [S : e]$ grupy $S = Se = eS$. Keďže G je zjednotením navzájom disjunktných, rovnako početných tried rozkladu $Sx \in G/S$ (prípadne xS), dokázali sme tým nasledujúcu vetu:

27.4.2. Veta. (Lagrange) *Nech S je podgrupa konečnej grupy G. Potom*

$$[G : e] = [G : S] \cdot [S : e],$$

teda rád aj index podgrupy S sú deliteľmi rádu grupy G.

Neskôr uvidíme, že pre daný deliteľ d rádu konečnej *abelovskej* grupy G existuje podgrupa grupy G rádu d (veta 28.3.8). Vo všeobecnosti však G nemusí obsahovať podgrupu rádu d pre každý deliteľ d svojho rádu $[G : e]$ (pozri cvičenie 28.20).

27.4.3. Dôsledok. (Malá veta Fermatova) *Nech p je prvočíslo. Potom pre každé celé číslo x, ktoré nie je násobkom čísla p, platí*

$$x^{p-1} \equiv 1 \mod p,$$

t. j. číslo x^{p-1} dáva po delení číslom p zvyšok 1.

Dôkaz. Označme z zvyšok, ktorý dáva x po delení p. Keďže x nie je násobkom p, $z \in \mathbb{Z}_p^\star$, čo je multiplikatívna grupa (nenulových prvkov) poľa $\mathbb{Z}_p$ s rádom $p - 1$. Potom rád r jej cyklickej podgrupy $\langle z \rangle$ je deliteľom čísla $p - 1$, teda $p - 1 = rk$ pre nejaké kladné celé číslo k. Podľa tvrdenia 27.2.6 v grupe $(\mathbb{Z}_p^\star, \cdot, 1)$ platí $z^r = 1$, z čoho vyplýva

$$z^{p-1} = z^{rk} = (z^r)^k = 1^k = 1.$$

Teda z^{p-1} (teraz už ako prvok $\mathbb{Z}$) dáva po delení číslom p zvyšok 1. Avšak čísla z^{p-1} a x^{p-1} dávajú po delení číslom p rovnaký zvyšok (pozri cvičenie 0.16).

Pre ľavú a pravú triedu rozkladu prvku x grupy G podľa jej podgrupy S môže vo všeobecnosti platiť $Sx \neq xS$. Napríklad rozklad dihedrálnej grupy

$\varDelta_3 = \mathcal{S}_3$ podľa cyklickej podgrupy $S = \langle\sigma_1\rangle = \{\iota, \sigma_1\}$ na ľavé triedy tvoria množiny

$$S\iota = S\sigma_1 = \{\iota, \sigma_1\}, \qquad S\varrho = S\sigma_2 = \{\varrho, \sigma_2\}, \qquad S\varrho^{-1} = S\sigma_3 = \left\{\varrho^{-1}, \sigma_3\right\},$$

a rozklad na pravé triedy zasa množiny

$$\iota S = \sigma_1 S = \{\iota, \sigma_1\}, \qquad \varrho S = \sigma_3 S = \{\varrho, \sigma_3\}, \qquad \varrho^{-1} S = \sigma_2 S = \left\{\varrho^{-1}, \sigma_2\right\}.$$

(Overte pomocou multiplikatívnej tabuľky tejto grupy v paragrafe 0.5.)

Za istých dodatočných podmienok kladených na podgrupu S však ľavé a pravé triedy rozkladu G podľa S splývajú.

27.4.4. Tvrdenie. *Nech S je podgrupa grupy G. Potom nasledujúce podmienky sú ekvivalentné:*
(i) $(\forall\, x \in G)(\forall\, s \in S)\big(x^{-1}sx \in S\big)$;
(ii) $(\forall\, x \in G)\big(x^{-1}Sx = S\big)$;
(iii) $(\forall\, x \in G)(Sx = xS)$;
(iv) $(\forall\, x, y \in G)(x \equiv_S y \;\Leftrightarrow\; x\, {}_S{\equiv}\, y)$.

Dôkaz. (i) $\Rightarrow$ (ii) Z predpokladu (i) vyplýva inklúzia $x^{-1}Sx \subseteq S$. Jej prenásobením prvkom x zľava a prvkom x^{-1} sprava dostaneme inklúziu $S \subseteq xSx^{-1}$; keďže $x \in G$ je ľubovoľný prvok, substitúciou x^{-1} miesto x získame $S \subseteq x^{-1}Sx$. Teda $x^{-1}Sx = S$.

(ii) $\Rightarrow$ (iii) Stačí prenásobiť rovnosť $x^{-1}Sx = S$ prvkom x zľava.

(iii) $\Rightarrow$ (iv) Ak $Sx = xS$, tak

$$x \equiv_S y \;\Leftrightarrow\; y \in Sx \;\Leftrightarrow\; y \in xS \;\Leftrightarrow\; x\, {}_S{\equiv}\, y.$$

(iv) $\Rightarrow$ (i) Podľa predpokladu pre všetky $x, y \in G$ platí $xy^{-1} \in S \Leftrightarrow x^{-1}y \in S$. Keďže pre ľubovoľné $s \in S$ je $x(sx)^{-1} = s^{-1} \in S$, stačí položiť $y = sx$ a hneď máme $x^{-1}sx \in S$.

Hovoríme, že *podgrupa* S grupy G je *normálna* alebo tiež *invariantná*, označenie $S \lhd G$, ak spĺňa niektorú (teda všetky) z navzájom ekvivalentných podmienok tvrdenia 27.4.4. V prípade normálnych podgrúp teda nemusíme rozlišovať medzi ľavými a pravými triedami rozladu ani medzi ekvivalenciami $\equiv_S$ a ${}_S{\equiv}$.

V každej grupe G platí $\{e\} \lhd G$ a $G \lhd G$. Zrejme v abelovskej grupe G je každá podgrupa normálna. Ako sme však videli pred chvíľou, netriválna vlastná podgrupa neabelovskej grupy už normálna byť nemusí.

Najdôležitejšími, a svojim spôsobom typickými príkladmi normálnych podgrúp sú jadrá grupových homomorfizmov.

27.4.5. Tvrdenie. *Nech* $\varphi\colon G \to H$ *je homomorfizmus grúp. Potom* $\operatorname{Ker}\varphi \lhd G$.

Dôkaz. Podľa tvrdenia 27.3.4 je $\operatorname{Ker}\varphi$ podgrupa grupy G. Overíme podmienku (i) tvrdenia 27.4.4, čím dokážeme jej normálnosť. Nech $a \in \operatorname{Ker}\varphi$, t. j. $\varphi(a) = e$, a $x \in G$ je ľubovoľný prvok. Potom

$$\varphi\big(x^{-1}ax\big) = \varphi\big(x^{-1}\big) \cdot \varphi(a) \cdot \varphi(x) = \varphi(x)^{-1} \cdot e \cdot \varphi(x) = e,$$

teda $x^{-1}ax \in \operatorname{Ker}\varphi$.

Každá podmnožina X grupy G generuje nielen najmenšiu podgrupu $\langle X\rangle \subseteq G$ takú, že $X \subseteq \langle X\rangle$, ale aj najmenšiu *normálnu* podgrupu $\langle\!|X|\!\rangle \subseteq G$, ktorá obsahuje X. Množinu $\langle\!|X|\!\rangle$ možno jednoducho popísať. Položme $X^{-1} = \{x^{-1};\ x \in X\}$, $\widehat{X} = \{gug^{-1};\ u \in X \cup X^{-1}\ \&\ g \in G\}$ a konečne

$$\langle\!|X|\!\rangle = \langle\widehat{X}\rangle,$$

čiže $\langle\!|X|\!\rangle$ je podgrupa generovaná množinou $\widehat{X}$. Jednoduchý dôkaz nasledujúceho tvrdenie prenechávame ako cvičenie čitateľovi.

27.4.6. Tvrdenie. *Nech* G *je grupa a* $X \subseteq G$. *Potom* $\langle\!|X|\!\rangle$ *je najmenšia normálna podgrupa grupy* G *taká, že* $X \subseteq \langle X|\!\rangle$.

Množinu $\langle\!|X|\!\rangle$ teda môžeme oprávnene nazývať *normálna podgrupa grupy* G *generovaná množinou* X. Zrejme platí $\langle X\rangle \subseteq \langle\!|X|\!\rangle$, a v abelovskej grupe G dokonca $\langle X\rangle = \langle\!|X|\!\rangle$.

27.5 Faktorové grupy

V tomto paragrafe sa zoznámime s konštrukciou *faktorovej grupy*. Ide o dôležitý príklad všeobecnejšej konštrukcie *faktorovej štruktúry*, s ktorou sa možno stretnúť v najrôznejších oblastiach matematiky. S prvou základnou myšlienkou tejto konštrukcie sme sa už zbežne zoznámili v paragrafe 0.6 – spočíva v nazeraní prvkov faktorovej množiny $X/{\sim}$ množiny X podľa ekvivalencie $\sim$ nie ako množín (t. j. tried príslušného rozkladu) ale ako jednotlivých prvkov. Druhou kľúčovou myšlienkou spomínanej konštrukcie je *prenos štruktúry* z pôvodnej množiny X na faktorovú množinu $X/{\sim}$ pomocou kanonickej projekcie $X \to X/{\sim}$ danej priradením $x \mapsto \tilde{x}$. Túto všeobecnú myšlienku možno asi najjednoduchšie ilustrovať práve na konštrukcii faktorovej grupy.

Nech $(G, \cdot)$ je grupa a $\sim$ je ekvivalencia na množine G. Na faktorovej množine $G/\sim$ hodláme zaviesť binárnu operáciu $\cdot$ tak, aby $(G/\sim, \cdot)$ bola grupa a kanonická projekcia $G \to G/{\sim}$ homomorfizmus grúp. To znamená, že pre všetky $x, y \in G$ musí platiť

$$\tilde{x} \cdot \tilde{y} = \widetilde{xy},$$

Inak povedané, podmienka homomorfnosti kanonickej projekcie $x \mapsto \tilde{x}$ už jednoznačne určuje príslušnú binárnu operáciu na $G/{\sim}$.

Aby však takto bola *korektne* definovaná binárna operácia na množine $G/{\sim}$, jej výsledok nesmie závisieť na konkrétnych reprezentantoch x, y tried $\tilde{x}, \tilde{y} \in G/{\sim}$. Pre prvky $x_1, x_2, y_1, y_2 \in G$ také, že $\tilde{x}_1 = \tilde{x}_2$ a $\tilde{y}_1 = \tilde{y}_2$, t. j. $x_1 \sim x_2$ a $y_1 \sim y_2$, musí totiž platiť $\tilde{x}_1 \cdot \tilde{y}_1 = \tilde{x}_2 \cdot \tilde{y}_2$. Teda podľa našej definície

$$\widetilde{x_1 y_1} = \tilde{x}_1 \cdot \tilde{y}_1 = \tilde{x}_2 \cdot \tilde{y}_2 = \widetilde{x_2 y_2},$$

v dôsledku čoho $x_1 y_1 \sim x_2 y_2$.

Hovoríme, že ekvivalencia $\sim$ na grupe G je *kongruencia*, ak pre ľubovoľné prvky $x_1, x_2, y_1, y_2 \in G$ platí

$$x_1 \sim x_2 \ \& \ y_1 \sim y_2 \ \Rightarrow \ x_1 y_1 \sim x_2 y_2.$$

Krátko povedané, aby rovnosťou $\tilde{x} \cdot \tilde{y} = \widetilde{xy}$ bola korektne definovaná binárna operácia na faktorovej množine $G/{\sim}$, ekvivalencia $\sim$ musí byť *kongruenciou* na grupe G. Táto nevyhnutná podmienka je aj postačujúca na dosiahnutie nášho vopred stanoveného cieľa.

27.5.1. Veta. *Nech $\sim$ je kongruencia na grupe $(G, \cdot)$. Potom faktorová množina $G/{\sim}$ s binárnou operáciou $\tilde{x} \cdot \tilde{y} = \widetilde{xy}$ tvorí grupu, ktorej jednotkovým prvkom je $\tilde{e}$ a inverzným prvkom k prvku $\tilde{x} \in G/{\sim}$ je prvok $\tilde{x}^{-1} = \widetilde{x^{-1}}$. Navyše, prirodzená projekcia $G \to G/{\sim}$ je surjektívny homomorfizmus grúp.*

Dôkaz. Na dôkaz prvej časti tvrdenia stačí overiť, že pre ľubovoľné prvky $x, y, z \in G$ platí

$$\tilde{x} \cdot (\tilde{y} \cdot \tilde{z}) = (\tilde{x} \cdot \tilde{y}) \cdot \tilde{z}, \qquad \tilde{e} \cdot \tilde{x} = \tilde{x} \cdot \tilde{e} = \tilde{x}, \qquad \tilde{x} \cdot \widetilde{x^{-1}} = \widetilde{x^{-1}} \cdot \tilde{x} = \tilde{e}.$$

Príslušné jednoduché výpočty prenechávame ako cvičenie čitateľovi.

Zrejme prirodzená projekcia $G \to G/{\sim}$ je surjekcia. Už len stačí pripomenúť, že násobenie na množine $G/{\sim}$ bolo definované práve tak, aby zobrazenie $x \mapsto \tilde{x}$ bolo homomorfizmom.

Ukazuje sa, že kongruencie na grupách úzko súvisia s normálnymi podgrupami.

27.5.2. Tvrdenie. *Nech $(G, \cdot, e)$ je grupa.*

(a) *Ak N je normálna podgrupa grupy G, tak ekvivalencia $\equiv_N$ je kongruenciou na grupe G.*

(b) *Ak $\sim$ je kongruencia na grupe G, tak $N = \tilde{e}$ je normálna podgrupa grupy G a pre všetky $x, y \in G$ platí*

$$x \sim y \;\Leftrightarrow\; xy^{-1} \in N,$$

t. j. $\sim$ splýva s ekvivalenciou $\equiv_N$ rozkladu grupy G podľa N a pre $x \in G$ platí $\tilde{x} = Nx = xN$.

Inými slovami, kongruencie na grupách sú práve ekvivalenciami rozkladu podľa ich *normálnych podgrúp.*

Dôkaz. (a) Nech $N \lhd G$. Už vieme, že $\equiv_N$ je ekvivalencia na G. Zvoľme v G prvky $x_1 \equiv_N x_2$, $y_1 \equiv_N y_2$; ukážeme, že platí $x_1y_1 \equiv_N x_2y_2$. Podľa predpokladu $x_1x_2^{-1} \in N$ a $y_1y_2^{-1} \in N$. Keďže N je normálna, platí $x_2\big(y_1y_2^{-1}\big)x_2^{-1} \in N$. Preto tiež

$$(x_1y_1)(x_2y_2)^{-1} = x_1\big(y_1y_2^{-1}\big)x_2^{-1} = \big(x_1x_2^{-1}\big)\big(x_2\big(y_1y_2^{-1}\big)x_2^{-1}\big) \in N,$$

teda $x_1y_1 \equiv_N x_2y_2$, takže $\equiv_N$ je kongruencia na G.

(b) Nech $\sim$ je kongruencia na G. Podľa vety 27.5.1 je $x \mapsto \tilde{x}$ homomorfizmus grúp $G \to G/{\sim}$. Jeho jadrom je zrejme $N = \tilde{e}$. Podľa tvrdenia 27.4.5 potom $N \lhd G$.

Keďže $(G/{\sim}, \cdot)$ je grupa, pre $x, y \in G$ máme

$$x \sim y \;\Leftrightarrow\; \tilde{x} = \tilde{y} \;\Leftrightarrow\; \tilde{x} \cdot \tilde{y}^{-1} = \tilde{e}.$$

Nakoľko $x \mapsto \tilde{x}$ je homomorfizmus, platí $\tilde{x} \cdot \tilde{y}^{-1} = \big(xy^{-1}\big)^{\sim}$, preto tretia podmienka je zrejme ekvivalentná so vzťahom $xy^{-1} \in N$. Zvyšok je dôsledkom tvrdenia 27.4.4.

Grupu $G/{\equiv_N}$ nazývame *faktorovou grupou* grupy G podľa jej normálnej podgrupy N a značíme ju G/N. Násobenie v G/N je definované rovnosťou

$$Nx \cdot Ny = Nxy,$$

jednotkovým prvkom v G/N je $N = Ne$ a inverzným prvkom k $Nx \in G/N$ je $(Nx)^{-1} = Nx^{-1}$.

Prirodzená projekcia $\zeta_N \colon G \to G/N$, daná predpisom $\zeta_N(x) = Nx$ pre $x \in G$, je surjektívny homomorfizmus s jadrom $\operatorname{Ker} \zeta_N = N$. To dokazuje platnosť aj obrátenej implikácie k tvrdeniu 27.4.5.

27.5.3. Veta. *Podgrupa N grupy G je normálna práve vtedy, keď existuje homomorfizmus grúp $\varphi \colon G \to H$ taký, že $N = \operatorname{Ker} \varphi$.*

27.5.4. Veta. (O homomorfizme) *Nech $\varphi\colon G \to H$ je homomorfizmus grúp; označme $N = \operatorname{Ker}\varphi$. Potom predpisom $\varphi_*(Nx) = \varphi(x)$ je definovaný injektívny homomorfizmus grúp $\varphi_*\colon G/N \to H$. V dôsledku toho*

$$\operatorname{Im}\varphi \cong G/\operatorname{Ker}\varphi;$$

ak φ je navyše surjektívny homomorfizmus, tak $H \cong G/\operatorname{Ker}\varphi$.

Dôkaz. Pre $x, y \in G$ platí $Nx = Ny \Leftrightarrow \varphi(x) = \varphi(y)$. Implikácia $\Rightarrow$ zaručuje, že priradením $Nx \mapsto \varphi(x)$ je korektne definované zobrazenie $\varphi_*\colon G/N \to H$. Obrátená implikácia $\Leftarrow$ vyjadruje injektívnosť tohto zobrazenia. Priamočiary výpočet

$$\varphi_*(Nx \cdot Ny) = \varphi_*(Nxy) = \varphi(xy) = \varphi(x) \cdot \varphi(y) = \varphi_*(Nx) \cdot \varphi_*(Ny)$$

ukazuje, že ide o homomorfizmus. Druhá časť vety je bezprostredným dôsledkom prvej.

Uvedený homomorfizmus φ_* spĺňa podmienku $\varphi = \varphi_* \circ \zeta_N$, t. j. zabezpečuje komutatívnosť diagramu

$$\begin{tikzcd} G \arrow[r, "\varphi"] \arrow[d, "\zeta_N"'] & H \\ G/N \arrow[ur, "\varphi_*"'] & \end{tikzcd}$$

Naopak, pokiaľ má zobrazenie $\varphi_*\colon G/N \to H$ vyhovovať tejto podmienke, musí byť nevyhnutne definované rovnosťou $\varphi_*(Nx) = \varphi(x)$.

Grupa H sa nazýva *homomorfným obrazom grupy G*, ak existuje surjektívny homomorfizmus $\varphi\colon G \to H$. Zrejme každá faktorová grupa grupy G je jej homomorfným obrazom. Ako vyplýva z vety o homomorfizme, tiež naopak, každý homomorfný obraz grupy G je izomorfný s faktorovou grupou G/N grupy G podľa niektorej jej normálnej podgrupy N. To znamená, že všetky homomorfné obrazy grupy G sú (až na izomorfizmus) skryte prítomné už v samotnej grupe G.

27.5.5. Príklad. Nech n je kladné celé číslo. Označme $\zeta_n(x)$ zvyšok po delení celého čísla x číslom n. Potom $\zeta_n\colon \mathbb{Z} \to \mathbb{Z}_n$ je homomorfizmus grúp (s operáciou $+$). Keďže ζ_n je surjektívny a $\operatorname{Ker}\zeta_n = n\mathbb{Z}$, podľa vety o homomorfizme platí $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$. Po stotožnení faktorovej grupy $\mathbb{Z}/n\mathbb{Z}$ s homomorfným obrazom $\mathbb{Z}_n = \zeta_n(\mathbb{Z})$ grupy $\mathbb{Z}$, daným uvedeným izomorfizmom, je každý prvok $x + n\mathbb{Z} = \{x + nk;\ k \in \mathbb{Z}\}$ faktorovej grupy $\mathbb{Z}/n\mathbb{Z}$, t. j. vlastne nekonečná trieda rozkladu podľa podgrupy $n\mathbb{Z} \subseteq \mathbb{Z}$, reprezentovaný ako *jediný prvok* $\zeta_n(x) \in \mathbb{Z}_n = \{0, 1, \dots, n-1\}$. Príslušná kongruencia $\equiv_{n\mathbb{Z}}$ zrejme splýva s kongruenciou $\equiv_n$ modulo n z cvičenia 0.16.

27.5.6. Príklad. Z príkladu 27.3.9 vieme, že zobrazenie $\boldsymbol{R}\colon (\mathbb{R},+) \to (\mathrm{GL}(2,\mathbb{R}),\cdot)$ je homomorfizmus grúp s jadrom $\operatorname{Ker}\boldsymbol{R} = 2\pi\mathbb{Z}$ a obrazom $\operatorname{Im}\boldsymbol{R} = \{\boldsymbol{R}_\alpha;\ \alpha \in \mathbb{R}\}$. Z príkladu 27.3.10 zasa vieme, že priradením $x \mapsto \mathrm{e}^{\mathrm{i}x}$ je definovaný homomorfizmus grúp $(\mathbb{R},+) \to (\mathbb{C}^\star,\cdot)$ s rovnakým jadrom $2\pi\mathbb{Z}$ a s obrazom $\mathrm{U}(1) = \{\mathrm{e}^{\mathrm{i}x};\ x \in \mathbb{R}\} = \{z \in \mathbb{C};\ |z| = 1\}$. Z vety 27.5.4 tak okamžite vyplýva $\operatorname{Im}\boldsymbol{R} \cong \mathbb{R}/2\pi\mathbb{Z} \cong \mathrm{U}(1)$.

Práve izomorfizmus aditívnej faktorovej grupy $\mathbb{R}/2\pi\mathbb{Z}$ grupy $\mathbb{R}$ a jej homomorfného obrazu, multiplikatívnej grupy $\mathrm{U}(1)$, je pre pochopenie všeobecnej konštrukcie faktorovej grupy veľmi poučný. Faktorová množina $\mathbb{R}/2\pi\mathbb{Z}$ je ním reprezentovaná ako jednotková kružnica – názorne si môžeme predstaviť, že vznikla priložením bodu 0 reálnej osi $\mathbb{R}$ na bod 1 jednotkovej kružnice $\mathrm{U}(1) \subseteq \mathbb{C}$ a následným namotaním kladnej polosi proti a zápornej v smere hodinových ručičiek. Body $x, y \in \mathbb{R}$ sa pritom ocitnú v tom istom bode kružnice $\mathrm{U}(1)$ práve vtedy keď $x-y \in 2\pi\mathbb{Z}$, t. j. práve vtedy, keď $(x-y)/2\pi$ je celé číslo. Celá nekonečná trieda rozkladu $x + 2\pi\mathbb{Z} \in \mathbb{R}/2\pi\mathbb{Z}$ prvku $x \in \mathbb{R}$ podľa podgrupy $2\pi\mathbb{Z} \subseteq \mathbb{R}$ je tak reprezentovaná *jediným bodom* $\mathrm{e}^{\mathrm{i}x} = \cos x + \mathrm{i}\sin x$ kružnice $\mathrm{U}(1)$.

Konštrukcia faktorovej grupy pripúšťa tiež abstraktnejší popis pomocou tzv. krátkych exaktných postupností. Hovoríme že *postupnosť*

$$G_0 \xrightarrow{\varphi_1} G_1 \xrightarrow{\varphi_2} \quad \cdots \quad \xrightarrow{\varphi_{n-1}} G_{n-1} \xrightarrow{\varphi_n} G_n$$

grupových homomorfizmov je *exaktná*, ak pre každé $1 \le i < n$ platí

$$\operatorname{Im}\varphi_i = \operatorname{Ker}\varphi_{i+1}.$$

Podobne možno definovať pojem exaktnosti aj pre (na jednu či na obe strany) nekonečné postupnosti na seba nadväzujúcich grupových homomorfizmov. *Krátkou exaktnou postupnosťou* nazývame exaktnú postupnosť tvaru

$$\{e\} \xrightarrow{\eta} F \xrightarrow{\varphi} G \xrightarrow{\psi} H \xrightarrow{\tau} \{e\},$$

kde $\eta\colon \{e\} \to F$ a $\tau\colon H \to \{e\}$ sú triviálne homomorfizmy. Exaktnosť v člene F znamená, že $\operatorname{Ker}\varphi = \operatorname{Im}\eta = \{e_F\}$, čiže injektívnosť φ. Podobne, exaktnosť v člene H znamená, že $\operatorname{Im}\psi = \operatorname{Ker}\tau = H$, čiže surjektívnosť ψ. Konečne exaktnosť v strednom člene G znamená, že platí $\operatorname{Im}\varphi = \operatorname{Ker}\psi \lhd G$, z čoho vzhľadom na vetu o homomorfizme a surjektívnosť ψ vyplýva $G/\operatorname{Im}\varphi \cong H$.

Triviálna grupa $\{e\}$ na okrajoch, ako aj triviálne homomorfizmy η, τ sa zvyknú pri zápise krátkej exaktnej postupnosti vynechávať. Ak teda hovoríme o krátkej exaktnej postupnosti

$$F \overset{\varphi}{\rightarrowtail} G \overset{\psi}{\twoheadrightarrow} H,$$

znamená to, že φ je injektívny a ψ je surjektívny homomorfizmus grúp, pričom platí $\operatorname{Im}\varphi = \operatorname{Ker}\psi$. Injektívnosť φ resp. surjektívnosť ψ sa vyznačuje modifikovanými šípkami $\rightarrowtail$ resp. $\twoheadrightarrow$. Ak existuje krátka exaktná postupnosť $F \overset{\varphi}{\rightarrowtail} G \overset{\psi}{\twoheadrightarrow} H$, hovoríme, že stredná grupa G je *rozšírením grupy H pomocou grupy F*.[3]

Paragraf uzatvárame zhrnutím úvah o krátkych exaktných postupnostiach a vety o homomorfizme do jedného celku.

27.5.7. Veta. *(a) Nech G je grupa a $N \lhd G$ je jej normálna podgrupa. Označme $\iota\colon N \to G$ vnorenie, t. j. identické zobrazenie N do G a $\zeta\colon G \to G/N$ kanonickú projekciu G na faktorovú grupu G/N. Potom $N \overset{\iota}{\rightarrowtail} G \overset{\zeta}{\twoheadrightarrow} G/N$ je krátka exaktná postupnosť homomorfizmov grúp.*

(b) Nech naopak $F \overset{\varphi}{\rightarrowtail} G \overset{\psi}{\twoheadrightarrow} H$ je krátka exaktná postupnosť homomorfizmov grúp. Označme $N = \operatorname{Im}\varphi = \operatorname{Ker}\psi$. Potom $F \cong N \lhd G$ a $H \cong G/N$.

27.6 Priamy súčin grúp

Konštrukcia priameho súčinu nám, spolu s konštrukciami podgrúp a faktorových grúp, umožňuje vytvárať z daných grúp nové. Taktiež naopak, niekedy nám umožňuje reprezentovať danú grupu v tvare priameho súčinu jednoduchších grúp, a tým sprehľadniť jej štruktúru. Opäť ide o špeciálny prípad veľmi dôležitej všeobecnej konštrukcie, ktorej pôsobnosť sa neobmedzuje len na teóriu grúp.

Priamym súčinom grúp $(G_1, \cdot, e_1)$, $(G_2, \cdot, e_2)$ nazývame karteziánsky súčin

$$G_1 \times G_2 = \{(x_1, x_2);\ x_1 \in G_1 \ \&\ x_2 \in G_2\},$$

množín G_1, G_2, s binárnou operáciou definovanou po zložkách, t. j.

$$(x_1, x_2) \cdot (y_1, y_2) = (x_1 y_1, x_2 y_2)$$

pre $x_1, y_1 \in G_1$, $x_2, y_2 \in G_2$. Roznásobením zvlášť v každej zložke sa možno ľahko presvedčiť, že táto operácia na množine $G_1 \times G_2$ je asociatívna, má (e_1, e_2) za neutrálny prvok a inverzným prvkom k prvku $(x_1, x_2) \in G_1 \times G_2$ je $(x_1, x_2)^{-1} = \left(x_1^{-1}, x_2^{-1}\right)$. Tým sme vlastne dokázali nasledujúcu vetu.

27.6.1. Veta. *Priamy súčin $G_1 \times G_2$ grúp G_1, G_2 je grupa.*

Zovšeobecnenie konštrukcie priameho súčinu, ako aj vety 27.6.1 na ľubovoľný konečný počet grúp prenechávame čitateľovi.

[3] Niektorí autori v takejto situácii hovoria, že grupa G je rozšírením grupy F pomocou grupy H.

Reprezentácia danej grupy G v tvare priameho súčinu $G \cong G_1 \times G_2$ prináša niečo nové, len ak sú obe grupy G_1, G_2 netriviálne. To nastoľuje zaujímavý problém: rozložiť danú grupu G na priamy súčin dvoch netriviálnych grúp, t. j. nájsť netriválne grupy G_1 a G_2 tak, že platí $G \cong G_1 \times G_2$, prípadne ukázať, že také grupy neexistujú – v tom prípade hovoríme, že *grupa* G je *priamo nerozložiteľná*. Rozumnú charakterizáciu priamo nerozložiteľných grúp nemáme naporúdzi. Jednako uvedieme niekoľko jednoduchých podmienok, ktoré nám umožnia rozložiť niektoré známe grupy na priamy súčin v istom zmysle jednoduchších faktorov.

Najprv si všimnime, že projekcie $\pi_1\colon G_1 \times G_2 \to G_1$, $\pi_2\colon G_1 \times G_2 \to G_2$, dané predpismi $\pi_1(x_1, x_2) = x_1$, $\pi_2(x_1, x_2) = x_2$, sú surjektívne grupové homomorfizmy a pre ich jadrá platí

$$\operatorname{Ker}\pi_1 = \{e_1\} \times G_2 \cong G_2, \qquad \operatorname{Ker}\pi_2 = G_1 \times \{e_2\} \cong G_1;$$

príslušné izomorfizmy sú dané priradeniami $(e_1, x_2) \mapsto x_2$ resp. $(x_1, e_2) \mapsto x_1$. To znamená, že grupa $G_1 \times G_2$ obsahuje dve normálne podgrupy $N_1 \cong G_1$ a $N_2 \cong G_2$. Navyše $N_1 \cap N_2 = \{(e_1, e_2)\}$ a pre ľubovoľný prvok $(x_1, x_2) \in G_1 \times G_2$ platí

$$(x_1, x_2) = (x_1, e_2) \cdot (e_1, x_2) = (e_1, x_2) \cdot (x_1, e_2),$$

z čoho vyplývajú rovnosti $G = N_1N_2 = N_2N_1$.

27.6.2. Tvrdenie. *Nech $(G, \cdot)$ je grupa a S, $T \subseteq G$ sú jej podgrupy. Označme $\varphi\colon S \times T \to G$ zobrazenie dané predpisom $\varphi(x, y) = xy$ pre $x \in S$, $y \in T$. Potom*

(a) φ je homomorfizmus grúp práve vtedy, keď pre všetky $x \in S$, $y \in T$ platí $xy = yx$;

(b) φ je injektívne práve vtedy, keď $S \cap T = \{e\}$;

(c) φ je surjektívne práve vtedy, keď $G = ST$.

Dôkaz. (a) Predpokladajme, že φ je homomorfizmus. Potom pre ľubovoľné $x \in S$, $y \in T$ platí

$$xy = \varphi(x, y) = \varphi\big((e, y) \cdot (x, e)\big) = \varphi(e, y) \cdot \varphi(x, e) = (ey)(xe) = yx.$$

Nech naopak pre všetky $x \in S$, $y \in T$ platí $xy = yx$. Zvoľme $x_1, x_2 \in S$, $y_1, y_2 \in T$. Potom

$$\begin{aligned}\varphi\big((x_1, y_1) \cdot (x_2, y_2)\big) &= \varphi(x_1x_2, y_1y_2) = (x_1x_2)(y_1y_2)\\ &= (x_1y_1)(x_2y_2) = \varphi(x_1, y_1) \cdot \varphi(x_2, y_2),\end{aligned}$$

čo znamená, že φ je homomorfizmus.

(b) Predpokladajme, že platí $S \cap T = \{e\}$. Nech $(x_1, y_1), (x_2, y_2) \in S \times T$ sú také, že $\varphi(x_1, y_1) = \varphi(x_2, y_2)$. Potom $x_1 y_1 = x_2 y_2$, a keďže S, T sú podgrupy a $x_1, x_2 \in S$, $y_1, y_2 \in T$, platí $x_2^{-1} x_1 = y_2 y_1^{-1} \in S \cap T$. Z toho vyplýva $x_2^{-1} x_1 = y_2 y_1^{-1} = e$, teda $(x_1, y_1) = (x_2, y_2)$, čo dokazuje injektívnosť zobrazenia φ.

Nech naopak φ je injektívne. Pre $x \in S \cap T$ platí tiež $x^{-1} \in S \cap T$, lebo $S \cap T$ je podgrupa grupy G. Preto $(x, x^{-1}) \in S \times T$ a

$$\varphi(x, x^{-1}) = xx^{-1} = e = \varphi(e, e).$$

Vďaka injektívnosti φ z toho vyplýva $x = e = x^{-1}$. Teda $S \cap T = \{e\}$.

(c) platí triviálne.

27.6.3. Dôsledok. *Nech G je grupa a S, $T \lhd G$ sú jej dve normálne podgrupy také, že $S \cap T = \{e\}$ a $G = ST$. Potom $G \cong S \times T$.*

Dôkaz. Stačí overiť, že za uvedených predpokladov je splnená aj podmienka (a) predchádzajúceho tvrdenia, čiže pre všetky $x \in S$, $y \in T$ platí $xy = yx$. Potom totiž priradenie $(x, y) \mapsto xy$ bude izomorfizmus grúp $S \times T \cong G$. Keďže S, $T \lhd G$, platí $yx^{-1}y^{-1} \in S$, $xyx^{-1} \in T$ a

$$x \cdot yx^{-1}y^{-1} = xyx^{-1} \cdot y^{-1} \in S \cap T,$$

teda $xyx^{-1}y^{-1} = e$, z čoho už vyplýva požadovaná rovnosť.

Nasleduje niekoľko aplikácií čerstvo dokázaných výsledkov.

27.6.4. Tvrdenie. *Nech m, n sú kladné celé čísla. Potom cyklická aditívna grupa $\mathbb{Z}_{mn}$ je izomorfná s priamym súčinom $\mathbb{Z}_m \times \mathbb{Z}_n$ cyklických aditívych grúp $\mathbb{Z}_m$ a $\mathbb{Z}_n$ práve vtedy, keď m a n sú nesúdeliteľné.*

Dôkaz. Nech m, n sú nesúdeliteľné. Označme S, T podmnožiny množiny $\mathbb{Z}_{mn}$ pozostávajúce z čísel deliteľných číslom n resp. m. Zrejme S aj T sú podgrupy grupy $\mathbb{Z}_{mn}$ a platí $\mathbb{Z}_m \cong S = \{0, n, 2n, \ldots, (m-1)n\}$, $\mathbb{Z}_n \cong T = \{0, m, 2m \ldots, (n-1)m\}$. Keďže grupa $\mathbb{Z}_{mn}$ je komutatívna, rovnosť $x + y = y + x$ platí pre všetky $x, y \in \mathbb{Z}_{mn}$, a nielen pre $x \in S$, $y \in T$. Z nesúdeliteľnosti m a n vyplýva $S \cap T = \{0\}$. Podľa častí (a), (b) tvrdenia 27.6.2 je predpisom $(x, y) \mapsto x + y$ definovaný injektívny homomorfizmus grúp $S \times T \to \mathbb{Z}_{mn}$. Keďže množina S má m prvkov a množina T má n prvkov, je to injekcia mn-prvkovej množiny $S \times T$ do množiny $\mathbb{Z}_{mn}$ s rovnakým počtom prvkov, teda zároveň surjekcia. Môžeme uzavrieť, že uvedené zobrazenie je izomorfizmus grúp, preto platí

$$\mathbb{Z}_{mn} \cong S \times T \cong \mathbb{Z}_m \times \mathbb{Z}_n.$$

Nech naopak m, n sú súdeliteľné s najväčším spoločným deliteľom $d > 1$. Potom m/d, n/d aj $k = mn/d < mn$ sú celé čísla. Pre každý prvok (x, y) grupy $\mathbb{Z}_m \times \mathbb{Z}_n$ platí

$$k(x, y) = (kx, ky) = \left(\frac{n}{d}\, mx, \frac{m}{d}\, ny\right) = (0, 0),$$

lebo $mx = 0$ v $\mathbb{Z}_m$ a $ny = 0$ v $\mathbb{Z}_n$ (výrazy ako kz označujú súčet k exemplárov prvku z v príslušnej grupe). To znamená, že rád každého prvku tejto grupy je nanajvýš k, teda ostro menší než mn. Z toho dôvodu podľa tvrdenia 27.2.6 grupa $\mathbb{Z}_m \times \mathbb{Z}_n$ nemôže byť izomorfná s cyklickou grupou rádu mn.

Poznámka. Metódami do istej miery podobnými niektorým metódam z kapitoly 21 možno dokázať, že každá konečne generovaná abelovská grupa je izomorfná s priamym súčinom konečného počtu cyklických grúp. V dôsledku toho je každá konečná abelovská grupa izomorfná s priamym súčinom $\mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_k}$ pre vhodné kladné celé čísla $n_1, \ldots, n_k$. Čísla $n_1, \ldots, n_k$ možno navyše vybrať tak, že každé z nich je mocninou nejakého prvočísla (primárny kanonický tvar), prípadne tak, že každé z nich je násobkom nasledujúceho (racionálny kanonický tvar). Tieto výsledky však už výrazne presahujú rámec nášho kurzu.

Nasledujúce jednoduché dôsledky tvrdenia 27.6.4 nám poslúžia pri dôkaze záverečného tvrdenia.

27.6.5. Dôsledok. *V ľubovoľnej grupe $(G, \cdot)$ platí:*

(a) *Ak prvky $x, y \in G$ komutujú, t.j. $xy = yx$, a majú konečné po dvoch nesúdeliteľné rády m resp. n, tak prvok xy má rád mn.*

(b) *Ak prvok $x \in G$ má konečný rád mn, kde m, n sú navzájom nesúdeliteľné kladné celé čísla, tak existujú prvky $u, v \in \langle x \rangle$ rádov m resp. n také, že $x = uv$.*

(c) *Ak prvky $x, y \in G$ komutujú a majú konečné rády m resp. n, tak existuje prvok $z \in \langle x, y \rangle$, ktorého rád je najmenším spoločným násobkom čísel m a n.*

Dôkaz. (a) Rád prvku $xy \in \langle x, y \rangle \cong \langle x \rangle \times \langle y \rangle$ je zrejme rovnaký ako rád prvku $(1, 1) \in \mathbb{Z}_m \times \mathbb{Z}_n$.

(b) Cyklická podgrupa $\langle x \rangle$ grupy G je izomorfná s priamym súčinom $\mathbb{Z}_m \times \mathbb{Z}_n$; nech $\varphi\colon \mathbb{Z}_m \times \mathbb{Z}_n \to \langle x \rangle$ je nejaký izomorfizmus a $(a, b) = \varphi^{-1}(x)$. Potom rád prvku $a \in \mathbb{Z}_m$ je m a rád prvku $b \in \mathbb{Z}_n$ je n (rozmyslite si prečo). Stačí položiť $u = \varphi(a, 0)$, $v = \varphi(0, b)$.

(c) Nech $m = p_1^{\alpha_1} \ldots p_k^{\alpha_k}$, $n = p_1^{\beta_1} \ldots p_k^{\beta_k}$, kde $p_1, \ldots, p_k$ sú navzájom rôzne prvočísla a $\alpha_i, \beta_i \in \mathbb{N}$. Keďže pre $i \neq j$ sú čísla $p_i^{\alpha_i}$, $p_j^{\alpha_j}$ navzájom nesúdeliteľné, podľa (b) (ktoré možno indukciou zovšeobecniť na ľubovoľný konečný počet činiteľov) existujú prvky $u_1, \dots, u_k \in \langle x \rangle$ také, že $x = u_1 \dots u_k$ a u_i má rád $p_i^{\alpha_i}$. Z rovnakého dôvodu existujú prvky $v_1, \dots, v_k \in \langle y \rangle$ také, že $y = v_1 \dots v_k$ a v_i má rád $p_i^{\beta_i}$. Zrejme ľubovoľné dva z prvkov $u_1, \dots, u_k, v_1, \dots, v_k$ komutujú. Položme $w_i = u_i$, ak $\alpha_i \geq \beta_i$ a $w_i = v_i$, ak $\alpha_i < \beta_i$. Podľa (a) (ktoré možno takisto zovšeobecniť indukciou na ľubovoľný konečný počet činiteľov) rád prvku $z = w_1 \dots w_k$ je $p_1^{\gamma_1} \dots p_k^{\gamma_k}$, kde $\gamma_i = \max(\alpha_i, \beta_i)$, čiže najmenší spoločný násobok čísel m a n.

Taktiež časť (c) uvedeného dôsledku možno indukciou zovšeobecniť na ľubovoľný konečný počet po dvoch komutujúcich prvkov grupy. Jeden špeciálny prípad tohto zovšeobecnenia si zaslúži našu pozornosť.

27.6.6. Dôsledok. *Nech G je konečná abelovská grupa. Potom existuje prvok $z \in G$, ktorého rád je násobkom rádu každého prvku grupy G.*

Nasledujúca veta odhaľuje zaujímavú súvislosť medzi štruktúrou polí a konečných cyklických grúp.

27.6.7. Veta. *Nech K je ľubovoľné pole. Potom každá konečná podgrupa multiplikatívnej grupy $K^\star$ jeho nenulových prvkov je cyklická.*

Dôkaz. Nech G je konečná podgrupa grupy $K^\star$. Podľa dôsledku 27.6.6 existuje prvok $z \in G$, ktorého rád $n = \#\langle z \rangle$ je násobkom rádu každého prvku grupy G. Potom n delí rád G, teda $n \leq \# G$, a pre každé $a \in G$ platí $a^n = 1$. Inak povedané, každý prvok grupy G je koreňom polynómu $x^n - 1 \in K[x]$. Ale, ako sme sa už zmienili v úvode paragrafu 19.1, polynóm n-tého stupňa má v poli K najviac n koreňov (dokonca vrátane násobnosti). Preto musí platiť $\# G \leq n$, teda $\# G = n$ a $G = \langle z \rangle$ je cyklická.

27.6.8. Dôsledok. *Nech K je konečné pole. Potom multiplikatívna grupa $K^\star$ jeho nenulových prvkov je cyklická.*

Ešte poznamenajme, že generátor cyklickej multiplikatívnej grupy $K^\star$ konečného q-prvkového poľa K sa nazýva $(q-1)$-*á primitívna odmocnina z jednotky.* Takýto prvok má totiž rád $q - 1$, t. j. vyhovuje rovnici $x^{q-1} = 1$, ale žiadnej z rovníc $x^k = 1$ pre $1 \leq k < q - 1$.

27.7 Voľné a konečne prezentované grupy

Nech X je ľubovoľná množina. Zamyslime sa nad otázkou, ako vyzerá „najvšeobecnejšia“ grupa generovaná množinou X. Predovšetkým treba upresniť, čo vlastne znamená ono „najvšeobecnejšia“. Označme F túto zatiaľ bližšie neurčenú grupu. Vieme, že $X \subseteq F = \langle X \rangle$. Podľa našich predstáv by v grupe

F nemali platiť „žiadne vzťahy navyše“ okrem tých, ktoré sú nevyhnutnými logickými dôsledkami axióm (a), (b) teórie grúp, uvedených na začiatku paragrafu 27.1.

Grupa F by teda okrem prvkov množiny X mala obsahovať aj nejaký (jediný) neutrálny prvok e, a keďže nás nič nenúti, aby pre nejaké $x \in X$ platilo $x = e$, tento prvok by mal byť rôzny od každého prvku $x \in X$. S každým prvkom $x \in X$ by F mala obsahovať aj k nemu inverzný prvok x^{-1}, ktorý by mal byť z rovnakého dôvodu rôzny od každého prvku $y \in X$ aj od prvku e. Na druhej strane, axiómy teórie grúp nás nútia položiť $e^{-1} = e$ a $(x^{-1})^{-1} = x$. Pôvodná množina X sa nám tak rozrástla do zjednotenia po dvoch disjunktných množín X, $\{e\}$ a $X^{-1} = \{x^{-1};\, x \in X\}$.

Grupa F však musí byť uzavretá aj vzhľadom na súčiny, teda musí popri prvku e obsahovať aj všetky výrazu tvaru $s_1 s_2 \dots s_n$, kde $1 \le n \in \mathbb{N}$ a $s_1, \dots, s_n \in X \cup X^{-1}$. Zátvorky si môžeme dovoliť vynechať, lebo v dôsledku asociatívnosti grupy F ich umiestnenie nemá vplyv na výsledok. Keď sa však v takomto „slove“ vyskytnú vedľa seba „písmená“ x a x^{-1}, je nevyhnutné ich navzájom vykrátiť a takto pokračovať, pokiaľ sa to len dá. Napr. zo slova $u^{-1}uxxyz^{-1}zy^{-1}zu^{-1}$, kde $x, y, z, u \in X$, tak najprv dostaneme $xxyy^{-1}zu^{-1}$ a potom $xxzu^{-1}$.

Ukážeme, že do množiny F netreba ďalej nič pridávať. Ak sa navyše dohodneme, že prvok e stotožníme s prázdnou postupnosťou prvkov množiny $X \cup X^{-1}$, máme

$$F = \bigl\{s_1 \dots s_n;\, n \in \mathbb{N} \;\&\; s_1, \dots, s_n \in X \cup X^{-1} \;\&\; (\forall\, i < n)(s_i^{-1} \ne s_{i+1})\bigr\}.$$

Prvky množiny F nazývame *redukované grupové slová nad abecedou* X. Ak k po sebe nasledujúcich rovnakých znakov x resp. x^{-1} napíšeme ako mocninu x^k resp. x^{-k}, možno každé redukované slovo jednoznačne vyjadriť v tvare $x_1^{k_1} \dots x_m^{k_m}$, kde $m \in \mathbb{N}$, $x_1, \dots, x_m \in X$, $k_1, \dots, k_m \in \mathbb{Z} \smallsetminus \{0\}$ a $x_i \ne x_{i+1}$ pre $1 \le i < m$.

Súčin redukovaných slov $r_1 \dots r_m$, $s_1 \dots s_n$ definujeme v dvoch krokoch: Najprv ich jednoducho napíšeme za sebou, t. j. utvoríme slovo $r_1 \dots r_m s_1 \dots s_n$, a potom ho zredukujeme podľa skôr spomínaného receptu.

Ľahko nahliadneme, že takto definovaná operácia násobenia je asociatívna. Prázdne slovo e je naozaj redukované a je jej neutrálnym prvkom. Inverzným prvkom k redukovanému slovu $s_1 \dots s_n$ je redukované slovo $s_n^{-1} \dots s_1^{-1}$, lebo redukciou zloženého slova $s_1 \dots s_n s_n^{-1} \dots s_1^{-1}$ zrejme dostaneme prázdne slovo e. Krátko povedané, množina F s práve definovanou operáciou tvorí grupu. Nazývame ju *voľná grupa* nad množinou generátorov X a prvkom množiny X hovoríme *voľné generátory*. Aby sme zvýraznili ich úlohu, používame označenie $F = \mathrm{FG}(X)$ (z anglického *free group*).

Voľné grupy však možno charakterizovať aj iným, abstraktným spôsobom.

27.7.1. Veta. *Nech X je množina a G je ľubovoľná grupa. Potom ku každému zobrazeniu $f\colon X \to G$ existuje práve jeden homomorfizmus $\varphi\colon \mathrm{FG}(X) \to G$ taký, že $\varphi\restriction X = f$, t.j. $\varphi(x) = f(x)$ pre každé $x \in X$. Naopak, ak nejaká grupa F má tiež uvedenú vlastnosť, t.j. $X \subseteq F$ a ku každému zobrazeniu $g\colon X \to G$ existuje jediný homomorfizmus grúp $\psi\colon F \to G$ taký, že $\psi\restriction X = g$, tak $F \cong \mathrm{FG}(X)$.*

Dôkaz. Nech $f\colon X \to G$ je nejaké zobrazenie z množiny X do grupy G. Ak má byť $\varphi\colon \mathrm{FG}(X) \to G$ homomorfizmus rozširujúci f, tak pre každé (redukované) slovo $x_1^{k_1}\dots x_n^{k_n} \in \mathrm{FG}(X)$, kde $x_i \in X$ a $k_i \in \mathbb{Z}$, musí platiť

$$\varphi\bigl(x_1^{k_1}\dots x_n^{k_n}\bigr) = \varphi(x_1)^{k_1}\dots\varphi(x_n)^{k_n} = f(x_1)^{k_1}\dots f(x_n)^{k_n},$$

teda φ možno definovať len jediným spôsobom. Naopak, ľahko nahliadneme, že uvedeným predpisom je naozaj definovaný homomorfizmus grúp rozširujúci f.

Predpokladajme, že $F \supseteq X$ je grupa a ku každému zobrazeniu $g\colon X \to G$ do ľubovoľnej grupy G existuje jediný homomorfizmus $\psi\colon F \to G$ rozširujúci g. Uvažujme identitu id_X ako zobrazenie $X \to F$ aj ako zobrazenie $X \to \mathrm{FG}(X)$. Existuje jediný homomorfizmus $\varphi\colon \mathrm{FG}(X) \to F$ rozširujúci id_X, a tiež jediný taký homomorfizmus $\chi\colon F \to \mathrm{FG}(X)$. Potom aj $\varphi\circ\chi\colon F \to F$ je homomorfizmus rozširujúci id_X, rovnako ako $\chi\circ\varphi\colon \mathrm{FG}(X) \to \mathrm{FG}(X)$. Preto nevyhnutne $\varphi\circ\chi = \mathrm{id}_F$ a $\chi\circ\varphi = \mathrm{id}_{\mathrm{FG}(X)}$, teda φ, χ sú navzájom inverzné izomorfizmy.

Uvedená vlastnosť voľných grúp pripomína vektorové priestory. Aj tie sú „voľné nad svojimi bázami“. Ak totiž X je (neusporiadaná) báza vektorového priestoru V nad poľom K, tak každé zobrazenie $f\colon X \to U$ do nejakého vektorového priestoru U možno jediným spôsobom rozšíriť do lineárneho zobrazenia $\varphi\colon V \to U$. Zrejme φ zobrazí lineárnu kombináciu $c_1x_1 + \ldots + c_nx_n$, kde $x_1, \ldots, x_n \in X$, $c_1, \ldots, c_n \in K$ do lineárnej kombinácie

$$\varphi(c_1x_1 + \ldots + c_nx_n) = c_1f(x_1) + \ldots + c_nf(x_n).$$

No na rozdiel od vektorových priestorov, zďaleka nie všetky grupy, dokonca ani tie konečne generované, sú voľné. Jednako – ako ukazuje nasledujúca veta – voľné grupy sú v istom zmysle predobrazmi všetkých grúp.

27.7.2. Veta. *Každá grupa je homomorfným obrazom nejakej voľnej grupy.*

Dôkaz. Nech G je ľubovovoľná grupa a $X \subseteq G$ je nejaká množina jej generátorov (vždy môžeme položiť napr. $X = G$). Nech $F = \mathrm{FG}(X)$ je voľná grupa nad množinou X. Uvažujme identitu id_X ako zobrazenie $\mathrm{id}_X\colon X \to G$. Podľa predchádzajúcej vety existuje (dokonca jediný) homomorfizmus grúp

$\varphi\colon F \to G$ taký, že $\varphi(x) = x$ pre všetky $x \in X$. Keďže X generuje G, zrejme φ je surjekcia.

Napríklad voľná grupa s jedným generátorom x je zrejme nekonečná cyklická grupa, teda $(\mathrm{FG}(x), \cdot) \cong (\mathbb{Z}, +)$. Podľa tvrdenia 27.3.1 je každý homomorfizmus $\mathbb{Z} \to G$ jednoznačne určený obrazom jej generátora $1 \in \mathbb{Z}$, ktorým môže byť každý prvok grupy G. Homomorfné obrazy grupy $\mathbb{Z}$ – grupy $\mathbb{Z}_n$, kde $2 \le n \in \mathbb{N}$, – však voľné nie sú. Voľné grupy s viac než jedným generátorom už nie sú komutatívne a majú podstatne zložitejšiu štruktúru. Nech $X \subseteq G$ je nejaká množina, ktorá generuje grupu G. Uvažujme (jednoznačne určený) surjektívny homomorfizmus $\varphi\colon \mathrm{FG}(X) \to G$ rozširujúci identické zobrazenie $\mathrm{id}_X\colon X \to G$. Potom jeho jadro $N = \operatorname{Ker}\varphi$ je normálna podgrupa voľnej grupy $\mathrm{FG}(X)$. Ľubovoľný prvok $s_1 \dots s_n \in N$ predpisuje rovnosť $s_1 \dots s_n = e$, ktorá musí platiť v grupe G. Ak napr. $xyx^{-1}y^{-1} \in N$, znamená to, že v G platí $xyx^{-1}y^{-1} = e$, t. j. generátory $x, y \in X$ komutujú. Podobne, ak $xxx \in N$, tak v G platí $x^3 = e$, t. j. generátor x má rád 3, prípadne 1. Keby sme vypísali všetky slová, ktoré sú prvkami grupy N, dostali by sme tak zoznam všetkých redukovaných grupových slov nad abecedou X, ktoré sa v grupe G rovnajú neutrálnemu prvku e. Keďže každú rovnosť medzi slovami $r_1 \dots r_m = s_1 \dots s_n$ možno prepísať na ekvivalentný tvar $r_1 \dots r_m s_n^{-1} \dots s_1^{-1} = e$, grupa N tak v sebe skrýva všetky možné rovnosti medzi (redukovanými) grupovými slovami nad abecedou X, ktoré platia v G.

Aby sme však mohli popísať grupu $G \cong \mathrm{FG}(X)/N$, nepotrebujeme explicitne poznať všetky rovnosti $s_1 \dots s_n = e$ spomínaného tvaru, ktoré platia v G. Stačí mať k dispozícii nejaký menší zoznam takých rovností, z ktorých všetky ostatné rovnosti v G logicky vyplývajú. Inak povedané, nepotrebujeme zoznam všetkých prvkov grupy $N \lhd \mathrm{FG}(X)$, stačí nejaká jej podmnožina $E \subseteq N$, taká, že $\langle\!\langle E \rangle\!\rangle = N$. Potom rovnosť $r_1 \dots r_m = s_1 \dots s_n$ platí v G práve vtedy, keď vyplýva z množiny rovností E, t. j. práve vtedy, keď $r_1 \dots r_m s_n^{-1} \dots s_1^{-1} \in \langle\!\langle E \rangle\!\rangle$.

Faktorovú grupu $\mathrm{FG}(X)/\langle\!\langle E \rangle\!\rangle$ značíme $\langle X \mid E \rangle$ a nazývame ju *grupa prezentovaná množinou generátorov X a množinou slov E*. $\langle X \mid E \rangle$ tak predstavuje „najvšeobecnejšiu“ grupu generovanú množinou X, v ktorej platia rovnosti $s_1 \dots s_n = e$, kde $s_1 \dots s_n \in E$. Akákoľvek iná rovnosť medzi slovami zostavenými z generátorov (t. j. z prvkov množiny X) v nej platí, len ak je logickým dôsledkom rovností z E a axióm teórie grúp – inak nie. Špeciálne $\mathrm{FG}(X) = \langle X \mid \emptyset \rangle$.

Niekedy slová $w \in E$ zapisujeme ako rovnosti $w = e$, prípadne slová tvaru $uv^{-1} \in E$ ako rovnosti $u = v$. Napr.

$$\big\langle x \mid x^n \big\rangle = \big\langle x \mid x^n = e \big\rangle,$$

predstavuje cyklickú grupu rádu n, teda izomorfnú s aditívnou grupou $\mathbb{Z}_n$. Podobne

$$\big\langle x, y \mid xyx^{-1}y^{-1} \big\rangle = \langle x, y \mid xy = yx \rangle$$

predstavuje „voľnú abelovskú grupu“ s dvoma generátormi. Možno ukázať, že je izomorfná s priamym súčinom $\mathbb{Z} \times \mathbb{Z}$.

Grupy tvaru $\langle X \mid E \rangle$, kde X aj $E \subseteq \mathrm{FG}(X)$ sú *konečné* množiny, nazývame *konečne prezentované grupy.* Ak $X = \{x_1, \ldots, x_n\}$, $E = \{w_1, \ldots, w_k\}$, píšeme

$$\langle X \mid E \rangle = \langle x_1, \ldots, x_n \mid w_1, \ldots, w_k \rangle.$$

Zrejme každá konečná grupa je konečne prezentovaná (rozmyslite si prečo). Ako sme však videli pred chvíľou, konečne prezentovaná grupa môže byť nekonečná.

Prezentácie grúp hrajú dôležitú úlohu pri popise na nich definovaných homomorfizmov.

27.7.3. Veta. *Nech $\langle X \mid E \rangle$ je prezentácia grupy G a H je ľubovoľná grupa. Potom zobrazenie $f\colon X \to H$ možno rozšíriť do homomorfizmu $\varphi\colon G \to H$ práve vtedy, keď pre každé $x_1^{k_1} \ldots x_n^{k_n} \in E$, kde $x_1, \ldots, x_n \in X$, $k_1, \ldots, k_n \in \mathbb{Z}$, platí $f(x_1)^{k_1} \ldots f(x_n)^{k_n} = e_H$. V tom prípade je homomorfizmus φ rozširujúci f určený jednoznačne.*

Dôkaz. Ak φ má byť homomorfizmus, tak jediný možný spôsob ako ho definovať zrejme je

$$\varphi\big(x_1^{k_1} \ldots x_n^{k_n}\big) = \varphi(x_1)^{k_1} \ldots \varphi(x_n)^{k_n} = f(x_1)^{k_1} \ldots f(x_n)^{k_n}$$

pre každý prvok $x_1^{k_1} \ldots x_n^{k_n}$ grupy G. Ľahko možno overiť, že takto definované zobrazenie je homomorfizmus grúp práve vtedy, keď $f(x_1)^k_1 \ldots f(x_n)^k_n = e_H$ pre každé $x_1^{k_1} \ldots x_n^{k_n} \in E$.

Ak si napríklad uvedomíme, že $\mathbb{Z}_n \cong \langle x \mid x^n \rangle$, pričom úlohu generátora x hrá prvok $1 \in \mathbb{Z}_n$, hneď vidíme, že priradenie $1 \mapsto a \in H$ možno rozšíriť do homomorfizmu grúp $\varphi\colon \mathbb{Z}_n \to H$ práve vtedy, keď v H platí $a^n = e$, čiže rád prvku a je deliteľom čísla n. Taktiež naopak, každý homomorfizmus $\varphi\colon \mathbb{Z}_n \to H$ je jednoznačne určený obrazom $a = \varphi(1)$ jediného prvku $1 \in \mathbb{Z}_n$. Úlohu prvku 1 môže samozrejme zohrať ľubovoľný generátor cyklickej grupy $\mathbb{Z}_n$, t. j. ľubovoľný prvok $k \in \mathbb{Z}_n$ nesúdeliteľný s n.

27.7.4. Príklad. *Metacyklickou grupou* nazývame každú konečne prezentovanú grupu tvaru

$$F_{mn}^k = \big\langle x, y \mid x^m = y^n = e,\ xyx^{-1} = y^k \big\rangle,$$

kde m, n sú prirodzené čísla a $k \in \mathbb{Z}_n$ je nesúdeliteľné s n také, že $k^m \equiv_n 1$. Pripúšťame aj prípady $m = 0$ alebo $n = 0$; vtedy kladieme $\mathbb{Z}_0 = \mathbb{Z}$, navyše podmienka nesúdeliteľnosti $k \in \mathbb{Z}$ s $n = 0$ znamená $k = \pm 1$. Grupy F_{mn}^k patria do širšej triedy grúp nazývaných *Frobeniovými grupami.*

Ukážeme, že každý prvok grupy F_{mn}^k má tvar $y^b x^a$ pre jednoznačne určené $a \in \mathbb{Z}_m$, $b \in \mathbb{Z}_n$. Uvedená prezentácia totiž umožňuje vypočítať všetky súčiny $y^b x^a \cdot y^d x^c$ a previesť ich na súčin vhodných mocnín generátorov y a x. Ak si totiž uvedomíme, že pre prvky ľubovoľnej grupy platí $gug^{-1} \cdot gvg^{-1} = guvg^{-1}$ a $(gh)u(gh)^{-1} = g(huh^{-1})g^{-1}$, môžeme počítať

$$y^b x^a \cdot y^d x^c = y^b \cdot x^a y^d x^{-a} \cdot x^{a+c} = y^b (x^a y x^{-a})^d x^{a+c} = y^b y^{k^a d} x^{a+c} = y^{b+k^a d} x^{a+c},$$

lebo $xyx^{-1} = y^k$ a $x^2 y x^{-2} = x y^k x^{-1} = (xyx^{-1})^k = (y^k)^k = y^{k^2}, \ldots,$ $x^a y x^{-a} = y^{k^a}$.

Pre $m = 2$, $k = n - 1$ dostávame dihedrálnu grupu Δ_n z príkladu 27.2.2. Keďže pre generátor y rádu n je $y^{n-1} = y^{-1}$, môžeme jej prezentáciu prepísať do obvyklej podoby

$$\Delta_n \cong \langle x, y \mid x^2 = y^n = e,\ xyx^{-1} = y^{-1} \rangle.$$

Zrejme v uvedenej prezentácii zodpovedá generátor x osovej súmernosti podľa niektorej z osí pravidelného n-uholníka a y otočeniu okolo jeho stredu o uhol $2\pi/n$.

27.7.5. Príklad. Nech p, q sú ľubovoľné prvočísla. Keďže podľa dôsledku 27.6.8 multiplikatívna grupa $\mathbb{Z}_q^\star$ poľa $\mathbb{Z}_q$ je cyklická rádu $q - 1$, prvok $k \in \mathbb{Z}_q^\star$ rádu p (t. j. $k \in \mathbb{Z}_q$, $k^p \equiv_q 1$ a $k \neq 1$) existuje práve vtedy, keď p delí $q - 1$. V takom prípade $p < q$ a

$$F_{pq}^k = \langle x, y \mid x^p = y^q = e,\ xyx^{-1} = y^k \rangle$$

je nekomutatívna grupa rádu pq. Ukážeme, že metacyklická grupa F_{pq}^k nezávisí na k; presnejšie, ak $l \in \mathbb{Z}_q^\star$ je iný prvok rádu p, tak $F_{pq}^k \cong F_{pq}^l$.

Množina $A = \{a \in \mathbb{Z}_q^\star;\ a^p = 1\}$ tvorí cyklickú podgrupu grupy $\mathbb{Z}_q^\star$ rádu p, preto $l \in A = \langle k \rangle$. Teda existuje $1 \le r \le p - 1$ také, že v $\mathbb{Z}_q$ platí $l = k^r$, čiže $l \equiv_q k^r$. Potom pre $z = x^r \in F_{pq}^k$ máme $\langle z \rangle = \langle x \rangle$, $z^p = e$ a

$$zyz^{-1} = x^r y x^{-r} = y^{k^r} = y^l.$$

To však podľa vety 27.7.3 znamená, že priradenie $x \mapsto x^r$, $y \mapsto y$ možno rozšíriť do homomorfizmu grupy

$$F_{pq}^l = \langle x, y \mid x^p = y^q = e,\ xyx^{-1} = y^l \rangle$$

do grupy F_{pq}^k. Čitateľ by si mal samostatne premyslieť, že ide naozaj o izomorfizmus.

27.8 Grupy homomorfizmov a charaktery abelovských grúp

V paragrafe 6.5 sme videli, že všetky lineárne zobrazenia z vektorového priestoru V do vektorového priestoru U nad daným poľom K opäť tvoria vektorový priestor $\mathcal{L}(V,U)$. Niečo podobné platí aj pre abelovské grupy.

Množinu všetkých homomorfizmov grupy G do grupy H označujeme $\mathrm{Hom}(G,H)$. Ak $\varphi,\psi \in \mathrm{Hom}(G,H)$, tak ich súčin po zložkách, t. j. zobrazenie $\varphi\psi\colon G \to H$ dané predpisom $(\varphi\psi)(x) = \varphi(x)\,\psi(x)$ pre $x \in G$, nemusí byť homomorfizmom. Situácia sa však zmení, ak grupa H je komutatívna.

Jednoduchý dôkaz nasledujúceho tvrdenia prenechávame čitateľovi ako cvičenie.

27.8.1. Tvrdenie. *Nech G je ľubovoľná grupa a H je abelovská grupa. Potom množina* $\mathrm{Hom}(G,H)$ *s operáciou súčinu po zložkách je abelovská grupa.*

Ďalej sa budeme zaoberať výlučne grupami homomorfizmov medzi abelovskými grupami.

27.8.2. Tvrdenie. *Nech A, B_1, B_2, resp. A_1, A_2, B sú abelovské grupy. Potom*
(a) $\mathrm{Hom}(A, B_1 \times B_2) \cong \mathrm{Hom}(A,B_1) \times \mathrm{Hom}(A,B_2)$;
(b) $\mathrm{Hom}(A_1 \times A_2, B) \cong \mathrm{Hom}(A_1,B) \times \mathrm{Hom}(A_2,B)$;

Dôkaz. len naznačíme. Budeme používať aditívny zápis grupových operácií.

(a) Označme $\pi_1\colon B_1 \times B_2 \to B_1$, $\pi_2\colon B_1 \times B_2 \to B_2$ projekcie priameho súčinu na prvý resp. druhý faktor, t. j. $\pi_1(b_1,b_2) = b_1$, $\pi_2(b_1,b_2) = b_2$ pre $(b_1,b_2) \in B_1 \times B_2$. Zrejme sú to surjektívne homomorfizmy. Niekoľkými priamočiarymi výpočtami sa možno presvedčiť, že priradením $\varphi \mapsto (\pi_1\circ\varphi, \pi_2\circ\varphi)$ je určený izomofizmus grúp $\mathrm{Hom}(A, B_1 \times B_2) \to \mathrm{Hom}(A,B_1) \times \mathrm{Hom}(A,B_2)$.

(b) Označme $\iota_1\colon A_1 \to A_1 \times A_2$, $\iota_2\colon A_2 \to A_1 \times A_2$ kanonické vnorenia prvého resp. druhého činiteľa do ich priameho súčinu, t. j. $\iota_1(a_1) = (a_1,0)$, $\iota_2(a_2) = (0,a_2)$ pre $a_1 \in A_1$, $a_2 \in A_2$. Zrejme sú to injektívne homomorfizmy. Opäť sa sa možno ľahko presvedčiť, že zobrazenie $\varphi \mapsto (\varphi\circ\iota_1, \varphi\circ\iota_2)$ je izomorfizmus grúp $\mathrm{Hom}(A_1 \times A_2, B) \to \mathrm{Hom}(A_1,B) \times \mathrm{Hom}(A_2,B)$.

Zvláštnu pozornosť si zaslúžia grupy homomorfizmov do (abelovskej) grupy komplexných jednotiek $\mathrm{U}(1)$. Jedným z dôvodov je ich súvis s diskrétnou Fourierovou transformáciou, ktorý si bližšie všimneme v paragrafe 30.2. *Charakterom abelovskej grupy G* nazývame ľubovoľný homomorfizmus $\alpha\colon G \to \mathrm{U}(1)$ a grupe

$$G^{\mathrm{d}} = \mathrm{Hom}(G, \mathrm{U}(1))$$

hovoríme *grupa charakterov* alebo tiež *duálna grupa* abelovskej grupy G. Jednotkou v grupe G^{d} je *triviálny charakter* $\varepsilon\colon G \to \mathrm{U}(1)$, t. j. $\varepsilon(x) = 1$ pre každé $x \in G$, a inverzný prvok k charakteru $\alpha \in G^{\mathrm{d}}$ je k nemu komplexne združený charakter $\alpha^{-1} = \overline{\alpha}$.

Tvrdenie 27.8.2 (b) má nasledujúci dôsledok.

27.8.3. Dôsledok. *Pre ľubovoľné abelovské grupy G_1, G_2 platí*

$$(G_1 \times G_2)^{\mathrm{d}} \cong G_1^{\mathrm{d}} \times G_2^{\mathrm{d}}.$$

27.8.4. Tvrdenie. *Pre ľubovoľné $n \geq 1$ je $\mathbb{Z}_n^{\mathrm{d}} \cong \mathbb{Z}_n$.*

Dôkaz. Keďže grupa $\mathbb{Z}_n$ je cyklická s generátorom 1, ľubovoľný charakter $\alpha\colon \mathbb{Z}_n \to \mathrm{U}(1)$ je podľa vety 27.7.3 jednoznačne určený svojou hodnotou $\alpha(1) \in \mathrm{U}(1)$. Naopak, číslo $c \in \mathrm{U}(1)$ môže byť hodnotou $c = \alpha(1)$ charakteru $\alpha\colon \mathbb{Z}_n \to \mathrm{U}(1)$ práve vtedy, keď vyhovuje binomickej rovnici $c^n = 1$. Možné hodnoty $\alpha(1)$ tak tvoria podgrupu

$$C_n = \{c \in \mathrm{U}(1);\ c^n = 1\} = \left\{\mathrm{e}^{2\pi \mathrm{i} a/n};\ a \in \mathbb{Z}_n\right\}$$

grupy $\mathrm{U}(1)$ izomorfnú so $\mathbb{Z}_n^{\mathrm{d}}$. Zrejme $a \mapsto \mathrm{e}^{2\pi \mathrm{i} a/n}$ je izomorfizmus $\mathbb{Z}_n \to C_n$, takže $\mathbb{Z}_n^{\mathrm{d}} \cong C_n \cong \mathbb{Z}_n$.

Na základe faktu, že každá konečná abelovská grupa G je izomorfná s nejakým priamym súčinom cyklických grúp $\mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_k}$, ktorý sme nútení prijať bez dôkazu, tak z dôsledku 27.8.3 a z tvrdenia 27.8.4 vyplýva nasledujúci výsledok.

27.8.5. Veta. *Pre ľubovoľnú konečnú abelovskú grupu G platí $G^{\mathrm{d}} \cong G$.*

Charaktery grupy $G = \mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_k}$ už vieme explicitne popísať: každý z nich má tvar

$$\alpha(x) = c_1^{x_1} \ldots c_k^{x_k} = \mathrm{e}^{2\pi \mathrm{i} a_1 x_1/n_1} \ldots \mathrm{e}^{2\pi \mathrm{i} a_k x_k/n_k} = \exp\left(2\pi \mathrm{i} \sum_{j=1}^{k} \frac{a_j x_j}{n_j}\right),$$

kde $x = (x_1, \ldots, x_k) \in G$ a $c_1, \ldots, c_k \in \mathrm{U}(1)$ sú pevné čísla také, že $c_j^{n_j} = 1$, t. j. $c_j = \mathrm{e}^{2\pi \mathrm{i} a_j/n_j}$ pre nejaké $a_j \in \mathbb{Z}_{n_j}$. Potom priradenie $\alpha \mapsto (c_1, \ldots, c_k)$ určuje izomorfizmus grúp $(\mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_k})^{\mathrm{d}} \cong C_{n_1} \times \ldots \times C_{n_k}$ a priradenie $\alpha \mapsto (a_1, \ldots, a_k)$ izomorfizmus grúp $(\mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_k})^{\mathrm{d}} \cong \mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_k}$.

Podotknime, že vetu 27.8.5 nemožno rozšíriť na nekonečné abelovské grupy. Napr. pre grupu celých čísel $\mathbb{Z}$ zrejme platí

$$\mathbb{Z}^{\mathrm{d}} \cong \mathrm{U}(1) \not\cong \mathbb{Z}.$$

Ďalej poznamenajme, že izomorfizmus $\mathbb{Z}_n \cong \mathbb{Z}_n^{\mathrm{d}}$ z tvrdenia 27.8.4 nie je *kanonický*. Závisí totiž od voľby generátora cyklickej grupy $\mathbb{Z}_n$. Tým môže byť popri prvku 1 ľubovoľný prvok $a \in \mathbb{Z}_n$ nesúdeliteľný s n. Z podobných dôvodov nie je kanonický izomorfizmus $G \cong G^{\mathrm{d}}$ z vety 27.8.5.

Konečná abelovská grupa G je však *kanonicky izomorfná* so svojim *druhým duálom* G^{dd} (a tento výsledok už možno – ale až s využitím vhodných topologických pojmov – zovšeobecniť aj na niektoré nekonečné, tzv. *lokálne kompaktné* abelovské grupy).

Ľahko nahliadneme, že pre každý prvok x abelovskej grupy G je predpisom

$$x^{\Delta}(\alpha) = \alpha(x),$$

kde $\alpha \in G^{\mathrm{d}}$, definovaný homomorfizmus $x^{\Delta}\colon G^{\mathrm{d}} \to \mathrm{U}(1)$. To znamená, že $x^{\Delta} \in G^{\mathrm{dd}}$ a $x \mapsto x^{\Delta}$ je zobrazenie $G \to G^{\mathrm{dd}}$.

27.8.6. Veta. *Nech G je konečná abelovská grupa. Potom $x \mapsto x^{\Delta}$ je izomorfizmus grúp $G \cong G^{\mathrm{dd}}$.*

Dôkaz. Označme $\varDelta(x) = x^{\Delta}$. Priamym výpočtom, ktorý prenechávame čitateľovi, možno overiť, že $\varDelta\colon G \to G^{\mathrm{dd}}$ je homomorfizmus. Keďže $G \cong G^{\mathrm{d}} \cong G^{\mathrm{dd}}$, všetky tieto grupy majú rovnaký počet prvkov. Stačí preto dokázať, že $\varDelta$ je injektívne, čiže $\operatorname{Ker}\varDelta = \{0\}$.

Zvoľme $x \in G$, $x \neq 0$. Keďže G je konečná, stačí sa zaoberať prípadom, keď $G = \mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_k}$ je priamym súčinom cyklických grúp. Potom $x = (x_1, \ldots, x_k)$, kde $x_j \in \mathbb{Z}_{n_j}$. Nakoľko $x \neq (0, \ldots, 0)$, $x_j \neq 0$ pre nejaké $j \leq k$. Definujme $\alpha \in G^{\mathrm{d}}$ predpisom $\alpha(y) = \mathrm{e}^{2\pi \mathrm{i} y_j / n_j}$ pre $y = (y_1, \ldots, y_k) \in G$. Zrejme $\alpha(x) \neq 1$, preto

$$\varDelta(x)(\alpha) = x^{\Delta}(\alpha) = \alpha(x) \neq 1,$$

t. j. $x^{\Delta} = \varDelta(x)$ nie je jednotkový prvok grupy G^{dd} (triviálny charakter $G^{\mathrm{d}} \to \mathrm{U}(1)$) a $x \notin \operatorname{Ker}\varDelta$. Teda $\operatorname{Ker}\varDelta = \{0\}$.

Na základe uvedeného kanonického izomorfizmu stotožníme konečnú abelovskú grupu G s grupou G^{dd} charakterov jej duálnej grupy G^{d}. Každý prvok $x = x^{\Delta}$ grupy G tak zároveň považujeme za charakter grupy G^{d} a pre $\alpha \in G^{\mathrm{d}}$ máme $x(\alpha) = \alpha(x)$.

Na budúce použitie ešte zaznamenáme jeden technický výsledok.

27.8.7. Lema. *Nech G je konečná abelovská grupa a $\alpha \in G^{\mathrm{d}}$. Potom*

$$\sum_{x \in G} \alpha(x) = \begin{cases} 0, & ak\ \alpha \neq \varepsilon, \\ \#\, G, & ak\ \alpha = \varepsilon. \end{cases}$$

Dôkaz. Opäť stačí uvažovať prípad $G = \mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_k}$. Ľubovoľné $\alpha \in G^{\mathrm{d}}$ má tvar
$$\alpha(x) = c_1^{x_1} \ldots c_k^{x_k},$$
kde $x = (x_1, \ldots, x_k) \in G$ a $c_1, \ldots, c_k \in \mathrm{U}(1)$ sú pevné čísla také, že $c_j^{n_j} = 1$. Zrejme $\alpha = \varepsilon$ práve vtedy, keď $c_1 = \ldots = c_k = 1$. Počítajme
$$\sum_{x \in G} \alpha(x) = \sum_{x_1=0}^{n_1-1} \ldots \sum_{x_k=0}^{n_k-1} c_1^{x_1} \ldots c_k^{x_k} = \left(\sum_{x_1=0}^{n_1-1} c_1^{x_1}\right) \ldots \left(\sum_{x_k=0}^{n_k-1} c_k^{x_k}\right).$$
V poslednom výraze je každý jednotlivý činiteľ súčtom prvých n_j členov geometrickej postupnosti s kvocientom c_j, t. j.
$$\sum_{x_j=0}^{n_j-1} c_j^{x_j} = \begin{cases} \dfrac{c_j^{n_j} - 1}{c_j - 1} = 0, & \text{ak } c_j \neq 1, \\ \displaystyle\sum_{x_j=0}^{n_j-1} 1 = n_j, & \text{ak } c_j = 1. \end{cases}$$
V dôsledku toho
$$\sum_{x \in G} \alpha(x) = \begin{cases} 0, & \text{ak } \alpha \neq \varepsilon, \\ n_1 \ldots n_k = \# G, & \text{ak } \alpha = \varepsilon. \end{cases}$$

Cvičenia

27.1. Pojem grupy možno definovať viacerými alternatívnymi spôsobmi.

(a) Nech G je neprázdna množina s asociatívnou binárnou operáciou $\cdot$. Dokážte, že nasledujúce podmienky sú ekvivalentné:

(i) $(G, \cdot)$ je grupa;

(ii) v $(G, \cdot)$ platia pravidlá o krátení zľava $(\forall\, x, y, z \in G)(xy = xz \Rightarrow y = z)$ a sprava $(\forall\, x, y, z \in G)(xy = zy \Rightarrow x = z)$ a pre každé $a, b \in G$ majú rovnice $ax = b$, $ya = b$ aspoň jedno riešenie, t. j. $(\forall\, a, b \in G)(\exists\, x, y \in G)(ax = b \;\&\; ya = b)$;

(iii) pre ľubovoľné $a, b \in G$ majú rovnice $ax = b$, $ya = b$ jediné riešenie v G, t. j. $(\forall\, a, b \in G)(\exists!\, x, y \in G)(ax = b \;\&\; ya = b)$;

(iv) $(G, \cdot)$ má ľavú jednotku a každý prvok v G má ľavý inverzný, t. j. $(\exists\, e \in G)(\forall\, x \in G)\big(ex = x \;\&\; (\exists\, x' \in G)(x'x = e)\big)$;

(v) $(G, \cdot)$ má pravú jednotku a každý prvok v G má pravý inverzný, t. j. $(\exists\, e \in G)(\forall\, x \in G)\big(xe = x \;\&\; (\exists\, x' \in G)(xx' = e)\big)$.

(*Návod*: V dôkaze implikácie (iv) $\Rightarrow$ (v), dokážte najprv, že ľavá jednotka v G je zároveň pravou jednotkou, a potom, že ľavý inverzný prvok k $x \in G$ je zároveň pravý inverzný prvok k x.)

(b) Na základe (iv) resp. (v) ukážte, že aj z podmienok (ii) resp. (iii) možno niečo vynechať, a $(G, \cdot)$ spĺňajúca takéto oslabené podmienky bude stále ešte grupou. Skúste nájsť viac možností.

(c) Nech G je neprázdna množina s asociatívnou binárnou operáciou $\cdot$, ktorá spĺňa pravidlo o krátení zľava, má ľavú jednotkou a ku každému prvku existuje pravý inverzný, t. j. $(\exists\, e \in G)(\forall\, x \in G)\big(ex = x \;\&\; (\exists\, x' \in G)(xx' = e)\big)$. Zostrojte príklad, z ktorého vyplýva, že $(G, \cdot)$ aj tak nemusí byť nevyhnutne grupa.
(*Návod*: Operáciu $\cdot$ zadefinujte tak, aby *každý* prvok množiny G bol ľavou jednotkou.)

27.2. (a) Dokážte tvrdenie 27.1.5.

(b) Doplňte vynechané podrobnosti v dôkazoch tvrdení 27.2.5, 27.2.6.

(c) Dokážte tvrdenia 27.3.2 – 27.3.6.

27.3. (a) V abelovskej grupe pre ľubovoľné prvky a, b a celé číslo n platí $(ab)^n = a^n b^n$. Dokážte, že uvedené tvrdenie platí v ľubovoľnej grupe za predpokladu, že prvky a, b komutujú.

(b) Nájdite príklad neabelovskej grupy a v nej prvkov a, b takých, že $(ab)^2 \neq a^2 b^2$.

27.4. Neprázdna podmnožina konečnej grupy $(G, \cdot)$ je jej podgrupou práve vtedy, keď je uzavretá na grupový súčin. Dokážte.

27.5. Dokážte nasledujúce tvrdenia o cyklických grupách:

(a) Každá podgrupa cyklickej grupy je cyklická.

(b) Každý homomorfný obraz cyklickej grupy je cyklická grupa.

(c) Grupa G je cyklická práve vtedy, keď je homomorfným obrazom grupy $(\mathbb{Z}, +)$.

(d) Ak celé číslo $k \geq 1$ delí rád konečnej *cyklickej* grupy G, tak G má práve jednu podgrupu rádu k.

27.6. Dokážte nasledujúce zovšeobecnenie Lagrangeovej vety: Ak S, T sú podgrupy konečnej grupy G a $S \subseteq T$, tak $[G : S] = [G : T]\,[T : S]$.

27.7. Nájdite všetky podgrupy grúp $\mathbb{Z}_4$ a a $\mathbb{Z}_2 \times \mathbb{Z}_2$. Odvoďte z toho, že uvedené grupy nie sú izomorfné.

27.8. (a) Zostrojte multiplikatívne tabuľky grúp Δ_4 a Δ_5.

(b) Nájdite všetky podgrupy týchto grúp. Ktoré z nich sú normálne?

27.9. Nech G je grupa a S, T sú jej podgrupy. Označme $S \vee T = \langle S \cup T \rangle$ podgrupu grupy G generovanú množinou $S \cup T$, t. j. najmenšiu podgrupu grupy G, ktorá obsahuje S aj T. Dokážte dve nasledujúce tvrdenia:

(a) Ak pre podgrupy S, T platí $ST = TS$, tak $S \vee T = ST = TS$.

(b) Ak aspoň jedna z podgrúp S, T je normálna, tak $ST = TS$, teda tiež $S \vee T = ST = TS$.

(c) Na príklade podgrúp vhodnej dihedrálnej grupy Δ_n ukážte, že množina ST nemusí byť podgrupou grupy G.

27.10. (a) Všetky otočenia roviny $\mathbb{R}^2$ okolo počiatku tvoria podgrupu grupy $\big(\mathcal{S}(\mathbb{R}), \circ\big)$ všetkých bijekcií $\mathbb{R} \to \mathbb{R}$ s operáciou kompozície zobrazení, izomorfnú s multiplikatívnou grupou všetkých komplexných jednotiek $\mathrm{U}(1)$. Dokážte.

(b) Dokážte, že všetky otočenia roviny $\mathbb{R}^2$ okolo počiatku o celočíselný počet stupňov tvoria podgrupu grupy všetkých otočení roviny okolo počiatku. Pre ktoré n je táto grupa izomorfná s cyklickou grupou $\mathbb{Z}_n$?

27.11. Dokážte postupne nasledujúce tvrdenia:

(a) Každá *konečná* podgrupa multiplikatívnej grupy $(\mathbb{C}^\star, \cdot)$ je nevyhnutne podgrupou grupy komplexných jednotiek $\mathrm{U}(1) \subseteq \mathbb{C}^\star$.

(b) Každá konečná podgrupa S grupy $\mathrm{U}(1)$ je cyklická, teda izomorfná s grupou $\mathbb{Z}_n$ pre $n = \# S$.

(c) Množina $\mathrm{e}^{2\pi\mathrm{i}\mathbb{Q}} = \left\{\mathrm{e}^{2\pi\mathrm{i}q};\ q \in \mathbb{Q}\right\}$ je podgrupou grupy $\mathrm{U}(1)$.

(d) Každá konečná podgrupa grupy $\mathrm{U}(1)$, teda aj $\mathbb{C}^\star$, je podgrupou grupy $\mathrm{e}^{2\pi\mathrm{i}\mathbb{Q}}$.

(e) Ak $a \in \mathbb{R}$ je iracionálne číslo, tak cyklická podgrupa $\left\langle \mathrm{e}^{2\pi\mathrm{i}a}\right\rangle$ grupy $\mathrm{U}(1)$ je nekonečná, teda izomorfná s grupou $(\mathbb{Z}, +)$.

(f) Podgrupa $S = \left\langle \mathrm{e}^{2\pi\mathrm{i}a}\right\rangle$, kde $a \in \mathbb{R} \smallsetminus \mathbb{Q}$, je navyše *hustá* v $\mathrm{U}(1)$, t. j. pre ľubovoľné $x \in \mathrm{U}(1)$ a $\varepsilon > 0$ existuje $s \in S$ také, že $|x - s| < \varepsilon$.

(g) Podgrupa $\mathrm{e}^{2\pi\mathrm{i}\mathbb{Q}}$ je takisto hustá v $\mathrm{U}(1)$.

27.12. Nech $N \lhd G$ je normálna podgrupa grupy G a $M \lhd N$ je normálna podgrupa grupy N. Vyplýva z toho, že M je normálna podgrupa grupy G?

27.13. Nech G je konečná grupa a H je jej podgrupa indexu 2. Dokážte, že H je normálna. Odvoďte z toho, že pre každé $n \geq 2$ je alternujúca grupa $\mathcal{A}_n$ normálnou podgrupou symetrickej grupy $\mathcal{S}_n$. Viete tento výsledok dokázať aj iným spôsobom?

27.14. Sformulujeme a dokážeme *Eulerovu vetu*, ktorá je zovšeobecnením malej Fermatovej vety (dôsledok 27.4.3).

(a) Ak $n \geq 2$ je celé číslo, tak prvky množiny $\mathbb{Z}_n$ nesúdeliteľné s n tvoria grupu vzhľadom na násobenie v $\mathbb{Z}_n$. Dokážte. Túto grupu značíme $(\mathbb{Z}_n^\star, \cdot)$ a jej rád $\phi(n)$. Funkciu $\phi\colon \mathbb{N} \smallsetminus \{0\} \to \mathbb{N} \smallsetminus \{0\}$ (doplnenú o hodnotu $\phi(1) = 1$) nazývame *Eulerovou funkciou.*

(b) Dokážte, že $\phi(p) = p - 1$ pre každé prvočíslo p.

(c) Nech $n = p_1^{m_1} \dots p_k^{m_k}$ je rozklad čísla n na prvočinitele (p_i sú rôzne prvočísla a $m_i \geq 1$). Odvoďe explicitné vyjadrenie $\phi(n) = p_1^{m_1 - 1}(p_1 - 1) \dots p_k^{m_k - 1}(p_k - 1)$.

(d) Dokážte Eulerovu vetu: Ak $n \geq 2$ je celé číslo, x je celé číslo nesúdeliteľné s n, tak $x^{\phi(n)} \equiv_n 1$.

(e) Počet prvkov cyklickej grupy rádu n, ktorých rád je rovný danému deliteľovi d čísla n je rovný číslu $\phi(d)$. Dokážte.

27.15. Nech $\varphi\colon G \to H$ je homomorfizmus grúp a $X \subseteq G$ je ľubovoľná množina.

(a) Dokážte rovnosť $\varphi\langle X\rangle = \langle \varphi X\rangle$.

(b) Dokážte inklúziu $\varphi\langle\!\langle X\rangle\!\rangle \subseteq \langle\!\langle \varphi X\rangle\!\rangle$ a na príklade ukážte, že rovnosť nemusí nastať.

(c) Dokážte, že ak homomorfizmus φ je surjektívny, tak $\varphi\langle\!\langle X\rangle\!\rangle = \langle\!\langle \varphi(X)\rangle\!\rangle$.

(d) Odvoďte z (b), že homomorfný obraz $\varphi(N)$ normálnej podgrupy $N \lhd G$ nemusí byť normálnou podgrupou v H.

(e) Odvoďte z (c), že homomorfný obraz $\varphi(N)$ normálnej podgrupy $N \lhd G$ v *surjektívnom* homomorfizme φ je normálna podgrupa v H.

(f) Predpokladajme, že φ je surjektívny homomorfizmus. Dokážte, že zobrazenie $N \mapsto \varphi(N)$ je bijekcia z množiny $\mathcal{N}_\varphi = \{N \lhd G;\, \operatorname{Ker}\varphi \subseteq N\}$ na množinu všetkých normálnych podgrúp grupy H a pre ľubovoľné $M, N \in \mathcal{N}_\varphi$ platí $M \subseteq N \Leftrightarrow \varphi(M) \subseteq \varphi(N)$.

27.16. Nech $N \lhd G$ je normálna podgrupa grupy G. Overte $Nx \cdot Ny = N(xy)$, $(Nx)^{-1} = Nx^{-1}$ ako *rovnosti množín.*

27.17. Zmodifikujte definíciu homomorfizmu $x \mapsto \mathrm{e}^{\mathrm{i}x}$ z príkladu 27.5.6 tak, aby ste získali izomorfizmus $\mathbb{R}/\mathbb{Z} \cong \mathrm{U}(1)$.

27.18. Dokážte nasledujúce vety o izomorfizmoch grúp tak, že zakaždým nájdete konkrétne kanonické izomorfizmy.

(a) *Prvá (kosoštvorcová) veta o izomorfizme*: Nech N, S sú podgrupy grupy G, pričom $N \lhd G$. Potom $N \cap S \lhd S$, NS je podgrupa v G a $S/(N \cap S) \cong NS/N$.

(b) *Druhá veta o (dvojnom) izomorfizme*: Nech M, N sú normálne podgrupy grupy G, pričom $M \subseteq N$. Potom $M \lhd N$, $N/M = \zeta_M(N)$ je normálna podgrupa faktorovej grupy G/N a platí $(G/M)/(N/M) \cong G/N$.

(c) *Tretia veta o izomorfizme*: Nech N_1 je normálna podgrupa grupy G_1 a N_2 je normálna podgrupa grupy G_2. Potom $N_1 \times N_2$ je normálna podgrupa grupy $G_1 \times G_2$ a platí $G_1/N_1 \times G_2/N_2 \cong (G_1 \times G_2)/(N_1 \times N_2)$.

27.19. Dokážte, že nasledujúce grupy sú izomorfné:

(a) $(\mathbb{R}^\star, \cdot) \cong (\mathbb{R}^+, \cdot) \times (\{-1, 1\}, \cdot) \cong (\mathbb{R}, +) \times (\mathbb{Z}_2, +)$;

(b) $(\mathbb{C}^\star, \cdot) \cong (\mathbb{R}^+, \cdot) \times (\mathrm{U}(1), \cdot) \cong (\mathbb{R}, +) \times (\mathrm{U}(1), \cdot)$.

27.20. Pripomeňme, že $\boldsymbol{Rh}_\theta = \left(\begin{smallmatrix} \operatorname{ch}\theta & \operatorname{sh}\theta \\ \operatorname{sh}\theta & \operatorname{ch}\theta \end{smallmatrix}\right)$ je označuje maticu *hyperbolickej rotácie* (pozri paragraf 16.9). Dokážte, že priradením $\theta \mapsto \boldsymbol{Rh}_\theta$ je definovaný injektívny grupový homomorfizmus $\boldsymbol{Rh}\colon (\mathbb{R}, +) \to \big(\mathrm{GL}(2, \mathbb{R}), \cdot\big)$.

27.21. (a) Podajte priamy dôkaz izomorfizmu grúp $\mathbb{Z}_n \times \mathbb{Z}_n \cong \mathbb{Z}_{mn}$ pre nesúdeliteľné celé čísla $m, n \geq 1$. (*Návod*: Určte rád prvku $(1, 1)$).

(b) Pre nesúdeliteľné celé čísla $m, n \geq 1$ existujú celé čísla a, b také, že $am + bn = 1$. Dokážte jednak priamo, jednak ako dôsledok tvrdenia 27.6.4.

(c) Odvoďte z (b), že pre ľubovoľné celé čísla $m, n \geq 1$ existujú $a, b \in \mathbb{Z}$ také, že $am + bn$ je najväčší spoločný deliteľ čísel m, n.

27.22. Uvažujme komutatívny diagram grupových homomorfizmov:

$$\begin{tikzcd}
\{e\} \arrow[r] & F_1 \arrow[r, "\varphi_1"] \arrow[d, "\alpha"] & G_1 \arrow[r, "\psi_1"] \arrow[d, "\beta"] & H_1 \arrow[r] \arrow[d, "\gamma"] & \{e\} \\
\{e\} \arrow[r] & F_2 \arrow[r, "\varphi_2"'] & G_2 \arrow[r, "\psi_2"'] & H_2 \arrow[r] & \{e\}
\end{tikzcd}$$

ktorého riadky tvoria krátke exaktné postupnosti. Dokážte, tzv. *lemu o piatich homomorfizmoch*:[4]

(a) Ak sú oba krajné zvislé homomorfizmy α, γ injektívne, tak aj stredný homomorfizmus β je injektívny.

[4]Názov je mierne povedané čudný, lebo v hre je minimálne sedem homomorfizmov. Zrejme ide o pokračovanie literárnej tradície v duchu Alexandra Dumasa: aj jeho *Traja mušketieri* boli v skutočnosti štyria.

(b) Ak sú oba krajné zvislé homomorfizmy α, γ surjektívne, tak aj stredný homomorfizmus β je surjektívny.

(c) Ak sú oba krajné zvislé homomorfizmy α, γ izomorfizmy, tak aj stredný homomorfizmus β je izomorfizmus.

27.23. *Komutátorom prvkov x, y v grupe* G nazývame výraz $[x,y] = xyx^{-1}y^{-1}$. Podgupu $[G,G] = \langle [x,y];\, x,y \in G\rangle$ grupy G generovanú všetkými komutátormi nazývame *komutantom* alebo tiež *deriváciou grupy* G. Dokážte nasledujúce tvrdenia:

(a) Prvky x, y komutujú práve vtedy, keď $[x,y] = 1$.

(b) Grupa G je komutatívna práve vtedy, keď jej komutant je triválna podgrupa, t. j. $[G,G] = \{e\}$.

(c) Komutant grupy je vždy jej normálna podgrupa, t. j. $[G,G] \lhd G$.

(d) Nech $N \lhd G$ je normálna podgrupa grupy G. Potom faktorová grupa G/N je komutatívna práve vtedy, keď $[G,G] \subseteq N$.

27.24. Dokážte nasledujúci dodatok k vete 27.7.3. Nech X je podmnožina grupy G a E je nejaká množina redukovaných grupových slov nad abecedou X. Ak $\varphi\colon G \to \langle X \mid E\rangle$ je prostý homomorfizmus grúp taký, že $\varphi(x) = x$ pre každé $x \in X$, tak φ je izomorfizmus G na $\langle X \mid E\rangle$.

27.25. (a) Podrobne zdôvodnite výpočet súčinu $y^b x^a \cdot y^d x^c$ v príklade 27.7.4 o metacyklických grupách F^k_{mn}.

(b) Na základe cvičenia 27.23 (c) doplňte ďalšie vynechané detaily v tomto príklade.

(c) Popíšte komutant metacyklickej grupy F^k_{mn} a dokážte, že $\left[F^k_{mn}, F^k_{mn}\right] \cong \mathbb{Z}_n$.

(d) Pomocou cvičenia 27.23 (c) odvoďte z (c), že grupa F^k_{mn} je rozšírením cyklickej grupy $\mathbb{Z}_m$ pomocou cyklickej grupy $\mathbb{Z}_n$. Popíšte príslušnú krátku exaktnú postupnosť.

27.26. Dokážte, že homomorfizmus $F^l_{pq} \to F^k_{pq}$ z príkladu 27.7.5 je naozaj izomorfizmus.

27.27. Dokážte tvrdenie 27.8.1.

27.28. (a) Nech H je abelovská grupa. Dokážte izomorfizmus grúp $\mathrm{Hom}(\mathbb{Z}, H) \cong H$.

(b) Odvoďte z (a), že duálna grupa ku grupe celých čísel $\mathbb{Z}$ je izomorfná s grupou komplexných jednotiek $\mathrm{U}(1)$, čiže $\mathbb{Z}^{\mathrm{d}} \cong \mathrm{U}(1)$.

27.29. Dokážte tvrdenie 27.8.1 a doplňte vynechané detaily z dôkazov tvrdenia 27.8.2 a vety 27.8.6.

27.30. (a) Nech $m, n \geq 1$ sú celé čísla. Dokážte izomorfizmus grúp $\mathrm{Hom}(\mathbb{Z}_m, \mathbb{Z}_n) \cong \mathbb{Z}_d$, kde d je najväčší spoločný deliteľ čísel m, n.

(b) Vychádzajúc z faktu, že každá konečná abelovská grupa je priamym súčinom cyklických grúp, a z tvrdenia 27.8.2 odvoďte z (a), že pre ľubovoľné konečné abelovské grupy G, H platí $\mathrm{Hom}(G,H) \cong \mathrm{Hom}(H,G)$.

28. Grupy transformácií

V tejto kapitole si trochu bližšie všimneme niektoré typy grúp transformácií. Nebudeme sa však systematicky venovať ich štúdiu. Začneme dôkazom tzv. Cayleyho vety o reprezentácii, podľa ktorej je každá grupa izomorfná s nejakou grupou transformácií, a sústredíme sa najmä na využitie tohto faktu pri štúdiu abstraktných grúp. Uvidíme, že práve reprezentácie abstraktných grúp ako grúp transformácií konkrétnych množín, často vybavených dodatočnou štruktúrou, sú veľmi účinným nástrojom, ktorý nám umožňuje hlbšie preniknúť do štruktúry pôvodných grúp a jasnejšie osvetliť ich stavbu.

Ani v tomto smere však nebudeme postupovať príliš ďaleko. Uspokojíme sa s vybudovaním niekoľkých základných pojmov a techník, ktoré využijeme v nasledujúcej kapitole venovanej maticovým grupám, a týmito elementárnymi prostriedkami sa pokúsime zodpovedať niektoré prirodzené otázky, ktoré by pri štúdiu tejto a predchádzajúcej kapitoly mohli či – lepšie povedané – mali napadnúť zvedavého čitateľa.

28.1 Cayleyho veta o reprezentácii

Nech X je ľubovoľná množina. Každú podgrupu $G \subseteq \mathcal{S}(X)$ grupy všetkých permutácií množiny X nazývame *grupou transformácií množiny* X. Hovoríme, že G je *grupa transformácií*, ak G je grupou transformácií nejakej množiny.

Zrejme všetky prvky ľubovoľnej grupy transformácií sú *bijektívne* transformácie príslušnej množiny. Podľa uvedenej definície je množina $G \subseteq \mathcal{S}(X)$ grupou transformácií množiny X práve vtedy, keď $\mathrm{id}_X \in G$ a pre všetky $f, g \in G$ platí $f \circ g \in G$ a $f^{-1} \in G$, t. j. G obsahuje identické zobrazenie na X, je uzavretá vzhľadom na kompozíciu zobrazení a s každým zobrazením obsahuje aj k nemu inverzné zobrazenie.

Ukazuje sa, že grupy transformácií, až na izomorfizmus, zahŕňajú vôbec všetky grupy. Inak povedané, každú abstraktnú grupu $(G, \cdot)$ možno reprezentovať konkrétnym spôsobom ako grupu transformácií nejakej množiny X a jej grupovú operáciu skladaním zobrazení.

28.1.1. Veta. (Cayley) *Každá grupa je izomorfná s nejakou grupou transformácií.*

Dôkaz. Vlastne stačí k danej grupe G nájsť nejakú vhodnú množinu X a injektívny homomorfizmus grúp $\Phi\colon G \to \mathcal{S}(X)$. Podľa vety 27.5.4 o homomorfizme G je potom izomorfná s grupou transformácií $\operatorname{Im}\Phi \subseteq \mathcal{S}(X)$.

Zvoľme $X = G$ a pre $g \in G$, $x \in X$ položme $\Phi_g(x) = gx$. Inak povedané, pre $g \in G$ je $\Phi_g\colon X \to X$ zobrazenie dané priradením $x \mapsto gx$. Každé zobrazenie Φ_g je zrejme bijektívne, s inverzným zobrazením $\Phi_g^{-1} = \Phi_{g^{-1}}\colon x \mapsto g^{-1}x$, takže $\Phi\colon G \to \mathcal{S}(X)$.

Dokážeme, že Φ je homomorfizmus grúp, t. j. $\Phi_{gh} = \Phi_g \circ \Phi_h$ pre všetky $g, h \in G$. Nato stačí overiť, že obe zobrazenia dávajú rovnaké výsledky pre každé $x \in X$:

$$\Phi_{gh}(x) = (gh)x = g(hx) = \Phi_g(\Phi_h(x)) = (\Phi_g \circ \Phi_h)(x).$$

Na dôkaz injektívnosti homomorfizmu Φ sa stačí presvedčiť, že má triválne jadro. Ak $g \in \operatorname{Ker}\Phi$, tak $\Phi_g = \mathrm{id}_X$, preto $gx = \Phi_g(x) = \mathrm{id}_X(x) = x$ pre všetky $x \in X$. Voľbou $x = e$ dostávame $g = ge = e$, teda $\operatorname{Ker}\Phi = \{e\}$.

28.1.2. Príklad. Nech $(V, +)$ je aditívna grupa nejakého vektorového priestoru (nad ľubovoľným poľom K). Pre každé $\boldsymbol{u} \in V$ označme $\Phi_{\boldsymbol{u}}\colon V \to V$ zobrazenie dané priradením $\boldsymbol{x} \mapsto \boldsymbol{x} + \boldsymbol{u}$ pre $\boldsymbol{x} \in V$, t. j. posunutie (transláciu) o vektor $\boldsymbol{u}$. Zrejme každé posunutie $\Phi_{\boldsymbol{u}}$ je bijektívne zobrazenie (k nemu inverzné zobrazenie je posunutie o vektor $-\boldsymbol{u}$) a pre $\boldsymbol{u}, \boldsymbol{v} \in V$ platí $\Phi_{\boldsymbol{u}+\boldsymbol{v}} = \Phi_{\boldsymbol{u}} \circ \Phi_{\boldsymbol{v}}$. Inak povedané, zložením posunutí o vektory $\boldsymbol{u}$ a $\boldsymbol{v}$ dostaneme posunutie o vektor $\boldsymbol{u} + \boldsymbol{v}$. Samozrejme, identickým zobrazením $\mathrm{id}_V\colon V \to V$ je jedine posunutie $\Phi_{\boldsymbol{0}}$ o nulový vektor $\boldsymbol{0}$. $\Phi\colon (V, +) \to (\mathcal{S}(V), \circ)$ je tak injektívny grupový homomorfizmus a $(V, +)$ je izomorfná s grupou $\operatorname{Im}\Phi \subseteq \mathcal{S}(V)$ všetkých posunutí priestoru V.

Výslovne upozorňujeme, že – napriek dojmu, ktorý by snáď mohol navodiť dôkaz vety 28.1.1 a príklad 28.1.2, – zďaleka nie pre každý injektívny homomorfizmus grúp $\Phi\colon G \to \mathcal{S}(X)$ musí množina X splývať so základnou množinou grupy G.

28.2 Akcie a reprezentácie grúp

Niektoré metódy, ktoré sa vyskytli v dôkaze Cayleyho vety, stoja za podrobnejšie preskúmanie vo všeobecnej polohe.

Nech $(G, \cdot, e)$ je grupa a X je ľubovoľná množina.

(a) *Akciou grupy G na množine X* nazývame binárnu operáciu $\cdot\colon G \times X \to X$, ktorá spĺňa podmienky

$$ex = x \qquad \text{a} \qquad g(hx) = (gh)x$$

pre ľubovoľné $g, h \in G$, $x \in X$. Uvedený typ akcie presnejšie nazývame *ľavou akciou grupy G na množine X*. Analogicky možno definovať aj *pravú akciu*

$\cdot\colon X \times G \to X$ *grupy* G *na množine* X; formuláciu príslušných podmienok prenechávame čitateľovi.

(b) *Reprezentáciou grupy* G *v množine* X nazývame ľubovoľný grupový homomorfizmus $\Phi\colon G \to \mathcal{S}(X)$. Keďže pre $g \in G$ je samotné $\Phi(g)\colon X \to X$ zobrazenie, budeme miesto $\Phi(g)$ dávať prednosť zápisu Φ_g; podmienka homomorfnosti Φ má potom tvar rovnosti $\Phi_{gh} = \Phi_g \circ \Phi_h$ pre všetky $g, h \in G$.

Rozdiel medzi pojmami akcie a reprezentácie je čiste formálny, sú to len dva mierne odlišné spôsoby ako hovoriť o tom istom. Ak je daná akcia $\cdot\colon G \times X \to X$, tak každé $g \in G$ určuje predpisom $\Phi_g(x) = gx$ bijekciu $\Phi_g \in \mathcal{S}(X)$; k nej inverzná bijekcia je daná prepisom $\Phi_{g^{-1}}(x) = g^{-1}x$. Zobrazenie $\Phi\colon G \to \mathcal{S}(X)$ je potom reprezentáciou grupy G v množine X. Z vlastností akcie totiž vyplýva podmienka homomorfnosti Φ:

$$\Phi_{gh}(x) = (gh)x = g(hx) = \Phi_g(\Phi_h(x)) = (\Phi_g \circ \Phi_h)(x),$$

pre všetky $g, h \in G$, $x \in X$. Naopak, každá reprezentácia $\Phi\colon G \to \mathcal{S}(X)$ určuje predpisom $gx = \Phi_g(x)$ akciu $\cdot\colon G \times X \to X$. Dôvody sú podobné: na dôkaz podmienky $g(hx) = (gh)x$ stačí prepísať uvedené rovnosti v mierne pozmenenom poradí; podmienka $ex = x$ vyplýva z vlastnosti reprezentácie $\Phi_e = \mathrm{id}_X$.

Nech grupa G má akciu na množine X. *Orbitou prvku* $x \in X$ (vzhľadom na túto akciu) naývame množinu

$$Gx = \{gx;\ g \in G\}.$$

Podobne ako v paragrafe 27.4, i teraz možno ľahko nahliadnuť, že pre ľubovoľné $x, y \in X$ platí $x = ex \in Gx$ a

$$Gx \cap Gy \neq \emptyset \ \Leftrightarrow\ Gx = Gy.$$

Z toho vyplýva, že systém orbít všetkých prvkov grupy G tvorí rozklad množiny X; hovoríme mu *orbitálny rozklad množiny* X a značíme ho X/G. K nemu prislúchajúcu ekvivalenciu možno vyjadriť nasledujúcimi štyrmi ekvivalentnými spôsobmi:

$$x \equiv_G y \ \Leftrightarrow\ Gx = Gy \ \Leftrightarrow\ x \in Gy \ \Leftrightarrow\ y \in Gx \ \Leftrightarrow\ Gx \cap Gy \neq \emptyset.$$

Hovoríme, že *množina* $T \subseteq X$ je *transverzálna*, ak $\#\,(Gx \cap T) = 1$ pre každé $x \in X$, t. j. T obsahuje práve jedného reprezentanta každej orbity $Gx \in X/G$.

Stabilizátorom prvku $x \in X$ nazývame množinu

$$\mathrm{Stb}(x) = \{g \in G;\ gx = x\}.$$

Ľahko sa možno presvedčiť, že stabilizátor $\mathrm{Stb}(x)$ je dokonca podgrupou grupy G. Podobne nazývame

$$\mathrm{Fix}(g) = \{x \in G;\ gx = x\}$$

množinou pevných bodov, prípadne *množinou fixpunktov* prvku $g \in G$. Zrejme pre $g = e$ máme $\mathrm{Fix}(e) = X$.

Nasledujúce dva výsledky dávajú do súvisu práve zavedené pojmy.

28.2.1. Veta. *Nech grupa G má akciu na množine X. Potom pre každé $x \in X$ je predpisom $g\,\mathrm{Stb}(x) \mapsto gx$ korektne definované bijektívne zobrazenie $G/\mathrm{Stb}(x) \to Gx$ medzi množinou pravých tried rozkladu grupy G podľa stabilizátora $\mathrm{Stb}(x)$ a orbitou Gx. Ak G alebo X je konečná, tak z toho vyplýva*

$$\#\,Gx = [G : \mathrm{Stb}(x)].$$

Ak X je konečná, tak navyše pre každú transverzálnu množinu $T \subseteq X$ platí

$$\#X = \sum_{x \in T} [G : \mathrm{Stb}(x)].$$

Dôkaz. Označme $S = \mathrm{Stb}(x)$. Predpisom $g \mapsto gx$ je zrejme definované surjektívne zobrazenie $G \to Gx$. Pre všetky $g, h \in G$ pritom platí

$$gx = hx \;\Leftrightarrow\; x = g^{-1}hx \;\Leftrightarrow\; g^{-1}h \in S \;\Leftrightarrow\; gS = hS,$$

To znamená, že predpisom $gS \mapsto gx$ je naozaj korektne definované bijektívne zobrazenie $G/S \to Gx$. To dokazuje prvú rovnosť. Keďže množina X je zjednotením po dvoch disjunktných orbít Gx, $x \in T$, druhá rovnosť je dôsledkom prvej.

Poznamenajme, že druhá z rovností vety 28.2.1 sa niekedy zvykne nazývať *rovnosť tried*. Podľa prvej rovnosti sa počet prvkov orbity Gx ľubovoľného prvku $x \in X$ rovná indexu jeho stabilizátora $\mathrm{Stb}(x)$ v grupe G. Ak G je konečná, tak z toho podľa Lagrangeovej vety 27.4.2 vyplýva rovnosť $\#\,\mathrm{Stb}(x) = (\#\,G)/(\#\,Gx)$, v dôsledku čoho majú všetky stabilizátory prvkov tej istej orbity rovnaký rád. No nielen to – tieto stabilizátory sú dokonca izomorfné.

28.2.2. Tvrdenie. *Nech grupa G má akciu na množine X. Potom stabilizátory ľubovoľných dvoch prvkov tej istej orbity sú izomorfné. Presnejšie, ak $x, y \in X$ a $y = hx$, kde $h \in G$, tak predpisom $g \mapsto hgh^{-1}$ je definovaný izomorfizmus grúp $\mathrm{Stb}(x) \cong \mathrm{Stb}(y)$.*

Dôkaz prenechávame ako cvičenie čitateľovi.

Ďalšia veta, známa pod názvom *Burnsideova lema*, vyjadruje počet orbít akcie ako aritmetický priemer počtov fixpunktov jednotlivých prvkov grupy. Ako vzápätí uvidíme Burnsideova lema je zovšeobecnením Lagrangeovej vety 27.4.2.

28.2.3. Veta. *Nech G je konečná grupa, ktorá má akciu na konečnej množine X. Potom*
$$\#\,(X/G) = \frac{1}{\#\,G}\sum_{g\in G}\#\,\mathrm{Fix}(g).$$

Dôkaz. Počet prvkov množiny $R = \{(g,x) \in G \times X;\ gx = x\}$ spočítame dvoma spôsobmi. Zrejme platí
$$\#R = \sum_{g\in G}\#\,\mathrm{Fix}(g) = \sum_{x\in X}\#\,\mathrm{Stb}(x).$$

S využitím predchádzajúceho tvrdenia a faktu, že X je zjednotením po dvoch disjunktných orbít $\omega \in X/G$, z toho dostaneme
$$\sum_{g\in G}\#\,\mathrm{Fix}(g) = \sum_{x\in X}\#\,\mathrm{Stb}(x) = \sum_{x\in X}\frac{\#\,G}{\#\,Gx} = \sum_{\omega\in X/G}\ \sum_{x\in\omega}\frac{\#\,G}{\#\,\omega} = \#\,(X/G)\cdot\#\,G.$$

S príkladom reprezentácie sme sa už stretli v predchádzajúcom paragrafe, v dôkaze vety 28.1.1. Tam zostrojené zobrazenie $\Phi\colon G \to \mathcal{S}(X)$ bolo injektívnou reprezentáciou grupy G v množine $X = G$. To je tiež dôvod, prečo uvedený výsledok nazývame Cayleyho veta *o reprezentácii.* Trochu všeobecnejšiu situáciu si teraz preberieme v reči akcií.

28.2.4. Príklad. Nech G je grupa a H je jej podgrupa. Potom grupa H má na množine G *ľavú akciu transláciou* $H \times G \to G$ danú priradením $(h,x) \mapsto hx$ pre $h \in H$, $x \in G$, ako aj *pravú akciu transláciou* $G \times H \to G$ danú priradením $(x,h) \mapsto xh$.

Orbitou prvku $x \in G$ v ľavej akcii transláciou je práve ľavá trieda rozkladu Hx prvku x podľa podgrupy H; jeho orbitou v pravej akcii transláciou je zasa pravá trieda rozkladu xH podľa podgrupy H. Označenie G/H množiny všetkých orbít (ľavej resp. pravej) akcie podgrupy H transláciou si tak zachováva svoj predošlý význam z paragrafu 27.4 množiny (ľavých resp. pravých) rozkladových tried grupy G podľa podgrupy H.

Ďalej sa sústreďme len na ľavú akciu podgrupy H na grupe G. Zrejme každé $x \in G$ má triviálny stabilizátor $\mathrm{Stb}(x) = \{e\}$, a taktiež množina pevných bodov každého $h \in H \setminus \{e\}$ je prázdna, čiže $\mathrm{Fix}(h) = \emptyset$; pre $h = e$ samozrejme $\mathrm{Fix}(e) = G$. Prvá časť vety 28.2.1 v tomto kontexte hovorí, že podgrupu $H \cong H/\{e\}$ možno prirodzeným spôsobom vzájomne jednoznačne zobraziť na každú orbitu (ľavú triedu rozkladu) Hx. Podobne, Burnsideova lema splýva v tomto prípade s Lagrangeovou vetou, podľa ktorej $\#(G/H) = (\#\,G)/(\#\,H)$.

28.2.5. Príklad. Tento príklad nadväzuje na príklad 27.3.11, v ktorom sme zaviedli jednoparametrické grupy matíc. Pre ľubovoľnú maticu $\boldsymbol{A}$ z $\mathbb{R}^{n\times n}$ resp. $\mathbb{C}^{n\times n}$ je priradením $(t,\boldsymbol{c}) \mapsto \mathrm{e}^{\boldsymbol{A}t}\cdot\boldsymbol{c}$ daná akcia grupy $(\mathbb{R},+)$ na vektorovom priestore $\mathbb{R}^n$ resp. $\mathbb{C}^n$. Orbitu bodu $\boldsymbol{c}$, t. j. množinu $\left\{\mathrm{e}^{\boldsymbol{A}t}\cdot\boldsymbol{c};\, t\in\mathbb{R}\right\}$ nazývame *dráhou* alebo *trajektóriou bodu* $\boldsymbol{c}$. Riešenie $\boldsymbol{x}(t) = \mathrm{e}^{\boldsymbol{A}t}\cdot\boldsymbol{c}$ autonómnej počiatočnej úlohy $\boldsymbol{x}' = \boldsymbol{A}\cdot\boldsymbol{x}$, $\boldsymbol{x}(0) = \boldsymbol{c}$ možno chápať ako *parametrizáciu* tejto trajektórie.

Poznámka o značení. V literatúre sa možno stretnúť so značne rôznorodým označením akcií, orbít, stabilizátorov a pod. Napr. výsledok akcie prvku $g \in G$ na prvok $x \in X$ sa často značí ako „mocnina“ x^g a orbita prvku x ako $x^G = \{x^g;\, g \in G\}$. Tento zápis je vhodný najmä pre pravú akciu grupy $(G,\cdot,1)$ na množine X, kedy definujúce podmienky akcie nadobúdajú tvar $x^1 = x$ a $(x^g)^h = x^{gh}$. Bežné označenie stabilizátora prvku $x \in X$ v akcii grupy G je G_x. Orbita prvku x sa občas značí $\mathrm{Orb}(x)$.

28.3 Grupy automorfizmov a konjugácia

Niektoré reprezentácie Φ grupy G v množine X môžu mať isté vlastnosti navyše – jednotlivé zobrazenia Φ_g nemusia byť len bijekcie $X \to X$, ale môžu to byť zobrazenia zachovávajúce nejakú štruktúru na množine X. Potom reprezentáciu Φ možno chápať ako homomorfizmus $\Phi\colon G \to A$, kde A je nejaká podgrupa grupy $\mathcal{S}(X)$. Jedným takým prípadom sa budeme zaoberať v tomto paragrafe: množina X bude opäť splývať s grupou G a spomínaná podgrupa $A \subseteq \mathcal{S}(G)$ bude pozostávať z izomorfizmov $G \to G$.

Homomorfizmus $\varphi\colon G \to G$ grupy G do seba sa nazýva *endormorfizmus grupy* G. Množinu všetkých endomorfizmov grupy G značíme $\operatorname{End} G$. Zrejme $\mathrm{id}_G \in \operatorname{End} G$ a pre $\varphi,\psi \in \operatorname{End} G$ platí $\varphi\circ\psi \in \operatorname{End} G$, čiže množina $\operatorname{End} G$ obsahuje identické zobrazenie a je uzavretá vzhľadom na skladanie zobrazení. Taktiež obsahuje ďalší význačný prvok – je ním triviálny (konštantný) endomorfizmus $x \mapsto e$.

Automorfizmom grupy G nazývame každý jej bijektívny endomorfizmus. Množinu všetkých automorfizmov grupy G značíme $\operatorname{Aut} G$.

28.3.1. Veta. *Nech* $(G,\cdot)$ *je grupa. Potom množina jej automorfizmov* $\operatorname{Aut} G$ *je podgrupou grupy* $\mathcal{S}(G)$*, teda je to grupa transformácií množiny* G*.*

Dôkaz. Zrejme $\operatorname{Aut} G \subseteq \mathcal{S}(G)$, $\mathrm{id}_G \in \operatorname{Aut} G$ a pre ľubovoľné $\varphi,\psi \in \operatorname{Aut} G$ platí $\varphi\circ\psi \in \operatorname{Aut} G$ a $\varphi^{-1} \in \operatorname{Aut} G$.

Grupu $\operatorname{Aut} G$ nazývame *grupou automorfizmov grupy* G.

Dôležitým príkladom automorfizmov sú tzv. konjugácie. Ak g je prvok grupy G, tak zobrazenie $\Gamma_g\colon G \to G$ dané predpisom $\Gamma_g(x) = gxg^{-1}$ pre $x \in G$ nazývame *konjugáciou* prvkom $g \in G$.

28.3.2. Tvrdenie. *Nech G je grupa. Potom*
(a) pre každé $g \in G$ platí $\Gamma_g \in \operatorname{Aut} G$, t. j. zobrazenie Γ_g je automorfizmus grupy G;
(b) samotné zobrazenie $\Gamma\colon G \to \operatorname{Aut} G$ je reprezentácia grupy G v množine G.

Dôkaz. (a) Zrejme každé zobrazenie Γ_g je bijektívne – k nemu inverzným zobrazením je $\Gamma_g^{-1} = \Gamma_{g^{-1}}$. Ukážeme, že je to tiež homomorfizmus. Pre ľubovoľné $x, y \in G$ platí

$$\Gamma_g(xy) = g(xy)g^{-1} = \big(gxg^{-1}\big)\big(gyg^{-1}\big) = \Gamma_g(x)\Gamma_g(y).$$

(b) Treba overiť rovnosť $\Gamma_{gh} = \Gamma_g \circ \Gamma_h$ pre všetky $g, h \in G$. Zvoľme $x \in G$ a počítajme

$$\Gamma_{gh}(x) = (gh)x(gh)^{-1} = g\big(hxh^{-1}\big)g^{-1} = \Gamma_g(\Gamma_h(x)) = (\Gamma_g \circ \Gamma_h)(x).$$

Automorfizmy grupy G, ktoré majú tvar Γ_g pre nejaké $g \in G$, nazývame jej *vnútornými automorfizmami*. Podgrupu $\operatorname{Im} \Gamma \subseteq \operatorname{Aut} G$ značíme $\operatorname{In} G$ a nazývame *grupou vnútorných (interných) automorfizmov* grupy G.

Akciu grupy G na množine G konjugáciou budeme značiť $(g, x) \mapsto gxg^{-1}$; v predchádzajúcom paragrafe zavedené označenie $(g, x) \mapsto gx$ by v tomto prípade zrejme viedlo k nedorozumeniam.

Orbitou prvku $x \in G$ v tejto akcii je množina $\{gxg^{-1};\ g \in G\}$; *prvky* $x, y \in G$ sa nazývajú *konjugované*, označenie $x \sim_G y$, ak existuje $g \in G$ také, že $y = gxg^{-1}$, t. j. ak ležia v tej istej orbite. Zrejme relácia konjugovanosti $\sim_G$ je ekvivalencia na množine G. Orbitálny rozklad tvorí faktorová množina $G/{\sim_G}$, príslušné orbity sa nazývajú *triedy konjugácie*. Výslovne upozorňujeme, že (s výnimkou komutatívnych grúp) ekvivalencia konjugovanosti $\sim_G$ *nie je kongruencia* na grupe G.

Stabilizátorom prvku $x \in G$ v akcii konjugáciou je podgrupa

$$C_G(x) = C(x) = \{g \in G;\ gxg^{-1} = x\} = \{g \in G;\ gx = xg\}$$

grupy G, nazývaná *centralizátor prvku* x; zrejme pre $g \in G$ platí $g \in C(x)$ práve vtedy, keď $gx = xg$, t. j. g *komutuje* s x. Podobne, množinou pevných bodov prvku $g \in G$ je opäť centralizátor

$$C_G(g) = C(g) = \{x \in G;\ gxg^{-1} = x\} = \{x \in G;\ gx = xg\}.$$

Jadrom reprezentácie $\Gamma\colon G \to \operatorname{Aut} G$ je množina

$$C(G) = \{g \in G;\ \Gamma_g = \mathrm{id}_G\} = \{g \in G;\ (\forall\, x \in G)(gx = xg)\}$$

nazývaná *centrum grupy* G, pozostávajúca zo všetkých prvkov grupy G, ktoré komutujú s každým jej prvkom. Zrejme $C(G)$ je abelovská grupa a G je abelovská práve vtedy, keď $C(G) = G$. Z vety 27.5.3 a vety 27.5.4 o homomorfizme okamžite vyplýva

28.3.3. Veta. *Centrum $C(G)$ grupy G je jej normálna podgrupa a faktorová grupa $G/C(G)$ je izomorfná s grupou $\operatorname{In} G$ vnútorných automorfimov grupy G; symbolicky $C(G) \lhd G$ a $G/C(G) \cong \operatorname{In} G$.*

Centrum hodne vypovedá o štruktúre pôvodnej grupy. Ako jednoduchú ukážku uvádzame nasledujúci zďaleka nie očividný výsledok.

28.3.4. Veta. *Nech G je ľubovoľná grupa. Ak faktorová grupa $G/C(G)$ je cyklická, tak G je komutatívna.*

Dôkaz. Označme $C = C(G)$ a predpokladajme, že G/C je cyklická rádu r s generátorom Ca, kde $0 < r \in \mathbb{N} \cup \{\infty\}$, $a \in G$. Potom množina $C \cup \{a\}$ generuje celú grupu G. Pre ľubovoľný prvok $x \in G$ totiž existuje $0 \le k < r$ také, že $Cx = (Ca)^k = Ca^k$. Preto $x \in Ca^k$, teda $x = ca^k \in \langle C \cup \{a\}\rangle$ pre nejaké $c \in C$. Ukážeme, že i samotná grupa G je komutatívna. Ľubovoľné $x, y \in G$ možno napísať v tvare $x = ca^k$, $y = da^l$ pre vhodné $c, d \in C$, $0 \le k, l < r$. V dôsledku toho platí $xy = ca^k \cdot da^l = da^l \cdot ca^k = yx$.

Prvok $x \in C(G)$ je konjugovaný len sám so sebou, t. j. jeho orbitou je jednoprvková množina $\{x\}$, preto pre každú transverzálnu množinu T platí $C(G) \subseteq T$. Ak $x \notin C(G)$, tak jeho orbita obsahuje aspoň dva prvky. Centralizátor prvku $x \in C(G)$ je celá grupa, t. j. $C(x) = G$; pre $x \notin C(G)$ je $C(x)$ vlastná podgrupa grupy G.

Špecifikáciou viet 28.2.1 a 28.2.3 na akciu konjugáciou dostávame nasledujúce dve vety – rovnosť v prvej z nich sa opäť nazýva *rovnosť tried*, druhá je zvláštnym prípadom Burnsideovej lemy.

28.3.5. Veta. *Nech G je konečná grupa. Potom počet prvkov grupy G konjugovaných s prvkom $x \in G$ sa rovná indexu $[G : C(x)]$ jeho centralizátora. Ak $T \subseteq G$ je transverzálna množina vzhľadom na akciu konjugáciou, tak $C(G) \subseteq T$ a platí*

$$\#\, G = \sum_{x \in T} [G : C(x)] = \#\, C(G) + \sum_{x \in T \smallsetminus C(G)} [G : C(x)].$$

28.3.6. Veta. *Nech G je konečná grupa. Potom počet tried konjugácie v grupe G je*

$$\#\,(G/{\sim_G}) = \frac{1}{\#\,G}\sum_{g\in G}\#\,C(g) = \#\,C(G) + \frac{1}{\#\,G}\sum_{g\in G\smallsetminus C(G)}\#\,C(g)$$

a počet netriviálnych (t. j. aspoň dvojprvkových) tried konjugácie v G je

$$\#\,(G/{\sim_G}) - \#\,C(G) = \frac{1}{\#\,G}\sum_{g\in G\smallsetminus C(G)}\#\,C(g).$$

Rovnosť tried ma celý rad zaujímavých dôsledkov. Jedným z nich je *Cauchyho veta*, ktorá je obrátením Lagrangeovej vety pre prvočíselné delitele rádu konečnej grupy.

28.3.7. Veta. *Nech G je konečná grupa a p je prvočíslo, ktoré delí jej rád. Potom G má podgrupu rádu p.*

Dôkaz. Najprv budeme predpokladať, že $G = \{x_1, \dots, x_n\}$ je komutatívna rádu n, pričom prvok x_i má rád r_i. Vďaka komutatívnosti G ľahko nahliadneme, že priradením $(k_1, \dots, k_n) \mapsto x_1^{k_1} \dots x_n^{k_n}$ je definovaný surjektívny homomorfizmus grúp

$$\varphi\colon \mathbb{Z}_{r_1} \times \dots \times \mathbb{Z}_{r_n} \to G.$$

Podľa vety o homomorfizme je $G \cong (\mathbb{Z}_{r_1} \times \dots \times \mathbb{Z}_{r_n})/\operatorname{Ker}\varphi$ a z Lagrangeovej vety vyplýva, že rád n grupy G delí rád $r_1 \dots r_n$ grupy $\mathbb{Z}_{r_1} \times \dots \times \mathbb{Z}_{r_n}$. Potom aj prvočíslo p delí $r_1 \dots r_n$, preto musí deliť niektorý z činiteľov r_i. Označme $k = r_i/p$. Prvok $y = x_i^k \in \langle x_i\rangle \subseteq G$ má zrejme rád p, teda $\langle y\rangle \subseteq G$ je cyklická podgrupa rádu p.

Predpokladajme teraz, že G je konečná grupa s najmenším možným rádom, ktorý je deliteľný číslom p, ale G nemá podgrupu rádu p. Potom G nie je komutatívna, takže centrum $C = C(G)$ je jej vlastná komutatívna podgrupa a z prvej časti dôkazu vyplýva, že p nemôže deliť ani rád centra C. Podľa rovnosti tried z vety 28.3.5 platí

$$\#\,G = \#\,C + \sum_{x\in T\smallsetminus C}[G : C(x)],$$

kde $T \subseteq G$ je nejaká transverzálna množina vzhľadom na akciu konjugáciou grupy G na sebe samej. Keby p nedelilo rád žiadneho z centralizátorov $C(x)$, $x \in T \smallsetminus C$, delilo by všetky ich indexy $[G : C(x)]$; potom by však muselo deliť aj rád $\#\,C$, čo je spor. Preto p delí rád $C(x)$ pre nejaké $x \notin C$, čo je vlastná podgrupa grupy G. Z minimality rádu G vyplýva, že $C(x)$, a tým aj G, obsahuje pogrupu rádu p, čo je opäť spor.

Pre konečné abelovské grupy je možné úplné obrátenie Lagrangeovej vety.

28.3.8. Veta. *Nech G je konečná abelovská grupa a d je prirodzené číslo, ktoré delí jej rád. Potom G má podgrupu rádu d.*

Dôkaz. Ak d je prvočíslo, tak potrebný záver vyplýva z predchádzajúcej vety. Predpokladajme, že d je najmenšie prirodzené číslo, pre ktoré existuje konečná abelovská grupa s rádom deliteľným číslom d, no bez podgrupy rádu d. Potom $d = pm$ pre nejaké prvočíslo p a prirodzené číslo $m > 1$. Nech G je spomínaná grupa. Keďže $m < d$ tiež delí rád grupy G, táto má podgrupu S rádu m. Faktorová grupa G/S má rád deliteľný číslom p, teda aj podgrupu H rádu p. Potom však $\zeta_S^{-1}(H)$ je podgrupa grupy G rádu $pm = d$, čo je spor. (Pripomíname, že $\zeta_S\colon G \to G/S$ označuje prirodzenú projekciu.)

Aj nasledujúci výsledok je jednoduchým dôsledkom rovnosti tried. Ide o slabšiu verziu výsledku dokázaného Burnsidom.

28.3.9. Veta. *Nech rád grupy G je kladnou mocninou prvočísla p. Potom G má netriviálne centrum.*

Dôkaz. Nech $T \subseteq G$ je transverzálna množina vzhľadom na akciu konjugáciou. Pre každé $x \in T \smallsetminus C(G)$ je centralizátor $C(x)$ vlastnou podgrupou grupy G, teda jeho index $[G : C(x)]$ je deliteľný číslom p. Potom však aj rád centra

$$\# C(G) = \# G - \sum_{x \in T \smallsetminus C(G)} [G : C(x)]$$

je deliteľný číslom p.

28.3.10. Dôsledok. *Nech G je grupa rádu p^2, kde p prvočíslo. Potom G je komutatívna.*

Dôkaz. Podľa predošlej vety G má netriviálne centrum; toto môže mať len p alebo p^2 prvkov. Prvý prípad však nemôže nastať, lebo potom by faktorová grupa $G/C(G)$ bola rádu p, teda cyklická. Podľa vety 28.3.4 by G bola komutatívna, čiže $C(G) = G$.

28.4 Polopriamy súčin grúp

Reprezentácie jednej grupy automorfizmami inej grupy umožňujú zaujímavé zovšeobecnenie konštrukcie priameho súčinu grúp. Zatiaľ čo priamy súčin abelovských grúp je abelovská grupa, pomocou tzv. polopriameho súčinu možno i z abelovských grúp vytvoriť grupy neabelovské. Naopak, rozkladom nejakej grupy na polopriamy súčin v istom zmysle jednoduchších grúp možno získať hlbší vhľad do jej štruktúry.

Nech G a X sú grupy a $\Phi\colon G \to \operatorname{Aut} X$ je homomorfizmus grúp, teda vlastne reprezentácia grupy G v množine X *automorfizmami grupy* X. *Polopriamym súčinom grúp* G a X vzhľadom na reprezentáciu Φ nazývame množinu $G \times X$ s binárnou operáciou definovanou predpisom

$$(g, x) \cdot (h, y) = (gh, x\Phi_g(y))$$

pre $g, h \in G$, $x, y \in X$. Polopriamy súčin grúp G, X budeme značiť $G \ltimes_\Phi X$, prípadne len $G \ltimes X$, ak reprezentácia Φ bude zrejmá z kontextu.

28.4.1. Veta. *Nech $(G, \cdot, e)$, $(X, \cdot, \varepsilon)$ sú grupy a $\Phi : G \to \operatorname{Aut} X$ je reprezentácia grupy G. Potom polopriamy súčin $G \ltimes_\Phi X$ grúp G a X je grupa.*

Dôkaz. Na základe definície operácie na polopriamom súčine $G \ltimes_\Phi X$ a vlastností homomorfizmov pre ľubovoľné $(g, x), (h, y), (f, z) \in G \times X$ platí

$$\begin{aligned}
(g, x) \cdot \big((h, y) \cdot (f, z)\big) &= (g, x) \cdot (hf, y\,\Phi_h(z)) = \big(g(hf), x\,\Phi_g(y\,\Phi_h(z))\big) \\
&= \big((gh)f, x\,\Phi_g(y)\,\Phi_{gh}(z)\big) = (gh, x\,\Phi_g(y)) \cdot (f, z) \\
&= \big((g, x) \cdot (h, y)\big) \cdot (f, z), \\
(e, \varepsilon) \cdot (g, x) &= (eg, \varepsilon\,\Phi_e(x)) = (g, x) \\
&= (ge, x\,\Phi_g(\varepsilon)) = (g, x) \cdot (e, \varepsilon), \\
(g, x) \cdot \big(g^{-1}, \Phi_{g^{-1}}(x^{-1})\big) &= \big(gg^{-1}, x\,\Phi_{gg^{-1}}(x^{-1})\big) = \big(e, xx^{-1}\big) = (e, \varepsilon) \\
&= \big(g^{-1}g, \Phi_{g^{-1}}(x)^{-1}\Phi_{g^{-1}}(x)\big) = \big(g^{-1}, \Phi_{g^{-1}}(x)^{-1}\big) \cdot (g, x).
\end{aligned}$$

To znamená, že príslušná binárna operácia je asociatívna, jej neutrálnym prvkom je (e, ε) a inverzným prvok k prvku (g, x) je

$$(g, x)^{-1} = \big(g^{-1}, \Phi_{g^{-1}}(x^{-1})\big) = \big(g^{-1}, \Phi_{g^{-1}}(x)^{-1}\big),$$

teda $G \ltimes_\Phi X$ je grupa.

Ak $\Phi\colon G \to \operatorname{Aut} X$ je triviálny homomorfizmus, čiže $\Phi_g = \mathrm{id}_X$ pre všetky $g \in G$, tak uvedená definícia operácie na množine $G \times X$ nadobúda tvar $(g, x) \cdot (h, y) = (gh, xy)$, teda polopriamy súčin $G \ltimes_\Phi X$ splýva s priamym súčinom $G \times X$ grúp G a X.

Iný dôležitý špeciálny prípad polopriameho súčinu dostaneme tak, že za grupu G vezmeme priamo grupu $\operatorname{Aut} X$ všetkých automorfizmov grupy X a za homomorfizmus Φ identické zobrazenie $\mathrm{id}_G\colon G = \operatorname{Aut} X \to \operatorname{Aut} X$. Príslušný polopriamy súčin značíme $\operatorname{Aut} X \ltimes X = \operatorname{Hol}(X)$ a nazývame *holomorf grupy* X. Bližší pohľad na holomorf nájde čitateľ v cvičeniach 28.10 a 28.11. Trochu všeobecnejšie možno za G vziať akúkoľvek podgrupu grupy automorfizmov $\operatorname{Aut} X$. Násobenie v takomto polopriamom súčine $G \ltimes X$ je dané formulou

$$(g, x) \cdot (h, y) = (g \circ h, xg(y)).$$

Konečne tretí špeciálny prípad možno dostať ako polopriamy súčin grupy G samej so sebou vzhľadom na reprezentáciu konjugáciou $\varGamma\colon G \to \operatorname{Aut} G$ (porovnaj s cvičením 28.11). I táto konštrukcia funguje za trochu všeobecnejších podmienok. Ak H a N sú podgrupy grupy G, pričom $N \lhd G$ (dokonca stačí, aby platilo $hxh^{-1} \in N$ pre všetky $h \in H$, $x \in N$), tak H má reprezentáciu $\varGamma : H \to \operatorname{Aut} N$ konjugáciou na grupe N. Násobenie na polopriamom súčine $H \ltimes N = H \ltimes_{\varGamma} N$ je dané predpisom

$$(g,x)\cdot(h,y) = \bigl(gh, xgyg^{-1}\bigr),$$

pre $g, h \in H$, $x, y \in N$. Pre tento polopriamy súčin nezavádzame osobitný názov práve preto, že – ako hneď uvidíme – ide svojim spôsobom o prípad typický.

Podobne ako v prípade priameho súčinu, aj v súvislosti s polopriamym súčinom prirodzene vzniká otázka rozložiteľnosti danej grupy na polopriamy súčin netriválnych, v istom zmysle jednoduchších faktorov. Tieto, ak existujú, možno opäť nájsť medzi jej vhodnými podgrupami.

Ľahko možno overiť, že s každým polopriamym súčinom $G \ltimes_{\varPhi} X$ grúp $(G, \cdot, e)$, $(X, \cdot, \varepsilon)$ sú zviazané tri grupové homomorfizmy

$$X \xrightarrow{\ \xi\ } G \ltimes_{\varPhi} X \underset{\gamma}{\overset{\pi}{\rightleftarrows}} G,$$

dané predpismi

$$\xi(x) = (e,x), \qquad \pi(g,x) = g, \qquad \gamma(g) = (g,\varepsilon),$$

pre $g \in G$, $x \in X$. Pritom ξ je injektívny, π je surjektívny a platí $\operatorname{Im}\xi = \operatorname{Ker}\pi$, teda $X \overset{\xi}{\rightarrowtail} G \ltimes_{\varPhi} X \overset{\pi}{\twoheadrightarrow} G$ je krátka exaktná postupnosť. Taktiež $\gamma\colon G \to G \ltimes_{\varPhi} X$ je injektívny a spĺňa podmienku $\pi\circ\gamma = \mathrm{id}_G$. Podotýkame, že ani jedno z ponúkajúcich sa zobrazení $(g,x) \mapsto x$ resp. $(g,x) \mapsto \varPhi_g(x)$ vo všeobecnosti *nie je homomorfizmom* $G \ltimes_{\varPhi} X \to X$.

V polopriamom súčine $G \ltimes_{\varPhi} X$ sme tak identifikovali dve podgrupy

$$H = \operatorname{Im}\gamma = \{(g,\varepsilon);\ g \in G\} \cong G,$$
$$N = \operatorname{Im}\xi = \operatorname{Ker}\pi = \{(e,x);\ x \in X\} \cong X$$

také, že $N \lhd G \ltimes_{\varPhi} X$ a $H \cap N = \{(e,\varepsilon)\}$. Navyše každý prvok $(g,x) \in G \ltimes_{\varPhi} X$ možno písať v tvare $(g,x) = (e,x)\cdot(g,\varepsilon) \in NH$, ako aj $(g,x) = (g,\varepsilon)\cdot(e,\varPhi_{g^{-1}}(x)) \in HN$, teda $G = NH = HN$. Ukazuje sa, že prítomnosť takýchto podgrúp v danej grupe už zabezpečuje jej rozklad na ich polopriamy súčin.

28.4.2. Veta. *Nech $(G, \cdot, e)$ je grupa a H, N sú jej podgrupy také, že $N \lhd G$, $H \cap N = \{e\}$ a $G = NH$. Potom G je izomorfná s polopriamym súčinom $H \ltimes_\Gamma N$ grúp H, N vzhľadom na reprezentáciu konjugáciou $\Gamma\colon H \to \operatorname{Aut} N$.*

Dôkaz. Ukážeme, že zobrazenie $\varphi(g, x) = xg$ je homomorfizmus grúp $\varphi\colon H \ltimes N \to G$. Pre $g, h \in H$, $x, y \in N$ jednoduchým výpočtom dostávame

$$\varphi\big((g, x)\cdot(h, y)\big) = \varphi(gh, xgyg^{-1}) = (xgyg^{-1})(gh) = (xg)(yh) = \varphi(g, x)\varphi(h, y).$$

Keďže $H \cap N = \{e\}$, z tvrdenia 27.6.2 (b) vyplýva, že φ je injekcia. Surjektívnosť φ je dôsledkom rovností $\operatorname{Im}\varphi = NH = G$. Teda $G \cong H \ltimes N$.

Hovoríme, že *krátka exaktná postupnosť* grúp $F \overset{\varphi}{\rightarrowtail} G \overset{\psi}{\twoheadrightarrow} H$ je *rozštiepená*, ak existuje homomorfizmus $\eta\colon H \to G$ taký, že $\psi \circ \eta = \mathrm{id}_H$.

Videli sme, že každý polopriamy súčin grúp určuje istú rozštiepenú krátku exaktnú postupnosť. Toto tvrdenie však možno aj obrátiť.

28.4.3. Dôsledok. *Nech $F \overset{\varphi}{\rightarrowtail} G \overset{\psi}{\twoheadrightarrow} H$ je krátka exaktná postupnosť grúp rozštiepená homomorfizmom $\eta\colon H \to G$. Potom grupa G je izomorfná s polopriamym súčinom $H \ltimes F$ grúp H a F.*

Dôkaz. Stačí overiť, že podgrupy $N = \operatorname{Im}\varphi$ a $H' = \operatorname{Im}\eta$ grupy G spĺňajú podmienky $N \lhd G$, $N \cap H' = \{e\}$ a $G = NH'$ predošlej vety. Detaily prenechávame čitateľovi.

28.4.4. Príklad. Nech m, n sú prirodzené čísla. Keďže $\mathbb{Z}_n \cong \langle x \mid x^n \rangle$, každý endomorfizmus $\varphi\colon \mathbb{Z}_n \to \mathbb{Z}_n$ je jednoznačne určený jediným prvkom $k \in \mathbb{Z}_n$, totiž obrazom $k = \varphi(1)$ generátora $1 \in \mathbb{Z}_n$. Pre $b \in \mathbb{Z}_n$ potom platí $\varphi(b) = kb$. Samostatne si rozmyslite, že $\varphi \in \operatorname{Aut}\mathbb{Z}_n$ práve vtedy, keď $k = \varphi(1)$ je nesúdeliteľné s n. Podobne, z dôvodu $\mathbb{Z}_m \cong \langle x \mid x^m \rangle$ je každý homomorfizmus $\Phi\colon \mathbb{Z}_m \to \operatorname{Aut}\mathbb{Z}_n$ jednoznačne určený jediným automorfizmom φ grupy $\mathbb{Z}_n$, totiž obrazom $\varphi = \Phi_1$ generátora $1 \in \mathbb{Z}_m$, ktorý však musí vyhovovať podmienke $\varphi^m = \mathrm{id}_{\mathbb{Z}_n}$ (pozri vetu 27.7.3). V konečnom dôsledku je tak každý homomorfizmus $\Phi\colon \mathbb{Z}_m \to \operatorname{Aut}\mathbb{Z}_n$ jednoznačne určený jediným prvkom $k = \Phi_1(1) \in \mathbb{Z}_n$, ktorý je nesúdeliteľný s n a vyhovuje podmienke $k^m \equiv_n 1$. Potom pre $a \in \mathbb{Z}_m$, $b \in \mathbb{Z}_n$ platí $\Phi_a(b) = k^a b$.

Nech teda $\Phi\colon \mathbb{Z}_m \to \operatorname{Aut}\mathbb{Z}_n$ je reprezentácia grupy $\mathbb{Z}_m$ automorfizmami grupy $\mathbb{Z}_n$ a $k = \Phi_1(1)$. Polopriamy súčin $\mathbb{Z}_m \ltimes_\Phi \mathbb{Z}_n$ budeme značiť $\mathbb{Z}_m \ltimes_k \mathbb{Z}_n$. Operácia v grupe $\mathbb{Z}_m \ltimes_k \mathbb{Z}_n$ je daná formulou

$$(a, b) * (c, d) = (a + c, b + k^a d),$$

pre $a, c \in \mathbb{Z}_m$, $b, d \in \mathbb{Z}_n$. Pri porovnaní s príkladom 27.7.4 vidíme, že priradením $(a, b) \mapsto y^b x^a$ je definovaný homomorfizmus grúp $\mathbb{Z}_m \ltimes_k \mathbb{Z}_n \to F^k_{mn}$. Čitateľ by si mal samostatne premyslieť, že ide dokonca o izomorfizmus. Teda

$$\mathbb{Z}_m \ltimes_k \mathbb{Z}_n \cong F^k_{mn} = \langle x, y \mid x^m = y^n = e,\, xyx^{-1} = y^k \rangle,$$

čím sme metacyklické grupy predstavili v tvare polopriamych súčinov cyklických grúp.

28.5 Štruktúra grúp jednoduchých rádov

Ako naznačuje naše doterajšie krátke zoznámenie so svetom grúp, štruktúra konečnej grupy do značnej miery závisí od jej rádu, presnejšie od štruktúry deliteľov tohto čísla. Jednako rád grupy (okrem istých singulárnych prípadov) ani zďaleka neurčuje jej štruktúru jednoznačne. Štruktúrou deliteľov rádu grupy je však daný akýsi predbežný rozvrh možností, ktoré u grúp daného rádu vôbec prichádzajú do úvahy. Naopak, ak je štruktúra deliteľov niektorého prirodzeného čísla dostatočne jednoduchá, umožňuje to popísať štruktúru všetkých grúp daného rádu.

Pomocou doteraz dokázaných výsledkov sme schopní jednoznačne až na izomorfizmus popísať všetky grupy rádov p, p^2 a pq, kde p, q sú rôzne prvočísla. Skúsme si najprv zosumarizovať to málo, čo už vieme:

1. Každá grupa G rádu p je cyklická, teda izomorfná s aditívnou grupou $\mathbb{Z}_p$.
2. Každá grupa G rádu p^2 je podľa dôsledku 28.3.10 komutatívna. Ak obsahuje prvok rádu p^2, tak je cyklická, teda izomorfná s grupou $\mathbb{Z}_{p^2}$. V opačnom prípade musí obsahovať aspoň dva prvky a, b rádu p také, že $b \notin \langle a \rangle$. Ľahko možno overiť, že potom $\langle a \rangle \cap \langle b \rangle = \{e\}$ a $\langle a \rangle \langle b \rangle = G$ (skúste sami). Z tvrdenia 27.6.2 tak vyplýva izomorfizmus grúp $G \cong \langle a \rangle \times \langle b \rangle \cong \mathbb{Z}_p \times \mathbb{Z}_p$.

Štruktúru všetkých grúp rádu pq popisuje nasledujúca veta.

28.5.1. Veta. *Nech $p < q$ sú dve prvočísla. Potom*

(a) *každá komutatívna grupa rádu pq je cyklická, teda izomorfná s grupou $\mathbb{Z}_{pq}$;*

(b) *nekomutatívna grupa rádu pq existuje práve vtedy, keď p delí $q - 1$, a v tom prípade je izomorfná s metacyklickou grupou F^k_{pq}, kde $k \in \mathbb{Z}_q^\star$ je prvok rádu p.*

Dôkaz. Nech G je grupa rádu pq. Podľa Cauchyho vety G obsahuje podgrupu A rádu p aj podgrupu B rádu q. Obe sú zrejme cyklické, teda $A = \langle a \rangle$, $B = \langle b \rangle$ pre nejaké prvky $a, b \in G$ rádov p resp. q. Keďže rád podgrupy $A \cap B$ musí deliť rád každej z grúp A, B, nevyhnutne $A \cap B = \{e\}$. Podľa tvrdenia 27.6.2 (b) je predpisom $(x, y) \mapsto yx$ definované injektívne zobrazenie

$A \times B \to G$. Keďže obe množiny $A \times B$ aj G majú zhodne pq prvkov, je to dokonca bijekcia, teda $G = BA$.

Podobným spôsobom ukážeme, že B je normálna podgrupa v G. Keby pre niektoré $g \in G$ platilo $gBg^{-1} \neq B$, podgrupa $B \cap gBg^{-1}$ by bola triviálna, teda predpisom $(x, y) \mapsto xy$ by bolo definované prosté zobrazenie $B \times gBg^{-1} \to G$. Ale množina $B \times gBg^{-1}$ má q^2 kým množina G len pq prvkov, čo je spor, teda $B \lhd G$.

Podľa vety 28.4.2 grupa G je izomorfná s polopriamym súčinom $A \ltimes B$ vzhľadom na reprezentáciu konjugáciou $\Gamma\colon A \to \operatorname{Aut} B$ grupy A automorfizmami grupy B. Nech $\Gamma_a(b) = aba^{-1} = b^k$, kde $k \in \mathbb{Z}_q^\star$. Z úvah vykonaných v príkladoch 27.7.4–5 a 28.4.4 je jasné, že G je izomorfná s metacyklickou grupou F_{pq}^k.

Ak $k = 1$, tak generátory a, b komutujú, teda i G je komutatívna a podľa tvrdenia 27.6.2 izomorfná s priamym súčinom svojich cyklických podgrúp $A \cong \mathbb{Z}_p$, $B \cong \mathbb{Z}_q$. Z tvrdenia 27.6.4 potom vyplýva, že $G \cong \mathbb{Z}_{pq}$ je cyklická grupa.

Ak $k \neq 1$, tak k je prvok rádu p multiplikatívnej grupy $\mathbb{Z}_q^\star$ poľa $\mathbb{Z}_q$. Keďže $\mathbb{Z}_q^\star$ má rád $q-1$, takéto k existuje vtedy a len vtedy, keď p delí $q-1$. Navyše, ako sme dokázali v príklade 27.7.5, metacyklická grupa $F_{pq}^k \cong G$ je nezávisle na $k \neq 1$ určená jednoznačne až na izomorfizmus.

Len pre informáciu a bez dôkazu uveďme, že pre každé prvočíslo p existujú až na izomorfizmus práve tri komutatívne grupy rádu p^3: $\mathbb{Z}_{p^3}$, $\mathbb{Z}_{p^2} \times \mathbb{Z}_p$ a $\mathbb{Z}_p \times \mathbb{Z}_p \times \mathbb{Z}_p$; a práve dve nekomutatívne grupy rádu p^3. Pre nepárne p sú to

$$\bigl\langle x, y, z \mid x^p = y^p = z^p = e,\ xz = zx,\ yz = zy,\ xyx^{-1}y^{-1} = z \bigr\rangle,$$
$$\bigl\langle x, y \mid x^p = y^{p^2} = e,\ xyx^{-1} = y^{p+1} \bigr\rangle.$$

Pre $p = 2$ máme tiež dve grupy

$$\begin{aligned} \Delta_4 &\cong \bigl\langle x, y \mid x^2 = y^4 = e,\ xyx^{-1} = y^{-1} \bigr\rangle, \\ Q_8 &\cong \bigl\langle x, y \mid x^4 = e,\ x^2 = y^2,\ xyx^{-1} = y^{-1} \bigr\rangle. \end{aligned}$$

Ešte poznamenajme, že Q_8 sa zvykne nazývať *kvaterniónová grupa*. Obsahuje osem prvkov, typicky označovaných ako ± 1, $\pm \mathrm{i}$, $\pm \mathrm{j}$, $\pm \mathrm{k}$, nazývaných tiež *kvaterniónové jednotky*. Násobenie v Q_8 je jednoznačne dané požiadavkami, že 1 je neutrálny prvok, $(-1)q = -q$ pre každé $q \in Q_8$, a rovnosťami

$$\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1,$$
$$\mathrm{ij} = -\mathrm{ji} = \mathrm{k}, \qquad \mathrm{jk} = -\mathrm{kj} = \mathrm{i}, \qquad \mathrm{ki} = -\mathrm{ik} = \mathrm{j}.$$

Kvaterniónmi sa budeme podrobnejšie zaoberať v paragrafoch 30.4–6.

Cvičenia

28.1. (a) Nech $\sigma = (i_1, \dots, i_k) \in \mathcal{S}_n$ je permutácia n-prvkovej množiny pozostávajúca z jediného cyklu (pozri cvičenie 0.14). Dokážte, že jej rád je práve k.

(b) Dokážte, že cykly párneho rádu sú nepárne permutácie a cykly párneho rádu sú nepárne permutácie.

(c) Nech $\sigma = \varrho_1 \circ \dots \circ \varrho_m$ je rozklad permutácie $\sigma \in \mathcal{S}_n$ na disjunktné cykly. Potom jej rád je najmenším spoločným násobkom rádov jednotlivých cyklov $\varrho_1, \dots, \varrho_m$; dokážte.

28.2. (a) Nech grupa G má akciu na množine X. Podrobne dokážte, že stabilizátor $\mathrm{Stb}(x)$ ľubovoľného prvku $x \in X$ je podgrupa grupy G.

(b) Dokážte tvrdenie 28.2.2.

28.3. Nech $G \times X \to X$ je akcia grupy G na množine X. Dokážte, že každé z nasledujúcich priradení určuje akciu grupy G na príslušnej množine:
(a) pre každé $n \geq 1$ je $\big(g, (x_1, \dots, x_n)\big) \mapsto (gx_1, \dots, gx_n)$ akcia grupy G na množine X^n;

(b) pre každé $n \geq 1$ je $\big(g, \{x_1, \dots, x_n\}\big) \mapsto \{gx_1, \dots, gx_n\}$ akcia grupy G na množine $\mathcal{P}_n(X)$ všetkých n-prvkových podmnožín množiny X;

(c) $(g, S) \mapsto gM = \{gx;\, x \in M\}$ je akcia grupy G na množine $\mathcal{P}(X)$ všetkých podmnožín množiny X.

28.4. Nech $G \times \mathcal{P}(X) \to \mathcal{P}(X)$ je akcia grupy G na množine všetkých podmnožín množiny X pochádzajúca z akcie $G \times X \to X$. *Stabilizátorom množiny* $M \subseteq X$ nazývame jej stabilizátor ako prvku $M \in \mathcal{P}(X)$, t. j. $\mathrm{Stb}(M) = \{g \in G;\, gM = M\}$. Jej *bodový stabilizátor* definujeme ako prienik stabilizátorov všetkých jej prvkov, čiže $\mathrm{Stb}^\bullet(M) = \bigcap_{x \in M} \mathrm{Stb}(x) = \big\{g \in G;\, (\forall\, x \in M)(gx = x)\big\}$. Dokážte, že stabilizátor $\mathrm{Stb}(M)$ množiny $M \subseteq X$ je podgrupa grupy G a jej bodový stabilizátor $\mathrm{Stb}^\bullet(M)$ je normálna podgrupa je stabilizátora $\mathrm{Stb}(S)$, teda $\mathrm{Stb}^\bullet(M) \lhd \mathrm{Stb}(M)$.

28.5. Nech $\varPhi\colon G \to \mathrm{Aut}\, X$ je reprezentácia grupy G automorfizmami grupy X. Potom predpisom $gS = \{\varPhi_g(x);\, x \in S\}$ je definovaná akcia grupy G na množine $\mathrm{Sub}(X)$ všetkých podgrúp grupy X. Dokážte.

28.6. V zmysle predchádzajúceho cvičenia označme $(g, S) \mapsto gSg^{-1}$ akciu grupy G na svojich podgrupách pochádzajúcu od reprezentácie konjugáciou $\varGamma\colon G \to \mathrm{Aut}\, G$.

(a) Podgrupa N grupy G je normálna práve vtedy, keď je pevným bodom uvedenej akcie; dokážte.

(b) Stabilizátor podgrupy $S \subseteq G$ nazývame *normalizátorom podgrupy* S v grupe G a značíme ho $N_G(S) = N(S) = \{g \in G;\, gSg^{-1} = S\}$. Dokážte, že normalizátor $N_G(S)$ je najväčšia podgrupa grupy G, pre ktorú platí $S \lhd N_G(S)$.

(c) Dokážte, že počet rôznych podgrúp grupy G konjugovaných s pogrupou $S \subseteq G$ sa rovná indexu jej normalizátora $[G : N(S)]$.

(d) Bodový stabilizátor podgrupy $S \subseteq G$ nazývame *centralizátorom podgrupy* S v grupe G a značíme ho $C_G(S) = C(S) = \{g \in G;\, (\forall\, x \in S)(gxg^{-1} = x)\}$. Dokážte, že centralizátor $C_G(S)$ je normálnou podgrupou normalizátora $N_G(S)$.

28.7. Nech X, Y sú ľubovoľné množiny; označme $F = Y^X$ množinu všetkých funkcií $f\colon X \to Y$.

(a) Predpokladajme, že $H \times Y \to Y$ je ľavá akcia grupy H na množine Y. Dokážte, že predpisom $(\eta, f) \mapsto \eta f$, kde $(\eta f)(x) = \eta f(x)$ pre $\eta \in H$, $f \in F$, $x \in X$, je definovaná ľavá akcia grupy H na množine F.

(b) Funkcia $f \in F$ je pevným bodom prvku $\eta \in H$ práve vtedy, keď pre každé $y \in f(X)$ platí $\eta y = y$, t. j. $\eta \restriction f(X) = \mathrm{id}_{f(X)}$. Dokážte.

(c) Predpokladajme, že $X \times G \to X$ je pravá akcia grupy G na množine X. Dokážte, že predpisom $(f, \gamma) \mapsto f\gamma$, kde $(f\gamma)(x) = f(\gamma x)$ pre $\gamma \in G$, $f \in F$, $x \in X$, je definovaná pravá akcia grupy G na množine F.

(d) Funkcia $f \in F$ je pevným bodom prvku $\gamma \in G$ práve vtedy, keď pre každé $x \in X$ platí $f(\gamma x) = x$, t. j. f je konštantná na každej orbite podgrupy $\langle\gamma\rangle \subseteq G$ v akcii $X \times \langle\gamma\rangle \to X$. Dokážte.

28.8. Kruh v rovine je rozdelený svojimi n polomermi na n rovnakých kruhových výsečí. Každú z týchto výsečí vyfarbíme niektorou z k rôznych farieb (pripúšťame, že niektoré susedné výseče možu mať rovnakú farbu). Pomocou Burnsideovej lemy určíme počet rôznych terčov, ktoré možno získať takýmto farbením (dva terče považujeme za rovnaké, ak jeden možno získať z druhého otočením o celočíselný násobok uhla $2\pi/n$).

(a) Vysvetlite ako zodpovedajú farbenia výsečí kruhu funkciám $f\colon \{1, \dots, n\} \to \{1, \dots, k\}$. Označíme $F = F(n, k)$ množinu všetkých takých funkcií (farbení) a $C_n = \langle\rho\rangle$ podgrupu symetrickej grupy $\mathcal{S}_n$ generovanú cyklickou permutáciou $\rho = (1, \dots, n)$. Zrejme $(C_n, \circ) \cong (\mathbb{Z}_n, +)$.

(b) Podľa cvičenia 28.7 (c) je priradením $(f, \gamma) \mapsto f \circ \gamma$ daná pravá akcia grupy C_n na množine F. Dokážte, že dve farbenia $f, g \in F$ určujú rovnaký terč práve vtedy, keď ležia v tej istej orbite tejto akcie $F \times C_n \to F$, teda počet terčov splýva s počtom orbít $\#(F/C_n)$

(c) Pre $j \in \mathbb{Z}_n$ označme d_j najväčšieho spoločného deliteľa čísel j a n (pre $j = 0$ kladieme $d_0 = n$). Dokážte, že množina $\mathrm{Fix}(\rho^j) = \{f \in F;\ f\rho^j = f\}$ všetkých pevných bodov permutácie $\rho^j \in C_n$ má práve k^{d_j} prvkov.

(d) Odvoďte z (c), že počet všetkých terčov, ktoré možno získať zafarbením kruhu rozdeleného na n výsečí k farbami je práve $\frac{1}{n}\sum_{j=0}^{n-1} k^{d_j}$.

(e) Nech $1 \le d \le n$ je ľubovoľný deliteľ čísla n. Dokážte, že počet všetkých prvkov $j \in \mathbb{Z}_n$, pre ktoré je najväčší spoločný deliteľ j a n rovný práve d, je $\phi(n/d)$, kde ϕ je Eulerova funkcia – pozri cvičenie 27.14.

(f) Odvoďte z (e), že hľadaný počet terčov je práve $\frac{1}{n}\sum_d k^d \phi(n/d)$, kde súčet beží cez všetky delitele d čísla n.

(g) Vypočítajte niekoľko konkrétnych hodnôt počtov terčov $\#(F(n,k)/C_n)$ pre $3 \le n \le 12$, $2 \le k \le 6$.

28.9. Nech G je ľubovoľná aspoň dvojprvková grupa.

(a) Dokážte, že priradením $(x, y) \mapsto (y, x)$ je definovaný autorfizmus grupy $G^2 = G \times G$, ktorý nie je vnútorný.

(b) Nech $n \ge 2$. Pre každú permutáciu $\sigma \in \mathcal{S}_n$ označme $\Phi_\sigma\colon G^n \to G^n$ zobrazenie dané predpisom $\Phi_\sigma(x_1, \dots, x_n) = (x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(n)})$. Dokážte, že $\Phi\colon \mathcal{S}_n \to \operatorname{Aut} G^n$ je injektívna reprezentácia symetrickej grupy $\mathcal{S}_n$ automorfizmami grupy G^n a platí $\operatorname{Im}\Phi \cap \operatorname{In} G^n = \{\mathrm{id}_{G^n}\}$.

(c) Vysvetlite, prečo sme v (b) nepoužili jednoduchšiu definíciu $\Phi_\sigma(x_1, \dots, x_n) = (x_{\sigma(1)}, \dots, x_{\sigma(n)})$.

28.10. Pripomeňte si definíciu *okruhu s jednotkou* z cvičení 1.8 a 2.10.

(a) Dokážte že všetky endomorfizmy ľubovoľnej *abelovskej* grupy $(A, +)$ tvoria okruh s jednotkou $(\operatorname{End} A, +, \circ, 0, \mathrm{id}_A)$ so sčítaním po zložkách, násobením daným kompozíciou zobrazení a identickým zobrazením id_A ako jednotkou.

(b) Zadefinujte pojem izomorfizmu pre okruhy s jednotkou a dokážte nasledujúce izomorfizmy okruhov s jednotkou: $\operatorname{End} \mathbb{Z}_n \cong \mathbb{Z}_n$ pre ľubovoľné $n \geq 1$ a $\operatorname{End} \mathbb{Z} \cong \mathbb{Z}$.

28.11. (a) Nech p je prvočíslo. Dokážte, že grupa $\operatorname{Aut} \mathbb{Z}_p$ všetkých automorfizmov cyklickej grupy $(\mathbb{Z}_p, +)$ je izomorfná s multiplikatívnou grupou $(\mathbb{Z}_p^\star, \cdot)$ nenulových prvkov poľa $\mathbb{Z}_p$ ako aj s cyklickou aditívnou grupou $(\mathbb{Z}_{p-1}, +)$.

(b) Nech $n \geq 2$. Dokážte, že grupa $\operatorname{Aut} \mathbb{Z}_n$ všetkých automorfizmov cyklickej grupy $(\mathbb{Z}_n, +)$ je izomorfná s multiplikatívnou grupou $(\mathbb{Z}_n^\star, \cdot)$ všetkých zvyškov modulo n nesúdeliteľných s n.

(c) Grupa $\operatorname{Aut} \mathbb{Z}$ všetkých automorfizmov nekonečnej cyklickej grupy $(\mathbb{Z}, +)$ je izomorfná s multiplikatívnou grupou $(\{1, -1\}, \cdot)$, teda tiež s cyklickou aditívnou grupou $(\mathbb{Z}_2, +)$. Dokážte.

28.12. *Holomorfizmom grupy* X nazývame ľubovoľnú bijekciu $h\colon X \to X$ takú, že pre ľubovoľné $x, y, z \in G$ platí $h\big(xy^{-1}z\big) = h(x)h(y)^{-1}h(z)$. Dokážte nasledujúce tvrdenia:

(a) Každý automorfizmus grupy X je jej holomorfizmom.

(b) Každá ľavá translácia, t. j. vynásobenie $\mu_a\colon x \mapsto ax$ konštantným prvkom $a \in X$, je holomorfizmus grupy X.

(c) Všetky holomorfizmy grupy X tvoria grupu $\operatorname{Hol}(X)$ vzhľadom na skladanie zobrazení.

(d) Grupa $\operatorname{Aut} G$ všetkých automorfizmov grupy G je podgrupou grupy $\operatorname{Hol}(G)$.

(e) Všetky translácie $\mu_a\colon G \to G$, $a \in X$, tvoria *normálnu* podgrupu grupy $\operatorname{Hol} X$ izomorfnú s grupou X.

(f) $\operatorname{Hol}(X) \cong \operatorname{Aut} X \ltimes X$.

28.13. (a) Podobným spôsobom ako v predchádzajúcom cvičení opíšte *vnútorný holomorf* $\operatorname{IHol}(X)$ grupy X ako polopriamy súčin $\operatorname{In} X \ltimes X$ grupy X s grupou jej vnútorných automorfizmov $\operatorname{In} X$.

(b) Dokážte, že vnútorný holomorf $\operatorname{IHol}(X)$ je izomorfný s homomorfným obrazom polopriameho súčinu $X \ltimes X$ grupy X so sebou samou vzhľadom na reprezentáciu konjugáciou $\Gamma\colon X \to \operatorname{Aut} X$. Popíšte explicitne prirodzený surjektívny homomorfizmus $X \ltimes X \to \operatorname{In} X \ltimes X = \operatorname{IHol}(X)$ aj jeho jadro.

28.14. Podrobne overte, že zobrazenia ξ, π a γ z rozštiepenej krátkej exaktnej postupnosti polopriameho súčinu v paragrafe 28.4 sú naozaj homomorfizmy. Na jednoduchých príkladoch predveďte, že zobrazenia $(g, x) \mapsto x$, $(g, x) \mapsto \Phi_g(x)$ nemusia byť homomorfizmy.

28.15. Doplňte vynechané detaily dôkazu dôsledku 28.4.3 a popíšte explicitne homomorfizmus $\Phi\colon H \to \operatorname{Aut} F$ pomocou homomorfizmov φ, ψ a η.

28.16. Doplňte vynechané podrobnosti v úvahe z bodu 2 na začiatku paragrafu 28.5.

28.17. Dokážte nasledujúce tvrdenia:

(a) Každá grupa rádu 15, 33, 35, resp. 77 je cyklická.

(b) Existuje jediná nekomutatívna grupa rádu 21, 55, resp. 155.

(c) Ak p je nepárne prvočíslo, tak $\varDelta_p$ je jediná nekomutatívna grupa rádu $2p$.

(d) Ak q je prvočíslo tvaru $3k+1$, tak existuje jediná nekomutatívna grupa rádu $3q$.

(e) Ak q je prvočíslo tvaru $3k-1$, tak každá grupa rádu $3q$ je cyklická.

28.18. (a) Dokážte, že grupy $\varDelta_4$ a Q_8 nie sú izomorfné.

(b) Dokážte, že každá nekomutatívna grupa rádu 8 je izomorfná s niektorou z grúp $\varDelta_4$ alebo Q_8.

28.19. (a) Dokážte, že jediná dvojprvková podgrupa kvaterniónovej grupy Q_8 je jej centrum $C(Q_8) = \{1, -1\}$.

(b) Odvoďte z (a), že hoci Q_8 nie je komutatívna, každá jej podgrupa je normálna.

28.20. (a) Znázornite si všetkých 12 prvkov alternujúcej grupy $\mathcal{A}_4$ ako permutácie štvorprvkovej množiny $\{1, 2, 3, 4\}$ umiestnenej do vrcholov pravidelného štvorstena. Uvedomte si, že všetky tieto permutácie možno reprezentovať rotáciami v euklidovskom priestore $\mathbb{R}^3$ okolo vhodných osí prechádzajúcich cez stred štvorstena o uhol π resp. $2\pi/3$.

(b) Nájdite všetky dvoj-, troj- a štvorprvkové podgrupy grupy $\mathcal{A}_4$.

(c) Dokážte, že $\mathcal{A}_4$ nemá podgrupu rádu 6.

(d) Dokážte, že žiadne dve z nasledujúcich nekomutatívnych dvanásťprvkových grúp nie sú izomorfné: alternujúca grupa $\mathcal{A}_4$, dihedrálna grupa $\varDelta_6$, priamy súčin $\varDelta_3 \times \mathbb{Z}_2$.

29. Lineárne a afinné grupy

Lineárne grupy, t. j. rôzne grupy regulárnych matíc resp. grupy automorfizmov vektorových priestorov, patria k vôbec najdôležitejším príkladom grúp. V tejto kapitole sa zoznámime so základnými typmi lineárnych grúp a ich afinnými rozšíreniami, ako aj s rôznymi vzťahmi medzi nimi. Využitie aparátu teórie grúp nám zároveň umožní uvidieť v novom, jednotiacom svetle niektoré otázky týkajúce sa zachovávania niektorých geometrických *invariantov*, napr. objemu, vzdialenosti či časopriestorovej odľahlosti. Lineárne grupy (najmä nad poľom $\mathbb{C}$) hrajú navyše kľúčovú úlohu pri objasnení štruktúry vôbec všetkých konečných grúp pomocou ich *lineárnych reprezentácií*, t. j. homomorfizmov danej grupy do rôznych lineárnych grúp. Teóriou takýchto reprezentácií sa však už v tomto kurze zaoberať nebudeme.

V dvoch záverečných paragrafoch podnikneme malý výlet do *topológie* lineárnych grúp. Zavedieme pojmy *súvislosti* a *jednoduchej súvislosti* a okrem iného ukážeme, ako súvisia *súvislé komponenty* niektorých maticových grúp s *orientáciou* (časo)priestoru. Obmedzený rozsah nám však nedovoľuje zaviesť všetky potrebné topologické pojmy a techniky, ktoré by nám umožnili pojednať o tejto téme uceleným a elegantnejším spôsobom. Preto budeme nútení urobiť zopár dôkazov takpovediac na kolene a trochu „zašmodrchať" niektoré detaily.

29.1 Všeobecná a špeciálna lineárna grupa

Ako sme už uviedli v príklade 27.1.4, všetky regulárne matice daného rozmeru $n \times n$ nad poľom K tvoria vzhľadom na násobenie matíc grupu, ktorú značíme $\mathrm{GL}(n, K)$ a nazývame *všeobecná lineárna grupa* stupňa n nad poľom K.

Podobne, pre ľubovoľný vektorový priestor V nad poľom K tvorí množina všetkých lineárnych izomorfizmov $\varphi\colon V \to V$ grupu vzhľadom na skladanie zobrazení. Značíme ju $\mathrm{GL}(V)$ a nazývame *všeobecná lineárna grupa*, prípadne tiež *grupa automorfizmov vektorového priestoru* V. Zrejme $\mathrm{GL}(V)$ je podgrupou grupy $\mathcal{S}(V)$ všetkých bijekcií $V \to V$ a taktiež grupy $\mathrm{Aut}(V, +)$ všetkých automorfizmov aditívnej grupy $(V, +)$ vektorového priestoru V.

Ak $\dim V = n$ a $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ je pevne zvolená báza vektorového priestoru V, tak priradenie $\varphi \mapsto (\varphi)_{\boldsymbol{\alpha}}$ určuje izomorfizmus grúp $\mathrm{GL}(V) \cong \mathrm{GL}(n, K)$. Pre inú bázu $\boldsymbol{\beta}$ dostávame iný izomorfizmus $\varphi \mapsto (\varphi)_{\boldsymbol{\beta}}$. Pre (stĎp-cový) vektorový priestor $V = K^n$ sa grupy $\mathrm{GL}(K^n)$ a $\mathrm{GL}(n, K)$ zvyknú často stotožňovať prostredníctvom izomorfizmu $\varphi \mapsto (\varphi)_{\boldsymbol{\varepsilon}}$ sprostredkovaného kanonickou bázou $\boldsymbol{\varepsilon}$ v K^n.

Lineárnou grupou stupňa n nad poľom K budeme nazývať ľubovoľnú podgrupu G všeobecnej lineárnej grupy $\mathrm{GL}(n,K)$. Podobne, *lineárnou grupou vektorového priestoru* V budeme nazývať ľubovoľnú podgrupu G všeobecnej lineárnej grupy $\mathrm{GL}(V)$.[1]

Lineárne grupy transformácií $G \subseteq \mathrm{GL}(V)$ a lineárne grupy matíc $G \subseteq \mathrm{GL}(n,K)$, nazývané tiež *maticovými grupami*, budeme študovať paralelne. Pritom budeme prednostne používať tú formuláciu, ktorá sa nám v danej situácii vidí názornejšia, prípadne technicky výhodnejšia. Spomínané izomorfizmy nám kedykoľvek umožnia prenášať získané výsledky oboma smermi.

Všeobecná lineárna grupa $\mathrm{GL}(n,K)$ má dve význačné *pravé akcie* na množine $K^{n\times n}$ všetkých matíc rozmeru $n\times n$ nad poľom K: sú dané priradeniami

$$(\boldsymbol{X},\boldsymbol{A}) \mapsto \boldsymbol{A}^{-1}\cdot\boldsymbol{X}\cdot\boldsymbol{A}, \qquad \text{resp.} \qquad (\boldsymbol{X},\boldsymbol{A}) \mapsto \boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{X}\cdot\boldsymbol{A},$$

pre $\boldsymbol{X}\in K^{n\times n}$, $\boldsymbol{A}\in\mathrm{GL}(n,K)$. V prípade poľa komplexných čísel $K=\mathbb{C}$ k druhú z nich často nahrádza niektorá z akcií

$$(\boldsymbol{X},\boldsymbol{A}) \mapsto \boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{X}\cdot\overline{\boldsymbol{A}}, \qquad \text{resp.} \qquad (\boldsymbol{X},\boldsymbol{A}) \mapsto \boldsymbol{A}^{*}\cdot\boldsymbol{X}\cdot\boldsymbol{A}.$$

To nám umožňuje zasadiť známe vzťahy podobnosti a kongruencie štvorcových matíc do kontextu orbitálnych rozkladov akcií grúp. Priamočiary dôkaz nasledujúceho tvrdenia tak možno prenechať čiateľovi.

29.1.1. Tvrdenie. *(a) Všetky štyri uvedené zobrazenia sú pravými akciami grupy* $\mathrm{GL}(n,K)$ *na vektorovom priestore* $K^{n\times n}$*, pričom pre pevné* $\boldsymbol{A}\in\mathrm{GL}(n,K)$ *je každé z priradení* $\boldsymbol{X}\mapsto\boldsymbol{A}^{-1}\cdot\boldsymbol{X}\cdot\boldsymbol{A}$, $\boldsymbol{X}\mapsto\boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{X}\cdot\boldsymbol{A}$, $\boldsymbol{X}\mapsto\boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{X}\cdot\overline{\boldsymbol{A}}$, $\boldsymbol{X}\mapsto\boldsymbol{A}^{*}\cdot\boldsymbol{X}\cdot\boldsymbol{A}$ *lineárnym izomorfizmom* $K^{n\times n}\to K^{n\times n}$.

(b) Matice $\boldsymbol{X},\boldsymbol{Y}\in K^{n\times n}$ *sú podobné, t. j.* $\boldsymbol{X}\approx\boldsymbol{Y}$*, práve vtedy, keď patria do tej istej orbity grupy* $\mathrm{GL}(n,K)$ *vzhľadom na akciu* $(\boldsymbol{X},\boldsymbol{A})\mapsto\boldsymbol{A}^{-1}\cdot\boldsymbol{X}\cdot\boldsymbol{A}$.

(c) Matice $\boldsymbol{X},\boldsymbol{Y}\in K^{n\times n}$ *sú kongruentné, t. j.* $\boldsymbol{X}\equiv\boldsymbol{Y}$*, práve vtedy, keď patria do tej istej orbity grupy* $\mathrm{GL}(n,K)$ *vzhľadom na akciu* $(\boldsymbol{X},\boldsymbol{A})\mapsto\boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{X}\cdot\boldsymbol{A}$.

(d) Matice $\boldsymbol{X},\boldsymbol{Y}\in\mathbb{C}^{n\times n}$ *sú hermitovsky kongruentné, t. j.* $\boldsymbol{X}\overset{*}{\equiv}\boldsymbol{Y}$*, práve vtedy, keď patria do tej istej orbity grupy* $\mathrm{GL}(n,\mathbb{C})$ *vzhľadom na akciu* $(\boldsymbol{X},\boldsymbol{A})\mapsto\boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{X}\cdot\overline{\boldsymbol{A}}$*, prípadne vzhľadom na akciu* $(\boldsymbol{X},\boldsymbol{A})\mapsto\boldsymbol{A}^{*}\cdot\boldsymbol{X}\cdot\boldsymbol{A}$*, keďže obe majú rovnaký orbitálny rozklad.*

[1] Kvôli vylúčeniu niektorých patologických príkladov sa od lineárnych (maticových) grúp (aspoň nad poľami $\mathbb{R}$ resp. $\mathbb{C}$) tradične žiada, aby boli navyše *uzavretými* podmnožinami príslušnej všeobecnej lineárnej grupy (chápanej – ak treba po vhodnom stožnení – ako podmnožina n^2-rozmerného priestoru $\mathbb{R}^{n\times n}$ resp. $2n^2$-rozmerného priestoru $\mathbb{C}^{n\times n}$). Keďže všetky konkrétne príklady lineárnych grúp, s ktorými sa v tejto kapitole stretneme, budú túto popmienku spĺňať, nepovažujeme za potrebné zahrnúť ju do našej definície.

Uvedené relácie, ako aj tvrdenie 29.1.1 však možno zovšeobecniť na ľubovoľnú lineárnu grupu $G \subseteq \mathrm{GL}(n, K)$. Matice $\boldsymbol{X}, \boldsymbol{Y} \in K^{n \times n}$ sa nazývajú *G-podobné*, označenie $\boldsymbol{X} \approx_G \boldsymbol{Y}$, ak existuje matica $\boldsymbol{A} \in G$ taká, že $\boldsymbol{Y} = \boldsymbol{A}^{-1} \cdot \boldsymbol{X} \cdot \boldsymbol{A}$. Hovoríme, že $\boldsymbol{X}, \boldsymbol{Y}$ sú *G-kongrunetné*, označenie $\boldsymbol{X} \equiv_G \boldsymbol{Y}$, ak existuje $\boldsymbol{A} \in G$ taká, že $\boldsymbol{Y} = \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{X} \cdot \boldsymbol{A}$. V prípade poľa $K = \mathbb{C}$ možno zaviesť aj vzťah *hermitovskej G-kongruencie*: $\boldsymbol{X} \overset{*}{\equiv}_G \boldsymbol{Y}$, ak existuje $\boldsymbol{A} \in G$ taká, že $\boldsymbol{Y} = \boldsymbol{A}^* \cdot \boldsymbol{X} \cdot \boldsymbol{A}$. Obzvlášť zaujímavá je situácia, keď pre všetky $\boldsymbol{A} \in G$ platí $\boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathsf{T}}$ (resp. v komplexnom prípade $\boldsymbol{A}^{-1} = \boldsymbol{A}^*$), kedy vzťahy G-podobnosti a (hermitovskej) G-kongruencie splývajú.

Ešte si všimnime, že akciu z bodu (c) možno uvažovať aj ako akciu len na lineárnom podpriestore všetkých symetrických matíc z $K^{n \times n}$ a akcie z bodu (d) ako akcie len na podpriestore všetkých hermitovských matíc z $\mathbb{C}^{n \times n}$.

Podľa Cauchyho vety 10.3.2 a jej dôsledku, vety 10.3.3, je determinant homomorfizmom všeobecnej lineárnej grupy $\mathrm{GL}(n, K)$ na multiplikatívnu grupu $K^\star$ poľa K. Jeho jadro

$$\mathrm{SL}(n, K) = \{\boldsymbol{A} \in \mathrm{GL}(n, K);\ \det \boldsymbol{A} = 1\}$$

je tak normálnou podgrupou grupy $\mathrm{GL}(n, K)$ – nazývame ju *špeciálna lineárna grupa* stupňa n nad poľom K.

Taktiež dôkaz nasledujúceho jednoduchého tvrdenia prenechávame čitateľovi.

29.1.2. Tvrdenie. *Pre $\boldsymbol{A} \in \mathrm{GL}(n, K)$ platí $\boldsymbol{A} \in \mathrm{SL}(n, K)$ práve vtedy, keď*

$$\det(\boldsymbol{A} \cdot \boldsymbol{X}) = \det \boldsymbol{X}$$

pre každú maticu $\boldsymbol{X} \in K^{n \times n}$.

Teda v prípade poľa $K = \mathbb{R}$ a euklidovského priestoru $\mathbb{R}^n$ vybaveného štandardným skalárnym súčinom špeciálna lineárna grupa $\mathrm{SL}(n, \mathbb{R})$ pozostáva práve zo všetkých matíc $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ takých, že lineárne zobrazenie $\mathbb{R}^n \to \mathbb{R}^n$, dané predpisom $\boldsymbol{x} \mapsto \boldsymbol{A} \cdot \boldsymbol{x}$, zachováva orientované objemy n-rozmerných rovnobežnostenov vytvorených usporiadanými n-ticami vektorov $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ z $\mathbb{R}^n$. Hovoríme, že orientovaný objem je *invariantom grupy* $\mathrm{SL}(n, \mathbb{R})$; všeobecnejšie, determinant je *invariantom grupy* $\mathrm{SL}(n, K)$.

Všeobecnú lineárnu grupu $\mathrm{GL}(n, K)$ možno tiež opísať ako polopriamy súčin multiplikatívnej grupy $K^\star$ nenulových prvkov poľa K a špeciálnej lineárnej grupy $\mathrm{SL}(n, K)$.

29.1.3. Tvrdenie. $\mathrm{GL}(n, K) \cong K^\star \ltimes \mathrm{SL}(n, K)$.

Dôkaz. Ľahko možno nahliadnuť, že grupové homomorfizmy

$$\mathrm{SL}(n, K) \overset{\iota}{\rightarrowtail} \mathrm{GL}(n, K) \overset{\delta}{\twoheadrightarrow} K^\star,$$

kde ι je identické zobrazenie a δ priradí matici $\boldsymbol{A} \in \mathrm{GL}(n, K)$ jej determinant, tvoria krátku exaktnú postupnosť, rozštiepenú napr. homomorfizmom $\eta\colon K^\star \to \mathrm{GL}(n, K)$, $a \mapsto \mathrm{diag}(\boldsymbol{I}_{n-1}, a)$. Potrebný záver tak vyplýva z dôsledku 28.4.3.

29.2 Afinné rozšírenia lineárnych grúp

Podľa vety 8.5.2 a bezprostredne za ňou nasledujúcej poznámky možno každé afinné zobrazenie $f\colon V \to V$ vektorového priestoru V do seba jednoznačne vyjadriť v tvare kompozície $f = \varphi + \boldsymbol{u} = \tau_{\boldsymbol{u}} \circ \varphi$ lineárneho zobrazenia $\varphi = f - f(\boldsymbol{0})\colon V \to V$ a posunutia $\tau_{\boldsymbol{u}}\colon V \to V$ o pevný vektor $\boldsymbol{u} = f(\boldsymbol{0}) \in V$. Navyše podľa tvrdenia 8.5.8 je f bijektívne práve vtedy, keď je bijektívne φ.

Afinným rozšírením lineárnej grupy $G \subseteq \mathrm{GL}(V)$ nad vektorovým priestorom V nazývame množinu G^{af} všetkých afinných zobrazení $f\colon V \to V$, ktorých lineárna časť $\varphi = f - f(\boldsymbol{0})$ patrí do grupy G – pozri paragraf 8.5. Keďže kompozícia $f \circ g$ afinných bijekcií $f, g\colon V \to V$ s lineárnymi časťami φ resp. ψ je opäť afinná bijekcia s lineárnou časťou $\varphi \circ \psi$, a inverzné zobrazenie k f je tiež afinná bijekcia s lineárnou časťou φ^{-1}, vidíme, že afinné rozšírenie G^{af} lineárnej grupy G je podgrupou grupy transformácií $\mathcal{S}(V)$.

Štruktúru afinného rozšírenia G^{af} teraz popíšeme pomocou polopriameho súčinu grupy G a aditívnej grupy $(V, +)$ vektorového priestoru V.

Každá lineárna grupa $G \subseteq \mathrm{GL}(V)$ nad vektorovým priestorom V je zároveň podgrupou grupy $\mathrm{Aut}(V, +)$ všetkých automorfizmov jeho aditívnej grupy $(V, +)$. Množina $G \times V$ tak prirodzene nesie štruktúru polopriameho súčinu týchto grúp s neutrálnym prvkom $(\mathrm{id}_V, \boldsymbol{0})$ a s operáciami definovanými pomocou rovností

$$\begin{aligned}(\varphi, \boldsymbol{u}) \circ (\psi, \boldsymbol{v}) &= (\varphi \circ \psi, \boldsymbol{u} + \varphi(\boldsymbol{v})),\\ (\varphi, \boldsymbol{u})^{-1} &= \big(\varphi^{-1}, -\varphi^{-1}(\boldsymbol{u})\big),\end{aligned}$$

pre $\varphi, \psi \in G$, $\boldsymbol{u}, \boldsymbol{v} \in V$ – pozri paragraf 28.4. Ak tieto rovnosti porovnáme s rovnosťou tesne za dôkazom tvrdenia 8.5.4 a s rovnosťou z tvrdenia 8.5.8, uvidíme, že sme vlastne dokázali nasledujúcu vetu:

29.2.1. Veta. *Nech V je vektorový priestor nad poľom K a $G \subseteq \mathrm{GL}(V)$ je ľubovoľná lineárna grupa nad V. Potom jej afinné rozšírenie G^{af} je izomorfné s polopriamym súčinom $G \ltimes V$; príslušný izomorfizmus $G \ltimes V \to G^{\mathrm{af}}$ je daný predpisom $(\varphi, \boldsymbol{u}) \mapsto \varphi + \boldsymbol{u}$.*

Z popisu polopriameho súčinu v paragrafe 28.4 tak dostávame tri grupové homomorfizmy

$$V \xrightarrow{\ \tau\ } G^{\mathrm{af}} \underset{\iota}{\overset{\lambda}{\rightleftarrows}} G,$$

kde $\tau_{\boldsymbol{u}}$ je posunutie o vektor $\boldsymbol{u} \in V$, ι je identické vnorenie grupy G do grupy G^{af} a $\lambda(f) = f - f(\mathbf{0})$ je lineárna časť afinného zobrazenia $f \in G$. Zrejme $V \overset{\tau}{\rightarrowtail} G^{\mathrm{af}} \overset{\lambda}{\twoheadrightarrow} G$ je krátka exaktná postupnosť a $\lambda \circ \iota = \mathrm{id}_G$.

Vzhľadom na vetu 29.2.1 a očividný izomorfizmus $\mathrm{GL}(K^n) \cong \mathrm{GL}(n, K)$ daný kanonickou bázou $\boldsymbol{\varepsilon}$ v K^n, *afinné rozšírenie lineárnej grupy* $G \subseteq \mathrm{GL}(n, K)$ priamo definujeme ako polopriamy súčin $G^{\mathrm{af}} = G \ltimes K^n$, t. j. jej prvky budeme reprezentovať ako usporiadané dvojice (blokové matice) $(\boldsymbol{A}, \boldsymbol{u}) = (\boldsymbol{A} \,|\, \boldsymbol{u}) \in \mathrm{GL}(n, K) \times K^n$, s násobením

$$(\boldsymbol{A}, \boldsymbol{u}) \cdot (\boldsymbol{B}, \boldsymbol{v}) = (\boldsymbol{A} \cdot \boldsymbol{B}, \boldsymbol{u} + \boldsymbol{A} \cdot \boldsymbol{v})$$

– porovnaj s tvrdením 8.5.9.

Afinné rozšírenie všeobecnej lineárnej grupy $\mathrm{GL}(n, K)$ značíme $\mathrm{GA}(n, K)$ a nazývame *všeobecnou afinnou grupou* stupňa n nad poľom K. Podobne, afinné rozšírenie špeciálnej lineárnej grupy $\mathrm{SL}(n, K)$ značíme $\mathrm{SA}(n, K)$ a nazývame *špeciálnou afinnou grupou* stupňa n nad poľom K.

Záverom tohto paragrafu si ukážeme, že všeobecnú afinnú grupu $\mathrm{GA}(n, K)$ možno prirodzene vnoriť do všeobecnej lineárnej grupy $\mathrm{GL}(n + 1, K)$. V dôsledku toho je i každé afinné rozšírenie ľubovoľnej lineárnej grupy izomorfné s nejakou lineárnou grupou.

29.2.2. Veta. *Zobrazenie dané predpisom*

$$(\boldsymbol{A}, \boldsymbol{u}) \mapsto \begin{pmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \mathbf{0}_n & 1 \end{pmatrix},$$

kde $(\boldsymbol{A}, \boldsymbol{u}) \in \mathrm{GA}(n, K)$ *a* $\mathbf{0}_n \in K^n$ *je nulový riadkový vektor, je injektívny homomorfizmus grúp* $\mathrm{GA}(n, K) \to \mathrm{GL}(n + 1, K)$.

Dôkaz. Nakoľko pre $(\boldsymbol{A}, \boldsymbol{u}) \in \mathrm{GA}(n, K)$ platí

$$\det \begin{pmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \mathbf{0}_n & 1 \end{pmatrix} = \det \boldsymbol{A} \neq 0,$$

je to naozaj zobrazenie do grupy $\mathrm{GL}(n + 1, K)$. Podmienku homomorfnosti overíme priamym výpočtom. Pre $(\boldsymbol{A}, \boldsymbol{u}), (\boldsymbol{B}, \boldsymbol{v}) \in \mathrm{GA}(n, K)$ dostávame

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \mathbf{0}_n & 1 \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{B} & \boldsymbol{v} \\ \mathbf{0}_n & 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} \cdot \boldsymbol{B} & \boldsymbol{A} \cdot \boldsymbol{v} + \boldsymbol{u} \\ \mathbf{0}_n & 1 \end{pmatrix},$$

čo je matica zodpovedajúca prvku $(\boldsymbol{A}, \boldsymbol{u}) \cdot (\boldsymbol{B}, \boldsymbol{v}) = (\boldsymbol{A} \cdot \boldsymbol{B}, \boldsymbol{u} + \boldsymbol{A} \cdot \boldsymbol{v}) \in \mathrm{GA}(n, K)$. Tento homomorfizmus je injektívny, lebo jeho jadro obsahuje jedine neutrálny prvok $(\boldsymbol{I}_n, \mathbf{0}) \in \mathrm{GA}(n, K)$.

Trik spočívajúci v nahradení dvojice $(\boldsymbol{A}, \boldsymbol{u})$ maticou $\left(\begin{smallmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \boldsymbol{0} & 1 \end{smallmatrix}\right)$ je výpočtovo nesmierne výhodný. Umožňuje totiž počítať s afinnými transformáciami pomocou obvyklého násobenia matíc. Preto je hojne využívaný napr. pri programovaní úloh počítačovej grafiky. Navyše funguje aj bez predpokladu regularity matice $\boldsymbol{A}$ (pozri cvičenie 28.5).

29.3 Izometrie

Pojem izometrie kvadratickej formy je zovšeobecnením pojmu zhodného zobrazenia známeho zo stredoškolskej geometrie.

Nech V je vektorový priestor nad ľubovoľným poľom K a $q\colon V \to K$ je kvadratická forma. Hovoríme, že zobrazenie $f\colon V \to V$ je *izometriou kvadratickej formy* q (prípadne len *izometriou*, ak q je jasná z kontextu), ak pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí

$$q(f\boldsymbol{x} - f\boldsymbol{y}) = q(\boldsymbol{x} - \boldsymbol{y}).$$

29.3.1. Príklad. Ak V je vektorový priestor s reálnym prípadne komplexným skalárnym súčinom $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ a

$$q(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{x} \rangle = \|\boldsymbol{x}\|^2$$

je ním indukovaná kvadratická forma, tak izometrie formy q sú práve tie zobrazenia $f\colon V \to V$, ktoré zachovávajú vzdialenosť vo V, t. j. pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí

$$\|f\boldsymbol{x} - f\boldsymbol{y}\| = \|\boldsymbol{x} - \boldsymbol{y}\|.$$

Hovoríme im tiež *zhodné zobrazenia* alebo *izometrie priestoru V*.

Väčšina čitateľov sa už asi stretla s príkladmi zhodných zobrazení v rovine a v trojrozmernom priestore. Pojem izometrie však má ten istý geometrický význam v ľubovoľnom vektorovom priestore so skalárnym súčinom (či už reálnym alebo komplexným).

29.3.2. Príklad. V Minkowského časopriestore V s pseudoskalárnym súčinom $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ sú izometriami kvadratickej formy $q(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{x} \rangle$ práve tie zobrazenia $f\colon V \to V$, ktoré zachovávajú štvorec časopriestorovej odľahlosti, t. j. platí pre ne

$$\langle f\boldsymbol{x} - f\boldsymbol{y}, f\boldsymbol{x} - f\boldsymbol{y} \rangle = \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle$$

pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$.

K obom uvedeným príkladom sa ešte niekoľkokrát vrátime.

Všeobecné štúdium izometrií začneme jednoduchým pozorovaním, ktorého dôkaz prenechávame čitateľovi. Pripomíname, že posunutie $\tau_{\boldsymbol{u}}$ vektorového priestoru V o pevný vektor $\boldsymbol{u} \in V$ je zobrazenie $V \to V$ dané predpisom $\tau_{\boldsymbol{u}}(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{u}$.

29.3.3. Tvrdenie. *Nech $q\colon V \to K$ je kvadratická forma na vektorovom priestore V nad poľom K.*
(a) Ak $\boldsymbol{u} \in V$ je pevný vektor, tak posunutie $\tau_{\boldsymbol{u}}\colon V \to V$ je izometria formy q.
(b) Ak $f\colon V \to V$ je izometria formy q, tak aj zobrazenie $\varphi\colon V \to V$ dané predpisom $\varphi(\boldsymbol{x}) = f(\boldsymbol{x}) - f(\mathbf{0})$ je izometria formy q, ktorá navyše fixuje počiatok, t. j. spĺňa podmienku $\varphi(\mathbf{0}) = \mathbf{0}$.

Pri štúdiu štruktúry všetkých izometrií danej kvadratickej formy $q\colon V \to K$ sa teda stačí obmedziť na izometrie, ktoré navyše fixujú počiatok. Každú izometriu totiž možno získať ako kompozíciu $f = \tau_{\boldsymbol{u}} \circ \varphi = \varphi + \boldsymbol{u}$ izometrie $\varphi = f - f(\mathbf{0})$ fixujúcej počiatok a posunutia $\tau_{\boldsymbol{u}}$ o vektor $\boldsymbol{u} = f(\mathbf{0})$. Táto súvislosť silne pripomína vzťah medzi afinnými a lineárnymi zobrazeniami. Ukážeme si, že to nie je náhoda.

Ďalej budeme predpokladať, že V je konečnorozmerný vektorový priestor nad ľubovoľným poľom K charakteristiky $\neq 2$ a $F\colon V \times V \to K$ je symetrická bilineárna, prípadne, pre $K = \mathbb{C}$, kososymetrická poldruhalineárna forma na V. Pripomeňme, že F je regulárna, práve vtedy, keď pre každé $\boldsymbol{x} \in V$ platí

$$(\forall \boldsymbol{y} \in V)(F(\boldsymbol{x}, \boldsymbol{y}) = 0) \;\Rightarrow\; \boldsymbol{x} = \mathbf{0};$$

potom takisto

$$(\forall \boldsymbol{x} \in V)(F(\boldsymbol{x}, \boldsymbol{y}) = 0) \;\Rightarrow\; \boldsymbol{y} = \mathbf{0},$$

pre ľubovoľné $\boldsymbol{y} \in V$ (pozri záver paragrafu 11.1).

29.3.4. Veta. *Nech V je konečnorozmerný vektorový priestor nad poľom K charakteristiky $\neq 2$, $F\colon V \times V \to K$ je regulárna symetrická bilineárna, prípadne kososymetrická poldruhalineárna forma a q je ňou indukovaná kvadratická forma. Potom pre ľubovoľné zobrazenie $\varphi\colon V \to V$ nasledujúce podmienky sú ekvivalentné:*
(i) φ zachováva formu F, t. j. pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí
$F(\varphi\boldsymbol{x}, \varphi\boldsymbol{y}) = F(\boldsymbol{x}, \boldsymbol{y})$;
(ii) φ je lineárne zobrazenie zachovávajúce formu q, t. j. pre každé $\boldsymbol{x} \in V$ platí
$q(\varphi\boldsymbol{x}) = q(\boldsymbol{x})$;
(iii) φ je izometria formy q zachovávajúca počiatok, t. j. pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí $q(\varphi\boldsymbol{x} - \varphi\boldsymbol{y}) = q(\boldsymbol{x} - \boldsymbol{y})$ a $\varphi(\mathbf{0}) = \mathbf{0}$.
Navyše každá lineárna transformácia $\varphi\colon V \to V$ spĺňajúca niektorú z uvedených podmienok je už nevyhnutne bijektívna, teda je to lineárny izomorfizmus.

Dôkaz. (i) ⇒ (ii): Nech φ zachováva F. Najprv dokážeme linearitu φ. Zvoľme ľubovoľnú bázu $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ priestoru V. Potom $[F]_{\boldsymbol{\alpha}} = \big(F(\boldsymbol{u}_i, \boldsymbol{u}_j)\big) \in K^{n\times n}$ je regulárna matica. Pre $i \leq n$ označme $\boldsymbol{v}_i = \varphi(\boldsymbol{u}_i)$ a položme $\boldsymbol{\beta} =$

$(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$. Podľa predpokladu $F(\boldsymbol{u}_i, \boldsymbol{u}_j) = F(\boldsymbol{v}_i, \boldsymbol{v}_j)$ pre všetky i, j. Z regularity matice $\big(F(\boldsymbol{v}_i, \boldsymbol{v}_j)\big)_{n\times n}$ vyplýva, že i $\boldsymbol{\beta}$ je báza priestoru V. Existuje jediné lineárne zobrazenie $\vartheta\colon V \to V$ také, že $\vartheta(\boldsymbol{u}_i) = \boldsymbol{v}_i$ pre $i \le n$ a, keďže $\boldsymbol{\beta}$ je báza, ϑ je nevyhnutne bijektívne. Z linearity ϑ a podmienok $F(\vartheta\boldsymbol{u}_i, \vartheta\boldsymbol{u}_j) = F(\boldsymbol{u}_i, \boldsymbol{u}_j)$ navyše vyplýva, že aj ϑ zachováva formu F. Dokážeme, že $\varphi = \vartheta$. Na to stačí overiť, že pre každé $\boldsymbol{x} \in V$ a pre každý vektor $\boldsymbol{v}_i$ bázy $\boldsymbol{\beta}$ platí $F(\varphi\boldsymbol{x} - \vartheta\boldsymbol{x}, \boldsymbol{v}_i) = 0$. Potom totiž $F(\varphi\boldsymbol{x} - \vartheta\boldsymbol{x}, \boldsymbol{y})$ pre každé $\boldsymbol{y} \in V$, a z regularity formy F vyplýva $\varphi(\boldsymbol{x}) = \vartheta(\boldsymbol{x})$. Počítajme

$$F(\varphi\boldsymbol{x} - \vartheta\boldsymbol{x}, \boldsymbol{v}_i) = F(\varphi\boldsymbol{x}, \varphi\boldsymbol{u}_i) - F(\vartheta\boldsymbol{x}, \vartheta\boldsymbol{u}_i) = F(\boldsymbol{x}, \boldsymbol{u}_i) - F(\boldsymbol{x}, \boldsymbol{u}_i) = 0.$$

Druhá časť tvrdenia (ii) je triviálnym dôsledkom predpokladu (i) a rovnosti $q(\boldsymbol{x}) = F(\boldsymbol{x}, \boldsymbol{x})$.

(ii) $\Rightarrow$ (iii): Nech φ je lineárne a zachováva q. Potom $\varphi(\boldsymbol{0}) = \boldsymbol{0}$ a pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ máme

$$q(\varphi\boldsymbol{x} - \varphi\boldsymbol{y}) = q(\varphi(\boldsymbol{x} - \boldsymbol{y})) = q(\boldsymbol{x} - \boldsymbol{y}).$$

(iii) $\Rightarrow$ (i): Nech φ je izometria, ktorá zachováva počiatok. Ak si uvedomíme, že pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí

$$F(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2}\big(q(\boldsymbol{x}) + q(\boldsymbol{y}) - q(\boldsymbol{x} - \boldsymbol{y})\big),$$

vyplýva potrebný záver z nasledujúceho jednoduchého výpočtu:

$$\begin{aligned} F(\varphi\boldsymbol{x}, \varphi\boldsymbol{y}) &= \frac{1}{2}\big(q(\varphi\boldsymbol{x} - \varphi\boldsymbol{0}) + q(\varphi\boldsymbol{y} - \varphi\boldsymbol{0}) - q(\varphi\boldsymbol{x} - \varphi\boldsymbol{y})\big) \\ &= \frac{1}{2}\big(q(\boldsymbol{x} - \boldsymbol{0}) + q(\boldsymbol{y} - \boldsymbol{0}) - q(\boldsymbol{x} - \boldsymbol{y})\big) = F(\boldsymbol{x}, \boldsymbol{y}). \end{aligned}$$

Dôkaz záverečného dovetku je už zahrnutý v dôkaze implikácie (i) $\Rightarrow$ (ii).

Zdôraznime, že linearita ani bijektívnosť zobrazenia φ nie sú zahrnuté vo východzích predpokladoch práve dokázaného tvrdenia. Pozoruhodné je, že v *konečnorozmerných* priestoroch sú už obe tieto vlastnosti dôsledkom zachovávania *regulárnej* bilineárnej (prípadne poldruhalineárnej) formy (všimnite si, že v príslušných častiach dôkazu sme nevyužívali ani žiadnu vlastnosť symetrie formy F). Vari ešte prekvapivejšia je skutočnosť, že linearita i bijektívnosť sú dôsledkom izometrie a fixovania počiatku. To však znamená, že každá izometria $f\colon V \to V$ je bijektívnym afinným zobrazením. Práve dokázaná veta tak má nasledujúci dôsledok.

29.3.5. Dôsledok. *Všetky zobrazenia zachovávajúce regulárnu symetrickú bilineárnu formu na konečnorozmernom vektorovom priestore V nad poľom charakteristiky $\neq 2$ (v prípade poľa $\mathbb{C}$ tiež kososymetrickú poldruhalineárnu na V) tvoria podgrupu všeobecnej lineárnej grupy $\mathrm{GL}(V)$. Všetky izometrie príslušnej kvadratickej formy tvoria podgrupu všeobecnej afinnej grupy $\mathrm{GA}(V)$, ktorá je jej afinným rozšírením.*

Zachovávanie (nie nutne symetrických) bilineárnych foriem (nie nutne bijektívnymi) lineárnymi transformáciami na konečnorozmernom vektorovom priestore možno jednoducho vyjadriť v reči ich matíc. To nám ďalej umožní stotožniť obe grupy izometrií z dôsledku 29.3.5 s istými podgrupami všeobecných lineárnych grúp matíc.

Nech $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ je nejaká báza vektorového priestoru V, $F\colon V \times V \to K$ je akákoľvek bilineárna forma a $\varphi\colon V \to V$ je lineárna transformácia. Označme $\boldsymbol{Z} = [F]_{\boldsymbol{\alpha}}$ maticu F a $\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha}}$ maticu φ v báze $\boldsymbol{\alpha}$. Potom pre všetky $\boldsymbol{x}, \boldsymbol{y} \in V$ platí

$$F(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x})_{\boldsymbol{\alpha}}^{\mathsf{T}} \cdot \boldsymbol{Z} \cdot (\boldsymbol{y})_{\boldsymbol{\alpha}},$$
$$F(\varphi\boldsymbol{x}, \varphi\boldsymbol{y}) = (\boldsymbol{A} \cdot (\boldsymbol{x})_{\boldsymbol{\alpha}})^{\mathsf{T}} \cdot \boldsymbol{Z} \cdot (\boldsymbol{A} \cdot (\boldsymbol{y})_{\boldsymbol{\alpha}}) = (\boldsymbol{x})^{\mathsf{T}} \cdot (\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{Z} \cdot \boldsymbol{A}) \cdot (\boldsymbol{y})_{\boldsymbol{\alpha}}.$$

To znamená, že φ zachováva F práve vtedy, keď

$$\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{Z} \cdot \boldsymbol{A} = \boldsymbol{Z}.$$

V prípade poldruhalineárnej formy $F\colon V \times V \to \mathbb{C}$ možno rovnako odvodiť, že φ zachováva F práve vtedy, keď

$$\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{Z} \cdot \overline{\boldsymbol{A}} = \boldsymbol{Z}.$$

V oboch prípadoch teda $\boldsymbol{A}$ musí patriť do stabilizátora matice $\boldsymbol{Z}$ v príslušnej akcii grupy $\mathrm{GL}(n, K)$ na priestore matíc $K^{n\times n}$. Ak $\boldsymbol{Z}$ je regulárna, t. j. $\det \boldsymbol{Z} \neq 0$, tak z toho vyplýva $(\det \boldsymbol{A})^2 = 1$, t. j. $\det \boldsymbol{A} = \pm 1$, prípadne pre poldruhalineárne formy $|\det \boldsymbol{A}|^2 = 1$, t. j. $|\det \boldsymbol{A}| = 1$. To má o. i. za dôsledok regularitu matice $\boldsymbol{A}$ a iným spôsobom dokazuje, že lineárna transformácia φ konečnorozmerného priestoru V zachovávajúca *regulárnu* bilineárnu formu F na V je nevyhnutne bijektívna. Tým sme vlastne dokázali nasledujúcu vetu.

29.3.6. Veta. *Nech V je vektorový priestor konečnej dimenzie n nad poľom K charakteristiky $\neq 2$, $F\colon V \times V \to K$ je regulárna symetrická bilineárna, prípadne, pre $K = \mathbb{C}$, hermitovská poldruhalineárna forma, q je ňou indukovaná kvadratická forma, $\boldsymbol{\alpha}$ je ľubovoľná báza priestoru V a $\boldsymbol{Z} = [F]_{\boldsymbol{\alpha}}$ je matica formy F v báze $\boldsymbol{\alpha}$. Potom priradením $\varphi \mapsto (\varphi)_{\boldsymbol{\alpha}}$ je daný izomorfizmus lineárnej grupy tvorenej všetkými zobrazeniami $\varphi\colon V \to V$ zachovávajúcimi F so stabilizátorom*

$$\mathrm{Stb}(\boldsymbol{Z}) = \left\{\boldsymbol{A} \in \mathrm{GL}(n, K);\ \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{Z} \cdot \boldsymbol{A} = \boldsymbol{Z}\right\}$$

matice $\boldsymbol{Z}$ *v akcii* $(\boldsymbol{X},\boldsymbol{A}) \mapsto \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{X} \cdot \boldsymbol{A}$ *všeobecnej lineárnej grupy* $\mathrm{GL}(n,K)$ *na priestore všetkých symetrických matíc* $\boldsymbol{X} \in \boldsymbol{K}^{n\times n}$*, prípadne s jej stabilizátorom*

$$\mathrm{Stb}(\boldsymbol{Z}) = \left\{\boldsymbol{A} \in \mathrm{GL}(n,\mathbb{C});\ \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{Z} \cdot \overline{\boldsymbol{A}} = \boldsymbol{Z}\right\}$$

v akcii $(\boldsymbol{X},\boldsymbol{A}) \mapsto \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{X} \cdot \overline{\boldsymbol{A}}$ *všeobecnej lineárnej grupy* $\mathrm{GL}(n,\mathbb{C})$ *na priestore všetkých hermitovských matíc* $\boldsymbol{X} \in \mathbb{C}^{n\times n}$*. Grupa všetkých izometrií kvadratickej formy* q *je potom izomorfná s afinným rozšírením* $\mathrm{Stb}(\boldsymbol{Z}) \ltimes K^n$ *stabilizátora* $\mathrm{Stb}(\boldsymbol{Z})$*, teda tiež s podgrupou všeobecnej lineárnej grupy* $\mathrm{GL}(n{+}1,K)$ *tvorenou všetkými maticami tvaru* $\left(\begin{smallmatrix}\boldsymbol{A} & \boldsymbol{u}\\ \boldsymbol{0} & 1\end{smallmatrix}\right)$*, kde* $\boldsymbol{A} \in \mathrm{Stb}(\boldsymbol{Z})$ *a* $\boldsymbol{u} \in K^n$*.*

Pri ďalšom štúdiu grúp izometrií sa preto zameriame len na stabilizátory (koso)symetrických regulárnych matíc $\boldsymbol{Z} \in K^{n\times n}$. Navyše, keďže podľa tvrdenia 28.2.2 sú stabilizátory kongruentných matíc navzájom izomorfné, stačí sa pritom obmedziť na diagonálne matice $\boldsymbol{Z}$; v prípade kvadratických foriem nad poľom $\mathbb{R}$ resp. hermitovských kvadratických foriem nad poľom $\mathbb{C}$ dokonca na matice tvaru $\mathrm{diag}(\boldsymbol{I}_k, -\boldsymbol{I}_l)$.

29.4 Ortogonálna, špeciálna ortogonálna a euklidovská grupa

Nech K je ľubovoľné pole. Potom stabilizátor jednotkovej matice $\boldsymbol{I}_n$ vzhľadom na akciu $(\boldsymbol{X},\boldsymbol{A}) \mapsto \boldsymbol{A}^{\mathsf{T}}\cdot\boldsymbol{X}\cdot\boldsymbol{A}$ všeobecnej lineárnej grupy $\mathrm{GL}(n,K)$ na priestore $K^{n\times n}$ tvorí podgrupu

$$\mathrm{O}(n,K) = \mathrm{Stb}(\boldsymbol{I}_n) = \left\{\boldsymbol{A} \in \mathrm{GL}(n,K);\ \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A} = \boldsymbol{I}_n\right\}$$

grupy $\mathrm{GL}(n,K)$ nazývanú *ortogonálna grupa* stupňa n nad poľom K. Zrejme pre $\boldsymbol{A} \in \mathrm{GL}(n,K)$ platí

$$\boldsymbol{A} \in \mathrm{O}(n,K) \ \Leftrightarrow\ \boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathsf{T}}.$$

Lineárnu grupu

$$\mathrm{SO}(n,K) = \mathrm{O}(n,K) \cap \mathrm{SL}(n,K) = \{\boldsymbol{A} \in \mathrm{O}(n,K);\ \det\boldsymbol{A} = 1\}$$

nazývame *špeciálna ortogonálna grupa* stupňa n nad poľom K. Zrejme $\mathrm{SO}(n,K)$, ako jadro homomorfizmu $\det\colon \mathrm{O}(n,K) \to \{1,-1\}$, je normálna podgrupa grupy $\mathrm{O}(n,K)$ s indexom 2 (okrem prípadu $\mathrm{char}\,K = 2$, kedy $\mathrm{SO}(n,K) = \mathrm{O}(n,K)$).

Uvedené grupy majú najväčší význam nad poľom reálnych čísel. Preto sa pre ne zaužívalo skrátené označenie $\mathrm{O}(n) = \mathrm{O}(n,\mathbb{R})$, resp. $\mathrm{SO}(n) = \mathrm{SO}(n,\mathbb{R})$, a názvy *ortogonálna*, resp. *špeciálna ortogonálna grupa* stupňa n (bez explicitnej zmienky o poli $\mathbb{R}$). Prvky grupy $\mathrm{O}(n)$ sa nazývajú *ortogonálne matice* a jednotlivo sme ich už dosť podrobne preštudovali v paragrafoch 13.5 a 23.5.

Ak uvažujeme $\mathbb{R}^n$ ako euklidovský priestor so štandardným skalárnym súčinom $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^\mathsf{T} \cdot \boldsymbol{y}$ a kanonickou ortonormálnou bázou $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$, tak každú maticu $\boldsymbol{A} \in \mathrm{GL}(n, \mathbb{R})$ možno stotožniť s lineárnou transformáciou $\boldsymbol{x} \mapsto \boldsymbol{A} \cdot \boldsymbol{x}$, teda považovať za prvok grupy $\mathrm{GL}(\mathbb{R}^n)$. Prvky grupy $\mathrm{O}(n)$ tak zodpovedajú izometriám euklidovského priestoru $\mathbb{R}^n$, ktoré fixujú počiatok. Jej podgrupa $\mathrm{SO}(n)$ je tvorená práve tými maticami z $\mathrm{O}(n)$, ktoré zachovávajú orientáciu danú kanonickou bázou $\boldsymbol{\varepsilon}$ (pozri paragraf 15.2). Všetky izometrie euklidovského priestoru $\mathbb{R}^n$ tvoria grupu izomorfnú s afinným rozšírením $\mathrm{O}(n) \ltimes \mathbb{R}^n$ ortogonálnej grupy $\mathrm{O}(n)$ pomocou grupy posunutí $\mathbb{R}^n$ – hovoríme jej *euklidovská grupa* stupňa n. Jej prvky nazývame tiež *zhodné zobrazenia* alebo len *zhodnosti* v $\mathbb{R}^n$. Prvky jej podgrupy $\mathrm{SO}(n) \ltimes \mathbb{R}^n$, t. j. afinného rozšírenia špeciálnej ortogonálnej grupy $\mathrm{SO}(n)$, nazývame *priamymi zhodnosťami*, ostatné prvky euklidovskej grupy nazývame *nepriamymi zhodnosťami.*

K vete 23.5.3, ktorá vyčerpávajúcim spôsobom popisuje štruktúru ortogonálnych matíc, ešte dodajme, že priame zhodnosti fixujúce počiatok, t. j. prvky grupy $\mathrm{SO}(n)$, sú práve tie, ktoré majú v nejakej ortonormálnej báze priestoru $\mathbb{R}^n$ maticu tvaru

$$\mathrm{diag}(\boldsymbol{R}_{\alpha_1}, \dots, \boldsymbol{R}_{\alpha_m}, 1),$$

ak $n = 2m + 1 \geq 1$ je nepárne, resp.

$$\mathrm{diag}(\boldsymbol{R}_{\alpha_1}, \dots, \boldsymbol{R}_{\alpha_m}),$$

ak $n = 2m \geq 2$ je párne. Zvyšné matice

$$\mathrm{diag}(\boldsymbol{R}_{\alpha_1}, \dots, \boldsymbol{R}_{\alpha_m}, -1) \qquad \text{resp.} \qquad \mathrm{diag}(\boldsymbol{R}_{\alpha_1}, \dots, \boldsymbol{R}_{\alpha_{m-1}}, 1, -1)$$

zodpovedajú nepriamym zhodnostiam.

Vzťah grúp $\mathrm{SO}(n, K)$ a $\mathrm{O}(n, K)$ nad ľubovoľným poľom K charakteristiky $\neq 2$ možno objasniť pomocou konštrukcie polopriameho súčinu (ešte raz pripomíname, že v prípade $\mathrm{char}\, K = 2$ platí $\mathrm{SO}(n, K) = \mathrm{O}(n, K)$, takže niet čo objasňovať). Označme

$$\boldsymbol{Q}_n = \mathrm{diag}(\boldsymbol{I}_{n-1}, -1) \in \mathrm{O}(n, K) \smallsetminus \mathrm{SO}(n, K)$$

a

$$\Delta = \{\boldsymbol{I}_n, \boldsymbol{Q}_n\}.$$

Potom Δ je zrejme podgrupa grupy $\mathrm{O}(n, K)$ izomorfná s aditívnou grupou $\mathbb{Z}_2$ ako aj s multiplikatívnou grupou $\{1, -1\}$. Ďalej platí $\Delta \cap \mathrm{SO}(n, K) = \{\boldsymbol{I}_n\}$, a keďže index podgrupy $\mathrm{SO}(n, K)$ v $\mathrm{O}(n, K)$ je 2, tak nevyhnutne $\mathrm{SO}(n, K) \cdot \Delta = \mathrm{O}(n, K)$. To podľa vety 28.4.2 dokazuje nasledujúci výsledok.

29.4.1. Veta. *Nech K je pole chrakteristiky $\neq 2$. Potom ortogonálna grupa $\mathrm{O}(n, K)$ je izomorfná s polopriamym súčinom $\Delta \ltimes \mathrm{SO}(n, K)$ svojich podgrúp $\Delta \cong \mathbb{Z}_2$ a $\mathrm{SO}(n, K)$.*

Ešte podotknime, že úlohu generátora $\boldsymbol{Q}_n$ grupy Δ by mohol zohrať ľubovoľný prvok $\boldsymbol{Q} \in \mathrm{O}(n, K) \smallsetminus \mathrm{SO}(n, K)$ taký, že $\boldsymbol{Q}^2 = \boldsymbol{I}_n$, napr. ľubovoľná matica tvaru $\boldsymbol{Q} = \mathrm{diag}(\pm 1, \ldots, \pm 1)$ s nepárnym počtom prvkov -1.

29.5 Pseudortogonálna, Lorentzova a Poincarého grupa

Nech K je ľubovoľné pole a $n = k + l$, kde k, l sú kladné celé čísla. Potom stabilizátor matice $\mathrm{diag}(\boldsymbol{I}_k, -\boldsymbol{I}_l)$ vzhľadom na akciu $(\boldsymbol{X}, \boldsymbol{A}) \mapsto \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{X} \cdot \boldsymbol{A}$ všeobecnej lineárnej grupy $\mathrm{GL}(n, K)$ na priestore $K^{n\times n}$ tvorí podgrupu

$$\begin{aligned}\mathrm{O}(k, l, K) &= \mathrm{Stb}\big(\mathrm{diag}(\boldsymbol{I}_k, -\boldsymbol{I}_l)\big)\\ &= \big\{\boldsymbol{A} \in \mathrm{GL}(n, K);\ \boldsymbol{A}^{\mathsf{T}} \cdot \mathrm{diag}(\boldsymbol{I}_k, -\boldsymbol{I}_l) \cdot \boldsymbol{A} = \mathrm{diag}(\boldsymbol{I}_k, -\boldsymbol{I}_l)\big\}\end{aligned}$$

grupy $\mathrm{GL}(n, K)$, nazývanú *pseudoortogonálna grupa* typu (k, l) nad poľom K. Jej normálnu podgrupu

$$\mathrm{SO}(k, l, K) = \mathrm{O}(k, l, K) \cap \mathrm{SL}(n, K) = \{\boldsymbol{A} \in \mathrm{O}(k, l, K);\ \det \boldsymbol{A} = 1\}$$

nazývame *špeciálna pseudoortogonálna grupa* typu (k, l) nad poľom K.

I tentokrát majú uvedené grupy majú najväčší význam nad poľom reálnych čísel. Z toho dôvodu ich označujeme len $\mathrm{O}(k, l) = \mathrm{O}(k, l, \mathbb{R})$, resp. $\mathrm{SO}(k, l) = \mathrm{SO}(k, l, \mathbb{R})$, a nazývame stručne *pesudoortogonálna*, resp. *špeciálna pseudoortogonálna grupa* typu (k, l) (bez explicitnej zmienky o poli $\mathbb{R}$). Keďže kongruentné matice majú izomorfné stabilizátory vzhľadom na uvedenú akciu a zrejme platí $\mathrm{Stb}(-\boldsymbol{I}_n) = \mathrm{Stb}(\boldsymbol{I}_n) = \mathrm{O}(n)$, stabilizátor každej regulárnej symetrickej matice $\boldsymbol{Z} \in \mathbb{R}^{n\times n}$ je už izomorfný s niektorou z grúp $\mathrm{O}(n)$ či $\mathrm{O}(k, l)$.

Poznamenajme, že štúdium grúp $\mathrm{O}(k, l, \mathbb{C})$ by bolo nadbytočné. Každá z matíc $\mathrm{diag}(\boldsymbol{I}_k, -\boldsymbol{I}_l)$ je totiž nad poľom komplexných čísel kongruentná s maticou $\boldsymbol{I}_n$. Preto platí $\mathrm{O}(k, l, \mathbb{C}) \cong \mathrm{O}(n, \mathbb{C})$ a $\mathrm{SO}(k, l, \mathbb{C}) \cong \mathrm{SO}(n, \mathbb{C})$. Ale ani samotné grupy $\mathrm{O}(n, \mathbb{C})$ a $\mathrm{SO}(n, \mathbb{C})$ nie sú až také dôležité. Významnejšie sú ich hermitovské varianty, unitárna grupa $\mathrm{U}(n)$ a špeciálna unitárna grupa $\mathrm{SU}(n)$, s ktorými sa letmo zoznámime v nasledujúcom paragrafe.

Na druhej strane nad inými poľami ako $\mathbb{R}$ a $\mathbb{C}$, napríklad nad poľom $\mathbb{Q}$ všetkých racionálnych čísel, nemusí byť každá symetrická matica kongruentná s niektorou z diagonálnych matíc $\boldsymbol{I}_n$, $-\boldsymbol{I}_n$ či $\mathrm{diag}(\boldsymbol{I}_k, \boldsymbol{I}_l)$. Preto stabilizátor nie každej regulárnej symetrickej matice $\boldsymbol{Z} \in \mathbb{Q}^{n\times n}$ je nevyhnutne izomorfný s niektorou z grúp $\mathrm{O}(n, \mathbb{Q})$ či $\mathrm{O}(k, l, \mathbb{Q})$.

Ani všetkými „dôležitými“ reálnymi pseudortogonálnymi grupami sa však v tomto kurze nebudeme osobitne zaoberať. Obmedzíme sa len na grupy $\mathrm{O}(1,n)$, ktoré úzko súvisia so špeciálnou teóriou relativity a štruktúrou Minkowského časopriestoru $\mathbb{R}^{(1,n)}$ so štandardným pseudoskalárnym súčinom $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = x_0y_0 - x_1y_1 - \ldots - x_ny_n$. Grupu

$$\Lambda(n) = \mathrm{O}(1,n)$$

budeme nazývať *Lorentzovou grupou* stupňa n. *Poincarého grupou* nazývame jej afinné rozšírenie $\Lambda(n) \ltimes \mathbb{R}^{n+1}$.[2]

V súlade s tým budeme *Lorentzovými transformáciami* Minkowského časopriestoru $\mathbb{R}^{(1,n)}$ nazývať lineárne zobrazenia $\varphi\colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$, ktorých matice vzhľadom na kanonickú inerciálnu bázu $\boldsymbol{\varepsilon} = (\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ (s ktorými ich budeme stotožňovať) patria do Lorentzovej grupy $\Lambda(n)$. Ako o chvíľu uvidíme, Lorentzove transformácie, tak ako sme ich zaviedli v paragrafe 16.9, tvoria istú významnú podgrupu grupy $\Lambda(n)$.

Na začiatok vyčleníme niekoľko dôležitých podgrúp Lorentzovej grupy $\Lambda(n)$: popri nám už známej grupe $\mathrm{SO}(1,n)$ všetkých Lorentzových transformácií s determinantom 1, je to grupa

$$\Lambda^{\uparrow}(n) = \{\boldsymbol{A} \in \Lambda(n);\ \langle \boldsymbol{e}_0, \boldsymbol{A}\cdot \boldsymbol{e}_0\rangle > 0\}$$

všetkých *ortochrónnych Lorentzových transformácií*, t. j. takých, ktoré zachovávajú orientáciu časového šípu $\boldsymbol{e}_0$. Pri číslovaní riadkov a stĺpcov matíc $\boldsymbol{A} \in \mathbb{R}^{(n+1)\times(n+1)}$ od 0 do n je $\langle \boldsymbol{e}_0, \boldsymbol{A}\cdot\boldsymbol{e}_0\rangle = a_{00}$, teda Lorentzova transformácia $\boldsymbol{A} \in \Lambda(n)$ je ortochrónna práve vtedy, keď $a_{00} > 0$ (podľa cvičenia 16.1 je toto číslo, ako pseudoskalárny súčin dvoch časových šípov, vždy $\neq 0$). Práve týmto typom Lorentzových transformácií sme sa zaoberali v paragrafe 16.9. Ďalej je to grupa

$$\Lambda_{+}(n) = \{\boldsymbol{A} \in \Lambda(n);\ \langle \boldsymbol{e}_0, \boldsymbol{A}\cdot \boldsymbol{e}_0\rangle \det \boldsymbol{A} > 0\}$$

všetkých *vlastných Lorentzových transformácií*, t. j. takých ktoré zachovávajú orientáciu euklidovského podpriestoru $[\boldsymbol{e}_0]^{\perp} \subseteq \mathbb{R}^{(1,n)}$,[3] a prienik uvedených grúp

$$\Lambda_{+}^{\uparrow}(n) = \Lambda_{+}(n) \cap \Lambda^{\uparrow}(n) = \Lambda_{+}(n) \cap \mathrm{SO}(1,n) = \Lambda^{\uparrow}(n) \cap \mathrm{SO}(1,n),$$

[2]Vo väčšine učebníc sa pod Lorentzovou či Poincarého grupou rozumejú len fyzikálne relevantné grupy $\Lambda(3)$ resp. $\Lambda(3) \ltimes \mathbb{R}^4$. Našou všeobecnejšou definíciou sa chceme zavďačiť najmä „malým“ Lorentzovým grupám $\Lambda(1)$ a $\Lambda(2)$, náramne užitočným pri ich štúdiu. Lorentzova grupa $\Lambda(4)$ sa vyskytuje ako grupa izometrií istého kozmologického modelu, tzv. *de Sitterovho univerza.* S vyššími Lorentzovými grupami $\Lambda(n)$ sa zasa možno stretnúť v niektorých strunových teóriách.

[3]Najmä vo fyzike sa pod pojmom „vlastná Lorentzova grupa“ často rozumie špeciálna pseudoortogonálna grupa $\mathrm{SO}(1,n)$, jej prvkom sa potom prirodzene hovorí vlastné Lorentzove transformácie.

čiže grupa všetkých *vlastných ortochrónnych Lorentzových transformácií.*

Konečne je to tzv. *štvorgrupa* Γ tvorená maticami $\boldsymbol{I}_{n+1}$, $\boldsymbol{P}_{n+1} = \operatorname{diag}(1, \boldsymbol{I}_{n-1}, -1)$, $\boldsymbol{T}_{n+1} = \operatorname{diag}(-1, \boldsymbol{I}_{n-1}, 1)$ a $\boldsymbol{P}_{n+1}\boldsymbol{T}_{n+1} = \operatorname{diag}(-1, \boldsymbol{I}_{n-1}, -1)$, izomorfná s aditívnou grupou $\mathbb{Z}_2 \times \mathbb{Z}_2$. Ak vynecháme indexy v označení týchto matíc, môžeme písať

$$\Gamma = \{\boldsymbol{I}, \boldsymbol{P}, \boldsymbol{T}, \boldsymbol{PT}\}.$$

29.5.1. Veta. *Lorentzova grupa* $\Lambda(n)$ *je izomorfná s polopriamym súčinom* $\Gamma \ltimes \Lambda_+^{\uparrow}(n)$ *svojich podgrúp* $\Gamma \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ *a* $\Lambda_+^{\uparrow}(n)$.

Dôkaz. Označme $H = \{1, -1\} \times \{1, -1\}$ priamy súčin dvoch exemplárov multiplikatívnej grupy $\{1, -1\} \subseteq \mathbb{R}^\star$. Zrejme H je generovaná svojimi prvkami $(1, -1)$, $(-1, 1)$ a (nezávisle na n) je $H \cong \mathbb{Z}_2 \times \mathbb{Z}_2 \cong \Gamma$. Nech ďalej $\iota\colon \Lambda_+^{\uparrow}(n) \to \Lambda(n)$ je identické zobrazenie a $\pi\colon \Lambda(n) \to H$ je dané predpisom

$$\pi(\boldsymbol{A}) = \left(\frac{\langle \boldsymbol{e}_0, \boldsymbol{A} \cdot \boldsymbol{e}_0 \rangle}{|\langle \boldsymbol{e}_0, \boldsymbol{A} \cdot \boldsymbol{e}_0 \rangle|}, \det \boldsymbol{A} \right) = \left(\frac{a_{00}}{|a_{00}|}, \det \boldsymbol{A} \right),$$

pre $\boldsymbol{A} \in \Lambda(n)$. Niekoľkými priamymi výpočtami možno jednoducho overiť, že

$$\Lambda_+^{\uparrow}(n) \overset{\iota}{\rightarrowtail} \Lambda(n) \overset{\pi}{\twoheadrightarrow} H$$

je krátka exaktná postupnosť grúp rozštiepená homomorfizmom $\eta\colon H \to \Lambda(n)$, ktorý je jednoznačne určený obrazmi generátorov $\eta(1, -1) = \boldsymbol{P}$, $\eta(-1, 1) = \boldsymbol{PT}$. Preto podľa dôsledku 28.4.3 je

$$\Lambda(n) \cong H \ltimes \Lambda_+^{\uparrow}(n) \cong \Gamma \ltimes \Lambda_+^{\uparrow}(n).$$

Detaily prenechávame ako cvičenie čitateľovi.

Teda Lorentzova grupa je zjednotením

$$\Lambda(n) = \Lambda_+^{\uparrow}(n) \cup \Lambda_-^{\uparrow}(n) \cup \Lambda_+^{\downarrow}(n) \cup \Lambda_-^{\downarrow}(n)$$

štyroch disjunktných tried rozkladu

$$\begin{aligned} \Lambda_+^{\uparrow}(n) &= \Lambda_+^{\uparrow}(n) \cdot \boldsymbol{I}, & \Lambda_-^{\uparrow}(n) &= \Lambda_+^{\uparrow}(n) \cdot \boldsymbol{P}, \\ \Lambda_+^{\downarrow}(n) &= \Lambda_+^{\uparrow}(n) \cdot \boldsymbol{T}, & \Lambda_-^{\downarrow}(n) &= \Lambda_+^{\uparrow}(n) \cdot \boldsymbol{PT} \end{aligned}$$

podľa svojej normálnej podgrupy $\Lambda_+^{\uparrow}(n)$.

Násobenie maticou $\boldsymbol{P}$ sprava reprezentuje zmenu orientácie v priestore a maticou $\boldsymbol{T}$ zasa obrátenie smeru plynutia času. Kým výber matice $\boldsymbol{T}$ je viac-menej kanonický, úlohu matice $\boldsymbol{P}$ by mohla zohrať ľubovoľná matica tvaru $\mathrm{diag}(1, \pm 1, \dots, \pm 1)$ s nepárnym počtom členov -1 (t. j. s determinantom -1).

V euklidovskom priestore $\mathbb{R}^n$ má špeciálna ortogonálna grupa $\mathrm{SO}(n)$ v ortogonálnej grupe $\mathrm{O}(n)$ index 2, na základe čoho možno hovoriť o jeho dvoch možných orientáciách. Naproti tomu v Minkowského časopriestore $\mathbb{R}^{(1,n)}$ má vlastná ortochrónna Lorentzova grupa $\Lambda_+^\uparrow(n)$ v Lorentzovej grupe $\Lambda(n)$ index 4, v dôsledku čoho Minkowského časopriestoru pripisujeme štyri orientácie, dané kombináciami dvoch možných orientácií priestoru a dvoch možných orientácií času.

Na objasnenie štruktúry grupy $\Lambda(n)$ tak stačí poznať štruktúru jej podgrupy $\Lambda_+^\uparrow(n)$ vlastných ortochrónnych Lorentzových transformácií. Najprv podrobnejšie preskúmame „malú“ Lorentzovu grupu $\Lambda(1)$ a jej podgrupu $\Lambda_+^\uparrow(1)$.

29.5.2. Veta. *Nech $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$. Potom $\boldsymbol{A} \in \Lambda_+^\uparrow(1)$ práve vtedy, keď*

$$\boldsymbol{A} = \begin{pmatrix} \frac{1}{\sqrt{1-v^2}} & \frac{v}{\sqrt{1-v^2}} \\ \frac{v}{\sqrt{1-v^2}} & \frac{1}{\sqrt{1-v^2}} \end{pmatrix} = \boldsymbol{L}_v$$

pre nejaké $v \in \mathbb{R}$ také, že $|v| < 1$. Ďalej $\boldsymbol{A} \in \Lambda(1)$ práve vtedy, keď $\boldsymbol{A}$ je tvaru $\boldsymbol{L}_v$, $\boldsymbol{L}_v \cdot \boldsymbol{P}$, $\boldsymbol{L}_v \cdot \boldsymbol{T}$ alebo $\boldsymbol{L}_v \cdot \boldsymbol{PT}$.

Dôkaz. Nech $\boldsymbol{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Potom

$$\boldsymbol{A}^\mathsf{T} \cdot \mathrm{diag}(1,-1) \cdot \boldsymbol{A} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 - c^2 & ab - cd \\ ba - dc & b^2 - d^2 \end{pmatrix},$$

teda podmienka $\boldsymbol{A} \in \Lambda(1)$, t. j. $\boldsymbol{A}^\mathsf{T} \cdot \mathrm{diag}(1,-1) \cdot \boldsymbol{A} = \mathrm{diag}(1,-1)$, je ekvivalentná s rovnosťami

$$\begin{aligned} a^2 - c^2 &= 1, \\ ab - cd &= 0, \\ b^2 - d^2 &= -1. \end{aligned}$$

Z prvej rovnosti vyplýva $a^2 = 1 + c^2 > c^2 \geq 0$, teda $a \neq 0$ a $|a| > |c|$. Podobne, z tretej vyplýva $d^2 = 1 + b^2 > b^2 \geq 0$, teda $d \neq 0$ a $|d| > |b|$. Druhú rovnosť preto možno upraviť na tvar

$$\frac{b}{d} = \frac{c}{a}.$$

Označme túto spoločnú hodnotu $v = b/d = c/a$. Zrejme $|v| < 1$ a $b = dv$, $c = av$. Dosadením do prvej a tretej rovnosti dostávame

$$a^2(1 - v^2) = d^2(1 - v^2) = 1.$$

Teda
$$a = \pm\frac{1}{\sqrt{1-v^2}}, \qquad\qquad d = \pm\frac{1}{\sqrt{1-v^2}}.$$
Matica $\boldsymbol{A}$ tak nadobúda niektorý z tvarov
$$\boldsymbol{L}_v = \begin{pmatrix} \frac{1}{\sqrt{1-v^2}} & \frac{v}{\sqrt{1-v^2}} \\ \frac{v}{\sqrt{1-v^2}} & \frac{1}{\sqrt{1-v^2}} \end{pmatrix}, \qquad \boldsymbol{L}_v \cdot \boldsymbol{P} = \begin{pmatrix} \frac{1}{\sqrt{1-v^2}} & -\frac{v}{\sqrt{1-v^2}} \\ \frac{v}{\sqrt{1-v^2}} & -\frac{1}{\sqrt{1-v^2}} \end{pmatrix},$$
$$\boldsymbol{L}_v \cdot \boldsymbol{T} = \begin{pmatrix} -\frac{1}{\sqrt{1-v^2}} & \frac{v}{\sqrt{1-v^2}} \\ -\frac{v}{\sqrt{1-v^2}} & \frac{1}{\sqrt{1-v^2}} \end{pmatrix}, \qquad \boldsymbol{L}_v \cdot \boldsymbol{PT} = \begin{pmatrix} -\frac{1}{\sqrt{1-v^2}} & -\frac{v}{\sqrt{1-v^2}} \\ -\frac{v}{\sqrt{1-v^2}} & -\frac{1}{\sqrt{1-v^2}} \end{pmatrix}.$$
Naopak, priamym výpočtom sa možno presvedčiť, že každá z uvedených matíc je prvkom grupy $\Lambda(1)$. Zrejme $\boldsymbol{A} \in \Lambda_+^{\uparrow}(1)$ práve vtedy, keď $\boldsymbol{A} = \boldsymbol{L}_v$; zvyšné tri tvary zodpovedajú ostatným triedam rozkladu grupy $\Lambda(1)$ podľa podgrupy $\Lambda_+^{\uparrow}(1)$.

Pripomeňme si, že Lorentzovu transformáciu $\boldsymbol{L}_v$ možno substitúciou $\theta = \frac{1}{2}\ln\frac{1+v}{1-v}$ previesť do tvaru hyperbolickej rotácie
$$\boldsymbol{Rh}_\theta = \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix}$$
(pozri paragraf 16.9). Zrejme tiež naopak, $\boldsymbol{Rh}_\theta \in \Lambda_+^{\uparrow}(1)$ pre každé $\theta \in \mathbb{R}$. Konečne z cvičenia 27.20 vieme, že zobrazenie $\boldsymbol{Rh}\colon (\mathbb{R},+) \to (\mathrm{GL}(2,\mathbb{R}),\cdot)$ je prostý homomorfizmus grúp. Práve dokázaná veta, tak má nasledujúci

29.5.3. Dôsledok. $\Lambda_+^{\uparrow}(1) = \{\boldsymbol{L}_v;\ v \in \mathbb{R}\ \&\ |v| < 1\} = \{\boldsymbol{Rh}_\theta;\ \theta \in \mathbb{R}\} \cong \mathbb{R}$.

Nech teda $n \geq 2$. Uvažujme všeobecnú vlastnú ortochrónnu Lorentozovu transformáciu $\boldsymbol{A} \in \Lambda_+^{\uparrow}(n)$. Inak povedané, stŤpce $\boldsymbol{a}_0, \boldsymbol{a}_1, \dots, \boldsymbol{a}_n \in \mathbb{R}^{n+1}$ matice $\boldsymbol{A}$ tvoria inerciálnu bázu $\boldsymbol{\alpha}$ Minkowského časopriestoru $\mathbb{R}^{(1,n)}$, pričom platí $a_{00} > 0$ a $\det\boldsymbol{A} = 1 > 0$. Potom $\boldsymbol{A} = \boldsymbol{P}_{\boldsymbol{\varepsilon},\boldsymbol{\alpha}}$ je maticou prechodu z bázy $\boldsymbol{\alpha}$ do kanonickej bázy $\boldsymbol{\varepsilon}$. Keďže časové šípy $\boldsymbol{e}_0$ a $\boldsymbol{a}_0 = \boldsymbol{A}\cdot\boldsymbol{e}_0$ sú zhodne orientované, z našich úvah vykonaných v paragrafe 16.9 vyplýva, že ich možno doplniť do inerciálnych báz $\boldsymbol{\beta} = (\boldsymbol{e}_0, \boldsymbol{b}_1, \boldsymbol{b}_2, \cdot, \boldsymbol{b}_n)$ a $\boldsymbol{\gamma} = (\boldsymbol{a}_0, \boldsymbol{c}_1, \boldsymbol{c}_2, \dots, \boldsymbol{c}_n)$ takých, že $\boldsymbol{b}_2 = \boldsymbol{c}_2$, …, $\boldsymbol{b}_n = \boldsymbol{c}_n$ a matica prechodu medzi nimi má tvar
$$\boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\gamma}} = \boldsymbol{P}_{\boldsymbol{e}_0,\boldsymbol{a}_0} = \mathrm{diag}(\boldsymbol{L}_v, \boldsymbol{I}_{n-1}),$$
kde
$$v = \sqrt{1 - \langle \boldsymbol{e}_0, \boldsymbol{a}_0\rangle^{-2}} = \sqrt{1 - \langle \boldsymbol{e}_0, \boldsymbol{A}\cdot\boldsymbol{e}_0\rangle^{-2}} = \sqrt{1 - a_{00}^{-2}}.$$
Keďže navyše $\det\boldsymbol{A} > 0$, vynásobením priestorového vektora vektora $\boldsymbol{b}_2 = \boldsymbol{c}_2$ skalárom -1 (ak treba) možno docieliť, že ortonormálne bázy $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$,

$(\boldsymbol{b}_1, \dots \boldsymbol{b}_n)$ euklidovského priestoru $[\boldsymbol{e}_0]^\perp$ budú súhlasne orientované, rovnako ako ortonormálne bázy $(\boldsymbol{c}_1, \dots, \boldsymbol{c}_n)$, $(\boldsymbol{a}_1, \dots \boldsymbol{a}_n)$ euklidovského priestoru $[\boldsymbol{a}_0]^\perp$. Príslušné matice prechodu majú preto tvar

$$\boldsymbol{P}_{\boldsymbol{\varepsilon},\boldsymbol{\beta}} = \operatorname{diag}(1, \boldsymbol{B}), \qquad \boldsymbol{P}_{\boldsymbol{\gamma},\boldsymbol{\alpha}} = \operatorname{diag}(1, \boldsymbol{C}),$$

kde $\boldsymbol{B}, \boldsymbol{C} \in \mathrm{SO}(n)$. Maticu $\boldsymbol{A}$ tak možno rozložiť na súčin

$$\begin{aligned}\boldsymbol{A} = \boldsymbol{P}_{\boldsymbol{\varepsilon},\boldsymbol{\alpha}} &= \boldsymbol{P}_{\boldsymbol{\varepsilon},\boldsymbol{\beta}} \cdot \boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\gamma}} \cdot \boldsymbol{P}_{\boldsymbol{\gamma},\boldsymbol{\alpha}} \\ &= \operatorname{diag}(1, \boldsymbol{B}) \cdot \operatorname{diag}(\boldsymbol{L}_v, \boldsymbol{I}_{n-1}) \cdot \operatorname{diag}(1, \boldsymbol{C})\end{aligned}$$

špeciálnej Lorentzovej transformácie $\operatorname{diag}(\boldsymbol{L}_v, \boldsymbol{I}_{n-1})$ a dvoch Lorentzových transformácií $\operatorname{diag}(1, \boldsymbol{B})$, $\operatorname{diag}(1, \boldsymbol{C})$, ktoré sú zároveň prvkami špeciálnej ortogonálnej (euklidovskej) grupy $\mathrm{SO}(n+1)$. Tým sme vlastne dokázali nasledujúcu vetu.

29.5.4. Veta. *Pre $n \geq 2$ možno každú vlastnú ortochrónnu Lorentzovu transformáciu $\boldsymbol{A} \in \Lambda_+^\uparrow(n)$ vyjadriť v tvare kompozície*

$$\boldsymbol{A} = \operatorname{diag}(1, \boldsymbol{B}) \cdot \operatorname{diag}(\boldsymbol{L}_v, \boldsymbol{I}_{n-1}) \cdot \operatorname{diag}(1, \boldsymbol{C})$$

špeciálnej Lorentzovej transformácie $\operatorname{diag}(\boldsymbol{L}_v, \boldsymbol{I}_{n-1})$, *kde* $v = \sqrt{1 - a_{00}^{-2}}$, *a dvoch Lorentzových transformácií* $\operatorname{diag}(1, \boldsymbol{B}), \operatorname{diag}(1, \boldsymbol{C}) \in \mathrm{SO}(n+1)$, *kde matice* $\boldsymbol{B}, \boldsymbol{C} \in \mathrm{SO}(n)$ *zodpovedajú priamym zhodnostiam n-rozmerných euklidovských podpriestorov $[\boldsymbol{e}_0]^\perp$ resp. $[\boldsymbol{A} \cdot \boldsymbol{e}_0]^\perp$ v $\mathbb{R}^{(1,n)}$.*

Špeciálne pre $n = 2$, resp. pre fyzikálne relevantný prípad $n = 3$ dostávame nasledujúci

29.5.5. Dôsledok. *(a) V Minkowského časopriestore $\mathbb{R}^{(1,2)}$ možno každú vlastnú ortochrónnu Lorentzovu transformáciu $\boldsymbol{A} \in \Lambda_+^\uparrow(2)$ vyjadriť v tvare kompozície*

$$\boldsymbol{A} = \operatorname{diag}(1, \boldsymbol{R}_\alpha) \cdot \operatorname{diag}(\boldsymbol{L}_v, 1) \cdot \operatorname{diag}(1, \boldsymbol{R}_\beta)$$

špeciálnej Lorentzovej transformácie $\operatorname{diag}(\boldsymbol{L}_v, 1)$, *kde* $v = \sqrt{1 - a_{00}^{-2}}$, *a dvoch Lorentzových transformácií* $\operatorname{diag}(1, \boldsymbol{R}_\alpha), \operatorname{diag}(1, \boldsymbol{R}_\beta) \in \mathrm{SO}(3)$, *kde matice* $\boldsymbol{R}_\alpha, \boldsymbol{R}_\beta \in \mathrm{SO}(2)$ *zodpovedajú otočeniam euklidovských rovín $[\boldsymbol{e}_1, \boldsymbol{e}_2]$ resp. $[\boldsymbol{A} \cdot \boldsymbol{e}_1, \boldsymbol{A} \cdot \boldsymbol{e}_2]$ v $\mathbb{R}^{(1,2)}$.*

(b) V Minkowského časopriestore $\mathbb{R}^{(1,3)}$ možno každú vlastnú ortochrónnu Lorentzovu transformáciu $\boldsymbol{A} \in \Lambda_+^\uparrow(3)$ vyjadriť v tvare kompozície

$$\boldsymbol{A} = \operatorname{diag}(1, \boldsymbol{R}_\alpha^{\boldsymbol{u}}) \cdot \operatorname{diag}(\boldsymbol{L}_v, 1, 1) \cdot \operatorname{diag}(1, \boldsymbol{R}_\beta^{\boldsymbol{w}})$$

špeciálnej Lorentzovej transformácie $\operatorname{diag}(\boldsymbol{L}_v, 1, 1)$, *kde* $v = \sqrt{1 - a_{00}^{-2}}$, *a dvoch Lorentzových transformácií* $\operatorname{diag}(1, \boldsymbol{R}_\alpha^{\boldsymbol{u}}), \operatorname{diag}(1, \boldsymbol{R}_\beta^{\boldsymbol{w}}) \in \mathrm{SO}(4)$, *kde matice* $\boldsymbol{R}_\alpha^{\boldsymbol{u}}, \boldsymbol{R}_\beta^{\boldsymbol{w}} \in \mathrm{SO}(3)$ *zodpovedajú otočeniam trojozmerných euklidovských podpriestorov $[\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]$ resp. $[\boldsymbol{A} \cdot \boldsymbol{e}_1, \boldsymbol{A} \cdot \boldsymbol{e}_2, \boldsymbol{A} \cdot \boldsymbol{e}_3]$ v $\mathbb{R}^{(1,3)}$.*

Transformáciám Minkovského časopriestoru $\mathbb{R}^{(1,3)}$, ktoré majú v nejakej inerciálnej báze $\boldsymbol{\alpha}$ maticu tvaru $\operatorname{diag}(\boldsymbol{L}_v, 1, 1)$, sa najmä v anglojazyčnej fyzikálnej literatúre zvykne hovoriť *boosty*. Taký *boost* má vzhľadom na kanonickú bázu $\boldsymbol{\varepsilon}$ maticu tvaru $\boldsymbol{P}_{\boldsymbol{\varepsilon},\boldsymbol{\alpha}} \cdot \operatorname{diag}(\boldsymbol{L}_v, 1, 1) \cdot \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\varepsilon}}$. Podľa dôsledku 29.5.5 (b) tak možno každú Lorentzovu transformáciu $\boldsymbol{A} \in \Lambda_+^{\uparrow}(3)$ vyjadriť ako kompozíciu *boostu* (výrazy vo veľkých zátvorkách) a euklidovskej rotácie

$$\begin{aligned}
\boldsymbol{A} &= \operatorname{diag}(1, \boldsymbol{B}) \cdot \operatorname{diag}(\boldsymbol{L}_v, 1, 1) \cdot \operatorname{diag}(1, \boldsymbol{C}) \\
&= \Big(\operatorname{diag}(1, \boldsymbol{B}) \cdot \operatorname{diag}(\boldsymbol{L}_v, 1, 1) \cdot \operatorname{diag}\big(1, \boldsymbol{B}^{-1}\big)\Big) \cdot \operatorname{diag}(1, \boldsymbol{BC}) \\
&= \operatorname{diag}(1, \boldsymbol{BC}) \cdot \Big(\operatorname{diag}\big(1, \boldsymbol{C}^{-1}\big) \cdot \operatorname{diag}(\boldsymbol{L}_v, 1, 1) \cdot \operatorname{diag}(1, \boldsymbol{C})\Big),
\end{aligned}$$

kde $\boldsymbol{B}, \boldsymbol{C} \in \mathrm{SO}(3)$. Keďže však kompozícia dvoch *boostov* nie je vo všeobecnosti *boostom*, z hľadiska teórie grúp nie je tento pojem až taký dôležitý.

Poznámka. Pri odvodzovaní matematickej podoby špeciálnej Lorentzovej transformácie môže v dôsledku nie príliš jasnej argumentácie niekedy vzniknúť dojem, že jej tvar je dôsledkom výlučne postulátu stálosti rýchlosti svetla vo všetkých inerciálnych sústavách (a šikovnej voľby inerciálnych báz prislúchajúcich uvažovaným pozorovateľom). Niekedy sa otvorene prijme aj predpoklad linearity hľadanej transformácie. Z nich sa potom nejako „vyargumentuje“ nevyhnutnosť zachovávania časopriestorovej odľahlosti. Ako sme videli, od tohto bodu už ide všetko hladko. Z postulátu stálosti rýchlosti svetla však bezprostredne vyplýva len požiadavka, aby hľadaná transformácia φ Minkowského časopriestoru V *zachovávala svetelný kužeľ*, t. j. spĺňala podmienku

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0 \;\Leftrightarrow\; \langle \varphi \boldsymbol{x}, \varphi \boldsymbol{x} \rangle = 0$$

pre každé $\boldsymbol{x} \in V$. V cvičení 29.12 si ukážeme, že už na Minkowského časopriamke $\mathbb{R}^{(1,1)}$ existujú bijektívne (lineárne aj nelineárne) transformácie, ktoré zachovávajú svetelný kužeľ, no nezachovávajú časopriestorovú odľahlosť ani pseudoskalárny súčin, teda nepatria do Lorentzovej grupy $\Lambda(1)$. Inak povedané, či už je matematický predpoklad zachovávania časopriestorovej odľahlosti fyzikálne dobre zdôvodniteľný alebo nie, bez neho to nejde, a predpoklad linearity je už po jeho prijatí nadbytočný.

29.6 Unitárna a špeciálna unitárna grupa

Nech $n = k + l > 0$, kde k, l sú kladné celé čísla. Potom stabilzátory matíc $\boldsymbol{I}_n$ a $\operatorname{diag}(\boldsymbol{I}_k, -\boldsymbol{I}_l)$ vzhľadom na akciu $(\boldsymbol{X}, \boldsymbol{A}) \mapsto \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{X} \cdot \overline{\boldsymbol{A}}$ všeobecnej lineárnej grupy $\mathrm{GL}(n,\mathbb{C})$ na priestore všetkých matic $\mathbb{C}^{n\times n}$ tvoria podgrupy

$$\begin{aligned}\mathrm{U}(n) &= \mathrm{Stb}(\boldsymbol{I}_n) = \big\{\boldsymbol{A}\in\mathrm{GL}(n,\mathbb{C});\ \boldsymbol{A}^{\mathsf{T}}\cdot\overline{\boldsymbol{A}} = \boldsymbol{I}_n\big\},\\ \mathrm{U}(k,l) &= \mathrm{Stb}\big(\mathrm{diag}(\boldsymbol{I}_k,-\boldsymbol{I}_l)\big)\\ &= \big\{\boldsymbol{A}\in\mathrm{GL}(n,\mathbb{C});\ \boldsymbol{A}^{\mathsf{T}}\cdot\mathrm{diag}(\boldsymbol{I}_k,-\boldsymbol{I}_l)\cdot\overline{\boldsymbol{A}} = \mathrm{diag}(\boldsymbol{I}_k,-\boldsymbol{I}_l)\big\}\end{aligned}$$

grupy $\mathrm{GL}(n,\mathbb{C})$, nazývané *unitárna grupa* stupňa n, resp. *pseudounitárna grupa* typu (k,l). Ich podgrupy

$$\begin{aligned}\mathrm{SU}(n) &= \mathrm{U}(n)\cap\mathrm{SL}(n,\mathbb{C}) = \{\boldsymbol{A}\in\mathrm{U}(n);\ \det\boldsymbol{A}=1\},\\ \mathrm{SU}(k,l) &= \mathrm{U}(k,l)\cap\mathrm{SL}(n,\mathbb{C}) = \{\boldsymbol{A}\in\mathrm{U}(k,l);\ \det\boldsymbol{A}=1\}\end{aligned}$$

nazývame *špeciálna unitárna grupa* stupňa n resp. *špeciálna pseudounitárna grupa* typu (k,l).

Zrejme pre $\boldsymbol{A}\in\mathrm{GL}(n,\mathbb{C})$ platí

$$\begin{aligned}\boldsymbol{A}\in\mathrm{U}(n) &\iff \boldsymbol{A}^{-1}=\boldsymbol{A}^*,\\ \boldsymbol{A}\in\mathrm{U}(k,l) &\iff \boldsymbol{A}^*\cdot\mathrm{diag}(\boldsymbol{I}_k,-\boldsymbol{I}_l)\cdot\boldsymbol{A} = \mathrm{diag}(\boldsymbol{I}_k,-\boldsymbol{I}_l),\end{aligned}$$

a stabilizátor každej regulárnej kososymetrickej matice $\boldsymbol{Z}\in\mathbb{C}^{n\times n}$ vzhľadom na uvedenú akciu je už izomorfný s niektorou z grúp $\mathrm{U}(n)$, $\mathrm{U}(k,l)$.

Lineárne grupy $\mathrm{U}(n)$, $\mathrm{U}(k,l)$, $\mathrm{SU}(n)$ a $\mathrm{SU}(k,l)$ sú tak komplexnou analógiou reálnych lineárnych grúp $\mathrm{O}(n)$, $\mathrm{O}(k,l)$, $\mathrm{SO}(n)$ a $\mathrm{SO}(k,l)$. Ani tentokrát sa im však nebudeme venovať všetkým. Obmedzíme sa len na jeden všeobecný výsledok o vzťahu unitárnej grupy $\mathrm{U}(n)$ a špeciálnej unitárnej grupy $\mathrm{SU}(n)$, ktorým doplníme paragrafy 17.4 a 23.4. Potom si podrobnejšie všimneme špeciálnu unitárnu grupu $\mathrm{SU}(2)$.

Začneme triviálnym pozorovaním, ktorým si pripomenieme, čo už vlastne vieme z príkladov 27.2.4 a 27.5.6: keďže pre $c\in\mathbb{C}$ platí $c^*=\bar{c}$ a $c\bar{c}=|c|^2$, unitárna grupa

$$\mathrm{U}(1) = \{c\in\mathbb{C};\ |c|=1\}$$

splýva s jednotkovou kružnicou v Gaussovej komplexnej rovine. Priradením $\mathrm{e}^{\mathrm{i}\alpha}\mapsto\boldsymbol{R}_\alpha$ je daný izomorfizmus grúp $\mathrm{U}(1)\cong\mathrm{SO}(2)$

Ak uvažujeme $\mathbb{C}^n$ ako unitárny priestor vybavený štandardným skalárnym súčinom $\langle\boldsymbol{x},\boldsymbol{y}\rangle = \boldsymbol{x}^{\mathsf{T}}\cdot\overline{\boldsymbol{y}}$, tak matica $\boldsymbol{A}\in\mathbb{C}^{n\times n}$ je zrejme unitárna práve vtedy, keď jej stĺpce tvoria ortonormálnu bázu v $\mathbb{C}^n$. Podľa tvrdenia 23.4.4 sú všetky vlastné hodnoty unitárnej matice $\boldsymbol{A}\in\mathbb{C}^{n\times n}$ komplexné jednotky (t. j. prvky $\mathrm{U}(1)$) a podľa vety 23.4.7 je $\boldsymbol{A}$ podobná s diagonálnou maticou prostredníctvom unitárnej matice prechodu.

Ukazuje sa, že grupu $\mathrm{U}(n)$ možno popísať ako polopriamy súčin grupy $\mathrm{SU}(n)$ a jednotkovej kružnice $\mathrm{U}(1)$. Úplne rovnako ako v dôkaze tvrdenia 29.1.3 možno totiž nahliadnuť, že grupové homomorfizmy

$$\mathrm{SU}(n)\overset{\iota}{\rightarrowtail}\mathrm{U}(n)\overset{\delta}{\twoheadrightarrow}\mathrm{U}(1),$$

kde ι je identické zobrazenie a δ priradí matici $\boldsymbol{A} \in \mathrm{U}(n)$ jej determinant, tvoria krátku exaktnú postupnosť, rozštiepenú napríklad homomorfizmom $\eta\colon \mathrm{U}(1) \to \mathrm{U}(n)$, $u \mapsto \mathrm{diag}(\boldsymbol{I}_{n-1}, u)$.

29.6.1. Veta. *Pre každé $n \geq 1$ je unitárna grupa $\mathrm{U}(n)$ izomorfná s polopriamym súčinom $\mathrm{U}(1) \ltimes \mathrm{SU}(n)$ jednotkovej kružnice $\mathrm{U}(1)$ a špeciálnej unitárnej grupy $\mathrm{SU}(n)$.*

Prejdime teraz k avizovanej grupe $\mathrm{SU}(2)$. Podľa vety 17.4.3 je matica $\boldsymbol{A} \in \mathbb{C}^{2\times 2}$ unitárna práve vtedy, keď má tvar

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ -\bar{b}u & \bar{a}u \end{pmatrix},$$

pre nejaké $a, b, u \in \mathbb{C}$ také, že $|a|^2 + |b|^2 = |u| = 1$. Potom $u = \det \boldsymbol{A}$. Ako špeciálny prípad pre $u = 1$ tak dostávame

29.6.2. Veta. *Nech $\boldsymbol{A} \in \mathbb{C}^{2\times 2}$. Potom $\boldsymbol{A} \in \mathrm{SU}(2)$ práve vtedy, keď existujú komplexné čísla a, b také, že $|a|^2 + |b|^2 = 1$ a*

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}.$$

Po rozklade komplexných čísel $a = a_0 + a_1\mathrm{i}$, $b = b_0 + b_1\mathrm{i}$ na reálnu a imaginárnu časť možno špeciálnu unitárnu grupu $\mathrm{SU}(2)$ priamo stotožniť s jednotkovou trojrozmernou sférou

$$S^{(3)} = \left\{(a_0, a_1, b_0, b_1) \in \mathbb{R}^4\,;\ a_0^2 + a_1^2 + b_0^2 + b_1^2 = 1\right\}$$

v štvorrozmernom euklidovskom priestore $\mathbb{R}^4$. Len ako zaujímavosť poznamenajme, že zo všetkých jednotkových sfér

$$S^{(n)} = \left\{(x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1}\,;\ x_0^2 + x_1^2 + \ldots + x_n^2 = 1\right\}$$

rôznych rozmerov $n \geq 0$ jedine sféry $S^{(0)} = \{1, -1\}$, $S^{(1)} \cong \mathrm{U}(1)$ a $S^{(3)} \cong \mathrm{SU}(2)$ nesú „prirodzenú" grupovú štruktúru. Posledný vzťah konkretizujeme do podoby grupového izomorfizmu v nasledujúcej kapitole v časti venovanej *kvaterniónom.*

29.7 Súvislé komponenty a orientácia

Pri našom sľúbenom výlete do topológie lineárnych grúp sa obmedzíme len na maticové grupy. Začneme však v mierne všeobecnejšom kontexte podmnožín priestoru $\mathbb{R}^n$.

Nech $X \subseteq \mathbb{R}^n$ pre nejaké $n \geq 1$. Pod *cestou* alebo *oblúkom* v množine X rozumieme ľubovoľné spojité zobrazenie $f\colon \langle a, b\rangle \to X$, kde $\langle a, b\rangle \subseteq \mathbb{R}$ je netriviálny uzavretý interval. Ak $f(a) = x$, $f(b) = y$, hovoríme, že f je cesta z x do y v množine X. *Množina* X sa nazýva *súvislá*, ak každé dva body $x, y \in X$ možno spojiť cestou v X.[4] Pri tom sa zrejme stačí obmedziť na cesty $f\colon \langle 0, 1\rangle \to X$.

Pre $x, y \in X$ položme $x \sim_X y$ práve vtedy, keď x a y možno spojiť cestou v X. Zrejme $\sim_X$ je relácia ekvivalencie na množine X. Triedu ekvivalencie bodu $x \in X$, t. j. množinu $\{y \in X;\ x \sim_X y\} \in X/{\sim_X}$, nazývame *komponentou súvislosti* alebo *súvislou komponentou* bodu x v množine X.

29.7.1. Príklad. (a) Každý interval v $\mathbb{R}$ (otvorený či uzavretý, ohraničený či neohraničený, možno dokonca prázdny alebo jednobodový) je zrejme súvislá množina. Naopak, ak množina $X \subseteq \mathbb{R}$ nie je interval, t. j. existujú body $x < y < z$ také, že $x, z \in X$ a $y \notin X$, tak X nie je súvislá. Každá cesta v $\mathbb{R}$ z x do z totiž musí prejsť cez y, teda nemôže celá ležať v X.

(b) Ak $X \subseteq \mathbb{R}^n$ je konečná množina, ktorá má aspoň dva prvky, tak X je nesúvislá. Všeobecnejšie, všetky komponenty súvislosti konečnej neprázdnej množiny sú jednoprvkové.

(c) Každý afinný podpriestor $M \subseteq \mathbb{R}^n$ je súvislá množina. Pre ľubovoľné $\boldsymbol{p}, \boldsymbol{q} \in M$ je totiž funkcia $f(t) = (1 - t)\boldsymbol{p} + t\boldsymbol{q}$, kde $t \in \langle 0, 1\rangle$, t. j. prirodzená parametrizácia úsečky spájajúcej body $\boldsymbol{p}$ a $\boldsymbol{q}$, zároveň cestou z $\boldsymbol{p}$ do $\boldsymbol{q}$ v M.

29.7.2. Tvrdenie. *Nech $X \subseteq \mathbb{R}^m$, $Y \subseteq \mathbb{R}^n$ a $F\colon X \to Y$ je surjektívne spojité zobrazenie. Ak množina X je súvislá, tak aj množina Y je súvislá; naopak, ak Y je nesúvislá, tak aj X je nesúvislá.*

Dôkaz. Stačí si uvedomiť, že ak je $f\colon \langle 0, 1\rangle \to X$ cesta z x do y, tak $F \circ f\colon \langle 0, 1\rangle \to Y$ je cesta z $F(x)$ do $F(y)$.

29.7.3. Tvrdenie. *Nech $X \subseteq \mathbb{R}^m$, $Y \subseteq \mathbb{R}^n$. Potom*
(a) množina $X \times Y \subseteq \mathbb{R}^{m+n}$ je súvislá práve vtedy, keď obe množiny X, Y sú súvislé;
(b) súvislá komponenta bodu $(x, y) \in X \times Y$ je karteziánskym súčinom súvislej komponenty bodu x v X a súvislej komponenty bodu y v X.

Dôkaz. (a) Keďže obe projekcie $(x, y) \mapsto x\colon X \times Y \to X$, $(x, y) \mapsto y\colon X \times Y \to Y$ sú spojité, súvislosť množiny $X \times Y$ implikuje súvislosť X aj Y podľa predchádzajúceho tvrdenia.

[4] Tento druh súvislosti sa v topológii zvykne nazývať *oblúkovou súvislosťou*, zatiaľ čo pod *súvislosťou* sa rozumie o niečo všeobecnejší (no menej názorný) pojem, definovaný ako nemožnosť rozkladu na dve otvorené podmnožiny. Keďže sa však mienime zaoberať výlučne oblúkovou súvislosťou, nebudeme si zbytočne komplikovať názvoslovie.

Nech naopak X aj Y sú súvislé. Zvoľme ľubovoľné $(x_0, y_0), (x_1, y_1) \in X \times Y$. Existujú cesty $f\colon \langle 0,1\rangle \to X$ z x_0 do x_1 a $g\colon \langle 0,1\rangle \to Y$ z y_0 do y_1. Potom $h(t) = (f(t), g(t))$, $t \in \langle 0,1\rangle$, je cesta z (x_0, y_0) do (x_1, y_1) v $X \times Y$.

(b) priamo vyplýva z (a).

Ďalej nám pôjde hlavne o súvislosť v maticových grupách nad poľami $\mathbb{R}$ alebo $\mathbb{C}$. Uvedomme si preto, že každá maticová grupa $G \subseteq \mathrm{GL}(n,\mathbb{R})$ je zároveň podmnožinou n^2-rozmerného priestoru $\mathbb{R}^{n\times n}$. Podobne, po stotožnení komplexných čísel s usporiadanými dvojicami reálnych čísel, možno každú grupu $G \subseteq \mathrm{GL}(n,\mathbb{C})$ považovať za podmnožinu $2n^2$-rozmerného priestoru $\mathbb{C}^{n\times n} \cong \mathbb{R}^{2n^2}$. Ďalej je užitočné si uvedomiť, že pre obe z uvedených polí, t. j. pre $K = \mathbb{R}$ aj $K = \mathbb{C}$, sú grupové operácie násobenia $\cdot : \mathrm{GL}(n,K) \times \mathrm{GL}(n,K) \to \mathrm{GL}(n,K)$ a inverzného prvku ${}^{-1}\colon \mathrm{GL}(n,K) \to \mathrm{GL}(n,K)$, ako i determinant $\det\colon \mathrm{GL}(n,K) \to K^\star$ spojité zobrazenia. Taktiež exponenciála $\boldsymbol{A} \mapsto \mathrm{e}^{\boldsymbol{A}}$ je spojité zobrazenie $K^{n\times n} \to \mathrm{GL}(n,K)$.

Nasledujúce tvrdenie je bezprostredným dôsledkom tvrdení 29.7.2 a 29.7.3 (a) a spojitosti násobenia v $\mathrm{GL}(n,K)$.

29.7.4. Tvrdenie. *Nech $G \subseteq \mathrm{GL}(n,K)$ je maticová grupa nad poľom $K \in \{\mathbb{R},\mathbb{C}\}$ a $\mathcal{A}, \mathcal{B} \subseteq G$ sú súvislé množiny. Potom aj množina*

$$\mathcal{A}\cdot\mathcal{B} = \{\boldsymbol{A}\cdot\boldsymbol{B};\ \boldsymbol{A} \in \mathcal{A}\ \&\ \boldsymbol{B} \in \mathcal{B}\}$$

je súvislá.

29.7.5. Veta. *Nech $G \subseteq \mathrm{GL}(n,K)$ je maticová grupa nad poľom $K \in \{\mathbb{R},\mathbb{C}\}$. Potom súvislá komponenta*

$$\Upsilon(G) = \{\boldsymbol{A} \in G;\ \boldsymbol{A} \sim_G \boldsymbol{I}_n\}$$

jednotkovej matice $\boldsymbol{I}_n$ je normálna podgrupa grupy G. Súvislá komponenta matice $\boldsymbol{A} \in G$ je trieda rozkladu $\Upsilon(G)\boldsymbol{A} \in G/\Upsilon(G)$ a index $[G : \Upsilon(G)]$ udáva počet komponent súvislosti grupy G.

Dôkaz. Zrejme $\boldsymbol{I}_n \in \Upsilon(G)$. Nech $\boldsymbol{A}, \boldsymbol{B} \in \Upsilon(G)$ a $f, g\colon \langle 0,1\rangle \to G$ sú cesty také, že $f(0) = g(0) = \boldsymbol{I}_n$, $f(1) = \boldsymbol{A}$, $g(1) = B$. Položme $h(t) = f(t)\,g(t)^{-1}$ pre $t \in \langle 0,1\rangle$. Potom h je cesta z $\boldsymbol{I}_n$ do $\boldsymbol{A}\boldsymbol{B}^{-1}$ v G. Ak $\boldsymbol{X} \in G$ je ľubovoľná matica, tak $j(t) = \boldsymbol{X}\,f(t)\boldsymbol{X}^{-1}$ definuje cestu j z $\boldsymbol{I}_n$ do $\boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^{-1}$ v G. Teda $\boldsymbol{A}\boldsymbol{B}^{-1}$ aj $\boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^{-1}$ patria do $\Upsilon(G)$, čo dokazuje, že $\Upsilon(G) \lhd G$. Druhé tvrdenie je bezprostredným dôsledkom prvého.

Preskúmajme teraz, ako je to so súvislosťou niektorých vybraných maticových grúp nad $\mathbb{R}$ resp. $\mathbb{C}$.

29.7.6. Veta. *Pre každé $n \geq 1$ nasledujúce maticové grupy sú súvislé:*

(a) špeciálna ortogonálna grupa $\mathrm{SO}(n)$*;*

(b) unitárna grupa $\mathrm{U}(n)$ *a špeciálna unitárna grupa* $\mathrm{SU}(n)$*;*

(c) všeobecná lineárna grupa $\mathrm{GL}(n, \mathbb{C})$ *a obe špeciálne lineárne grupy* $\mathrm{SL}(n, \mathbb{C})$ *a* $\mathrm{SL}(n, \mathbb{R})$*;*

(d) vlastná ortochrónna Lorentzova grupa $\Lambda_+^{\uparrow}(n)$*.*

Dôkaz. V každom z uvedených prípadov stačí dokázať, že ľubovoľnú maticu $\boldsymbol{A}$ z danej grupy možno v nej spojiť cestou s jednotkovou maticou $\boldsymbol{I} = \boldsymbol{I}_n$.

(a) Nech napr. $n = 2m + 1$ je nepárne. Zvoľme $\boldsymbol{A} \in \mathrm{SO}(n)$. Podľa vety 23.5.3 (pozri tiež paragraf 29.4) existuje ortogonálna matica $\boldsymbol{P}$ a reálne čísla $\alpha_1, \ldots, \alpha_m$ také, že

$$\boldsymbol{A} = \boldsymbol{P} \cdot \operatorname{diag}(\boldsymbol{R}_{\alpha_1}, \ldots, \boldsymbol{R}_{\alpha_m}, 1) \cdot \boldsymbol{P}^{-1}.$$

Položme

$$f(t) = \boldsymbol{P} \cdot \operatorname{diag}(\boldsymbol{R}_{\alpha_1 t}, \ldots, \boldsymbol{R}_{\alpha_m t}, 1) \cdot \boldsymbol{P}^{-1}$$

pre $0 \leq t \leq 1$. Potom f je cesta v $\mathrm{SO}(n)$, $f(0) = \boldsymbol{I}$ a $f(1) = \boldsymbol{A}$. Podobne možno postupovať aj pre párne n.

(b) Nech $\boldsymbol{A} \in \mathrm{U}(n)$. Podľa vety 23.4.9 existuje hermitovská matica $\boldsymbol{H}$ taká, že $\boldsymbol{A} = \mathrm{e}^{\mathrm{i}\boldsymbol{H}}$. Potom $f(t) = \mathrm{e}^{\mathrm{i}t\boldsymbol{H}}$ je hľadaná cesta z $\boldsymbol{I}$ do $\boldsymbol{A}$ v $\mathrm{U}(n)$.

Ak $\boldsymbol{A} \in \mathrm{SU}(n)$, tak $\boldsymbol{H}$ možno podľa vety 23.4.9 zvoliť tak, aby navyše platilo $\operatorname{tr} \boldsymbol{H} = 0$. Potom však $\operatorname{tr} t\boldsymbol{H} = 0$ a $\det f(t) = 1$ pre každé $t \in \langle 0, 1 \rangle$, teda f je cesta v $\mathrm{SU}(n)$.

(c) Nech $\boldsymbol{A} \in \mathrm{GL}(n, \mathbb{C})$. Podľa vety 25.1.3 o polárnom rozklade $\boldsymbol{A} = \boldsymbol{R} \cdot \boldsymbol{U}$ pre nejakú kladne definitnú hermitovskú maticu $\boldsymbol{R}$ a unitárnu maticu $\boldsymbol{U}$. Z (b) už vieme, že množina $\mathrm{U}(n) \subseteq \mathrm{GL}(n, \mathbb{C})$ je súvislá. Vzhľadom na tvrdenie 29.7.4 stačí dokázať, že aj množina $\mathcal{H}^+$ všetkých kladne definitných hermitovských matíc $\boldsymbol{R} \in \mathbb{C}^{n \times n}$ je súvislá. Podľa vety 23.3.2 (a) má každá taká matica tvar $\boldsymbol{R} = \boldsymbol{P} \cdot \boldsymbol{D} \cdot \boldsymbol{P}^{-1}$, kde $\boldsymbol{P} \in \mathrm{U}(n)$ a $\boldsymbol{D} \in \mathbb{R}^{n \times n}$ je diagonálna matica s diagonálou tvorenou kladnými číslami. Potom predpis

$$f(t) = \boldsymbol{P} \cdot \big((1-t)\boldsymbol{I} + t\boldsymbol{D}\big) \cdot \boldsymbol{P}^{-1},$$

pre $t \in \langle 0, 1 \rangle$, určuje cestu z $\boldsymbol{I}$ do $\boldsymbol{R}$ v $\mathcal{H}^+$.

Ak $\boldsymbol{A} \in \mathrm{SL}(n, \mathbb{C})$, tak z podmienok $\det \boldsymbol{R} > 0$ a $\det \boldsymbol{R} \det \boldsymbol{U} = \det \boldsymbol{A} = 1$ vyplýva $\det \boldsymbol{U} = 1$, čiže $\boldsymbol{U} \in \mathrm{SU}(n)$, a následne aj $\det \boldsymbol{R} = 1$. Vzhľadom na súvislosť grupy $\mathrm{SU}(n)$ opäť stačí dokázať, že existuje cesta z $\boldsymbol{I}$ do $\boldsymbol{R}$ v množine $\mathcal{H}_1^+$ všetkých kladne definitných hermitovských matíc $\boldsymbol{R} \in \mathbb{C}^{n \times n}$ s determinantom 1. Diagonálnu maticu $\boldsymbol{D}$ možno zapísať v tvare $\boldsymbol{D} = \operatorname{diag}\big(\mathrm{e}^{a_1}, \ldots, \mathrm{e}^{a_n}\big)$, kde $a_1, \ldots, a_n \in \mathbb{R}$ a $\sum_j a_j = 0$. Potom

$$g(t) = \boldsymbol{P} \cdot \operatorname{diag}\big(\mathrm{e}^{a_1 t}, \ldots, \mathrm{e}^{a_n t}\big) \cdot \boldsymbol{P}^{-1},$$

kde $0 \leq t \leq 1$, je cesta v $\mathcal{H}_1^+$ z $\boldsymbol{I}$ do $\boldsymbol{R}$.

Ak $\boldsymbol{A} \in \mathrm{SL}(n, \mathbb{R})$, tak matice $\boldsymbol{R}$, $\boldsymbol{U}$ možno zvoliť reálne. Presnejšie, $\boldsymbol{R}$ je symetrická, kladne definitná a $\boldsymbol{U}$ je ortogonálna. Potom aj $\boldsymbol{P}$ je reálna (teda ortogonálna) matica. Preto $g(t)$ je symetrická kladne definitná matica a $\det g(t) = 1$ pre každé t.

(d) Podľa vety 29.5.4 možno každú maticu $\boldsymbol{A} \in \Lambda_+^{\uparrow}(n)$ zapísať v tvare

$$\boldsymbol{A} = \mathrm{diag}(1, \boldsymbol{B}) \cdot \mathrm{diag}(\boldsymbol{L}_v, \boldsymbol{I}_{n-1}) \cdot \mathrm{diag}(1, \boldsymbol{C}),$$

kde $v \in \langle -1, 1 \rangle$ a $\boldsymbol{B}, \boldsymbol{C} \in \mathrm{SO}(n)$. Nech $f, g \colon \langle 0, 1 \rangle \to \mathrm{SO}(n)$ sú cesty také, že $f(0) = g(0) = \boldsymbol{I}_n$ a $f(1) = \boldsymbol{B}$, $g(1) = \boldsymbol{C}$. Položme

$$h(t) = \mathrm{diag}(1, f(t)) \cdot \mathrm{diag}(\boldsymbol{L}_{vt}, \boldsymbol{I}_{n-1}) \cdot \mathrm{diag}(1, g(t)).$$

Potom h je cesta z $\boldsymbol{I}$ do $\boldsymbol{A}$ v $\Lambda_+^{\uparrow}(n)$.

Na druhej strane pre každé $n \geq 1$ sú maticové grupy $\mathrm{GL}(n, \mathbb{R})$, $\mathrm{O}(n)$ a $\Lambda(n) = \mathrm{O}(1, n)$ nesúvislé. Presnejšie platí nasledujúca veta.

29.7.7. Veta. *Nech $n \geq 1$. Potom*

(a) súvislú komponentu jednotkovej matice vo všeobecnej lineárnej grupe $\mathrm{GL}(n, \mathbb{R})$ tvorí jej podgrupa $\mathrm{GL}^+(n, \mathbb{R}) = \{\boldsymbol{A} \in \mathrm{GL}(n, \mathbb{R});\ \det \boldsymbol{A} > 0\}$;

(b) súvislou komponentou jednotkovej matice v ortogonálnej grupe $\mathrm{O}(n)$ je špeciálna ortogonálna grupa $\mathrm{SO}(n)$;

(c) súvislou komponentou jednotkovej matice v Lorentzovej grupe $\Lambda(n)$ je vlastná ortochrónna Lorentzova grupa $\Lambda_+^{\uparrow}(n)$.

Dôkaz. (a) Podľa tvrdenia 29.1.3 platí $\mathrm{GL}(n, \mathbb{R}) \cong \mathbb{R}^\star \times \mathrm{SL}(n, \mathbb{R})$ a z dôkazu vyplýva, že oba navzájom inverzné izomorfizmy sú zároveň spojité zobrazenia. Keďže grupa $\mathrm{SL}(n, \mathbb{R})$ je podľa vety 29.7.6 (c) súvislá a súvislá komponenta jednotky v grupe $\mathbb{R}^\star$ je zrejme $\mathbb{R}^+$, súvislou komponentou jednotky v tomto priamom súčine je podľa tvrdenia 29.7.3 (b) podgrupa $\mathbb{R}^+ \times \mathrm{SL}(n, \mathbb{R})$. Tej zodpovedá v príslušnom izomorfizme podgrupa $\mathrm{GL}^+(n, \mathbb{R})$.

(b) Obrazom grupy $\mathrm{O}(n)$ v spojitom zobrazení $\boldsymbol{A} \mapsto \det \boldsymbol{A}$ je nesúvislá množina $\{-1, 1\}$. To podľa tvrdenia 29.7.2 znamená, že $\mathrm{O}(n)$ je nesúvislá. Rovnako je nesúvislá každá jej podmnožina, ktorá obsahuje aspoň jednu maticu s deteminantom 1 a aspoň jednu maticu s determinantom -1. Preto súvislá komponenta jednotkovej matice v $\mathrm{O}(n)$ je podmnožinou podgrupy $\mathrm{SO}(n)$. Na zavŕšenie dôkazu si stačí uvedomiť, že táto podgrupa je súvislá podľa vety 29.7.6 (a).

(c) možno dokázať podobne ako (b), len namiesto determinantu treba použiť surjektívne spojité zobrazenie $\boldsymbol{A} \mapsto (a_{00}/|a_{00}|, \det \boldsymbol{A}) \colon \Lambda(n) \to \{1, -1\} \times \{1, -1\}$ z dôkazu vety 29.5.1.

Z viet 29.7.6 a 29.7.7 zároveň vyplýva, že postavenie normálnej podgrupy $\mathrm{SU}(n)$ v grupe $\mathrm{U}(n)$ sa diametrálne líši od postavenia normálnej podgrupy $\mathrm{SO}(n)$ v grupe $\mathrm{O}(n)$: kým špeciálna unitárna grupa aj unitárna grupa sú obe súvislé, ortogonálna grupa je nesúvislá a špeciálna ortogonálna je jej komponentou súvislosti.

Afinné rozšírenia lineárnych grúp $G \subseteq \mathrm{GL}(n, K)$ sme v paragrafe 29.2 opísali ako polopriame súčiny $G^{\mathrm{af}} = G \ltimes K^n$ grupy G a aditívnej grupy $(K^n, +)$ vektorového priestoru K^n. Kedže základná množina polopriameho súčinu splýva s karteziánskym súčinom základných množín oboch činiteľov (pozri paragraf 28.4), z tvrdenia 29.7.3 a súvislosti priestorov $\mathbb{R}^n$, $\mathbb{C}^n$ vyplýva

29.7.8. Veta. *Nech K je jedno z polí $\mathbb{R}$, $\mathbb{C}$ a $G \subseteq \mathrm{GL}(n, K)$ je maticová grupa nad K. Potom*
(a) G je súvislá práve vtedy, keď jej afinné rozšírenie $G^{\mathrm{af}} = G \ltimes K^n$ je súvislé;
(b) súvislá komponenta jednotky $(\boldsymbol{I}, \mathbf{0})$ v afinnom rozšírení grupy G splýva s afinným rozšírením súvislej komponenty jednotkovej matice $\boldsymbol{I}$ v grupe G, čiže
$$\Upsilon\big(G^{\mathrm{af}}\big) = \Upsilon(G)^{\mathrm{af}} = \Upsilon(G) \ltimes K^n.$$

Teraz už konečne môžeme podporiť ďalšími argumentmi naše úvahy o dvoch orientáciách priestoru z paragrafov 15.2 a 29.5, resp. o štyroch orientáciách časopriestoru z paragrafu 29.5. Číslo 2 sme dostali ako index špeciálnej ortogonálnej grupy $\mathrm{SO}(n)$ v ortogonálnej grupe $\mathrm{O}(n)$ a číslo 4 ako index vlastnej ortochrónnej Lorentzovej grupy $\Lambda_+^\uparrow(n)$ v Lorentzovej grupe $\Lambda(n) = \mathrm{O}(1, n)$. Vyhli sme sa však otázke, prečo by o orientácii mali rozhodovať práve tieto podgrupy. Na túto otázku teraz môžeme poskytnúť jednotnú odpoveď. Príslušné číslo sa totiž objavuje ako index $[G : \Upsilon(G)]$ *súvislej komponenty* $\Upsilon(G)$ v grupe G „prirodzených“ transformácií (časo)priestoru.

Na „náš“ trojrozmerný priestor, no taktiež na priestory iných rozmerov n sa môžeme dívať rôznymi spôsobmi:

(a) Ako na (iba) vektorový, prípadne afinný priestor $\mathbb{R}^n$. „Prirodzenými“ transformáciami takého priestoru sú transformácie zachovávajúce jeho lineárnu resp. afinnú štrutúru. Tie tvoria (po stotožnení transformácie s jej maticou vzhľadom na kanonickú bázu $\boldsymbol{\varepsilon}$ resp. kanonický repér $(\mathbf{0}, \boldsymbol{\varepsilon})$) všeobecnú lineárnu grupu $\mathrm{GL}(n, \mathbb{R})$, so súvislou komponentou jednotky $\mathrm{GL}^+(n, \mathbb{R})$, prípadne všeobecnú afinnú grupu $\mathrm{GA}(n, \mathbb{R}) = \mathrm{GL}(n, \mathbb{R}) \ltimes \mathbb{R}^n$, so súvislou komponentou jednotky $\mathrm{GA}^+(n, \mathbb{R}) = \mathrm{GL}^+(n, \mathbb{R}) \ltimes \mathbb{R}^n$.

(b) Ako na (vektorový alebo afinný) *euklidovský* priestor $\mathbb{R}^n$. V tom prípade úlohu „prirodzených“ transformácií preberajú transformácie zachovávajúce dĺžku (skalárny súčin) vektorov, prípadne euklidovskú vzdialenosť bodov (čiže izometrie). Tie tvoria ortogonálnu grupa $\mathrm{O}(n)$, so súvislou komponentou jednotky $\mathrm{SO}(n)$, prípadne euklidovskú grupu $\mathrm{O}(n) \ltimes \mathbb{R}^n$, so súvislou komponentou jednotky $\mathrm{SO}(n) \ltimes \mathbb{R}^n$.

V každom z uvedených prípadov je index súvislej komponenty $\Upsilon(G)$ v celej grupe G rovný 2, presnejšie

$$G/\Upsilon(G) \cong (\{1,-1\},\cdot\,) \cong (\mathbb{Z}_2,+).$$

Pritom od jednej transformácie sa dá spojitým prechodom dostať k druhej transformácii práve vtedy, keď ležia v tej istej komponente. Celá grupa sa tak rozpadá na dva od seba oddelené súvislé bloky, zodpovedné za dve orientácie priestoru.

Komplexné analógy uvedených grúp sú napospol súvislé. Baviť sa o orientácii príslušných komplexných priestorov preto nedáva príliš zmysel – tieto priestory sú neorientovateľné, alebo, ak silou-mocou chceme, majú len jednu orientáciu.

V prípade „nášho“ Minkowského časopriestoru $\mathbb{R}^{(1,3)}$, no aj iných Minkowského časopriestorov $\mathbb{R}^{(1,n)}$ sa ako „prirodzené“ grupy transformácií ukazujú Lorentzova grupa $\Lambda(n)$, so súvislou komponentou jednotky $\Lambda_+^\uparrow(n)$, prípadne Poincarého grupa $\Lambda(n) \ltimes \mathbb{R}^{n+1}$, so súvislou komponentou jednotky $\Lambda_+^\uparrow(n) \ltimes \mathbb{R}^{n+1}$. V oboch prípadoch je index súvislej komponenty $\Upsilon(G)$ v celej grupe G rovný 4, presnejšie

$$G/\Upsilon(G) \cong \big(\{1,-1\}^2,\cdot\,\big) \cong \big(\mathbb{Z}_2^2,+\big).$$

Celá grupa sa tak rozpadá na štyri súvislé, no spojitými prechodmi navzájom neprepojiteľné časti, zodpovedné za štyri orientácie časopriestoru.

Len na doplnenie a bez dôkazu poznamenajme, že aj vo všeobecnom prípade je pseudoortogonálna grupa $\mathrm{O}(k,l)$ izomorfná s polopriamym súčinom štvorgrupy $\Gamma = \{\boldsymbol{I},\boldsymbol{P},\boldsymbol{T},\boldsymbol{PT}\} \cong \mathbb{Z}_4$ a svojej súvislej komponenty $\mathrm{SO}^+(k,l)$ jednotkovej matice. Pritom *vlastná špeciálna pseudoortogonálna grupa* $\mathrm{SO}^+(k,l)$ je tvorená maticami, ktoré zachovávajú orientáciu (vhodne zvoleného) maximálneho kladne definitného podpriestoru $S \subseteq \mathbb{R}^{(k,l)}$ aj jeho ortokomplementu $S^\perp$, alebo, ekvivalentne, maticami s determinantom 1, ktoré zachovávajú orientáciu jedného z nich. Špeciálne teda $\Lambda_+^\uparrow(n) = \mathrm{SO}^+(1,n)$. V tomto zmysle môžeme aj v prípade pseudoeuklidovského priestoru $\mathbb{R}^{(k,l)}$ hovoriť o štyroch orientáciách.

29.8 Homotopia a jednoduchá súvislosť

Aj medzi súvislými grupami existujú viaceré rozdiely: niektoré sú súvislejšie než iné. Akási „súvislosť vyššieho rádu“ spočíva v možnosti spojitého prechodu medzi istými uzavretými cestami v danej grupe.

Homotopiou medzi cestami $f,g\colon \langle 0,1\rangle \to X$ v množine $X \subseteq \mathbb{R}^n$ nazývame ľubovoľné spojité zobrazenie $H\colon \langle 0,1\rangle \times \langle 0,1\rangle \to X$ také, že pre všetky $t \in \langle 0,1\rangle$ platí $H(0,t) = f(t)$, $H(1,t) = g(t)$. Intuitívne určuje homotopia H

systém ciest $H_s = H(s, -)\colon \langle 0,1\rangle \to X$, kde $s \in \langle 0,1\rangle$, na ktorý sa možno dívať ako na jedinú cestu spájajúcu body $f = H_0$, $g = H_1$ v množine ciest $\mathcal{C}(\langle 0,1\rangle, X)$. Ak $\mathcal{F} \subseteq \mathcal{C}(\langle 0,1\rangle, X)$ je nejaká množina ciest a $H_s \in \mathcal{F}$ pre každé $s \in \langle 0,1\rangle$, hovoríme, že H je homotopia v množine $\mathcal{F}$.

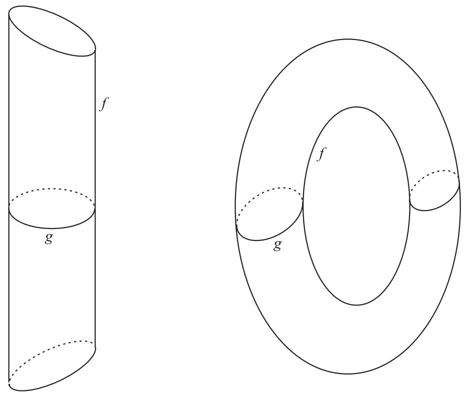
Hovoríme, že *cesta* $f\colon \langle 0,1\rangle \to X$ je *uzavretá*, ak $f(0) = f(1)$. Uzavreté cesty nazývame tiež *slučkami*. Množinu všetkých slučiek f v X takých, že $f(0) = f(1) = x_0 \in X$, značíme $\Omega(X, x_0)$. *Množina* $X \subseteq \mathbb{R}^n$ sa nazýva *jednoducho súvislá*, ak je súvislá a taktiež jej množina slučiek $\Omega(X, x_0)$ je súvislá. To druhé znamená, že pre každú slučku $f \in \Omega(X, x_0)$ existuje homotopia H v $\Omega(X, x_0)$ taká, že H_0 je konštantná slučka x_0 (t. j. $H(0,t) = x_0$ pre každé t) a $H_1 = f$. Názorne hovoríme, že každú slučku v X možno stiahnuť do bodu x_0. Dá sa dokázať, že uvedená definícia jednoduchej súvislosti nezávisí na voľbe bodu x_0 (pozri cvičenie 29.18).

29.8.1. Príklad. (a) Grupa $\mathrm{U}(1)$, t. j. kružnica, je zrejme súvislá, no nie je jednoducho súvislá. Slučku, ktorá obieha celú kružnicu práve raz (takou je napr. funkcia $f(t) = \exp(2\pi i t)$) totiž nemožno vnútri kružnice stiahnuť do bodu 1 bez toho, aby sme ju niekde pretrhli. Také niečo by sme mohli urobiť vovnútri kruhu $\{c \in \mathbb{C};\ |c| \le 1\}$ – ten však nemáme k dispozícii. Podobne nie je jednoducho súvislá napr. (konečná či nekonečná) valcová plocha $\big\{(x,y,z) \in \mathbb{R}^3;\ x^2 + y^2 = 1\ \&\ z \in J\big\}$, kde $\emptyset \ne J \subseteq \mathbb{R}$ je ľubovoľný (ohraničený či neohraničený) interval. Gumičku na ňu natiahnutú totiž opäť nemožno bez pretrhnutia stiahnuť do bodu.

Treba si však uvedomiť, že i v množinách, ktoré nie sú jednoducho súvislé, možno niektoré slučky stiahnuť do bodu. Takou je napríklad slučka f v kružnici $\mathrm{U}(1)$, ktorá beží z bodu 1 v kladnom zmysle dokola do ľubovoľného pevného bodu $c \in \mathrm{U}(1)$, tam sa obráti a beží v opačnom smere nazad. Napíšte si sami vzorec pre nejakú takú funkciu f a nájdite homotopiu, ktorá z nej spraví konštantnú slučku $g(x) = 1$.

(b) Na druhej strane každý interval $J \subseteq \mathbb{R}$, rovina $\mathbb{R}^2$ (resp. $\mathbb{C}$), ako aj dvojrozmerná sféra $S^{(2)}$ sú jednoducho súvislé. Z názoru je totiž zrejmé, že každú slučku v takejto množine možno vnútri nej stiahnuť do bodu. Z podobných príčin je jednoducho súvislá aj plocha, ktorá vznikne pridaním aspoň jednej z podstáv k ohraničenej valcovej ploche $\big\{(x,y,z) \in \mathbb{R}^3;\ x^2 + y^2 = 1\ \&\ 0 \le z \le 1\big\}$.

(c) Keď však (po jej potrebnom natiahnutí) zlepíme voľné okraje valcovej plochy $\big\{(x,y,z) \in \mathbb{R}^3;\ x^2 + y^2 = 1\ \&\ 0 \le z \le 1\big\}$, dostaneme niečo ako pneumatiku, čiže plochu učene nazývanú *torus*. Táto opäť nie je jednoducho súvislá. Dokonca vieme uviesť dva základné príklady slučiek na tore, nestiahnuteľných do bodu. Jedna vznikne napr. z úsečky spájajúcej body $(1,0,0)$ a $(1,0,1)$ zlepením práve týchto koncových bodov – môžeme si ju predstaviť ako gumičku znútra obopínajúcu vnútorný obvod toru (na obrázku ju predstavuje krivka f). Druhou je napr. slučka, ktorá práve raz obehne množinu $\{(x,y,1/2)\in\mathbb{R}^3;\ x^2+y^2=1\}$, čiže „gumička“ vopred natiahnutá na plášť valca ešte pred zlepením jeho okrajov (na obrázku krivka g).



Obr. 29.1. Valcová plocha a torus so slučkami f, g

Prechod od kružnice k valcovej ploche v bode (a) predchádzajúceho príkladu pripúšťa nasledujúce zovšeobecnenie.

29.8.2. Tvrdenie. *Nech $X\subseteq\mathbb{R}^n$ je súvislá množina a $J\subseteq\mathbb{R}$ je neprázdny interval. Potom množina $X\times J\subseteq\mathbb{R}^{n+1}$ je jednoducho súvislá práve vtedy, keď množina X je jednoducho súvislá.*

Dôkaz. Keďže interval J je súvislý, aj $X\times J$ je súvislá. Ďalej si zrejme môžeme dovoliť bez ujmy na všeobecnosti predpokladať, že $0\in J$. Zvoľme ľubovoľné $x_0\in X$.

Predpokladajme, že X je jednoducho súvislá. Dokážeme, že aj $X\times J$ má rovnakú vlastnosť. Nech $f\colon\langle 0,1\rangle\to X\times J$ je slučka taká, že $f(0)=f(1)=(x_0,0)$. Potom $f(t)=\big(g(t),h(t)\big)$ pre vhodné slučky $g\colon\langle 0,1\rangle\to X$ a $h\colon\langle 0,1\rangle\to J$. Keďže X je jednoducho súvislá, existuje homotopia $H\colon\langle 0,1\rangle^2\to X$ v $\Omega(X,x_0)$ taká, že $H(0,t)=g(t)$ a $H(1,t)=x_0$ pre každé $t\in\langle 0,1\rangle$. Potom predpisom

$$\widetilde{H}(s,t)=\big(H(s,t),\,(1-s)h(t)\big)$$

je určená homotopia $\widetilde{H}\colon\langle 0,1\rangle^2\to X\times J$ v $\Omega\big(X\times J,(x_0,0)\big)$ spájajúca slučku f s konštantnou slučkou $(x_0,0)$.

Nech naopak $X\times J$ je jednoducho súvislá. Treba dokázať, že aj ľubovoľnú slučku $f\colon\langle 0,1\rangle\to X$, $f(0)=f(1)=x_0$ možno v X stiahnuť do bodu x_0. Potom $\widetilde{f}(t)=(f(t),0)$ je slučka v $X\times J$. Z jednoduchej súvislosti množiny $X\times J$ vyplýva existencia homotopie $\widetilde{H}\colon\langle 0,1\rangle^2\to X\times J$, ktorá stiahne

$\widetilde{f}$ do bodu $(x_0, 0)$. $\widetilde{H}$ má tvar $\widetilde{H}(s,t) = \bigl(H(s,t),\, H'(s,t)\bigr)$ pre vhodné homotopie H v $\Omega(X, x_0)$, H' v $\Omega(J, 0)$. Potom H je hľadaná homotopia, ktorá stiahne f do x_0.

Ak $G \subseteq \mathrm{GL}(n, K)$, kde $K \in \{\mathbb{R}, \mathbb{C}\}$, je maticová grupa, tak miesto $\Omega(G, \boldsymbol{I})$ píšeme len $\Omega(G)$. Ľahko možno nahliadnuť, že $\Omega(G)$ obsahuje konštantnú slučku $\boldsymbol{I}(t) = \boldsymbol{I}$ a je uzavretá na operácie súčinu a inverzného prvku definované po zložkách. Presnejšie, ak $f, g \in \Omega(G)$, tak aj funkcie $f \cdot g$ a f^{-1} definované predpismi

$$(f \cdot g)(t) = f(t) \cdot g(t), \qquad \text{resp.} \qquad f^{-1}(t) = f(t)^{-1}$$

patria do $\Omega(G)$. To je vlastne obsahom prvej časti nasledujúcej vety. Jej druhú časť možno dokázať podobne ako vetu 29.7.5.

29.8.3. Veta. *Pre ľubovoľnú maticovú grupu G nad poľom $\mathbb{R}$ alebo $\mathbb{C}$ množina slučiek $\Omega(G)$ tvorí tiež grupu. Množina $\Upsilon\Omega(G)$ všetkých slučiek v $\Omega(G)$ stiahnuteľných do bodu $\boldsymbol{I}$, t. j. vlastne súvislá komponenta konštantnej slučky $\boldsymbol{I}$ v grupe $\Omega(G)$, je normálna podgrupa grupy $\Omega(G)$.*

Teda maticová grupa G je jednoducho súvislá práve vtedy, keď $\Upsilon(G) = G$ a $\Upsilon\Omega(G) = \Omega(G)$.

Pre nás sú dôležité dve skupiny jednoducho súvislých grúp.

29.8.4. Veta. *Pre každé $n \geq 1$ sú grupy $\mathrm{SU}(n)$ a $\mathrm{SL}(n, \mathbb{C})$ jednoducho súvislé.*

Náznak dôkazu. Podľa vety 29.7.6 sú obe uvedené grupy súvislé pre každé n.

Nech f je slučka v $\mathrm{SU}(n)$ so začiatkom i koncom v $\boldsymbol{I}$. Podľa vety 23.4.9 môžeme pre každé $t \in \langle 0, 1\rangle$ nájsť hermitovskú maticu $h(t) \in \mathbb{C}^{n\times n}$ s nulovou stopou, takú že $f(t) = \exp\bigl(\mathrm{i}h(t)\bigr)$. Navyše možno zaručiť, že h je spojité zobrazenie intervalu $\langle 0, 1\rangle$ do množiny takých matíc (čo nie je celkom samozrejmé, no nechávame na uverenie). Položme

$$H(s,t) = \exp\bigl(\mathrm{i}sh(t)\bigr).$$

Potom H je homotopia, H_0 je konštantná slučka $\boldsymbol{I}$, každé H_s je slučka v $\mathrm{SU}(n)$ a $H_1 = f$.

Ak f je slučka v $\mathrm{SL}(n, \mathbb{C})$, $f(0) = f(1) = \boldsymbol{I}$, tak podobným spôsobom ako v dôkaze bodu (c) vety 29.7.6 a v predošlom odstavci možno zaručiť, že existujú spojité funkcie $p\colon \langle 0, 1\rangle \to \mathrm{U}(n)$, $g_1, \dots, g_n\colon \langle 0, 1\rangle \to \mathbb{R}$ a h z $\langle 0, 1\rangle$ do množiny všetkých hermitovských matíc rádu n s nulovou stopou také, že $\sum_j g_j(t) = 0$ a

$$f(t) = p(t) \cdot \mathrm{diag}\bigl(\mathrm{e}^{g_1(t)}, \dots, \mathrm{e}^{g_n(t)}\bigr) \cdot p(t)^{-1} \cdot \mathrm{e}^{\mathrm{i}h(t)}$$

pre každé t. Potom

$$H(s,t) = p(t) \cdot \operatorname{diag}\bigl(\mathrm{e}^{sg_1(t)}, \dots, \mathrm{e}^{sg_n(t)}\bigr) \cdot p(t)^{-1} \cdot \mathrm{e}^{\mathrm{i}sh(t)}$$

je hľadaná homotopia.

Na druhej strane grupy $\mathrm{SO}(3)$ a $\Lambda_+^{\uparrow}(3)$, tak náramne dôležité z geometrického resp. z fyzikálneho hľadiska, jednoducho súvislé nie sú. Dôvody tejto skutočnosti len veľmi približne naznačíme v nasledujúcej kapitole. To dodáva zvláštny význam surjektívnym *dvojnásobne nakrývajúcim* homomorfizmom $\mathrm{SU}(2) \to \mathrm{SO}(3)$ a $\mathrm{SL}(2,\mathbb{C}) \to \Lambda_+^{\uparrow}(3)$, ktoré ich dávajú do súvislosti s jednoducho súvislými grupami $\mathrm{SU}(2)$ resp. $\mathrm{SL}(2,\mathbb{C})$.

Cvičenia

29.1. (a) Sformulujte zovšeobecnenia častí (a) – (d) tvrdenia 29.1.1 pre ľubovoľnú lineárnu grupu $G \subseteq \mathrm{GL}(n,K)$ a dokážte ich.

(b) Vysvetlite, prečo používame pravé a nie ľavé akcie. Zmeňte uvedené zobrazenia tak, aby ste dostali ľavé akcie.

29.2. (a) Dokážte tvrdenie 29.1.2.

(b) Popíšte grupu všetkých matíc $\boldsymbol{P} \in \mathrm{GL}(n,\mathbb{R})$ takých, že lineárne zobrazenie $\boldsymbol{x} \mapsto \boldsymbol{P}\cdot\boldsymbol{x}$, $\mathbb{R}^n \to \mathbb{R}^n$ zachováva (neorientované) objemy n-rozmerných rovnobežnostenov vytvorených usporiadanými n-ticami vektorov $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ z $\mathbb{R}^n$.

29.3. Nech V je n-rozmerný vektorový priestor nad poľom K. Dokážte, že všeobecné afinné grupy $\mathrm{GA}(V)$ a $\mathrm{GA}(n,K)$ sú izomorfné. Využite pritom akýkoľvek repér vo V (pozri cvičenia 8.7, 8.8).

29.4. Nech $G \subseteq \mathrm{GL}(n,K)$ je lineárna grupa a $S \subseteq G$ je jej podgrupa. Potom afinné rozšírenie S^{af} je podgrupou afinného rozšírenia G^{af}. Dokážte pre grupy matíc aj lineárnych zobrazení.

29.5. Nasledujúce označenie umožňuje rozlíšiť body a vektory v priestore K^n a afinné zobrazenia $K^n \to K^m$ reprezentovať pomocou lineárnych zobrazení $K^{n+1} \to K^{m+1}$, t. j. maticami z $K^{(m+1)\times(n+1)}$. Každý vektor $\boldsymbol{u} \in K^n$ nahradíme vektorom $\binom{\boldsymbol{u}}{0} \in K^{n+1}$ a každý bod $\boldsymbol{p} \in K^n$ nahradíme bodom $\binom{\boldsymbol{p}}{1} \in K^{n+1}$. Dvojicu $(\boldsymbol{A},\boldsymbol{u})$ pozostávajúcu z matice $\boldsymbol{A} \in K^{m\times n}$ a vektora $\boldsymbol{u} \in K^m$ nahradíme maticou $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \boldsymbol{0}_n & 1 \end{pmatrix} \in K^{(m+1)\times(n+1)}$.

(a) Dokážte, že v takejto reprezentácii skutočne platí, že súčet bodu a vektora je bod, rozdiel dvoch bodov je vektor a afinná kombinácia bodov je bod.

(b) Pre ľubovoľný bod $\binom{\boldsymbol{x}}{1} \in K^{n+1}$ a maticu $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \boldsymbol{0}_n & 1 \end{pmatrix} \in K^{(m+1)\times(n+1)}$ platí $\binom{\boldsymbol{A}\cdot\boldsymbol{x}+\boldsymbol{u}}{1} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \boldsymbol{0}_n & 1 \end{pmatrix} \cdot \binom{\boldsymbol{x}}{1}$. Dokážte.

(c) Dokážte, že kompozícia $f \circ g$ afinných zobrazení $g\colon K^p \to K^n$, $f\colon K^n \to K^m$, tvaru $g(\boldsymbol{x}) = \boldsymbol{B}\cdot\boldsymbol{x} + \boldsymbol{v}$, $f(\boldsymbol{y}) = \boldsymbol{A}\cdot\boldsymbol{y} + \boldsymbol{u}$ zodpovedá súčinu matíc $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \boldsymbol{0}_n & 1 \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{B} & \boldsymbol{v} \\ \boldsymbol{0}_p & 1 \end{pmatrix}$.

29.6. (a) Nech K je ľubovoľné pole. Uvažujme akciu $(\boldsymbol{X},\boldsymbol{A}) \mapsto \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{X} \cdot \boldsymbol{A}$ všeobecnej lineárnej grupy $\mathrm{GL}(n,K)$ na priestore matíc $K^{n\times n}$. Ak $\boldsymbol{Z} \in K^{n\times n}$ je regulárna, tak pre každú maticu $\boldsymbol{A} \in \mathrm{Stb}(\boldsymbol{Z})$ platí $\det \boldsymbol{A} = \pm 1$. Dokážte.

(b) Odvoďte z (a), že ak $\operatorname{char} K \neq 2$, tak normálna podgrupa $\mathrm{SO}(n, K)$ má v grupe $\mathrm{O}(n, K)$ index 2. Ako je to v prípade $\operatorname{char} K = 2$?

(c) Uvažujme akciu $(\boldsymbol{X}, \boldsymbol{A}) \mapsto \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{X} \cdot \overline{\boldsymbol{A}}$ všeobecnej lineárnej grupy $\mathrm{GL}(n, \mathbb{C})$ na priestore matíc $\mathbb{C}^{n \times n}$. Nech $\boldsymbol{Z} \in \mathbb{C}^{n \times n}$ je ľubovoľná regulárna matica. Potom pre každé $\boldsymbol{A} \in \mathrm{Stb}(\boldsymbol{Z})$ platí $|\det \boldsymbol{A}| = 1$. Dokážte.

(d) Odvoďte z (c), že $\mathrm{U}(n)/\mathrm{SU}(n) \cong \mathrm{U}(1)$.

29.7. Nech V je vektorový priestor nad poľom K, F je regulárna bilineárna (prípadne pre $K = \mathbb{C}$ tiež poldruhalineárna) forma na V a $\varphi\colon V \to K$ je ľubovoľné zobrazenie.

(a) Pre kladne resp. záporne definitnú formu F nad poľom $\mathbb{R}$ resp. $\mathbb{C}$ dokážte injektívnosť zobrazenia φ zachovávajúceho F aj bez predpokladu konečnorozmernosti priestoru V.

(b) Pre každé $n \geq 2$ nájdite príklad (neregulárnej) symetrickej bilineárnej formy na $\mathbb{R}^n$ a nebijektívneho lineárneho zobrazenia $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$, ako aj bijektívneho nelineárneho zobrazenia $\psi\colon \mathbb{R}^n \to \mathbb{R}^n$ fixujúceho počiatok, takých, že φ aj ψ zachovávavajú F (teda sú zároveň izometriami kvadratickej formy $q(\boldsymbol{x}) = F(\boldsymbol{x}, \boldsymbol{x})$).

(c) Odvoďte z (b), že všetky bijektívne izometrie kvadratickej formy q na vektorovom priestore V tvoria podgrupu grupy $\mathcal{S}(V)$, nie však nutne afinnej grupy $\mathrm{GA}(V)$.

29.8. Zameňte vo vete 29.3.4 vete predpoklad konečnorozmernosti priestoru V predpokladom surjektívnosti lineárnej transformácie φ a dokážte jej takto pozmenenú verziu.

29.9. Dokážte, že prienik ľubovoľných dvoch z grúp $\mathrm{SO}(1, n)$, $\Lambda_+(n)$, $\Lambda^{\uparrow}(n)$ je podgrupou tretej z nich. Odvoďte z toho rovnosti $\Lambda_+(n) \cap \Lambda^{\uparrow}(n) = \Lambda_+(n) \cap \mathrm{SO}(1, n) = \Lambda^{\uparrow}(n) \cap \mathrm{SO}(1, n)$.

29.10. Pre reálne čísla u, v z otvoreného intervalu $J = (-1, 1)$ (chápané ako veľkosti rýchlostí) definujme ich „relativistický súčet“ $u \oplus v = \frac{u+v}{1+uv}$. Dokážte, že interval J s takto definovanou operáciou $\oplus$ tvorí grupu a nájdite grupové izomorfizmy $\left(\Lambda_+^{\uparrow}(1), \cdot\right) \cong (J, \oplus) \cong (\mathbb{R}, +)$.

29.11. Nájdite explicitný tvar prvkov „malej“ Lorentzovej grupy $\Lambda_+^{\uparrow}(2)$.

29.12. Ukážeme, že zo zachovávania svetelného kužeľa v Minkowského časopriestore $\mathbb{R}^{(1,1)}$ nejakou bijektívnou transformáciou $\mathbb{R}^{(1,1)} \to \mathbb{R}^{(1,1)}$ ešte nevyplýva, že táto transformácia zachováva časopriestorovú odľahlosť resp. pseudoskalárny súčin, a dokonca ani jej linearita.

(a) Nech $\mathcal{F}$ označuje množinu všetkých rýdzo monotónnych (t. j. rastúcich alebo klesajúcich) surjektívnych funkcií $\mathbb{R} \to \mathbb{R}$ takých, že $f(0) = 0$ (také funkcie sú už nevyhnutne spojité a bijektívne). Dokážte, že $\mathcal{F}$ je podgrupa grupy $\mathcal{S}(\mathbb{R})$ všetkých bijekcií $\mathbb{R} \to \mathbb{R}$ s operáciou skladania zobrazení.

(b) Prvky priameho súčinu grúp $\mathcal{F} \times \mathcal{F}$ budeme značiť $\boldsymbol{f} = (f_0, f_1)$. Pre $\boldsymbol{f} = (f_0, f_1) \in \mathcal{F} \times \mathcal{F}$ označme $\boldsymbol{f}^*\colon \mathbb{R}^{(1,1)} \to \mathbb{R}^{(1,1)}$ zobrazenie dané predpisom

$$\boldsymbol{f}^*(\boldsymbol{x}) = \frac{1}{2} \begin{pmatrix} f_0(x_0 + x_1) + f_1(x_0 - x_1) \\ f_0(x_0 + x_1) - f_1(x_0 - x_1) \end{pmatrix},$$

pre $\boldsymbol{x} = (x_0, x_1)^{\mathsf{T}} \in \mathbb{R}^{(1,1)}$. Dokážte, že $\boldsymbol{f} \mapsto \boldsymbol{f}^*$ je prostý grupový homomorfizmus

$\mathcal{F} \times \mathcal{F} \to \mathcal{S}\big(\mathbb{R}^2\big)$.

(c) Overte, že pre každé $\boldsymbol{f} \in \mathcal{F} \times \mathcal{F}$ a $\boldsymbol{x} \in \mathbb{R}^{(1,1)}$ platí $\boldsymbol{f}^*(\mathbf{0}) = \mathbf{0}$ a $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0 \Leftrightarrow \langle \boldsymbol{f}^*(\boldsymbol{x}), \boldsymbol{f}^*(\boldsymbol{x}) \rangle = 0$, t. j. transformácia $\boldsymbol{f}^*$ zachováva svetelný kužeľ $\mathrm{LC}(\mathbf{0})$ v $\mathbb{R}^{(1,1)}$.

(d) Označme $\mathcal{G} = (\mathcal{F} \times \mathcal{F})^*$ obraz grupy $\mathcal{F} \times \mathcal{F}$ v homomorfizme $\boldsymbol{f} \mapsto \boldsymbol{f}^*$. Dokážte, že $\Lambda(1) \subseteq \mathcal{G}$ (pri stotožnení lineárnych transformácií $\mathbb{R}^2 \to \mathbb{R}^2$ s ich maticami vzhľadom na bázu $(\boldsymbol{e}_0, \boldsymbol{e}_1)$).

(e) Dokážte, že grupa $\mathcal{G} \cap \mathrm{GL}(2, \mathbb{R})$ (pri rovnakom stotožnení ako v (d)) pozostáva práve zo všetkých nenulových skalárnych násobkov $a\varphi$ Lorentzových transformácií $\varphi \in \Lambda(1)$. Nájdite príklady lineárnych transformácií $\boldsymbol{f}^* \in \mathcal{G}$, ktoré nezachovávajú časopriestorovú odľahlosť a pseudoskalárny súčin, teda nepatria do Lorentzovej grupy $\Lambda(1)$.

(f) Nájdite príklady nelineárnych transformácií $\boldsymbol{f}^* \in \mathcal{G}$. Nelineárne transformácie $\boldsymbol{f}^* \in \mathcal{G}$ nemôžu podľa vety 29.3.4 zachovávať pseudoskalárny súčin ani časopriestorovú odľahlosť v $\mathbb{R}^{(1,1)}$.

29.13. (a) Dokážte vetu 29.6.2 priamo, bez použitia vety 17.4.3.

(b) Odvoďte vetu 17.4.3 z viet 29.6.1 a 29.6.2.

29.14. Dokážte, že pseudounitárna grupa $\mathrm{U}(k, l)$ je izomorfná s polopriamym súčinom $\mathrm{U}(1) \ltimes \mathrm{SU}(k, l)$ kružnice $\mathrm{U}(1)$ a špeciálnej pseudounitárnej grupy $\mathrm{SU}(k, l)$.

29.15. *Projektívnou priamkou* nad poľom K nazývame množinu $K \cup \{\infty\}$, t. j. jednorozmerný vektorový priestor K rozšírený o *nevlastný bod* ∞. Dokážte nasledujúce tvrdenia:

(a) Pre ľubovoľnú maticu $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(2, K)$ je predpisom $F_{\boldsymbol{A}} = \frac{ax+b}{cx+d}$ definovaná bijektívna transformácia $F_{\boldsymbol{A}} \colon K \cup \{\infty\} \to K \cup \{\infty\}$ (kladieme $F_{\boldsymbol{A}}(x) = \infty$, ak $cx + d = 0$, a $F_{\boldsymbol{A}}(\infty) = \infty$ pre $c = 0$ resp. $F_{\boldsymbol{A}}(\infty) = a/c$ pre $c \neq 0$). Zobrazenia tvaru $F_{\boldsymbol{A}}$ nazývame *racionálne lineárne transformácie*.

(b) Priradením $\boldsymbol{A} \mapsto f_{\boldsymbol{A}}$ je daná reprezentácia grupy $\mathrm{GL}(2, K)$ v množine $K \cup \{\infty\}$. Obraz $\operatorname{Im} F \subseteq \mathcal{S}(K \cup \{\infty\})$ nazývame (najmä v prípade poľa $\mathbb{C}$) *Möbiovou grupou* alebo grupou *grupou racionálnych lineárnych transformácií.*

29.16. (a) Dokážte, že centrum všeobecnej lineárnej grupy $\mathrm{GL}(n, K)$ pozostáva práve zo všetkých nenulových skalárnych násobkov $a\boldsymbol{I}$ jednotkovej matice $\boldsymbol{I} = \boldsymbol{I}_n$, teda je izomorfné s multiplikatívnou grupou $K^\star$, s ktorou ho možno stotožniť.

(b) *Všeobecná projektívna grupa* rádu n nad poľom K je definovaná ako faktorová grupa $\mathrm{PGL}(n, K) \cong \mathrm{GL}(n, K)/K^\star$ všeobecnej lineárnej grupy podľa jej centra. Dokážte, že projektívna grupa $\mathrm{PGL}(2, K)$ je izomorfná s Möbiovou grupou všetkých racionálnych lineárnych transormácií $F_{\boldsymbol{A}} \colon K \cup \{\infty\} \to K \cup \{\infty\}$.

29.17. Doplňte vynechané podrobnosti v dôkaze vety 29.7.7 (c).

29.18. Nech $X \subseteq \mathbb{R}^n$ je súvislá množina. Ukážte nezávislosť definície jednoduchej súvislosti množiny X na vybranom bode $x_0 \in X$.

29.19. Nech V je vektorový priestor nad poľom $\mathbb{R}$ (prípadne zreálnenie vektorového priestoru nad poľom $\mathbb{C}$). *Množina* $X \subseteq V$ sa nazýva *konvexná*, ak pre všetky $\boldsymbol{x}, \boldsymbol{y} \in X$ celá úsečka spájajúca body $\boldsymbol{x}$, $\boldsymbol{y}$ leží v X, t. j. $(1-t)\boldsymbol{x} + t\boldsymbol{y} \in X$ pre každé reálne číslo $0 \leq t \leq 1$. Dokážte nasledujúce tvrdenia:

(a) Každá konvexná množina $X \subseteq V$ je jednoducho súvislá.

(b) Každý interval $J \subseteq \mathbb{R}$ je konvexná, teda jednoducho súvislá množina.

29.20. Nech K je jedno z polí $\mathbb{R}$, $\mathbb{C}$ a $G \subseteq \mathrm{GL}(n, K)$ je súvislá lineárna grupa. Potom jej afinné rozšírenie $G \ltimes K^n$ je jednoducho súvislé práve vtedy, keď G je jednoducho súvislá. Dokážte. (*Návod*: Použite tvrdenie 29.8.2 a matematickú indukciu.)

29.21. S využitím sférických súradníc (pozri paragraf 14.4) dokážte jednoduchú súvislosť viacrozmerných sfér $S^{(n)} = \{(x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1};\, x_0^2 + x_1^2 + \ldots + x_n^2 = 1\}$ pre $n \geq 2$.

30. Lineárne algebry

Lineárne algebry sú vektorové priestory vybavené dodatočnou bilineárnou operáciou. Spomedzi rôznych typov lineárnych algebier sa v tejto kapitole budeme zaoberať najmä asociatívnymi algebrami. Nepôjde nám však o ich systematické štúdium, a tak sa popri nevyhnutnom minime všeobecných poznatkov a krátkej zmienke o graduovaných algebrách obmedzíme len na dve ukážky: grupové algebry konečných abelovských grúp a ich súvis s diskrétnou Fourierovou transformáciou, a na kvaternióny (teda akúsi štvorrozmernú analógiu komplexných čísel), ktoré využijeme na reprezentáciu rotácií v trojrozmernom euklidovskom priestore $\mathbb{R}^3$. Na dôvažok predvedieme, ako možno túto reprezentáciu rozšíriť do reprezentácie Lorentzových transformácií v Minkowského časopriestore $\mathbb{R}^{(1,3)}$ maticami zo špeciálnej lineárnej grupy $\mathrm{SL}(2,\mathbb{C})$. Ďalšie aspekty reprezentácií grúp, ktoré tvoria hlavnú náplň teórie asociatívnych algebier, si už všímať nebudeme.

30.1 Algebry a štruktúrne konštanty

Lineárnou algebrou alebo len *algebrou* nad poľom K, prípadne K-*algebrou*, nazývame vektorový priestor $\mathcal{A}$ nad týmto poľom vybavený bilineárnou binárnou operáciou $\cdot\colon \mathcal{A}\times\mathcal{A}\to\mathcal{A}$. Bilinearita rozmenená na drobné znamená, že pre všetky $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in\mathcal{A}$, $c\in K$ platí

$$(\boldsymbol{x}+\boldsymbol{y})\boldsymbol{z}=\boldsymbol{x}\boldsymbol{z}+\boldsymbol{y}\boldsymbol{z},\qquad\qquad \boldsymbol{x}(\boldsymbol{y}+\boldsymbol{z})=\boldsymbol{x}\boldsymbol{y}+\boldsymbol{x}\boldsymbol{z},$$
$$(c\boldsymbol{x})\boldsymbol{y}=c(\boldsymbol{x}\boldsymbol{y})=\boldsymbol{x}(c\boldsymbol{y}).$$

Operáciu $\cdot$ často nazývame *násobením* alebo *súčinom v algebre* $\mathcal{A}$. Niekedy jej prítomnosť zvýraznime označením algebry v tvare $(\mathcal{A},\cdot)$. Hovoríme, že *(lineárna) algebra* $\mathcal{A}$ je *asociatívna* resp. *komutatívna*, ak binárna operácia $\cdot$ má príslušnú vlastnosť. Ak súčin $\cdot$ má neutrálny prvok, t. j. existuje $\boldsymbol{e}\in\mathcal{A}$ také, že

$$\boldsymbol{e}\boldsymbol{x}=\boldsymbol{x}=\boldsymbol{x}\boldsymbol{e}$$

pre každé $\boldsymbol{x}\in\mathcal{A}$, hovoríme, že $\mathcal{A}$ je *algebra s jednotkou* a (jednoznačne určený) prvok $\boldsymbol{e}$ nazývame *jednotkou algebry* $\mathcal{A}$.

Hovoríme, že *lineárna algebra* je *konečnorozmerná*, ak je konečnorozmerný jej podkladový vektorový priestor. Podobne, *bázou lineárnej algebry* nazývame ľubovoľnú bázu jej vektorového priestoru a jej *dimenziou* dimenziu tohto priestoru.

30.1.1. Príklad. Nech K je pole a n je ľubovoľné prirodzené číslo. (Len aby sme sa vyhli trivialitám budeme predpokladať, že $n \geq 1$.)

(a) Vektorový priestor K^n všetkých (pre určitosť riadkových, no rovnako aj stĺpcových) usporiadaných n-tíc s násobením definovaným po zložkách, t. j.

$$\boldsymbol{xy} = (x_1y_1, \dots, x_ny_n)$$

pre $\boldsymbol{x} = (x_1, \dots, x_n)$, $\boldsymbol{y} = (y_1, \dots, y_n) \in K^n$, tvorí komutatívnu a asociatívnu K-algebru. Jej jednotkou je usporiadaná n-tica $\mathbf{1} = (1, \dots, 1) \in K^n$.

(b) Vektorový priestor $K[x]$ všetkých polynómov v premennej x, s obvyklým násobením polynómov (pozri príklad 1.6.3), tvorí komutatívnu a asociatívnu K-algebru s jednotkou $1 \in K \subseteq K[x]$.

(c) Vektorový priestor $K^{n\times n}$ všetkých štvorcových matíc rozmeru $n \times n$ s operáciou násobenia matíc tvorí asociatívnu K-algebru s jednotkou $\boldsymbol{I}_n$. Táto algebra je komutatívna práve vtedy, keď $n = 1$.

Neskôr sa stretneme aj s príkladmi neasociatívnych algebier.

Podalgebrou (lineárnej) K-algebry $\mathcal{A}$ nazývame jej lineárny podpriestor $\mathcal{S}$, ktorý je zároveň uzavretý na súčin v $\mathcal{A}$, t. j. pre ľubovoľné $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}$ platí $\boldsymbol{xy} \in \mathcal{S}$.

Lineárne zobrazenie $\varphi\colon \mathcal{A} \to \mathcal{B}$ K-algebry $\mathcal{A}$ do K-algebry $\mathcal{B}$ nazývame *homomorfizmom (lineárnych) algebier*, ak φ zachováva aj operáciu súčinu, t. j. pre ľubovoľné $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{A}$ navyše platí $\varphi(\boldsymbol{xy}) = \varphi(\boldsymbol{x})\,\varphi(\boldsymbol{y})$. Bijektívny homomorfizmus algebier nazývame *izomorfizmom*. Hovoríme, že *algebry* $\mathcal{A}$, $\mathcal{B}$ sú *izomorfné*, označenie $\mathcal{A} \cong \mathcal{B}$, ak existuje nejaký izomorfimus algebier $\varphi\colon \mathcal{A} \to \mathcal{B}$.

Ukazuje sa, že celá štruktúra *konečnorozmernej* algebry je pri pevne zvolenej báze jednoznačne zakódovaná v istej „trojrozmernej“, presnejšie *trojindexovej* matici. *Systémom štruktúrnych konštánt algebry $\mathcal{A}$* vzhľadom na jej bázu $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ nazývame trojindexovú maticu $\boldsymbol{C} = \left(c_{ij}^k\right) \in K^{n\times n\times n}$ takú, že pre všetky $i, j \leq n$ platí

$$\boldsymbol{v}_i\boldsymbol{v}_j = \sum_{k=1}^{n} c_{ij}^k \boldsymbol{v}_k. \tag{$*$}$$

Inak povedané, pre ľubovoľnú dvojicu vektorov $\boldsymbol{v}_i, \boldsymbol{v}_j$ tvorí vektor $\left(c_{ij}^1, \dots, c_{ij}^n\right)^\mathsf{T} \in K^n$ súradnice ich súčinu $\boldsymbol{v}_i\boldsymbol{v}_j$ v báze $\boldsymbol{\beta}$, t. j.

$$(\boldsymbol{v}_i\boldsymbol{v}_j)_{\boldsymbol{\beta}} = \left(c_{ij}^1, \dots, c_{ij}^n\right)^\mathsf{T}.$$

Píšeme tiež $\boldsymbol{C} = \{\mathcal{A}\}_{\boldsymbol{\beta}}$.

Taktiež naopak, pri danej báze $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ vektorového priestoru V nad poľom K, ľubovoľná trojindexová matica $\boldsymbol{C} = \left(c_{ij}^k\right) \in K^{n\times n\times n}$ jednoznačne určuje bilineárnu operáciu na V vyhovujúcu rovnostiam $(*)$ pre

$i, j \leq n$, a tým robí z V algebru nad K. Podmienka bilinearity si totiž vynucuje, že súčin vektorov $\boldsymbol{x}, \boldsymbol{y} \in V$ so súradnicami $(\boldsymbol{x})_{\boldsymbol{\beta}} = (x_1, \ldots, x_n)^{\mathsf{T}}$, $(\boldsymbol{y})_{\boldsymbol{\beta}} = (y_1, \ldots, y_n)^{\mathsf{T}}$ musí mať tvar

$$\boldsymbol{xy} = \sum_{i=1}^{n} x_i \boldsymbol{v}_i \cdot \sum_{j=1}^{n} y_j \boldsymbol{v}_j = \sum_{i,j \leq n} x_i y_j\, \boldsymbol{v}_i \boldsymbol{v}_j = \sum_{i,j,k \leq n} c_{ij}^k x_i y_j\, \boldsymbol{v}_k.$$

Inak povedané, jednotlivé zložky súradníc $(\boldsymbol{xy})_{\boldsymbol{\beta}} = (z_1, \ldots, z_n)^{\mathsf{T}}$ súčinu $\boldsymbol{xy}$ vzhľadom na bázu $\boldsymbol{\beta}$ sú

$$z_k = \sum_{i,j \leq n} c_{ij}^k x_i y_j.$$

Nasledujúca veta vyjadruje jednoznačnú určenosť algebry systémom jej štruktúrnych konštánt abstraktnejším spôsobom. Dôkaz prenechávame čitateľovi ako cvičenie.

30.1.2. Veta. *Nech $\mathcal{A}, \mathcal{B}$ sú konečnorozmerné lineárne algebry nad poľom K. Potom $\mathcal{A} \cong \mathcal{B}$ práve vtedy, keď $\dim \mathcal{A} = \dim \mathcal{B}$ a existujú bázy $\boldsymbol{\alpha}$ v $\mathcal{A}$ a $\boldsymbol{\beta}$ v $\mathcal{B}$ také, že algebra $\mathcal{A}$ má vzhľadom na bázu $\boldsymbol{\alpha}$ rovnaký systém štruktúrnych konštánt ako algebra $\mathcal{B}$ vzhľadom na bázu $\boldsymbol{\beta}$, čiže $\{\mathcal{A}\}_{\boldsymbol{\alpha}} = \{\mathcal{B}\}_{\boldsymbol{\beta}}$.*

30.1.3. Príklad. V každom z príkladov 30.1.1 (a) – (c) určíme štruktúrne konštanty príslušnej algebry vzhľadom na (v istom zmysle) kanonickú bázu. Vo všetkých prípadoch v nich podstatným spôsobom vystupuje Kroneckerov delta-symbol.

(a) Nech $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ je kanonická báza v K^n. Zrejme platí $\boldsymbol{e}_i \boldsymbol{e}_j = \delta_{ij} \boldsymbol{e}_i = \delta_{ij} \boldsymbol{e}_j$, teda štruktúrne konštanty algebry K^n s po zložkách definovaným súčinom sú

$$c_{ij}^k = \delta_i^k \delta_j^k = \begin{cases} 1, & \text{ak } i = j = k, \\ 0, & \text{ak } i \neq k \text{ alebo } j \neq k. \end{cases}$$

(b) Kanonickou bázou v $K[x]$ je postupnosť $1, x, x^2, \ldots, x^n, \ldots$ mocnín premennej x. Ide síce o nekonečnorozmernú algebru, ale (ako vzápätí uvidíme) hodnoty jej štruktúrnych konštánt c_{mn}^k sú i tak dobre definované. Zo zrejmej rovnosti $x^m\, x^n = x^{m+n}$ vyplýva

$$c_{mn}^k = \delta_{m+n}^k = \begin{cases} 1, & \text{ak } k = m + n, \\ 0, & \text{ak } k \neq m + n. \end{cases}$$

(c) Kanonickú bázu v algebre matíc $K^{n \times n}$ s bežným maticovým súčinom tvorí systém matíc $\boldsymbol{E}_{kl} = \left(\delta_k^i \delta_l^j\right)$, kde $k, l \leq n$. S trochou námahy možno nahliadnuť, že pre ľubovoľné $k, l, p, q \leq n$ platí $\boldsymbol{E}_{kl} \cdot \boldsymbol{E}_{pq} = \delta_{lp} \boldsymbol{E}_{kq}$, teda

$$c_{kl,pq}^{ij} = \delta_{lp} \delta_k^i \delta_q^j = \begin{cases} 1, & \text{ak } p = l,\ i = k \text{ a } j = q, \\ 0, & \text{ak } l \neq p \text{ alebo } k \neq i \text{ alebo } q \neq j. \end{cases}$$

(Treba si uvedomiť, že „jednotlivý index“ štruktúrnej konštanty tvorí v tomto prípade dvojica čísel z množiny $\{1, \ldots, n\}$.)

30.2 Graduované algebry

Niektoré algebry sa prirodzene štiepia do akýchsi „vrstiev“, ktorých štruktúra tesne súvisí so súčinom v algebre.

Pripomeňme si z cvičenia 6.19, že vektorový priestor V je (izomorfný s) *priamym súčtom* systému $(V_i)_{i\in I}$ svojich lineárnych podpriestorov, ak každý prvok $\boldsymbol{x} \in V$ možno jednoznačne až na poradie sčítancov a nulové členy vyjadriť v tvare konečného súčtu

$$\boldsymbol{x} = \boldsymbol{x}_1 + \ldots + \boldsymbol{x}_n,$$

kde $\boldsymbol{x}_k \in V_{i_k}$ a $i_1, \ldots, i_n \in I$ sú navzájom rôzne indexy. V takom prípade píšeme $V = \bigoplus_{i\in I} V_i$.

Hovoríme, že systém $(A_k)_{k\in\mathbb{N}}$ lineárnych podpriestorov lineárnej algebry $\mathcal{A}$ je jej *rozvrstvením* alebo *graduáciou*, ak

$$\mathcal{A} = \bigoplus_{k\in\mathbb{N}} A_k,$$

t. j. $\mathcal{A}$ je priamym súčtom svojich lineárnych podpriestorov A_k, a pre všetky $\boldsymbol{x} \in A_k$, $\boldsymbol{y} \in A_l$ platí

$$\boldsymbol{xy} \in A_{k+l}.$$

Samotnú algebru $\mathcal{A}$ (spolu so systémom vyznačených podpriestorov A_k) nazývame *graduovanou (lineárnou) algebrou.*

Operácia súčinu v graduovanej algebre sa tak rozpadá na čiastkové bilineárne operácie

$$\cdot\colon A_k \times A_l \to A_{k+l}.$$

Výsledná operácia $\cdot\colon \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ je jednoznačne určená týmito čiastkovými operáciami a požiadavkou bilineárnosti. Ak je napr. algebra $\mathcal{A}$ nekonečnorozmerná, tak jej rozvrstvenie do konečnorozmerných podpriestorov často umožňuje lepší vhľad do jej štruktúry. Naopak, niekedy môže byť užitočné zahrnúť systém vektorových priestorov $(V_k)_{k\in\mathbb{N}}$ vybavený bilineárnymi operáciami $V_k \times V_l \to V_{k+l}$ do jedinej graduovanej lineárnej algebry s podkladovým priestorom $\bigoplus_{k\in\mathbb{N}} V_k$.

Každá podalgebra $\mathcal{B}$ graduovanej algebry $\mathcal{A}$ je sama graduovaná systémom svojich lineárnych podpriestorov $B_k = B \cap A_k$.

Hovoríme, že zobrazenie $\varphi\colon \mathcal{A} \to \mathcal{B}$ K-algebry $\mathcal{A}$ s graduáciou (A_k) do K-algebry $\mathcal{B}$ s graduáciou (B_k) je *homomorfizmus graduovaných algebier*, ak φ je homomorfizmus K-algebier, ktorý zachováva graduáciu, t. j. pre $\boldsymbol{x} \in A_k$

platí $\varphi(\boldsymbol{x}) \in B_k$. Bijektívny homomorfizmus graduovaných algebier nazývame *izomorfizmom graduovaných algebier.* Z podmienky rozkladu graduovanej algebry na priamy súčet svojich vrstiev možno odvodiť, že každé zúženie $\varphi \restriction A_k \colon A_k \to B_k$ izomorfizmu graduovaných algebier $\varphi\colon \mathcal{A} \to \mathcal{B}$ je lineárnym izomorfizmom vrstvy A_k na vrstvu B_k.

30.2.1. Príklad. Tento príklad v sebe zahŕňa viacero tvrdení, ktoré by si čitateľ mal dokázať samostatne.

(a) Všetky polynómy v n komutujúcich premenných $x_1, \ldots, x_n$ nad poľom K tvoria komutatívnu a asociatívnu algebru, ktorú značíme $K[x_1, \ldots, x_n]$. Každý taký polynóm možno jednoznačne až na poradie sčítancov a nulové členy vyjadriť v tvare súčtu

$$f(x_1, \ldots, x_n) = \sum_{k_1, \ldots, k_n} a_{k_1 \ldots k_n} x_1^{k_1} \ldots x_n^{k_n},$$

kde $k_1, \ldots, k_n$ sú nezáporné celé, čísla a len konečne mnoho koeficientov $a_{k_1 \ldots k_n} \in K$ je nenulových. Dva polynómy

$$f(x_1, \ldots, x_n) = \sum a_{k_1 \ldots k_n} x_1^{k_1} \ldots x_n^{k_n}, \qquad g(x_1, \ldots, x_n) = \sum b_{k_1 \ldots k_n} x_1^{k_1} \ldots x_n^{k_n}$$

sa rovnajú práve vtedy, keď pre každú n-ticu $(k_1, \ldots, k_n)$ platí $a_{k_1 \ldots k_n} = b_{k_1 \ldots k_n}$. Operácie súčtu a skalárneho násobku takýchto polynómov sú definované zrejmým spôsobom po zložkách. Ich súčin je jednoznačne určený požiadavkou bilineárnosti a pravidlom pre súčin jednočlenov

$$x_1^{k_1} \ldots x_n^{k_n} \cdot x_1^{l_1} \ldots x_n^{l_n} = x_1^{k_1+l_1} \ldots x_n^{k_n+l_n}.$$

Stupňom jednočlena $x_1^{k_1} \ldots x_n^{k_n}$ nazývame súčet jeho exponentov $k_1 + \ldots + k_n$ a *stupňom polynómu* $f(x_1, \ldots, x_n)$ je maximum stupňov jednočlenov $x_1^{k_1} \ldots x_n^{k_n}$, ktoré sa v ňom vyskytujú s nenulovým koeficientom $a_{k_1 \ldots k_n}$.

Hovoríme, že f je *homogénny polynóm stupňa* p, ak všetky jednočleny, ktoré sa v ňom vyskytujú s nenulovým koeficientom, majú ten istý stupeň p. Všetky homogénne polynómy stupňa p (spolu s nulovým polynómom) tvoria lineárny podpriestor $K[x_1, \ldots, x_n]_p$ algebry $K[x_1, \ldots, x_n]$ a algebra $K[x_1, \ldots, x_n]$ je priamym súčtom systému svojich podpriestorov $\big(K[x_1, \ldots, x_n]_p\big)_{p \in \mathbb{N}}$. Pre $f \in K[x_1, \ldots, x_n]_p$, $g \in K[x_1, \ldots, x_n]_q$ navyše platí

$$f \cdot g \in K[x_1, \ldots, x_n]_{p+q}.$$

Inak povedané, $K[x_1, \ldots, x_n]$ je graduovaná algebra, rozvrstvená do podpriestorov homogénnych polynómov jednotlivých stupňov $K[x_1, \ldots, x_n]_p$.

(b) Nech V je konečnorozmerný vektorový priestor nad poľom K s bázou $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$. Označme $s_i \colon V \to K$ lineárny funkcionál, ktorý vektoru

$\boldsymbol{x} \in V$ priradí jeho i-tú súradnicu v báze $\boldsymbol{\beta}$, t. j. $(\boldsymbol{x})_{\boldsymbol{\beta}} = \big(s_1(\boldsymbol{x}), \dots, s_n(\boldsymbol{x})\big)^{\mathsf{T}}$. Označme $K[V]$ množinu všetkých funkcií $\Phi\colon V \to K$, ktoré možno vyjadriť v tvare $\Phi = f(s_1, \dots, s_n)$, t. j.

$$\Phi(\boldsymbol{x}) = f\big(s_1(\boldsymbol{x}), \dots, s_n(\boldsymbol{x})\big),$$

pre nejaký polynóm $f \in K[x_1, \dots, x_n]$. Ľahko sa možno presvedčiť, že odpoveď na otázku či nejaká funkcia $\Phi : V \to K$ patrí alebo nepatrí do $K[V]$, nezávisí od bázy $\boldsymbol{\beta}$. Zrejme $K[V]$ s operáciami definovanými po zložkách tvorí komutatívnu a asociatívnu K-algebru – nazývame ju *algebrou polynomických funkcií* na vektorovom priestore V alebo aj *súradnicovým okruhom* priestoru V.

Hovoríme, že *polynomická funkcia* $\Phi\colon V \to K$ je *homogénna stupňa* p, ak pre všetky $\boldsymbol{x} \in V$, $c \in K$ platí $\Phi(c\boldsymbol{x}) = c^p\Phi(\boldsymbol{x})$. Homogénne polynomické funkcie stupňov 0, 1 a 2 sú postupne konštantné funkcie, lineárne funkcionály resp. kvadratické formy. Homogénne funkcie stupňa p, nazývané tiež *p-homogénnymi formami*, tvoria lineárny podpriestor $K[V]_p$ algebry $K[V]$. Pre $\Phi \in K[V]_p$, $\Psi \in K[V]_q$ zrejme platí $\Phi \cdot \Psi \in K[V]_{p+q}$.

Zobrazenie $K[x_1, \dots, x_n] \to K[V]$, ktoré polynómu $f \in K[x_1, \dots, x_n]$ priradí polynomickú funkciu $f(s_1, \dots, s_n) \in K[V]$, je surjektívny homomorfizmus K-algebier, ktorý zobrazí každý z podpriestorov $K[x_1, \dots, x_n]_p$ p-homogénnych polynómov na príslušný podpriestor $K[V]_p$ p-homogénnych foriem. Ak pole K má viac než p prvkov, tak toto zúženie je bijekcia $K[x_1, \dots, x_n]_p \to K[V]_p$ (pozri cvičenie 30.9). Preto ak pole K je nekonečné, tak systém podpriestorov $\big(K[V]_p\big)$ je rozvrstvením graduovanej algebry $K[V]$ a zobrazenie $K[x_1, \dots, x_n] \to K[V]$ (závislé na báze $\boldsymbol{\beta}$) je izomorfizmus graduovaných algebier.

Na druhej strane, ak pole K pole má q prvkov, tak rôzne nenulové polynómy $x \in K[x]_1$, $x^q \in K[x]_q$ určujú tú istú polynomickú funkciu $K \to K$ (pozri dôsledok 27.6.8). Takže pre konečné pole K systém podpriestorov $\big(K[V]_p\big)$ nie je graduáciou algebry $K[V]$.

Ak zafixujeme bázu $\boldsymbol{\beta}$ a – ako je to bežné – funkciu i-tej súradnice vzhľadom na bázu $\boldsymbol{\beta}$ budeme značiť x_i, môžeme polynóm aj príslušnú polynomickú funkciu značiť tým istým výrazom $f = f(x_1, \dots, x_n)$. V konečnom poli K sa nám môže akurát tak stať, že dva rôzne polynómy f a g budú označovať tú istú funkciu. Také niečo interpretujeme ako platnosť netriviálnej identity $f(x_1, \dots, x_n) = g(x_1, \dots, x_n)$ v K.

Graduácia algebry nemusí byť nutne indexovaná len prirodzenými číslami. V kapitole o tenzoroch sa napr. stretneme s graduáciou indexovanou dvojicami prirodzených čísel. V takom prípade je algebra $\mathcal{A}$ priamym súčtom systému $\big(A_l^k\big)_{k,l \in \mathbb{N}}$ svojich lineárnych podpriestorov a pre $\boldsymbol{x} \in A_l^k$, $\boldsymbol{y} \in A_n^m$ platí

$$\boldsymbol{x}\boldsymbol{y} \in A_{l+n}^{k+m}.$$

Ešte s inak indexovanou graduáciou sa stretneme v nasledujúcom paragrafe.

30.3 Grupové algebry a diskrétna Fourierova transformácia

Vektorový priestor K^G všetkých funkcií $f\colon G \to K$ z konečnej grupy G do poľa K s operáciou súčinu definovanou po zložkách, t. j. $(fg)(x) = f(x)\,g(x)$ pre $f, g \in K^G$, $x \in G$, tvorí asociatívnu K-algebru $\bigl(K^G, \cdot\bigr)$. Táto algebra je zrejme izomorfná s algebrou $(K^n, \cdot)$ z príkladu 30.1.1 (a) pre $n = \#\,G$. Popri nej však K^G nesie ešte jednu štruktúru algebry, tesne zviazanú so štruktúrou grupy G.

Grupovou algebrou konečnej grupy $(G, \cdot, 1)$ nad poľom K nazývame vektorový priestor K^G všetkých funkcií $f\colon G \to K$ (s operáciami definovanými po zložkách), vybavený operáciou násobenia, definovanou ako *konvolúcia*

$$(f * g)(x) = \sum_{st=x} f(s)\,g(t) = \sum_{s\in G} f(s)\,g\bigl(s^{-1}x\bigr) = \sum_{t\in G} f\bigl(xt^{-1}\bigr)\,g(t),$$

pre $f, g\colon G \to K$, $x \in G$.[1]

Ak prvok $x \in G$ stotožníme s δ-funkciou $\delta_x \in K^G$, kde

$$\delta_x(y) = \begin{cases} 1, & \text{ak } x = y,\\ 0, & \text{ak } x \neq y,\end{cases}$$

pre $y \in G$, tak ľubovoľnú funkciu $f \in K^G$ možno chápať ako lineárnu kombináciu

$$f = \sum_{x\in G} f(x)\,\delta_x = \sum_{x\in G} f(x)\,x\,.$$

Množina $G \subseteq K^G$ potom tvorí (neusporiadnú) bázu vektorového priestoru K^G a samotná funkcia $f \in K^G$ splýva so svojimi súradnicami $f = (f)_G$ vzhľadom na bázu G (pozri paragraf 5.5). Ľahko sa presvedčíme, že pre $x, y \in G$ platí $(\delta_x * \delta_y)(z) = \delta_{(xy)}(z)$, v dôsledku čoho

$$x * y = \delta_x * \delta_y = \delta_{(xy)} = xy.$$

Inými slovami, konvolúcia prvkov bázy G splýva s ich súčinom v grupe G.

Grupovú algebru $\bigl(K^G, *\bigr)$ tak možno alternatívne popísať ako množinu všetkých (konečných) formálnych lineárnych kombinácií $\sum_{x\in G} a_x x$ s koeficientmi $a_x \in K$. Skalárne násobky a súčty sú definované zrejmým spôsobom

[1] Grupovú algebru možno definovať aj pre nekonečné grupy, treba sa však obmedziť na priestor $K^{(G)}$ funkcií $f\colon G \to K$ s *konečným suportom* $\operatorname{supp} f = \{x \in G;\ f(x) \neq 0\}$.

po zložkách, zatiaľ čo tvar súčinu výrazov $\sum_{s\in G} a_s s$, $\sum_{t\in G} b_t t$ je jednoznačne vynútený požiadavkou bilinearity a zachovávania násobenia v G

$$\sum_{s\in G} a_s s \cdot \sum_{t\in G} b_t t = \sum_{s,t\in G} a_s b_t st = \sum_{x\in G} c_x x,$$

kde koeficienty výsledného výrazu sú opäť dané konvolúciou $c_x = \sum_{st=x} a_s b_t$. Grupová algebra $\left(K^G, *\right)$, najmä pri reprezentácii jej prvkov v uvedenom tvare, sa obvykle značí KG. [2]

Ako zaujímavosť si všimnime, že grupová algebra KG je priamym súčtom svojich jednorozmerných podpriestorov $Kx = \{ax;\, a \in K\}$, kde $x \in G$, a pre $ax \in Kx$, $by \in Ky$ platí

$$ax \cdot by = abxy \in Kxy.$$

Inak povedané, KG je *graduovaná algebra* rozvrstvená do jednorozmerných lineárnych podpriestorov Kx, $x \in G$. Ide teda o graduáciu indexovanú prvkami grupy G.[3]

30.3.1. Tvrdenie. *Nech G je konečná grupa. Potom jej grupová algebra $(KG, \cdot) \cong \left(K^G, *\right)$ je asociatívna lineárna algebra nad poľom K; jej jednotkou je funkcia δ_1, ktorú stotožňujeme s jednotkovým prvkom 1 grupy G. Ak G je navyše komutatívna, tak aj jej grupová algebra je komutatívna.*

Náčrt dôkazu. Naznačíme len dôkaz asociatívnosti konvolúcie a zvyšok prenecháme čitateľovi ako cvičenie. Jednoduchým výpočtom sa možno presvedčiť, že pre ľubovoľné $f, g, h \in K^G$, $x \in G$ platí

$$\big(f * (g * h)\big)(x) = \sum_{uvw=x} f(u)\, g(v)\, h(w) = \big((f * g) * h\big)(x).$$

Grupové algebry sú jedným zo základných objektov *teórie reprezentácií* konečných grúp. V špeciálnom a najdôležitejšom prípade poľ $K = \mathbb{C}$ je takáto algebra zároveň unitárnym priestorom so skalárnym súčinom

$$\langle f, g\rangle = \sum_{x\in G} f(x)\, \overline{g}(x),$$

vzhľadom na ktorý je kanonická báza $G = \{\delta_x;\, x \in G\}$ zrejme ortonormálna. My sa však vo zvyšku tohto paragrafu budeme venovať výhradne grupovým

[2] V prípade nekonečnej grupy G splýva KG s algebrou $\left(K^{(G)}, *\right)$.

[3] Táto úvaha zostáva v platnosti, aj keď grupa G je nekonečná. V takom prípade G je *Hamelovou bázou* algebry KG (pozri paragraf 5.5 a cvičenie 6.20).

algebrám konečných *abelovských* grúp nad poľom $\mathbb{C}$. Tie majú popri kanonickej báze $G \subseteq \mathbb{C}^G$ ešte jednu významnú ortogonálnu bázu tvorenú *charaktermi* grupy G. Práve ony nám umožnia pozrieť sa na diskrétnu Fourierovu transformáciu z istého nadhľadu a zovšeobecniť tak výsledky paragrafu 17.5 z grúp $(\mathbb{Z}_n, +)$ (skryte v nich prítomných) na ľubovoľné konečné abelovské grupy.

Pripomeňme, že charakterom abelovskej grupy G nazývame ľubovoľný grupový homomorfizmus $\alpha\colon G \to \mathrm{U}(1)$; množina $G^{\mathrm{d}} = \mathrm{Hom}(G, \mathrm{U}(1))$ všetkých charakterov grupy G s násobením definovaným po zložkách tvorí abelovskú grupu. Pre konečnú G okrem toho platí $G \cong G^{\mathrm{d}}$ a taktiež $G^{\mathrm{d}} \subseteq \mathbb{C}^G$ (pozri paragraf 27.8).

30.3.2. Tvrdenie. *Pre ľubovoľné charaktery α, β konečnej abelovskej grupy G platí*

$$\langle \alpha, \beta \rangle = (\#\, G)\, \delta_{\alpha,\beta} = \begin{cases} \#\, G, & \text{ak } \alpha = \beta, \\ 0, & \text{ak } \alpha \neq \beta. \end{cases}$$

Inak povedané, všetky charaktery grupy G majú rovnakú normu $\sqrt{\#\, G}$ a tvoria ortogonálnu bázu unitárneho priestoru $\mathbb{C}^G$.

Dôkaz. Ak si uvedomíme, že $\alpha\beta^{-1} = \alpha\overline{\beta}$ je tiež charakter grupy G a $\alpha\beta^{-1} = \varepsilon$ (triviálny charakter) práve vtedy, keď $\alpha = \beta$, tak s použitím lemy 27.8.7 dostávame

$$\langle \alpha, \beta \rangle = \sum_{x \in G} \alpha(x)\, \overline{\beta}(x) = \sum_{x \in G} \big(\alpha\beta^{-1}\big)(x) = \begin{cases} \#\, G, & \text{ak } \alpha = \beta, \\ 0, & \text{ak } \alpha \neq \beta. \end{cases}$$

Ako sme už spomínali, súradnice funkcie $f \in \mathbb{C}^G$ vzhľadom na ortonormálnu bázu $G \subseteq \mathbb{C}^G$ tvoria priamo funkčné hodnoty $f(x) = \langle f, \delta_x \rangle$. Podobne, jej súradnice vzhľadom na ortogonálnu bázu $G^{\mathrm{d}} \subseteq \mathbb{C}^G$ možno vyjadriť pomocou skalárneho súčinu v tvare

$$\frac{\langle f, \alpha \rangle}{\langle \alpha, \alpha \rangle} = \frac{1}{\#\, G} \langle f, \alpha \rangle.$$

Samotné priradenie $\alpha \mapsto \langle f, \alpha \rangle$ určuje funkciu $\widehat{f}\colon G^{\mathrm{d}} \to \mathbb{C}$ čiže vektor z unitárneho priestoru $\mathbb{C}^{G^{\mathrm{d}}}$. Pri takomto označení možno súradnice funkcie $f \in \mathbb{C}^G$ vzľadom na bázu charakterov G^{d} vyjadriť ako

$$(f)_{G^{\mathrm{d}}} = \frac{1}{\#\, G}\, \widehat{f}.$$

Diskrétnou Fourierovou transformáciou (*DFT*) na konečnej abelovskej grupe G nazývame zobrazenie $F\colon f \mapsto \widehat{f}$ unitárneho priestoru $\mathbb{C}^G$ do unitárneho priestoru $\mathbb{C}^{G^{\mathrm{d}}}$ dané predpisom

$$F(f)(\alpha) = \widehat{f}(\alpha) = \langle f, \alpha \rangle = \sum_{x \in G} f(x)\, \overline{\alpha}(x),$$

pre $f \in \mathbb{C}^G$, $\alpha \in G^{\mathrm{d}}$.

Každú funkciu $f \in \mathbb{C}^G$ možno teda vyjadriť ako lineárnu kombináciu charakterov s koeficientmi určenými hodnotami Fourierovej transformácie $\widehat{f}$:

$$f = \frac{1}{\# G} \sum_{\alpha \in G^{\mathrm{d}}} \widehat{f}(\alpha)\, \alpha,$$

čo v jednotlivých zložkách $x \in G$ dáva

$$f(x) = \frac{1}{\# G} \sum_{\alpha \in G^{\mathrm{d}}} \widehat{f}(\alpha)\, \alpha(x).$$

30.3.3. Veta. *Nech G je konečná abelovská grupa. Potom pre všetky $f, g \in \mathbb{C}^G$ platí*

$$\begin{aligned} \big\langle F(f), F(g) \big\rangle &= \big\langle \widehat{f}, \widehat{g} \big\rangle = (\# G)\, \langle f, g \rangle, \\ F(f * g) &= (f * g)\widehat{\ } = F(f)\, F(g) = \widehat{f}\, \widehat{g}. \end{aligned}$$

V dôsledku toho DFT na G je izomorfizmus algebier $F\colon \big(\mathbb{C}^G, *\big) \to \big(\mathbb{C}^{G^{\mathrm{d}}}, \cdot\big)$.

Dôkaz. Nech $f, g \in \mathbb{C}^G$. Označme $n = \# G$. Ak si uvedomíme, že násobky charakterov $n^{-1/2}\alpha$ tvoria ortonormálnu bázu v $\mathbb{C}^G$, tak z Parsevalovej rovnosti dostávame

$$\begin{aligned} \big\langle \widehat{f}, \widehat{g} \big\rangle &= \sum_{\alpha \in G^{\mathrm{d}}} \widehat{f}(\alpha)\, \overline{\widehat{g}}(\alpha) = \sum_{\alpha \in G^{\mathrm{d}}} \langle f, \alpha \rangle\, \langle \alpha, g \rangle \\ &= n \sum_{\alpha \in G^{\mathrm{d}}} \big\langle f, n^{-1/2}\alpha \big\rangle\, \big\langle n^{-1/2}\alpha, g \big\rangle = n\, \langle f, g \rangle. \end{aligned}$$

To dokazuje prvú rovnosť. Keďže F je očividne lineárne zobrazenie, vyplýva z toho tiež, že F je lineárny izomorfizmus.

Zvoľme ľubovoľný charakter $\beta \in G^{\mathrm{d}}$. Keďže $\overline{\beta}$ je grupový homomorfizmus, platí

$$\begin{aligned} F(f * g)(\beta) = \langle f * g, \beta \rangle &= \sum_{x \in G} (f * g)(x)\, \overline{\beta}(x) = \sum_{x \in G} \sum_{st = x} f(s)\, g(t)\, \overline{\beta}(st) \\ &= \sum_{s \in G} f(s)\, \overline{\beta}(s) \cdot \sum_{t \in G} g(t)\, \overline{\beta}(t) = \langle f, \beta \rangle\, \langle g, \beta \rangle = \widehat{f}(\beta)\, \widehat{g}(\beta) = \big(\widehat{f}\, \widehat{g}\big)(\beta). \end{aligned}$$

Teda $(f * g)\widehat{\ } = \widehat{f}\, \widehat{g}$. To znamená, že $F\colon \big(\mathbb{C}^G, *\big) \to \big(\mathbb{C}^{G^{\mathrm{d}}}, \cdot\big)$ je izomorfizmus algebier.

Nakoľko algebra $\big(\mathbb{C}^{G^{\mathrm{d}}}, \cdot\big)$ si z grupy G^{d} „pamätá“ len počet jej prvkov, presnejšie $\big(\mathbb{C}^{G^{\mathrm{d}}}, \cdot\big) \cong (\mathbb{C}^n, \cdot)$ pre $n = \# G^{\mathrm{d}} = \# G$, vyplýva z práve dokázanej vety obdobný, dosť nečakaný výsledok aj pre grupové algebry konečných abelovských grúp.

30.3.4. Dôsledok. *Nech G, H sú konečné abelovské grupy. Potom grupové algebry $\left(\mathbb{C}^G, *\right)$, $\left(\mathbb{C}^H, *\right)$ sú izomorfné práve vtedy, keď G a H majú rovnaký počet prvkov.*

Ešte poznamenajme, že vďaka stotožneniu konečnej abelovskej grupy G s jej druhým duálom G^{dd} prostredníctvom kanonického izomorfizmu $x \mapsto x^{\Delta}$ (pozri vetu 27.8.6 a text okolo), má aj inverzné zobrazenie $F^{-1}\colon \mathbb{C}^{G^{\mathrm{d}}} \to \mathbb{C}^G$ k DFT F tvar diskrétnej Fourierovej transformácie (až na násobok konštantou $(\# G)^{-1}$ a zámenu $\overline{\alpha} \leftrightarrow \alpha$), konkrétne

$$F^{-1}(h)(x) = \frac{1}{\# G} \left\langle h, x^{-1} \right\rangle = \frac{1}{\# G} \sum_{\alpha \in G^{\mathrm{d}}} h(\alpha)\, \alpha(x),$$

pre $h \in \mathbb{C}^{G^{\mathrm{d}}}$, $x \in G$.

Keďže každá konečná abelovská grupa je izomorfná s priamym súčinom cyklických grúp, stačí sa obmedziť na DFT na takýchto grupách. Z paragrafu 27.8 vieme, že každý charakter α grupy $G = \mathbb{Z}_{n_1} \times \ldots \mathbb{Z}_{n_k}$ má tvar

$$\alpha(x) = \exp\left(2\pi \mathrm{i} \sum_{j=1}^{k} \frac{a_j x_j}{n_j}\right),$$

kde $x = (x_1, \ldots, x_k) \in G$, pre jednoznačne určený prvok $a = (a_1, \ldots, a_k) \in G$ a priradenie $\alpha \mapsto a$ určuje izomorfizmus grúp $G^{\mathrm{d}} \cong G$. Označme $n = n_1 \ldots n_k = \# G$. Po takomto stotožnení grúp G^{d} a G sa z DFT na G stáva tzv. *k-rozmerná diskrétna Fourierova transformácia* $F\colon \mathbb{C}^G \to \mathbb{C}^G$ definovaná predpisom

$$F(f)(a) = \widehat{f}(a) = \langle f, \alpha \rangle = \sum_{x \in G} f(x) \exp\left(-2\pi \mathrm{i} \sum_{j=1}^{k} \frac{a_j x_j}{n_j}\right)$$

pre $f \in \mathbb{C}^G$, $a \in G$, čo dáva vyjadrenie f v ortogonálnej báze charakterov

$$f(x) = \frac{1}{n} \sum_{a \in G} \widehat{f}(a) \exp\left(2\pi \mathrm{i} \sum_{j=1}^{k} \frac{a_j x_j}{n_j}\right).$$

Inverzná k-rozmerná DFT má potom tvar

$$F^{-1}(h)(x) = \frac{1}{n} \widehat{h}(-x) = \frac{1}{n} \sum_{a \in G} h(a) \exp\left(2\pi \mathrm{i} \sum_{j=1}^{k} \frac{a_j x_j}{n_j}\right)$$

pre $h \in \mathbb{C}^G$, $x \in G$.

30.4 Algebra kvaterniónov

Začiatkom 19. stor. sa vyjasnilo, že komplexné čísla možno algebraicky reprezentovať ako usporiadané dvojice reálnych čísel a geometricky znázorniť ako body v rovine, pričom násobenie komplexnou jednotkou $\cos\alpha + \mathrm{i}\sin\alpha$ zodpovedá otočeniu roviny okolo počiatku o uhol α. Prirodzene sa tak vynorila otázka, či niečo podobné nie je možné aj v priestore. Na $\mathbb{R}^3$ máme síce k dispozícii vektorový súčin – ten však nie je komutatívny ani asociatívny, nemá jednotkový prvok a neumožňuje delenie.

Snahe objaviť násobenie trojíc reálnych čísel alias bodov priestoru, ktoré by (povedané moderným jazykom) spolu s obvyklým vektorovým sčítaním urobilo z $\mathbb{R}^3$ pole a zároveň algebru nad poľom $\mathbb{R}$ rozširujúcu $\mathbb{C}$, venoval Sir William Rowan Hamilton dlhých trinásť rokov. Postupne v ňom začali hlodať pochybnosti, či jeho cieľ je vôbec dosiahnuteľný. Podľa jeho vlastného vyjadrenia ho riešenie napadlo 16. októbra 1843, keď s manželkou prechádzali popri *Royal Canal* v Dubline. Zrazu si uvedomil, že násobenie, ktoré hľadá, nie je komutatívne a vyžaduje si nie tri ale štyri rozmery. V náhlom „záblesku génia“ zahliadol základnú formulu pre súčiny troch imaginárnych jednotiek i, j, k. Bol taký nadšený svojim objavom, že ho vyryl nožíkom do kameňa blízkeho mosta *Broombridge*. Dnes je na tom mieste pamätná tabuľa s nápisom:

Here as he walked by
on the 16th of October 1843
Sir William Rowan Hamilton
in a flash of genius discovered
the fundamental formula for
quaternion multiplication
$\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{ijk} = -1$
& cut it on a stone of this bridge

Kvaternióny definujeme ako lineárnu algebru $\mathbb{H}$ nad poľom $\mathbb{R}$ s bázou tvorenou jednotkou $1 \in \mathbb{R}$ a troma imaginárnymi jednotkami i, j, k, pričom súčiny bázických prvkov sú dané tabuľkou

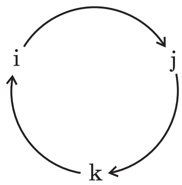
$\cdot$	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-$j
j	j	$-$k	-1	i
k	k	j	$-$i	-1

ktorú si však netreba pamätať, lebo všetky jej výsledky vyplývajú z uvedenej

Hamiltonovej formuly vďaka bilinearite a asociatívnosti (ktorú zatiaľ mlčky predpokladáme a dokážeme neskôr). Napr.

$$\begin{aligned} \mathrm{ij} &= (\mathrm{ij})(-\mathrm{k}^2) = -(\mathrm{ijk})\mathrm{k} = \mathrm{k}, \\ \mathrm{ji} &= (\mathrm{ji})(-\mathrm{ijk}) = -\mathrm{ji}^2\mathrm{jk} = \mathrm{j}^2\mathrm{k} = -\mathrm{k}, \\ \mathrm{kj} &= (\mathrm{ij})\mathrm{j} = \mathrm{ij}^2 = -\mathrm{i}, \end{aligned}$$

atď. Ešte lepšou mnemotechnickou pomôckou pre súčiny dvojíc rôznych imaginárnych jednotiek je nasledujúca cyklická schéma



Súčin dvojice rôznych imaginárnych jednotiek v smere šípky je tretia jednotka, súčin proti smeru šípky je tá ista jednotka so znamienkom mínus.

Kvaternióny teda možno zaviesť aj ako výrazy tvaru

$$q = q_0 + q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k},$$

kde $q_0, q_1, q_2, q_3 \in \mathbb{R}$, pričom rovnosť kvaterniónov $p = p_0 + p_1\mathrm{i} + p_2\mathrm{j} + p_3\mathrm{k}$, $q = q_0 + q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k}$ je definovaná podmienkou

$$p = q \;\Leftrightarrow\; p_0 = q_0 \;\&\; p_1 = q_1 \;\&\; p_2 = q_2 \;\&\; p_3 = q_3$$

a ich súčet a súčin sú dané formulami

$$\begin{aligned} p + q &= (p_0 + q_0) + (p_1 + q_1)\mathrm{i} + (p_2 + q_2)\mathrm{j} + (p_3 + q_3)\mathrm{k}, \\ p\,q &= (p_0q_0 - p_1q_1 - p_2q_2 - p_3q_3) + (p_0q_1 + p_1q_0 + p_2q_3 - p_3q_2)\mathrm{i} \\ &\quad + (p_0q_2 + p_2q_0 + p_3q_1 - p_1q_3)\mathrm{j} + (p_0q_3 + p_3q_0 + p_1q_2 - p_2q_1)\mathrm{k}. \end{aligned}$$

Toto je v skutočnosti obvyklá definícia algebry kvaterniónov $\mathbb{H}$.

Keďže reálne čísla komutujú so všetkými kvaterniónmi, môžeme tiež písať $q = q_0 + \mathrm{i}q_1 + \mathrm{j}q_2 + \mathrm{k}q_3$ ako aj

$$\frac{1}{a}\,q = q\,\frac{1}{a} = \frac{q}{a} = \frac{q_0}{a} + \frac{q_1}{a}\,\mathrm{i} + \frac{q_2}{a}\,\mathrm{j} + \frac{q_3}{a}\,\mathrm{k}$$

pre $0 \neq a \in \mathbb{R}$.

Skalár $q_0 \in \mathbb{R}$ nazývame *skalárnou* prípadne *reálnou časťou kvaterniónu* q a vektor $\vec{q} = q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k} \in \mathbb{R}^3$ nazývame jeho *vektorovou* prípadne *imaginárnou časťou.* Každý kvaternión teda môžeme rozložiť na súčet jeho skalárnej a vektorovej časti

$$q = q_0 + q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k} = q_0 + \vec{q}.$$

Množinu všetkých *skalárnych kvaterniónov*, t. j. takých, pre ktoré platí $q = q_0$, stotožňujeme s poľom $\mathbb{R} \subseteq \mathbb{H}$ a množinu všetkých *vektorových kvaterniónov*, t. j. takých, pre ktoré platí $q = \vec{q}$, stotožňujeme s vektorovým priestorom $\mathbb{R}^3 \subseteq \mathbb{H}$. Čitateľ už v tejto chvíli určite uhádol, odkiaľ pochádza označenie $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ pre kanonickú ortonormálnu bázu v euklidovskom priestore $\mathbb{R}^3$.

Združeným kvaterniónom ku kvaterniónu q nazývame kvaternión

$$q^* = q_0 - q_1\mathrm{i} - q_2\mathrm{j} - q_3\mathrm{k} = q_0 - \vec{q}.$$

Zrejme platí $(p+q)^* = p^* + q^*$; overenie nasledujúcej identity prenechávame ako cvičenie čitateľovi:

$$(p\,q)^* = q^* p^*.$$

Na kvaterniónoch možno definovať obvyklý euklidovský skalárny súčin

$$\langle p, q\rangle_E = p_0q_0 + p_1q_1 + p_2q_2 + p_3q_3,$$

ako aj Minkowského pseudoskalárny súčin

$$\langle p, q\rangle_M = \langle p, q^*\rangle_E = p_0q_0 - p_1q_1 - p_2q_2 - p_3q_3.$$

Ak si ešte spomenieme, že vektorový súčin vektorov $\vec{p}, \vec{q} \in \mathbb{R}^3$ je daný formulou

$$\vec{p} \times \vec{q} = \begin{vmatrix} p_1 & p_2 & p_3 \\ q_1 & q_2 & q_3 \\ \mathrm{i} & \mathrm{j} & \mathrm{k} \end{vmatrix},$$

zistíme, že súčin kvaterniónov možno vyjadriť v jednom z podstatne kompaktnejších tvarov

$$\begin{aligned} p\,q &= -\langle p, q\rangle_E + p_0q + q_0p + (\vec{p} \times \vec{q}) \\ &= p_0q_0 - \langle \vec{p}, \vec{q}\rangle_E + p_0\vec{q} + q_0\vec{p} + (\vec{p} \times \vec{q}) \\ &= \langle p, q\rangle_M + p_0\vec{q} + q_0\vec{p} + (\vec{p} \times \vec{q}), \end{aligned}$$

ktoré odhaľujú jeho tesnú spätosť so skalárnym resp. pseudoskalárnym a vektorovým súčinom. Špeciálne pre vektorové kvaternióny $p = p_1\mathrm{i} + p_2\mathrm{j} + p_3\mathrm{k}$, $q = q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k}$ dostávame

$$p\,q = -\langle p, q\rangle_E + (p \times q).$$

Absolútnou hodnotou kvaterniónu q nazývame jeho euklidovskú normu

$$|q| = \langle q, q\rangle_E^{1/2} = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}.$$

Prvú z nasledujúcich dvoch formúl možno ľahko nahliadnuť a s jej pomocou overiť i druhú:

$$|q|^2 = \langle q, q\rangle_E = \langle q, q^*\rangle_M = q\,q^* = q^*q = |q^*|^2,$$
$$|p\,q| = |p|\,|q|\,.$$

Z toho okrem iného vyplýva, že pre nenulový kvaternión q platí

$$q\,\frac{q^*}{|q|^2} = \frac{q^*}{|q|^2}\,q = \frac{|q|^2}{|q|^2} = 1.$$

Ale to znamená, že $q^{-1} = q^*/|q|^2$ je inverzný prvok ku kvaterniónu $q \neq 0$ vzhľadom na operáciu násobenia.

Kvaternióny jednotkovej dĺžky nazývame *jednotkovými kvaterniónmi* alebo *kvaterniónovými jednotkami.* Množinu všetkých jednotkových kvaterniónov

$$S^{(3)} = \{q \in \mathbb{H};\ |q| = 1\} = \bigl\{(x_0, x_1, x_2, x_3) \in \mathbb{R}^4;\ x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1\bigr\}$$

z očividných dôvodov stotožňujeme s trojrozmernou jednotkovou sférou v $\mathbb{R}^4$. Pre $p, q \in S^{(3)}$ zrejme platí $p\,q \in S^{(3)}$, ako aj $q^{-1} = q^* \in S^{(3)}$.

Kvaternióny možno taktiež reprezentovať komplexnými maticami tvaru

$$\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix},$$

kde a, b sú ľubovoľné komplexné čísla. Množinu všetkých takýchto matíc označíme $\mathcal{U}$. Pre matice z $\mathcal{U}$ platí

$$\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} + \begin{pmatrix} c & d \\ -\overline{d} & \overline{c} \end{pmatrix} = \begin{pmatrix} a + c & b + d \\ -(\overline{b} + \overline{d}) & \overline{a} + \overline{c} \end{pmatrix},$$

$$\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \cdot \begin{pmatrix} c & d \\ -\overline{d} & \overline{c} \end{pmatrix} = \begin{pmatrix} ac - b\overline{d} & ad + b\overline{c} \\ -(\overline{a}\overline{d} + \overline{b}c) & \overline{a}\,\overline{c} - \overline{b}d \end{pmatrix},$$

teda ich súčet i súčin sú opäť matice uvedeného typu. Zrejme i nulová matica $\mathbf{0} \in \mathbb{C}^{2\times 2}$ a jednotková matica $\boldsymbol{I} \in \mathbb{C}^{2\times 2}$ patria do $\mathcal{U}$ a pre $\boldsymbol{A} \in \mathcal{U}$, $c \in \mathbb{R}$ platí $c\boldsymbol{A} \in \mathcal{U}$. Tým sme dokázali nasledujúce tvrdenie.

30.4.1. Tvrdenie. *Množina* $\mathcal{U}$ *všetkých komplexných matíc tvaru* $\left(\begin{smallmatrix} a & b \\ -\overline{b} & \overline{a} \end{smallmatrix}\right)$, *kde* $a, b \in \mathbb{C}$, *tvorí asociatívnu algebru nad poľom* $\mathbb{R}$ *s jednotkou* $\boldsymbol{I}$.

Pre ľubovoľný kvaternión $q = q_0 + q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k}$ položme

$$\boldsymbol{U}(q) = \begin{pmatrix} q_0 + q_1\mathrm{i} & q_2 + q_3\mathrm{i} \\ -q_2 + q_3\mathrm{i} & q_0 - q_1\mathrm{i} \end{pmatrix}.$$

Potom $\boldsymbol{U}\colon \mathbb{H} \to \mathcal{U}$ je zrejme bijektívne $\mathbb{R}$-lineárne zobrazenie.

30.4.2. Tvrdenie. *Zobrazenie $\boldsymbol{U}\colon \mathbb{H} \to \mathcal{U}$ je izomorfizmus $\mathbb{R}$-algebier $\mathbb{H}$ a $\mathcal{U}$.*

Dôkaz. Zostáva dokázať, že $\boldsymbol{U}$ zachováva súčin. Na to stačí podľa vety 30.1.2 overiť, že pre ľubovoľnú dvojicu $x, y \in \{1, \mathrm{i}, \mathrm{j}, \mathrm{k}\}$ platí $\boldsymbol{U}(xy) = \boldsymbol{U}(x) \cdot \boldsymbol{U}(y)$ (pozri tiež cvičenie 30.30). Napr.

$$\boldsymbol{U}\big(\mathrm{i}^2\big) = \boldsymbol{U}(-1) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}^2 = \boldsymbol{U}(\mathrm{i})^2,$$

$$\boldsymbol{U}(\mathrm{ij}) = \boldsymbol{U}(\mathrm{k}) = \begin{pmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \boldsymbol{U}(\mathrm{i}) \cdot \boldsymbol{U}(\mathrm{j}).$$

Zvyšné priamočiare výpočty prenechávame čitateľovi (nakoľko $\boldsymbol{U}(1) = \boldsymbol{I}$, stačí sa obmedziť na $x, y \neq 1$).

Keďže $\mathcal{U}$ je asociatívna a $\mathbb{H} \cong \mathcal{U}$, je i $\mathbb{H}$ asociatívna. Ak chce čitateľ doceniť spôsob, akým sme práve dokázali asociatívnosť násobenia kvaterniónov, nech skúsi overiť identitu $p\,(qr) = (p\,q)r$ priamym výpočtom.

30.4.3. Dôsledok. *Množina $\mathbb{H}^\star = \mathbb{H} \smallsetminus \{0\}$ všetkých nenulových kvaterniónov s operáciou súčinu tvorí grupu. Množina $S^{(3)}$ všetkých jednotkových kvaterniónov je jej normálna podgrupa.*

Za povšimnutie stojí i to, že pre matice z $\mathcal{U}$ platí

$$\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix}^* = \begin{pmatrix} \overline{a} & -b \\ \overline{b} & a \end{pmatrix},$$

$$\begin{vmatrix} a & b \\ -\overline{b} & \overline{a} \end{vmatrix} = |a|^2 + |b|^2,$$

z čoho pre ľubovoľné $q \in \mathbb{H}$ vyplýva

$$\begin{aligned} \boldsymbol{U}(q)^* &= \boldsymbol{U}(q^*) \\ \boldsymbol{U}(q) \cdot \boldsymbol{U}(q)^* &= \boldsymbol{U}(qq^*) = |q|^2\boldsymbol{I}, \end{aligned}$$

a konečne

$$\det \boldsymbol{U}(q) = |q_0 + q_1\mathrm{i}|^2 + |q_2 + q_3\mathrm{i}|^2 = |q|^2 = \frac{1}{2}\,\mathrm{tr}\big(\boldsymbol{U}(q)\cdot\boldsymbol{U}(q)^*\big).$$

Ďalej si všimnime, že špeciálna unitárna grupa $\mathrm{SU}(2)$ splýva podľa vety 29.6.2 s podmnožinou algebry $\mathcal{U}$ tvorenou maticami s determinantom 1. Zúženie zobrazenia $\boldsymbol{U}$ na množinu jednotkových kvaterniónov je tak izomorfizmom grupy $S^{(3)}$ na grupu $\mathrm{SU}(2)$. Tým sme dokázali ďalší výsledok.

30.4.4. Veta. *Špeciálna unitárna grupa* $\mathrm{SU}(2)$ *je izomorfná s multiplikatívnou grupou jednotkových kvaterniónov alias jednotkovou trojrozmernou sférou* $S^{(3)}$.

Ak z definície poľa z paragrafu 1.2 vypustíme podmienku komutatívnosti násobenia, získame definíciu štruktúry, ktorej hovoríme *teleso*, prípadne *kosé pole* alebo *okruh s delením*. Pole teda je komutatívne teleso (t. j. teleso s komutatívnym násobením). Kvaternióny sú tak príkladom telesa, ktoré nie je poľom.

Algebra s jednotkou $\mathcal{A}$ nad poľom K sa nazýva *algebra s delením*, ak ku každému prvku $\mathbf{0} \neq \boldsymbol{a} \in \mathcal{A}$ existuje (obojstranný) inverzný prvok $\boldsymbol{a}^{-1} \in \mathcal{A}$. S využitím práve zavedenej terminológie teda môžeme povedať, že kvaternióny $\mathbb{H}$ tvoria nekomutatívnu asociatívnu $\mathbb{R}$-algebru s delením. Ako dokázal Georg Frobenius v r. 1878, okrem nej existujú (až na izomorfizmus) len dve asociatívne $\mathbb{R}$-algebry s delením: samotné $\mathbb{R}$ a $\mathbb{C}$, ktoré sú obe komutatívne. Ak oželieme aj podmienku asociatívnosti, môžeme dostať ešte jednu $\mathbb{R}$-algebru s delením: tzv. *Cayleyho čísla* alebo *oktonióny* $\mathbb{O}$, objavené vzápätí po Hamiltonovom objave jeho priateľom Johnom Gravesom a krátko nato a nezávisle na ňom Arthurom Cayleym. Ako naznačuje ich názov, ide o osemrozmernú algebru nad $\mathbb{R}$. Teda $\mathbb{R}$-algebry s delením existujú iba v dimenziách 1, 2, 4 a 8.

30.5 Goniometrický tvar kvaterniónu

Pre každý vektorový kvaternión $u = u_1\mathrm{i} + u_2\mathrm{j} + u_3\mathrm{k}$ jednotkovej dĺžky $|u| = 1$ platí

$$u^2 = -\langle u, u\rangle_E + (u \times u) = -|u|^2 = -1,$$

rovnako ako pre imaginárne jednotky i, j, k. Z toho dôvodu budeme všetky takéto kvaternióny nazývať *imaginárnymi jednotkami*. Množinu všetkých imaginárnych jednotiek stotožníme s jednotkovou *dvojrozmernou sférou*

$$S^{(2)} = \big\{(x_1, x_2, x_3) \in \mathbb{R}^3;\ x_1^2 + x_2^2 + x_3^2 = 1\big\}.$$

Obdobou vyjadrenia komplexného čísla $z = x+y\mathrm{i}$ v goniometrickom tvare

$$z = |z|\,(\cos\alpha + \mathrm{i}\sin\alpha)$$

je vyjadrenie kvaterniónu $q = q_0 + q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k}$ v tvare

$$q = |q|\,(\cos\alpha + u\sin\alpha), \tag{$*$}$$

kde $\alpha \in \mathbb{R}$ a u a je imaginárna jednotka. Pre $q = 0$ môžu byť α aj u ľubovoľné, pre $0 \neq q = q_0 \in \mathbb{R}$ je

$$q = q_0 = \begin{cases} |q|\cos 0, & \text{ak } q_0 > 0,\\ |q|\cos\pi, & \text{ak } q_0 < 0, \end{cases}$$

a u môže byť stále ľubovoľné. No pre $q \in \mathbb{H} \smallsetminus \mathbb{R}$ máme $|q| \neq 0 \neq |\vec{q}\,|$, v dôsledku čoho

$$q = q_0 + \vec{q} = |q|\left(\frac{q_0}{|q|} + \frac{\vec{q}}{|\vec{q}\,|}\frac{|\vec{q}\,|}{|q|}\right),$$

pričom $\vec{q}/|\vec{q}\,|$ je zrejme imaginárna jednotka. Preto musí platiť

$$\cos\alpha = \frac{q_0}{|q|}$$

a zároveň

$$\left.\begin{aligned} u &= \frac{\vec{q}}{|\vec{q}\,|},\\ \sin\alpha &= \frac{|\vec{q}\,|}{|q|}, \end{aligned}\right\} \qquad \text{alebo} \qquad \left\{\begin{aligned} u &= -\frac{\vec{q}}{|\vec{q}\,|},\\ \sin\alpha &= -\frac{|\vec{q}\,|}{|q|}. \end{aligned}\right.$$

Teda stále zostáva istá dvojznačnosť, nakoľko imaginárna jednotka u a uhol α určujú ten istý kvaternión ako dvojica $-u$, $-\alpha$. Keď na tom záleží, môžeme zafixovať napr. prvú možnosť dodatočnou podmienkou $0 \leq \alpha \leq \pi$.

Každú imaginárnu jednotku $u = u_1\mathrm{i} + u_2\mathrm{j} + u_3\mathrm{k} \in S^{(2)}$ možno vyjadriť pomocou jej sférických súradníc

$$u = \mathrm{i}\cos\theta\cos\omega + \mathrm{j}\sin\theta\cos\omega + \mathrm{k}\sin\omega$$

– pozri paragraf 14.4. Dosadením do ($*$) dostávame konečné vyjadrenie *kvaterniónu v goniometrickom tvare*

$$q = |q|\,\big(\cos\alpha + \sin\alpha\,(\mathrm{i}\cos\theta\cos\omega + \mathrm{j}\sin\theta\cos\omega + \mathrm{k}\sin\omega)\big).$$

No z praktických dôvodov dávame prednosť stručnejšiemu vyjadreniu ($*$). To v prípade kvaterniónovej jednotky $q \in S^{(3)}$ budeme zapisovať pomocou *zovšeobecneného Eulerovho vzťahu*

$$q = \cos\alpha + u\sin\alpha = \mathrm{e}^{u\alpha}.$$

Pre $n \in \mathbb{Z}$ (dokonca pre $n \in \mathbb{R}$) zostáva v platnosti aj obdoba Moivreovej vety

$$\left(\mathrm{e}^{u\alpha}\right)^n = (\cos\alpha + u\sin\alpha)^n = \cos n\alpha + u\sin n\alpha = \mathrm{e}^{un\alpha}\,.$$

Všimnime si však, že imaginárne jednotky u, v spĺňajú identitu

$$\mathrm{e}^{u\alpha}\,\mathrm{e}^{v\beta} = \mathrm{e}^{(u\alpha+v\beta)}$$

pre všetky $\alpha, \beta \in \mathbb{R}$ vtedy a len vtedy, keď $uv = vu$ (pozri cvičenie 30.18).

30.6 Kvaternióny a rotácie

V komplexnej rovine predstavuje násobenie $z \mapsto az$ komplexnou jednotkou $a = \mathrm{e}^{\mathrm{i}\alpha}$ otočenie okolo počiatku o uhol α. Súvis kvaterniónov s rotáciami v $\mathbb{R}^3$ a priamymi zhodnosťami v $\mathbb{R}^4$ nie je až taký priamočiary.

Nech $q = q_0 + q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k} \in \mathbb{H}$ je pevne zvolený kvaternión. Zobrazenia $L_q, R_q \colon \mathbb{H} \to \mathbb{H}$, dané predpismi

$$L_q(x) = qx \qquad \text{resp.} \qquad R_q(x) = xq$$

pre $x \in \mathbb{H}$, budeme nazývať *ľavou* resp. *pravou transláciou kvaterniónom* q. Očividne lineárne zobrazenia L_q, R_q budeme stotožňovať s ich maticami vzhľadom na kanonickú bázu $(1, \mathrm{i}, \mathrm{j}, \mathrm{k})$, čiže

$$L_q = \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix}, \qquad R_q = \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & q_3 & -q_2 \\ q_2 & -q_3 & q_0 & q_1 \\ q_3 & q_2 & -q_1 & q_0 \end{pmatrix}.$$

Keďže vďaka asociatívnosti pre všetky $p, q, x \in \mathbb{H}$ platí $L_p(R_q(x)) = pxq = R_q(L_p(x))$, zobrazenia (matice) L_p, R_q komutujú, t. j. $L_pR_q = R_qL_p$. Ďalej zrejme

$$|L_q(x)| = |R_q(x)| = |q|\,|x|$$

pre každé $x \in \mathbb{H}$, a jednoduchý výpočet dáva

$$\det L_q = \det R_q = \left(q_0^2 + q_1^2 + q_2^2 + q_3^2\right)^2 = |q|^4 \geq 0.$$

Špeciálne pre *kvaterniónovú jednotku* q a ľubovoľné $x \in \mathbb{H}$ platí

$$|L_q(x)| = |R_q(x)| = |x| \qquad \text{a} \qquad \det L_q = \det R_q = 1,$$

čiže ľavá aj pravá translácia kvaterniónom $q \in S^{(3)}$ sú *priame zhodnosti euklidovského priestoru* $\mathbb{H} \cong \mathbb{R}^4$, a ako matice, $L_q, R_q \in \mathrm{SO}(4)$. Poznamenajme, že každú maticu $\boldsymbol{A} \in \mathrm{SO}(4)$ možno zapísať v tvare

$$\boldsymbol{A} = L_pR_q = R_qL_p$$

pre vhodné kvaterniónové jednotky $p, q \in S^{(3)}$. Inak povedané, každá priama zhodnosť v euklidovskom priestore $\mathbb{R}^4 \cong \mathbb{H}$ fixujúca počiatok má tvar $x \mapsto pxq$ pre nejaké $p, q \in S^{(3)}$ (pozri cvičenie 30.31).

Nás však väčšmi zaujímajú rotácie v euklidovskom priestore $\mathbb{R}^3$, ktorý stotožňujeme s lineárnym podpriestorom $[\mathrm{i}, \mathrm{j}, \mathrm{k}] \subseteq \mathbb{H}$. *Konjugáciou kvaterniónom* $q \neq 0$ nazývame lineárne zobrazenie $\boldsymbol{\Gamma}_q \colon \mathbb{H} \to \mathbb{H}$ dané predpisom

$$\boldsymbol{\Gamma}_q(x) = qxq^{-1}$$

pre $x \in \mathbb{H}$. Zrejme $\boldsymbol{\Gamma}_q = L_q \circ R_{q^{-1}} = R_{q^{-1}} \circ L_q$ a pre reálne $a \neq 0$ platí $\boldsymbol{\Gamma}_{aq} = \boldsymbol{\Gamma}_q$. Ale to znamená, že pri konjugácii kvaterniónmi sa stačí obmedziť na kvaterniónové jednotky, lebo každá takáto konjugácia má tvar

$$\boldsymbol{\Gamma}_q = \boldsymbol{\Gamma}_s$$

kde $s \in S^{(3)}$ je kvaterniónová jednotka taká, že $q = \pm |q|\, s$.

30.6.1. Tvrdenie. *Konjugácia kvaterniónom $q \in \mathbb{H}^\star$ je priama zhodnosť $\boldsymbol{\Gamma}_q \colon \mathbb{H} \to \mathbb{H}$ fixujúca počiatok s invariantnými podpriestormi $\mathbb{R}, \mathbb{R}^3 \subseteq \mathbb{H}$.*

Dôkaz. Z našich predchádzajúcich úvah vyplýva, že $\boldsymbol{\Gamma}_q$ je izometria a jej matica v kanonickej báze $(1, \mathrm{i}, \mathrm{j}, \mathrm{k})$ má determinant

$$\det \boldsymbol{\Gamma}_q = \det L_q \det R_{q^{-1}} = |q|^4 \left|q^{-1}\right|^4 = \left|qq^{-1}\right|^4 = 1.$$

Stačí teda overiť, že $\mathbb{R}^3 = [\mathrm{i}, \mathrm{j}, \mathrm{k}]$ je $\boldsymbol{\Gamma}_q$-invariantný podpriestor. Pre $a \in \mathbb{R}$ platí

$$\boldsymbol{\Gamma}_q(a) = qaq^{-1} = aqq^{-1} = a,$$

teda podpriestor $\mathbb{R}$ všetkých skalárnych kvaterniónov je invariantný. Preto aj jeho ortokomplement $\mathbb{R}^3 = \mathbb{R}^\perp$ je invariantný – pozri tvrdenie 23.4.5.

Uvedený záver vyplýva aj z tvaru matice zobrazenia $\boldsymbol{\Gamma}_q$ vzhľadom na kanonickú bázu $(1, \mathrm{i}, \mathrm{j}, \mathrm{k})$. Po obvyklom stotožnení zobrazení a matíc môžeme písať

$$\begin{aligned}
\boldsymbol{\Gamma}_q &= L_q R_{q^{-1}} = L_q R_{q^*/|q|^2} = |q|^{-2} L_q R_{q^*} \\
&= |q|^{-2} \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix} \cdot \begin{pmatrix} q_0 & q_1 & q_2 & q_3 \\ -q_1 & q_0 & -q_3 & q_2 \\ -q_2 & q_3 & q_0 & -q_1 \\ -q_3 & -q_2 & q_1 & q_0 \end{pmatrix} \\
&= |q|^{-2} \begin{pmatrix} |q|^2 & 0 & 0 & 0 \\ 0 & q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 0 & 2(q_1q_2 + q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 0 & 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix}.
\end{aligned}$$

Vynechaním prvého riadku a stĺpca dostaneme maticu zúženia $\boldsymbol{\Phi}_q = \boldsymbol{\Gamma}_q \restriction \mathbb{R}^3$ konjugácie $\boldsymbol{\Gamma}_q$ na invariantný podpriestor $\mathbb{R}^3$ vzhľadom na kanonickú bázu $(\mathrm{i}, \mathrm{j}, \mathrm{k})$. Ak sa navyše obmedzíme len na kvaterniónové jednotky $q \in S^{(3)}$ a zobrazenie $\boldsymbol{\Phi}_q$ opäť stotožníme s jeho maticou, dostaneme

$$\boldsymbol{\Phi}_q = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix}.$$

Determinant tejto matice zrejme je $\det \boldsymbol{\Phi}_q = |q|^6 = 1$. Z toho dôvodu $\boldsymbol{\Phi}_q \in \mathrm{SO}(3)$, inak povedané $\boldsymbol{\Phi}_q$ predstavuje otočenie v $\mathbb{R}^3$ okolo nejakej osi prechádzajúcej počiatkom o určitý uhol. Štvorici reálnych čísel q_0, q_1, q_2, q_3, t. j. zložkám kvaterniónu q, hovoríme *Eulerove parametre* príslušného otočenia. Zostáva vyjasniť súvis medzi osou a uhlom daného otočenia a jeho Eulerovými parametrami. Začneme jednoduchým príkladom.

30.6.2. Príklad. Nájdeme matice $\boldsymbol{\Phi}_q$, ak q je niektorá z kvaterniónových jednotiek $\mathrm{e}^{\mathrm{i}\alpha} = \cos\alpha + \mathrm{i}\sin\alpha$, $\mathrm{e}^{\mathrm{j}\alpha} = \cos\alpha + \mathrm{j}\sin\alpha$, $\mathrm{e}^{\mathrm{k}\alpha} = \cos\alpha + \mathrm{k}\sin\alpha$, kde $\alpha \in \mathbb{R}$. Priamym dosadením a jednoduchými úpravami dostaneme

$$\boldsymbol{\Phi}_{\mathrm{e}^{\mathrm{i}\alpha}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2\alpha & -\sin 2\alpha \\ 0 & \sin 2\alpha & \cos 2\alpha \end{pmatrix} = \boldsymbol{R}^{\mathrm{i}}_{2\alpha},$$

$$\boldsymbol{\Phi}_{\mathrm{e}^{\mathrm{j}\alpha}} = \begin{pmatrix} \cos 2\alpha & 0 & \sin 2\alpha \\ 0 & 1 & 0 \\ -\sin 2\alpha & 0 & \cos 2\alpha \end{pmatrix} = \boldsymbol{R}^{\mathrm{j}}_{2\alpha},$$

$$\boldsymbol{\Phi}_{\mathrm{e}^{\mathrm{k}\alpha}} = \begin{pmatrix} \cos 2\alpha & -\sin 2\alpha & 0 \\ \sin 2\alpha & \cos 2\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{R}^{\mathrm{k}}_{2\alpha}.$$

Porovnanie s príkladom 23.5.6 ukazuje, že uvedené matice predstavujú postupne otočenia v $\mathbb{R}^3$ okolo osí $[\mathrm{i}]$, $[\mathrm{j}]$, $[\mathrm{k}]$ o uhol 2α v kladnom zmysle vzhľadom na vektory $\mathrm{i} = \boldsymbol{e}_1$, $\mathrm{j} = \boldsymbol{e}_2$, resp. $\mathrm{k} = \boldsymbol{e}_3$.

Nahraďme teraz základné imaginárne jednotky i, j, k všeobecnou imaginárnou jednotkou $u = u_1\mathrm{i} + u_2\mathrm{j} + u_3\mathrm{k} \in S^{(2)}$ a preskúmajme otočenie $\boldsymbol{\Phi}_q$ pre jednotkový kvaternión $q = \mathrm{e}^{u\alpha} = \cos\alpha + u\sin\alpha$. Predovšetkým si všimnime, že $qu = uq$, teda

$$\boldsymbol{\Phi}_q(u) = quq^{-1} = u,$$

čiže u je vlastný vektor lineárneho operátora $\boldsymbol{\Phi}_q$ prislúchajúci k jeho vlastnému číslu 1. $\boldsymbol{\Phi}_q$ teda bude otočenie okolo osi $[u]$ o zatiaľ neznámy uhol.

Ten zistíme z pôsobenia $\boldsymbol{\Phi}_q$ na vektory z ortokomplementu $[u]^{\perp}$ (vzhľadom na $\mathbb{R}^3$). Pritom stačí uvažovať vektorové kvaternióny $x \perp u$ také, že $|x| = 1$, t. j. opäť imaginárne jednotky. Pre také x platí $\langle u, x\rangle_E = 0$, teda

$$ux = -\langle u, x\rangle_E + (u \times x) = u \times x \qquad \text{a} \qquad |ux| = |u|\,|x| = 1,$$

a tiež

$$qx = \mathrm{e}^{u\alpha}\, x = x\cos\alpha + ux\sin\alpha.$$

Podľa vety 15.4.1 (c) tvoria vektory u, x a $ux = u \times x$ kladne orientovanú ortonormálnu bázu priestoru $\mathbb{R}^3$. Vektory x a ux tvoria kladne orientovanú ortonormálnu bázu roviny $[u]^{\perp}$, rovnako ako vektory 1 a i v komplexnej rovine $\mathbb{C} \cong \mathbb{R}^2$. Posledná rovnosť preto znamená, že vektor qx vznikne otočením vektora x o uhol α okolo osi $[u]$ v kladnom zmysle vzhľadom na vektor u.

Ak si ďalej uvedomíme, že $x \perp u$ má tiež za dôsledok $xu = x \times u = -(u \times x) = -ux$, dostaneme

$$xq^* = x\,\mathrm{e}^{-u\alpha} = x\cos\alpha - xu\sin\alpha = x\cos\alpha + ux\sin\alpha = qx$$

a konečne

$$\boldsymbol{\Phi}_q(x) = qxq^* = q^2x = x\cos 2\alpha + ux\sin 2\alpha.$$

Rovnako ako v predchádzajúcom odstavci dospejeme k záveru, že vektor $\boldsymbol{\Phi}_q(x) = qxq^*$ vznikne otočením vektora x o uhol 2α okolo tej istej osi $[u]$ v kladnom zmysle. S použitím označenia z paragrafu 23.5 môžeme písať

$$\boldsymbol{\Phi}_q = \boldsymbol{\Phi}_{\mathrm{e}^{u\alpha}} = \boldsymbol{R}^u_{2\alpha}$$

a výsledky našich úvah zhrnúť do vety.

30.6.3. Veta. *Nech $\alpha \in \mathbb{R}$, $u \in S^{(2)}$ je imaginárna jednotka a $q = |q|\,(\cos\alpha + u\sin\alpha)$ je nenulový kvaternión. Potom zobrazenie $x \mapsto qxq^{-1}\colon \mathbb{R}^3 \to \mathbb{R}^3$ predstavuje otočenie $\boldsymbol{R}^u_{2\alpha}$ euklidovského priestoru $\mathbb{R}^3 = [\mathrm{i}, \mathrm{j}, \mathrm{k}]$. Naopak, každé otočenie $\boldsymbol{R}^v_{\beta}$, kde $\beta \in \mathbb{R}$, $v \in S^{(2)}$, euklidovského priestoru $\mathbb{R}^3$ má tvar $\boldsymbol{R}^v_{\beta}(x) = pxp^{-1} = pxp^*$ pre jednotkový kvaternión $p = \mathrm{e}^{v\beta/2} = \cos(\beta/2) + v\sin(\beta/2)$.*

Maticu $\boldsymbol{R}^u_{\alpha}$ otočenia v $\mathbb{R}^3$ o uhol α okolo osi so smerovým jednotkovým vektorom $u = u_1\mathrm{i} + u_2\mathrm{j} + u_3\mathrm{k}$ v kladnom zmysle vzhľadom na vektor u teda získame dosadením jednotlivých zložiek jednotkového kvaterniónu

$$q = \mathrm{e}^{u\alpha/2} = \cos(\alpha/2) + u\sin(\alpha/2) = \cos(\alpha/2) + (u_1\mathrm{i} + u_2\mathrm{j} + u_3\mathrm{k})\sin(\alpha/2)$$

ako Eulerových parametrov do matice $\boldsymbol{\Phi}_q$. S využitím goniometrických vzorcov a vzťahu $u_1^2 + u_2^2 + u_3^2 = 1$ po jednoduchých úpravách dostaneme

$$\boldsymbol{R}_\alpha^u = \boldsymbol{\Phi}_{\mathrm{e}^{u\alpha/2}} =$$
$$\begin{pmatrix} u_1^2 + (u_2^2+u_3^2)\cos\alpha & u_1u_2(1-\cos\alpha) - u_3\sin\alpha & u_1u_3(1-\cos\alpha) + u_2\sin\alpha \\ u_1u_2(1-\cos\alpha) + u_3\sin\alpha & u_2^2 + (u_1^2+u_3^2)\cos\alpha & u_2u_3(1-\cos\alpha) - u_1\sin\alpha \\ u_1u_3(1-\cos\alpha) - u_2\sin\alpha & u_2u_3(1-\cos\alpha) + u_1\sin\alpha & u_3^2 + (u_1^2+u_2^2)\cos\alpha \end{pmatrix}.$$

30.7 Nakrývajúci homomorfizmus $\mathrm{SU}(2) \to \mathrm{SO}(3)$

Pozrime sa teraz na grupovo-teoretický aspekt reprezentácie rotácií v $\mathbb{R}^3$ pomocou kvaterniónov. Priradenie $q \mapsto \boldsymbol{\Phi}_q$ je samo zobrazením $\boldsymbol{\Phi}\colon \mathbb{H}^\star \to \mathrm{SO}(3)$. Na základe znalostí o konjugácii z paragrafu 28.3 ľahko nahliadneme (prípadne aj bez nich overíme priamym výpočtom), že

$$\boldsymbol{\Phi}_p \cdot \boldsymbol{\Phi}_q = \boldsymbol{\Phi}_{pq}$$

pre všetky $p, q \in \mathbb{H}^\star$. Inými slovami, $\boldsymbol{\Phi}$ je homomorfizmus grúp. Podľa vety 30.6.3 je dokonca zúženie $\boldsymbol{\Psi} = \boldsymbol{\Phi} \restriction S^{(3)}$ homomorfizmu $\boldsymbol{\Phi}$ na podgrupu $S^{(3)} \subseteq \mathbb{H}^\star$ jednotkových kvaterniónov surjektívne. Zrejme $\operatorname{Ker} \boldsymbol{\Phi} = \mathbb{R}^\star$ a $\operatorname{Ker} \boldsymbol{\Psi} = \{-1, 1\}$. Z vety 27.5.4 tak vyplývajú izomorfizmy grúp $\mathbb{H}^\star/\mathbb{R}^\star \cong \mathrm{SO}(3)$ a $S^{(3)}/\{\pm 1\} \cong \mathrm{SO}(3)$, a na základe izomorfizmu $\mathrm{SU}(2) \cong S^{(3)}$ z vety 30.4.4 tiež izomorfizmus $\mathrm{SU}(2)/\{\pm \boldsymbol{I}\} \cong \mathrm{SO}(3)$. Tým sme dokázali nasledujúcu vetu.

30.7.1. Veta. $\mathrm{SO}(3) \cong S^{(3)}/\{1, -1\} \cong \mathrm{SU}(2)/\{\boldsymbol{I}_2, -\boldsymbol{I}_2\}$.

Každá matica $\boldsymbol{A} \in \mathrm{SO}(3)$ má v homomorfizme $\boldsymbol{\Psi}\colon S^{(3)} \to \mathrm{SO}(3)$ dva vzory, rovnako ako jednotková matica $\boldsymbol{I}_3$. Tieto dva vzory majú navyše v $S^{(3)}$ také disjunktné okolia P resp. Q, že zúženie homomorfizmu $\boldsymbol{\Psi}$ na každé z nich je spojité bijektívne zobrazenie na to isté okolie M matice $\boldsymbol{A}$ v $\mathrm{SO}(3)$, a tiež obe inverzné zobrazenia $(\boldsymbol{\Psi} \restriction P)^{-1}\colon M \to P$, $(\boldsymbol{\Psi} \restriction Q)^{-1}\colon M \to Q$ sú spojité (v topologickej terminológii sa hovorí, že $\boldsymbol{\Psi} \restriction P$, $\boldsymbol{\Psi} \restriction Q$ sú *homeomorfizmy*). To je dôvod, prečo sa homomorfizmus $\boldsymbol{\Psi}$ nazýva *dvojnásobným nakrytím* špeciálnej ortogonálnej grupy $\mathrm{SO}(3)$ grupou kvaterniónových jednotiek $S^{(3)}$, prípadne len *nakrývajúcim homomorfizmom.* Rovnako dobre možno hovoriť o dvojnásobnom nakrytí grupy $\mathrm{SO}(3)$ špeciálnou unitárnou grupou $\mathrm{SU}(2)$.

Prečo týmto homomorfizmom a samotnej možnosti nakrytia grupy $\mathrm{SO}(3)$ grupou $S^{(3)}$ resp. $\mathrm{SU}(2)$ prikladáme taký význam? Podľa vety 29.8.4 sú všetky grupy $\mathrm{SU}(n)$ jednoducho súvislé. Na základe názoru a analógie s dvojrozmernou sférou $S^{(2)}$ sme zasa náchylní uveriť, že aj všetky sféry $S^{(n)}$ pre $n \geq 2$ sú jednoducho súvislé (v cvičení 29.21 je návod na presný dôkaz). Hoci zvodný izomorfizmus $\mathrm{SU}(n) \cong S^{(n+1)}$ ani nič podobné sa pre $n \neq 3$ nekoná – sféry

$S^{(2)}$ a $S^{(n)}$ pre $n > 3$ dokonca vôbec nie sú grupami. Každopádne však jednoduchá súvislosť grúp $S^{(3)}$ resp. SU(2) naznačuje, že majú v istom zmysle prehľadnú a jednoduchú štruktúru.

Na druhej strane, grupa SO(3) – hoci je súvislá – nie je jednoducho súvislá. Pokúsme sa aspoň približne naznačiť prečo. To si však vyžiada krátke odbočenie, v ktorom sa letmo zoznámime s pojmom projektívneho priestoru.

Izomorfizmus $\mathrm{SO}(3) \cong S^{(3)}/\{\pm 1\}$ o. i. hovorí, že grupa SO(3) vznikne zlepením dvojíc navzájom protiľahlých bodov q a $-q$ sféry $S^{(3)}$. Takúto konštrukciu možno urobiť aj pre sféry $S^{(n)}$ ľubovoľného iného rozmeru n. Výsledný tvar sa nazýva n-rozmerný *projektívny priestor* nad poľom $\mathbb{R}$ a značí $\mathbb{R}P^n$. Všeobecnejšie možno *projektívny priestor* KP^n nad poľom K popísať ako množinu všetkých priamok prechádzajúcich počiatkom (t. j. jednorozmerných lineárnych podpriestorov) vo vektorovom priestore K^{n+1}. V $\mathbb{R}^{n+1}$ pretína každá taká priamka jednotkovú sféru $S^{(n)}$ v dvoch protiľahlých bodoch, ktoré tak predstavujú jediný bod v priestore $\mathbb{R}P^n$.

V paragrafe 17.8 sme sa dohodli, že stavy kvantovomechanického systému zodpovedajúceho klasickej sústave s n stavmi budeme reprezentovať *lúčmi*, t. j. jednorozmernými podpriestormi $[u] = \mathbb{C}u$ v unitárnom priestore $\mathbb{C}^n$. S použitím práve zavedenej terminológie tak môžeme povedať, že stavový priestor takého kvantovomechanického systému je projektívny priestor $\mathbb{C}P^{n-1}$.

Historicky vznikla projektívna rovina $\mathbb{R}P^2$ pridaním jedného nekonečne vzdialeného *nevlastného bodu* ku každej priamke prechádzajúcej počiatkom v rovine $\mathbb{R}^2$ (tým sa „oba konce“ každej nekonečnej priamky a všetkých s ňou rovnobežných priamok akosi spojili do tohto nevlastného bodu). Všetky nevlastné body v projektívnej rovine tvoria tzv. *nevlastnú priamku*, ktorá ohraničuje či uzatvára („kompaktifikuje“) pôvodnú euklidovskú rovinu $\mathbb{R}^2$. Podobne možno dostať projektívny priestor $\mathbb{R}P^3$ pridaním nevlastného bodu ku každej priamke prechádzajúcej počiatkom v $\mathbb{R}^3$; všetky nevlastné body potom tvoria *nevlastnú rovinu*. Uvedené zovšeobecnenie pre ľubovoľné pole a ľubovoľnú dimenziu je až neskoršou záležitosťou.

Zamyslime sa, ako vyzerajú slučky v $\mathbb{R}P^n$. Tie vzniknú (po vhodnom preparametrizovaní) zo systémov ciest $f_1, \dots, f_k \colon \langle 0, 1\rangle \to S^{(n)}$, ktoré sa prepoja a uzavrú zlepením protiľahlých bodov sféry, t. j. spĺňajú podmienky $f_i(1) = -f_{i+1}(0)$ pre $i < k$ a $f_k(1) = \pm f_1(0)$. Také systémy ciest zahŕňajú okrem pôvodných slučiek v $S^{(n)}$, aj cesty $f \colon \langle 0, 1\rangle \to S^{(n)}$ také, že $f(0) = -f(1)$, dvojice ciest f_1, f_2 také, že $f_1(1) = -f_2(0)$ a $f_2(1) = \pm f_1(0)$, atď.

Pôvodné slučky sa dajú stiahnuť do bodu vďaka jednoduchej súvislosti $S^{(n)}$ pre $n \geq 2$ (tak isto je tomu so slučkami, ktoré vznikli párnym počtom lepení). Ale slučky, ktoré vznikli zlepením koncov neuzavretej cesty v $S^{(n)}$ (všeobecne slučky, ktoré potrebovali nepárny počet lepení), sa do bodu stiahnuť nedajú. V prípade sféry $S^{(2)}$ je takou napr. cesta po poludníku od jedného pólu k druhému. Teda projektívne priestory $\mathbb{R}P^n$ pre $n \geq 2$, vrátane $\mathbb{R}P^3$, ktorý nesie štruktúru grupy SO(3), nie sú jednoducho súvislé.

Podobne je tomu napr. so slučkami na plášti valca, ktorý vznikol zlepením dvoch protiľahlých strán obdĺžnika. Slučku na plášti pochádzajúcu zo slučky v pôvodnom obdĺžniku možno stiahnuť do bodu; slučka, ktorá vznikla z cesty spájajúcej dva body protiľahlých strán, čo sa majú zlepiť, sa však do bodu stiahnuť nedá – pozri obrázok 29.1.

Konkrétnym príkladom slučky v SO(3) vychádzajúcej z a končiacej v $\boldsymbol{I}$, ktorá sa nedá stiahnuť do bodu (konštantnej slučky) $\boldsymbol{I}$, je napr. funkcia $g(t) = \boldsymbol{R}^{\boldsymbol{u}}_{2\pi t}$, $t \in \langle 0, 1\rangle$, kde $\boldsymbol{u} \in S^{(2)}$ je ľubovoľný pevný vektor. Túto jej nestiahnuteľnosť si dokonca môžeme navodiť telesným pocitom. Ak sa postavíme a pripažíme pravú ruku k telu s dlaňou otočenou dozadu, môžeme ňou opísať kruh, dvíhajúc ju dopredu do vzpaženej polohy a potom ju spustiac zadom nadol. Pri *prirodzenom* pohybe skončíme s rukou opäť pripaženou no s dlaňou obrátenou *dopredu* (niečo ako „obrátenie spinu“). Ak si to skúsime ešte raz overiť a budeme sa snažiť skončiť s dlaňou obrátenou dozadu, zistíme, že v porovnaní s prvým otočením musíme ruku vo vrchnej fáze omnoho výraznejšie vychýliť zo zvislej roviny doprava (teda vlastne dosť švindlovať).

Prvky špeciálnej ortogonálnej grupy SO(3) možno chápať aj ako všetky možné pohyby sféry $S^{(2)}$ (zemského glóbu), pri ktorých jej stred zostáva nehybný (také niečo možno docieliť upevnením glóbusu pomocou gyroskopu). Zložitosť tvaru $\mathrm{SO}(3) \cong \mathbb{R}P^3$ sa prejavuje okrem iného v tom, že dvojrozmernú sféru nemožno „spojito učesať“. V dôsledku toho neexistujú spojité zobrazenia $f, g\colon S^{(2)} \to S^{(2)}$ také, že pre každé $\boldsymbol{x} \in S^{(2)}$ tvorí trojica vektorov $(\boldsymbol{x}, f(\boldsymbol{x}), g(\boldsymbol{x}))$ kladne orientovanú ortonormálnu bázu v $\mathbb{R}^3$. Napr. pokus umiestniť do každého bodu zemského glóbu trojicu (myslených) jednotkových šípiek ukazujúcich nahor, na sever a na západ stroskotá na oboch póloch. Túto skutočnosť treba brať do úvahy pri leteckej navigácii.

Ku každej zo špeciálnych ortogonálnych grúp SO(n) pre $n \geq 3$ existuje jednoducho súvislá maticová grupa Spin(n), nazývaná *spinorová grupa*, a dvojnásobne nakrývajúci homomorfizmus $\mathrm{Spin}(n) \to \mathrm{SO}(n)$. Z nich sme sa zatiaľ zoznámili iba s grupou Spin(3) a nakrývajúcim homomorfizmom $\mathrm{Spin}(3) \to \mathrm{SO}(3)$. Ako napovedá ich názov, spinorové grupy súvisia so spinom častíc, kvôli ktorému je často účelné rozlišovať medzi rotáciami okolo tej istej osi o uhly líšiace sa o nepárny násobok plného uhla 2π. Rotácia o 2π totiž síce zachováva geometrickú polohu, ale obracia spin. Okrem štyroch výnimiek

$$\begin{aligned}
&\mathrm{Spin}(3) \cong \mathrm{SU}(2) \cong S^{(3)},\\
&\mathrm{Spin}(4) \cong \mathrm{SU}(2) \times \mathrm{SU}(2) \cong S^{(3)} \times S^{(3)},\\
&\mathrm{Spin}(5) \cong \mathrm{Sp}(2),\\
&\mathrm{Spin}(6) \cong \mathrm{SU}(4)
\end{aligned}$$

sa však spinorové grupy nenachádzajú medzi grupami, s ktorými sa v tejto knihe stihneme čo len letmo zoznámiť. Aj k uvedenému zoznamu treba dodať, že tzv. *symplektická grupa* $\mathrm{Sp}(n)$ je kvaterniónová obdoba unitárnej grupy $\mathrm{U}(n)$, t. j. grupa všetkých kvaterniónových matíc $\boldsymbol{A} \in \mathbb{H}^{n\times n}$ takých, že $\boldsymbol{A} \cdot \boldsymbol{A}^* = \boldsymbol{I}_n$, kde $\boldsymbol{A}^*$ označuje maticu, ktorú dostaneme nahradením každého prvku matice $\boldsymbol{A}$ k nemu združeným kvaterniónom a následnou transpozíciou. Z nich sme sa doteraz, aj to len inkognito, stretli iba s prvou v poradí – symplektickou grupou $\mathrm{Sp}(1) = S^{(3)} \cong \mathrm{SU}(2)$. Ešte poznamenajme, že druhý z uvedených izomorfizmov, teda vlastne $\mathrm{Spin}(4) \cong \mathrm{Sp}(1) \times \mathrm{Sp}(1)$, súvisí s už spomínanou reprezentáciou matíc zo špeciálnej ortogonálnej grupy $\mathrm{SO}(4)$ v tvare $\boldsymbol{A} = L_p \cdot R_q$ pre $p, q \in \mathrm{Sp}(1)$. (pozri cvičenie 30.31). Na druhej strane hrá nakrývajúci homomorfizmus $\mathrm{Sp}(1)\times\mathrm{Sp}(1) \to \mathrm{SO}(4)$ významnú úlohu pri prepojení *štvorrozmerných variet*, *Yangových-Millsových konexií* a časticovej fyziky.

30.8 Nakrývajúci homomorfizmus $\mathrm{SL}(2,\mathbb{C}) \to \Lambda_+^{\uparrow}(3)$

Rovnako ako špeciálna ortogonálna grupa $\mathrm{SO}(3)$, ani vlastná ortochrónna Lorentzova grupa $\Lambda_+^{\uparrow}(3)$ nie je jednoducho súvislá. Z topologického hľadiska má totiž $\Lambda_+^{\uparrow}(3)$ rovnakú štruktúru ako množina $\mathrm{SO}(3) \times \mathbb{R}^3$ (hoci ako grupy nie sú izomorfné). To možno nahliadnuť na základe reprezentácie každej matice $\boldsymbol{A} \in \Lambda_+^{\uparrow}(3)$ v tvare súčinu *boostu* daného vhodným časovým vektorom $\boldsymbol{v} = (1, v_1, v_2, v_3)$ a euklidovskej rotácie v trojrozmernom priestore $[\boldsymbol{v}]^{\perp}$ (pozri komentár za dôsledkom 29.5.5 a cvičenie 29.10). Trojnásobným použitím tvrdenia 29.8.2 na interval $J = \mathbb{R}$ tak dostávame, že $\mathrm{SO}(3)\times\mathbb{R}^3$ nie je jednoducho súvislá.

Nakrývajúci homomorfizmus $\mathrm{SU}(2) \to \mathrm{SO}(3)$ však možno rozšíriť do nakrývajúceho homomorfizmu $\mathrm{SL}(2,\mathbb{C}) \to \Lambda_+^{\uparrow}(3)$, pri ktorom je vlastná ortochrónna Lorentzova grupa $\Lambda_+^{\uparrow}(3)$ dvojnásobne nakrytá špeciálnou lineárnou grupou $\mathrm{SL}(2,\mathbb{C})$. O tej už vieme, že – podobne ako grupa $\mathrm{SU}(2)$ – jednoducho súvislá je (pozri vetu 29.8.4). Navyše zúženie avizovaného homomorfizmu $\mathrm{SL}(2,\mathbb{C}) \to \Lambda_+^{\uparrow}(3)$ na podgrupu $\mathrm{SU}(2) \subseteq \mathrm{SL}(2,\mathbb{C})$ dáva (len v mierne pozmenenej podobe) nám už známy homomorfizmus $\mathrm{SU}(2) \to \mathrm{SO}(3)$.

Pripomeňme si, že pri reprezentácii kvaterniónu $q = q_0 + q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k}$ maticou

$$\boldsymbol{U}(q) = \begin{pmatrix} q_0 + q_1\mathrm{i} & q_2 + q_3\mathrm{i} \\ -q_2 + q_3\mathrm{i} & q_0 - q_1\mathrm{i} \end{pmatrix}$$

z paragrafu 30.4 pre jeho euklidovskú normu platí

$$\|q\|^2 = qq^* = q_0^2 + q_1^2 + q_2^2 + q_3^2 = \det \boldsymbol{U}(q),$$

v dôsledku čoho je podmienka $\|q\| = 1$, t. j. $q \in S^{(3)}$, ekvivalentná s $\boldsymbol{U}(q) \in \mathrm{SU}(2)$. Pokúsme sa teraz ten istý kvaternión reprezentovať mierne pozmenenou maticou $\boldsymbol{V}(q)$ tak, aby výraz $\det \boldsymbol{V}(q)$ reprezentoval indefinitnú kvadratickú formu signatúry $(1, 3)$ v *Minkovského časopriestore*, t. j. aby platilo

$$\det \boldsymbol{V}(q) = \langle q, q\rangle_M = q_0^2 - q_1^2 - q_2^2 - q_3^2.$$

Také niečo možno dosiahnuť napr. definíciou

$$\boldsymbol{V}(q) = \begin{pmatrix} q_0 + q_3 & q_1 - q_2\mathrm{i} \\ q_1 + q_2\mathrm{i} & q_0 - q_3 \end{pmatrix}.$$

Potom každé $\boldsymbol{V}(q) \in \mathbb{C}^{2\times 2}$ je zrejme hermitovská matica. Množina $\mathcal{H}$ všetkých hermitovských matíc rozmeru 2×2 tvorí vektorový priestor nad poľom $\mathbb{R}$ a samotné $\boldsymbol{V}\colon \mathbb{H} \to \mathcal{H}$ je lineárny izomorfizmus (na rozdiel od $\boldsymbol{U}\colon \mathbb{H} \to \mathcal{U}$ však $\boldsymbol{V}$ nezachováva súčin, ani $\mathcal{H}$ nie je podalgebra $\mathbb{R}$-algebry $\bigl(\mathbb{C}^{2\times 2}, \cdot\bigr)$).

Jednu z báz vektorového priestoru $\mathcal{H}$ tvoria tzv. *Pauliho matice*

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$
$$\sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

ktoré zaviedol Wolfgang Pauli, jeden z tvorcov kvantovej mechaniky, pri popise spinu elektrónu. Pri takomto označení

$$\boldsymbol{V}(q) = q_0\sigma_0 + q_1\sigma_1 + q_2\sigma_2 + q_3\sigma_3.$$

Ako cvičenie prenechávame čitateľovi, aby overil, že polárna forma ku kvadratickej forme $\langle \boldsymbol{X}, \boldsymbol{X}\rangle_M = \det \boldsymbol{X}$ na $\mathcal{H}$ je daná výrazom

$$\langle \boldsymbol{X}, \boldsymbol{Y}\rangle = \langle \boldsymbol{X}, \boldsymbol{Y}\rangle_M = \frac{1}{2}\bigl(\operatorname{tr}\boldsymbol{X}\operatorname{tr}\boldsymbol{Y} - \operatorname{tr}(\boldsymbol{X}\boldsymbol{Y})\bigr)$$

pre $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{H}$. Vzhľadom na tento pseudoskalárny súčin platí

$$\langle \sigma_0, \sigma_0\rangle = -\langle \sigma_1, \sigma_1\rangle = -\langle \sigma_2, \sigma_2\rangle = -\langle \sigma_3, \sigma_3\rangle = 1,$$
$$\langle \sigma_i, \sigma_j\rangle = 0$$

pre $i \neq j$, inak povedané, Pauliho matice tvoria inerciálnu bázu Minkowského časopriestoru $\mathcal{H} \cong \mathbb{R}^{(1,3)}$.

Obrazy reálnych kvaterniónov $q = q_0 \in \mathbb{R}$ tvoria kladne definitný podpriestor $[\sigma_0]$, obrazy vektorových kvaterniónov $q = q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k} \in \mathbb{R}^3$ tvoria jeho ortokomplement, teda maximálny záporne definitný podpriestor

$$[\sigma_0]^\perp = [\sigma_1, \sigma_2, \sigma_3] = \{\boldsymbol{X} \in \mathcal{H};\ \operatorname{tr}\boldsymbol{X} = 0\},$$

ktorý splýva s podpriestorom hermitovských matíc s nulovou stopou.

Keďže pre $\boldsymbol{A} \in \mathrm{GL}(2,\mathbb{C})$, $\boldsymbol{X} \in \mathcal{H}$ je i matica $\boldsymbol{A} \cdot \boldsymbol{X} \cdot \boldsymbol{A}^*$ hermitovská, je predpisom

$$\boldsymbol{\Theta}_{\boldsymbol{A}}(\boldsymbol{X}) = \boldsymbol{A} \cdot \boldsymbol{X} \cdot \boldsymbol{A}^*$$

definované lineárne zobrazenie $\boldsymbol{\Theta}_{\boldsymbol{A}} \colon \mathcal{H} \to \mathcal{H}$. Za predpokladu $\boldsymbol{A} \in \mathrm{SL}(2,\mathbb{C})$, t. j. $\det \boldsymbol{A} = 1$, navyše platí

$$\det\bigl(\boldsymbol{A} \cdot \boldsymbol{X} \cdot \boldsymbol{A}^*\bigr) = \det \boldsymbol{X},$$

čiže $\boldsymbol{\Theta}_A$ zachováva kvadratickú formu $\langle \boldsymbol{X}, \boldsymbol{X} \rangle = \det \boldsymbol{X}$ a tým podľa vety 29.3.4 aj pseudoskalárny súčin $\langle \boldsymbol{X}, \boldsymbol{Y} \rangle$. Preto po stotožnení lineárneho zobrazenia $\boldsymbol{\Theta}_{\boldsymbol{A}}$ s jeho maticou v kanonickej báze $\boldsymbol{\sigma} = (\sigma_0, \sigma_1, \sigma_2, \sigma_3)$, môžeme tvrdiť, že $\boldsymbol{\Theta}_{\boldsymbol{A}} \in \Lambda(3) = \mathrm{O}(1,3)$. Priradenie $\boldsymbol{A} \mapsto \boldsymbol{\Theta}_{\boldsymbol{A}}$ je očividne homomorfizmus grúp $\boldsymbol{\Theta} \colon \mathrm{SL}(2,\mathbb{C}) \to \Lambda(3)$. Okrem toho pre $\boldsymbol{A} \in \mathrm{SU}(2)$ platí $\boldsymbol{\Theta}_{\boldsymbol{A}}(\boldsymbol{X}) = \boldsymbol{A} \cdot \boldsymbol{X} \cdot \boldsymbol{A}^{-1}$, teda ak $\boldsymbol{A} = \boldsymbol{V}(q)$ pre $q \in S^{(3)}$, tak $\boldsymbol{\Theta}_{\boldsymbol{A}}$ funguje na $\mathcal{H}$ rovnako ako $\boldsymbol{\Gamma}_q$ z paragrafu 30.6 na $\mathbb{H}$. Platí však ešte o čosi viac.

30.8.1. Veta. *Zobrazenie* $\boldsymbol{\Theta} \colon \mathrm{SL}(2,\mathbb{C}) \to \mathrm{O}(1,3)$, *kde* $\boldsymbol{\Theta}_{\boldsymbol{A}}(\boldsymbol{X}) = \boldsymbol{A} \cdot \boldsymbol{X} \cdot \boldsymbol{A}^*$ *pre* $\boldsymbol{A} \in \mathrm{SL}(2,\mathbb{C})$, $\boldsymbol{X} \in \mathcal{H}$, *je spojitý homomorfizmus grúp s jadrom* $\mathrm{Ker}\, \boldsymbol{\Theta} = \{\pm \boldsymbol{I}\}$ *a obrazom* $\mathrm{Im}\, \boldsymbol{\Theta} = \Lambda_+^\uparrow(3)$. *Navyše, pre* $\boldsymbol{A} \in \mathrm{SU}(2)$ *je zúženie* $\boldsymbol{\Theta}_A \restriction [\sigma_1, \sigma_2, \sigma_3] \in \mathrm{SO}(3)$.

Inak povedané, $\boldsymbol{\Theta}$ je dvojnásobne nakrývajúci homomorfizmus vlastnej ortochrónnej Lorentzovej grupy $\Lambda_+^\uparrow(3)$ špeciálnou lineárnou grupou $\mathrm{SL}(2,\mathbb{C})$, v dôsledku čoho $\Lambda_+^\uparrow(3) \cong \mathrm{SL}(2,\mathbb{C})/\{\pm \boldsymbol{I}\}$. Zúženie izomorfizmu $\boldsymbol{\Theta}$ na špeciálnu unitárnu grupu $\mathrm{SU}(2)$ dáva nám už známy nakrývajúci homomorfizmus $\mathrm{SU}(2) \to \mathrm{SO}(3)$.

Dôkaz. Zostáva overiť, že pre jadro resp. obraz homomorfizmu $\boldsymbol{\Theta}$ skutočne platí $\mathrm{Ker}\, \boldsymbol{\Theta} = \{\pm \boldsymbol{I}\}$ resp. $\mathrm{Im}\, \boldsymbol{\Theta} = \Lambda_+^\uparrow(3)$ a invariantnosť podpriestoru $[\sigma_1, \sigma_2, \sigma_3]$ matíc $\boldsymbol{X} \in \mathcal{H}$ s nulovou stopou vzhľadom na zobrazenia $\boldsymbol{\Theta}_{\boldsymbol{A}}$ pre $\boldsymbol{A} \in \mathrm{SU}(2)$.

Začnime posledným bodom. Pre $\boldsymbol{A} \in \mathrm{SU}(2)$ a $\boldsymbol{X} \in \mathcal{H}$ s použitím tvrdenia 18.1.3 a rovností $\boldsymbol{A}^* \cdot \boldsymbol{A} = \boldsymbol{I}$, $\mathrm{tr}\, \boldsymbol{X} = 0$ vyplýva

$$\mathrm{tr}(\boldsymbol{A} \cdot \boldsymbol{X} \cdot \boldsymbol{A}^*) = \mathrm{tr}(\boldsymbol{A}^* \cdot \boldsymbol{A} \cdot \boldsymbol{X}) = \mathrm{tr}\, \boldsymbol{X} = 0.$$

Nech $\boldsymbol{A} \in \mathrm{Ker}\, \boldsymbol{\Theta}$. Potom $\boldsymbol{A} \cdot \sigma_i \cdot \boldsymbol{A}^* = \sigma_i$ pre $i = 0,1,2,3$. Pre $i = 0$ to znamená, že $\boldsymbol{A} \cdot \boldsymbol{A}^* = \boldsymbol{I}$, teda $\boldsymbol{A} \in \mathrm{SU}(2)$. Ďalšie tri podmienky znamenajú, že zúženie $\boldsymbol{\Theta}_{\boldsymbol{A}}$ na euklidovský podpriestor $[\sigma_1, \sigma_2, \sigma_3]$ je identické zobrazenie, z čoho na základe úvah predchádzajúcich vetu 30.7.1. vyplýva, že $\boldsymbol{A} = \pm \boldsymbol{I}$.

Keďže grupa $\mathrm{SL}(2,\mathbb{C})$ je podľa vety 29.7.6 (c) súvislá a $\boldsymbol{\Theta}$ je spojité zobrazenie, je aj obraz $\mathrm{Im}\, \boldsymbol{\Theta}$ súvislá množina. Preto $\mathrm{Im}\, \boldsymbol{\Theta}$ je podmnožinou súvislej komponenty jednotky $\Lambda_+^\uparrow(3)$ v $\Lambda(3)$. Naopak, každú maticu $\boldsymbol{M} \in \Lambda_+^\uparrow(3)$ možno podľa vety 29.5.4 napísať v tvare súčinu

$$\boldsymbol{M} = \mathrm{diag}(1, \boldsymbol{B}) \cdot \mathrm{diag}(\boldsymbol{L}_v, 1, 1) \cdot \mathrm{diag}(1, \boldsymbol{C}),$$

kde $\boldsymbol{B}, \boldsymbol{C} \in \mathrm{SO}(3)$ a $|v| < 1$. Stačí teda ukázať, že matice uvedených typov ležia v $\operatorname{Im} \boldsymbol{\Theta}$.

Ak $\boldsymbol{B} \in \mathrm{SO}(3)$, tak podľa výsledkov predchádzajúceho paragrafu existuje matica $\boldsymbol{A} \in \mathrm{SU}(2)$ taká, že matica zúženia $\boldsymbol{\Theta}_{\boldsymbol{A}} \restriction [\sigma_1, \sigma_2, \sigma_3]$ v báze $(\sigma_1, \sigma_2, \sigma_3)$ je $\boldsymbol{B}$. Keďže $\boldsymbol{\Theta}_{\boldsymbol{A}}(\sigma_0) = \sigma_0$, platí $\boldsymbol{\Theta}_{\boldsymbol{A}} = \operatorname{diag}(1, \boldsymbol{B})$ (presnejšie, matica zobrazenia $\boldsymbol{\Theta}_{\boldsymbol{A}}$ v báze $\boldsymbol{\sigma}$ je $\operatorname{diag}(1, \boldsymbol{B})$).

Ak $|v| < 1$, tak – ako sme sa presvedčili v paragrafe 16.9 (pozri tiež dôsledok 29.5.3) – existuje reálne číslo γ také, že $\boldsymbol{L}_v = \boldsymbol{Rh}_\gamma$. Hľadáme teda maticu $\boldsymbol{A} \in \mathrm{SL}(2, \mathbb{C})$, pre ktorú platí

$$\begin{aligned} \boldsymbol{A} \cdot \sigma_0 \cdot \boldsymbol{A}^* &= (\cosh \gamma)\, \sigma_0 + (\sinh \gamma)\, \sigma_1, & \boldsymbol{A} \cdot \sigma_2 \cdot \boldsymbol{A}^* &= \sigma_2, \\ \boldsymbol{A} \cdot \sigma_1 \cdot \boldsymbol{A}^* &= (\sinh \gamma)\, \sigma_0 + (\cosh \gamma)\, \sigma_1, & \boldsymbol{A} \cdot \sigma_3 \cdot \boldsymbol{A}^* &= \sigma_3. \end{aligned}$$

Prvá podmienka je vlastne rovnosť $\boldsymbol{A} \cdot \boldsymbol{A}^* = \boldsymbol{Rh}_\gamma$, ktorej vďaka tomu, že zobrazenie $\boldsymbol{Rh} \colon \mathbb{R} \to \Lambda_+^\uparrow(1)$ je homomorfizmus grúp, vyhovuje reálna symetrická matica $\boldsymbol{A} = \boldsymbol{Rh}_{\gamma/2} \in \mathrm{SL}(2, \mathbb{C})$. Priamymi výpočtami, ktoré prenechávame čitateľovi, možno overiť, že $\boldsymbol{A}$ spĺňa aj zvyšné tri podmienky.

Cvičenia

30.1. Pripomeňte si definíciu okruhu s jednotkou (cvičenia 1.8 a 2.10) a dokážte nasledujúce tvrdenia:

(a) Každá asociatívna algebra s jednotkou je zároveň okruh s jednotkou.

(b) Táto algebra je komutatívna práve vtedy, keď príslušný okruh je komutatívny.

30.2. Kompozícia homomorfizmov algebier je homomorfizmus algebier a inverzné zobrazenie k izomorfizmu algebier je tiež izomorfizmus algebier. Dokážte.

30.3. Nech $\mathcal{A}$, $\mathcal{B}$ sú K-algebry, $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ je báza v $\mathcal{A}$.

(a) Dokážte, že lineárne zobrazenie $\varphi \colon \mathcal{A} \to \mathcal{B}$ je homomorfizmus algebier práve vtedy, keď pre ľub. dvojicu $\boldsymbol{u}_i$, $\boldsymbol{u}_j$ prvkov bázy $\boldsymbol{\alpha}$ platí $\varphi(\boldsymbol{u}_i \boldsymbol{u}_j) = \varphi(\boldsymbol{u}_i)\, \varphi(\boldsymbol{u}_j)$; homomorfizmus φ je izomorfizmus práve vtedy, keď navyše $\varphi(\boldsymbol{\alpha}) = (\varphi \boldsymbol{u}_1, \ldots, \varphi \boldsymbol{u}_n)$ je báza v $\mathcal{B}$.

(b) Na základe (a) dokážte vetu 30.1.2.

30.4. Vyjadrite podmienky komutatívnosti resp. asociatívnosti lineárnej algebry algebry v reči štruktúrnych konštánt.

30.5. (a) V zhode s definíciou kongruencie na grupe (pozri paragraf 27.5) zadefinujte pojem *kongruencie na lineárnej algebre* $\mathcal{A}$ ako ekvivalancie na množine $\mathcal{A}$, ktorá sa zachováva pri operáciach súčtu, skalárneho násobku a súčinu v $\mathcal{A}$.

(b) Nech $\sim$ je kongruencia na K-algebre $\mathcal{A}$. Na faktorovej množine $\mathcal{A}/\!\sim$ zadefinujte operácie súčtu, skalárneho násobku a súčinu tak, aby kanonické zobrazenie $\boldsymbol{x} \mapsto \widetilde{\boldsymbol{x}} \colon \mathcal{A} \to \mathcal{A}/\!\sim$ bolo homomorfizmom K-algebier. Uvedomte si, že to možno urobiť jediným spôsobom, a overte korektnosť svojej definície, aj to, že $\mathcal{A}/\!\sim$ je naozaj K-algebra.

(c) Dokážte, že ak pôvodná algebra $\mathcal{A}$ je asociatívna, komutatívna, resp. algebra s jednotkou, tak aj faktorová algebra $\mathcal{A}/\!\sim$ má príslušnú vlastnosť.

30.6. Lineárny podpriestor $J \subseteq \mathcal{A}$ nazývame *ideálom K-algebry* $\mathcal{A}$, ak pre všetky $\boldsymbol{a} \in \mathcal{A}$, $\boldsymbol{x} \in J$ platí $\boldsymbol{ax}, \boldsymbol{xa} \in J$. Dokážte nasledujúce tvrdenia:

(a) Nech $\sim$ je kongruencia na algebre $\mathcal{A}$. Potom množina $J_\sim = \widetilde{\mathbf{0}} = \{\boldsymbol{x} \in \mathcal{A};\, \boldsymbol{x} \sim \mathbf{0}\}$ je ideál algebry $\mathcal{A}$.

(b) Nech J je ideál algebry $\mathcal{A}$. Potom relácia $\boldsymbol{x} \sim_J \boldsymbol{y} \Leftrightarrow \boldsymbol{x}-\boldsymbol{y} \in J$ je kongruencia na algebre $\mathcal{A}$ a triedou ekvivalencie prvku $\boldsymbol{x} \in \mathcal{A}$ v kongruencii $\sim_J$ je afinný podriestor $\boldsymbol{x} + J \subseteq \mathcal{A}$ so zameraním J (porovnaj s cvičením 9.8).

(c) Priradenia popísané v (a), (b) tvoria dvojicu navzájom inverzných bijekcií medzi množinou všetkých kongruencií na algebre $\mathcal{A}$ a množinou všetkých ideálov algebry $\mathcal{A}$.

(d) Prienik ľubovoľného (konečného či nekonečného) systému ideálov algebry $\mathcal{A}$ je opäť ideál v $\mathcal{A}$.

(e) Pre každú množinu $X \subseteq \mathcal{A}$ existuje najmenší ideál $I(X)$ algebry $\mathcal{A}$ taký, že $X \subseteq I(X)$, nazývaný *ideál generovaný množinou* X.

(f) Ak $\mathcal{A}$ je asociatívna algebra, tak $I(X)$ splýva s lineárnym obalom množiny $\{\boldsymbol{axb};\, \boldsymbol{x} \in X \;\&\; \boldsymbol{a}, \boldsymbol{b}, \in \mathcal{A}\}$ vo vektorovom priestore $\mathcal{A}$.

(g) Upravte popis ideálu $I(X)$ z bodu (f) tak, aby fungoval aj v neasociatívnych algebrách.

Ideály v algebrách sú obdobou normálnych podgrúp v grupách. Faktorovú algebru $\mathcal{A}/\!\sim_J$ budeme odteraz značiť $\mathcal{A}/J$ a nazývať *faktorovou algebrou* algebry $\mathcal{A}$ *podľa ideálu* J.

30.7. Nech $\varphi\colon \mathcal{A} \to \mathcal{B}$ je homomorfizmus K-algebier.

(a) Dokážte, že jeho jadro $\operatorname{Ker}\varphi = \varphi^{-1}(\mathbf{0})$ je ideál algebry $\mathcal{A}$ a jeho obraz $\operatorname{Im}\varphi = \varphi(\mathcal{A})$ je podalgebrou algebry $\mathcal{B}$.

(b) Nájdite kanonický izomorfizmu K-algebier $\mathcal{A}/\operatorname{Ker}\varphi \cong \operatorname{Im}\varphi$.

(c) Odvoďte z (b), že $\mathcal{A} \cong \operatorname{Im}\varphi$, ak φ je injektívny, a $\mathcal{A}/\operatorname{Ker}\varphi \cong \mathcal{B}$, ak φ je surjektívny.

30.8. Nech $\boldsymbol{\alpha}, \boldsymbol{\beta}$ sú bázy n-rozmerného vektorového priestoru V nad poľom K a $r_i, s_i\colon V \to K$ sú lineárne funkcionály také, že pre ľubovoľné $\boldsymbol{x} \in V$ platí $(\boldsymbol{x})_{\boldsymbol{\alpha}} = \big(r_1(\boldsymbol{x}), \dots, r_n(\boldsymbol{x})\big)^{\mathsf{T}}$, $(\boldsymbol{x})_{\boldsymbol{\beta}} = \big(s_1(\boldsymbol{x}), \dots, s_n(\boldsymbol{x})\big)^{\mathsf{T}}$.

(a) Pre daný jednočlen $\mu(x_1, \dots, x_n) = x_1^{m_1} \dots x_n^{m_n}$ nájdite polynóm $g(x_1, \dots, x_n) = \sum_{k_1,\dots,k_n} c_{k_1\dots k_n} x_1^{k_1} \dots x_n^{k_n} \in K[x_1, \dots, x_n]$ tak, aby pre každý vektor $\boldsymbol{x} \in V$ platilo

$$g\big(s_1(\boldsymbol{x}), \dots, s_n(\boldsymbol{x})\big) = \mu\big(r_1(\boldsymbol{x}), \dots, r_n(\boldsymbol{x})\big) = r_1(\boldsymbol{x})^{m_1} \dots r_n(\boldsymbol{x})^{m_n}.$$

(b) Na základe (a) dokážte, že ak platí $\mu \in K[x_1, \dots, x_n]_p$ (t. j. $m_1 + \ldots + m_n = p$), tak aj $g \in K[x_1, \dots, x_n]_p$. Odvoďte z toho, že definícia algebry polynomických funkcií $K[V]$ ani jednotlivých p-homogénnych vrstiev $K[V]_p$ nezávisí na báze priestoru V.

30.9. V označení predchádzajúceho cvičenia zafixujme bázu $\boldsymbol{\beta}$ vektorového priestoru V a uvažujme zobrazenie $K[x_1, \dots, x_n] \to K[V]$ dané priradením $f(x_1, \dots, x_n) \mapsto f_{\boldsymbol{\beta}} = f(s_1, \dots, s_n)$.

(a) Dokážte, že uvedené zobrazenie je surjektívny homomorfizmus K-algebier a pre $f \in K[x_1, \ldots, x_n]_p$ platí $f_{\boldsymbol{\beta}} \in K[V]_p$.

(b) Za predpokladu, že pole K má viac ako p prvkov, dokážte, že priradenie $f \mapsto f_{\boldsymbol{\beta}}$ určuje bijekciu vrstvy $K[x_1, \ldots, x_n]_p$ na vrstvu $K[V]_p$. (*Návod*: Najprv overte prípady $n = 0$ a $n = 1$ a ďalej postupujte indukciou.)

(c) Odvoďte z (b), že nad nekonečným poľom K určuje priradenie $f \mapsto f_{\boldsymbol{\beta}}$ izomorfizmus graduovaných algebier $K[x_1, \ldots, x_n] \cong K[V]$.

30.10. Dokážte, že $\dim K[x_1, \ldots, x_n]_p = \binom{n+p-1}{p}$.

30.11. Uvažujme algebru polynomických funkcií $K[V]$ pre (stĺpcový) vektorový priestor $V = K^n$. Každý polynóm $f(x_1, \ldots, x_n) \in K[x_1, \ldots, x_n]$ určuje rovnako značenú funkciu $f \in K[V]$ takú, že $f(\boldsymbol{x}) = f(x_1, \ldots, x_n)$ pre $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathsf{T}} \in K^n$. Pre $f \in K[V]$ a ľubovoľnú maticu $\boldsymbol{A} \in K^{n\times n}$ označme $f^{\boldsymbol{A}}\colon V \to K$ funkciu danú predpisom $f^{\boldsymbol{A}}(\boldsymbol{x}) = f(\boldsymbol{A} \cdot \boldsymbol{x})$. Dokážte nasledujúce tvrdenia:

(a) Pre každé $\boldsymbol{A} \in K^{n\times n}$ je $f \mapsto f^{\boldsymbol{A}}$ homomorfizmus K-algebier $K[V] \to K[V]$, ktorý zachovávava rozvrstvenie, t. j. pre $f \in K[V]_p$ platí $f^{\boldsymbol{A}} \in K[V]_p$.

(b) Priradením $(f, \boldsymbol{A}) \mapsto f^{\boldsymbol{A}}$ je definovaná (pravá) akcia všeobecnej lineárnej grupy $\mathrm{GL}(n, K)$ (aj každej jej podgrupy) na algebre polynomických funkcií $K[V]$.

(c) Nech $G \subseteq \mathrm{GL}(n, K)$ je ľubovoľná maticová grupa. Polynomická funkcia $f \in K[V]$ sa nazýva *invariantom grupy* G, ak $f^{\boldsymbol{A}} = f$ pre každú maticu $\boldsymbol{A} \in G$, t. j. ak f je pevným bodom (fixpunktom) každej matice $\boldsymbol{A} \in G$. Inak povedané, f je invariantom grupy G práve vtedy, keď stabilizátorm funkcie f v príslušnej akcii $K[V] \times G \to K[V]$ je celá grupa G (pozri paragraf 28.2). Dokážte, že všetky invarianty grupy G tvoria podalgebru $K[V]^G$ algebry $K[V]$ – nazývame ju *algebrou* alebo *okruhom invariantov* grupy G.

(d) Pre podgrupy $G_1 \subseteq G_2 \subseteq \mathrm{GL}(n, K)$ platí $K[V]^{G_2} \subseteq K[V]^{G_1}$, t. j. algebra invariantov $K[V]^{G_2}$ je podalgebrou algebry invariantov $K[V]^{G_1}$.

(e) Polynomická funkcia $\|\boldsymbol{x}\|^2 = x_1^2 + \ldots + x_n^2$ je invariantom ortogonálnej grupy $\mathrm{O}(n)$.

(f) Polynomická funkcia $s^2(\boldsymbol{x}) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2$ je invariantom pseudoortogonálnej grupy $\mathrm{O}(p, q)$.

30.12. Nech G je konečná grupa, $\# G = n$, a K je pole.

(a) Overte, že pre $x, y, z \in G$ v $KG = K^G$ platí $(\delta_x * \delta_y)(z) = \delta_{(xy)}(z)$ a nájdite štruktúrne konštanty konvolúcie v grupovej algebre KG vzhľadom na bázu $G \subseteq KG$.

(b) Za predpokladu, že G je navyše komutatívna a $K = \mathbb{C}$, overte, že pre charaktery $\alpha, \beta, \gamma \in G^{\mathrm{d}} \subseteq \mathbb{C}^G = K^G$ platí $\langle \alpha * \beta, \gamma \rangle = n^2 \delta_{\alpha,\gamma}\delta_{\beta,\gamma}$. Na základe toho nájdite štruktúrne konštanty konvolúcie v grupovej algebre $\mathbb{C}^G$ vzhľadom na bázu $G^{\mathrm{d}} \subseteq \mathbb{C}^G$.

30.13. Dokážte podrobne tvrdenie 30.3.1.

30.14. Nech G je konečná grupa. Pre $g \in \mathbb{C}^G$ označme $\widetilde{g}$ funkciu danú predpisom $\widetilde{g}(x) = \overline{g}(x^{-1})$. Dokážte nasledujúce tvrdenia:

(a) Pre každý homomorfizmus φ grupy G do multiplikatívnej grupy $\mathbb{C}^\star$ platí $\varphi(G) \subseteq \mathrm{U}(1)$ (pozri cvičenie 27.11), v dôsledku čoho $\widetilde{\varphi} = \varphi$.

(b) Pre $f, g \in \mathbb{C}^G$ platí $\widetilde{\widetilde{g}} = g$ a $\langle f, g\rangle = (f * \widetilde{g})(1)$.

30.15. Nech G je konečná abelovská grupa a $F_G\colon \mathbb{C}^G \to \mathbb{C}^{G^{\mathrm{d}}}$, $F_{G^{\mathrm{d}}}\colon \mathbb{C}^{G^{\mathrm{d}}} \to \mathbb{C}^G$ označujú diskrétne Fourierove transformácie $F_G(f)(\alpha) = \langle f, \alpha\rangle = \sum_{x\in G} f(x)\,\overline{\alpha}(x)$ pre $f \in \mathbb{C}^G$, $\alpha \in G^{\mathrm{d}}$ resp. $F_{G^{\mathrm{d}}}(h)(x) = \langle h, x\rangle = \sum_{\alpha\in G^{\mathrm{d}}} h(\alpha)\,\overline{\alpha}(x)$ pre $h \in \mathbb{C}^{G^{\mathrm{d}}}$, $x \in G$.

(a) V označení cvičenia 30.14 overte, že pre každé $f \in \mathbb{C}^G$, $\alpha \in G^{\mathrm{d}}$ platí

$$F_G\big(\widetilde{f}\,\big)(\alpha) = F_G\big(\bar{f}\,\big)\big(\alpha^{-1}\big), \quad F_G(f)^{\sim}(\alpha) = F_G\big(\bar{f}\,\big)(\alpha),$$
$$(F_{G^{\mathrm{d}}} \circ F_G)(f) = (\#\, G)\widetilde{f}, \qquad (F_{G^{\mathrm{d}}} \circ F_G \circ F_{G^{\mathrm{d}}} \circ F_G)(f) = (\#\, G)^2 f.$$

(b) Overte tvar inverznej DFT $F_G^{-1}\colon \mathbb{C}^{G^{\mathrm{d}}} \to \mathbb{C}^G$, konkrétne $F_G^{-1}(h) = \frac{1}{\#\, G} F_{G^{\mathrm{d}}}\big(\widetilde{h}\big)$ pre $h \in \mathbb{C}^{G^{\mathrm{d}}}$, $x \in G$.

30.16. Pozrite sa na látku paragrafu 17.5 zo zorného uhla paragrafu 30.3.

(a) Vysvetlite, v akom zmysle možno vektory $\boldsymbol{f}_k$ považovať za charaktery grupy $(\mathbb{Z}_n, +)$.

(b) Vysvetlite, v akom zmysle splývajú dva na oko rôzne popisy diskrétnej Fourierovej transformácie na grupe $\mathbb{Z}_n$ z paragrafov 17.5 a 30.3.

30.17. Nech p, q sú ľubovoľné kvaternióny. Overte nasledujúce podmienky:

(a) $(q^*)^* = q$ a $q \in \mathbb{R} \Leftrightarrow q = q^*$;

(b) $|q^*| = |q|$ a $\left|q^{-1}\right| = |q|^{-1}$ pre $q \neq 0$;

(c) $(pq)^* = q^*p^*$ a nájdite príklad, keď $(pq)^* \neq p^*q^*$.

30.18. (a) Dokážte, že pre kvaternióny p, q platí $p\,q = qp$ práve vtedy, keď ich vektorové časti $\vec{p}$, $\vec{q}$ sú lineárne závislé. Odvoďte z toho, že jediné kvaternióny, ktoré komutujú zo všetkými kvaterniónmi, sú reálne čísla.

(b) Dokážte, že imaginárne jednotky u, v spĺňajú identitu $\mathrm{e}^{u\alpha}\,\mathrm{e}^{v\beta} = \mathrm{e}^{(u\alpha+v\beta)}$ pre všetky $\alpha, \beta \in \mathbb{R}$ práve vtedy, keď $uv = vu$. Charakterizujte bližšie takéto dvojice u, v.

30.19. (a) Zdôvodnite aspoň dvoma spôsobmi, prečo platí $S^{(3)} \lhd \mathbb{H}^\star$?

(b) Dokážte, že faktorová grupa $\mathbb{H}^\star/S^{(3)}$ je izomorfná s multiplikatívnou grupou kladných reálnych čísel $\mathbb{R}^+$.

(c) Dokáže, že grupa $\mathbb{H}^\star$ je izomorfná s priamym súčinom grúp $S^{(3)} \times \mathbb{R}^+$.

30.20. Dokážte, že na $\mathbb{R}^3$ nemožno definovať bilineárnu operáciu $\boldsymbol{xy}$ tak, aby všetky súčiny dvojíc vektorov kanonickej bázy $(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$ boli celé čísla a pre všetky $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$ platilo $\|\boldsymbol{xy}\| = \|\boldsymbol{x}\|\,\|\boldsymbol{y}\|$, kde $\|\cdot\|$ označuje euklidovskú normu. (*Návod*: Ukážte, že číslo $(1^2+1^2+1^2)(0^2+1^2+2^2) = 15$ nie je súčtom žiadnych troch druhých mocnín celých čísel; prípadne, ak vás ruší 0, urobte to isté pre číslo $(1^2+1^2+1^2)(1^2+2^2+4^2) = 63$.)

30.21. Dokážte, že množina všetkých celých čísel, ktoré sa dajú napísať ako súčet štvorcov štyroch celých čísel, je uzavretá na súčet aj súčin. (Skúste najprv priamo, ak to nepôjde, použite kvaternióny.) Odvoďte z toho, že každé prirodzené číslo možno rozložiť na súčet štyroch štvorcov prirodzených čísel.

30.22. Dokážte priamym výpočtom, že každá imaginárna jednotka $u \in \mathbb{H}$ vyhovuje rovnici $x^2 + 1 = 0$. Teda polynóm konečného stupňa $n \geq 2$ môže mať v telese – na rozdiel od poľa – nekonečne mnoho koreňov!

30.23. (a) Porovnajte goniometrický tvar vektorového kvaterniónu so sférickými súradnicami v $\mathbb{R}^3$ (pozri paragraf 14.4).

(b) Opakujte úlohu (a) pre goniometrický tvar všeobecného kvaterniónu a sférické súradnice v $\mathbb{R}^4$. Vysvetlite názorne v geometrických pojmoch, v čom a prečo sa líšia. Zmodifikujte definíciu sférických súradníc v $\mathbb{R}^4$ tak, aby sa zhodovali.

30.24. Pripomeňme, že zobrazenia $L_p, R_p\colon \mathbb{H} \to \mathbb{H}$ označujú ľavú resp. pravú transláciu kvaterniónom p. Dokážte, že pre všetky $p, q \in \mathbb{H}$ platí $L_{pq} = L_p \circ L_q$, $R_{pq} = R_q \circ R_p$ a $L_p \circ R_q = R_q \circ L_p$.

30.25. Pripomeňme, že $\boldsymbol{\Phi}_q\colon \mathbb{H} \to \mathbb{H}$ označuje zúženie konjugácie $\boldsymbol{\Gamma}_q = L_q \circ R_{q^{-1}}$ kvaterniónom $q \neq 0$ na invariantný podpriestor $\mathbb{R}^3$ a zároveň jeho maticu v kanonickej báze $(\mathrm{i}, \mathrm{j}, \mathrm{k})$. Dokážte bez priameho výpočtu, že $\det \boldsymbol{\Phi}_q = 1$ pre každé $q \in \mathbb{H}^\star$.

30.26. Nech $q = q_0 + q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k}$ je jednotkový kvaternión. Nájdite explicitné vzťahy medzi Eulerovými parametrami q_0, q_1, q_2, q_3 otočenia $\boldsymbol{\Phi}_q$ na jednej strane, a jeho uhlom α a sférickými súradnicami θ, ω jednotkového smerového vektora u jeho osi na strane druhej.

30.27. Nájdite Eulerove parametre vybraných otočení z cvičení 23.7 a 23.8.

30.28. (a) Dokážte, že spojitý homomorfizmus $z \mapsto z^2$ je dvojnásobným nakrytím grupy $\mathrm{U}(1)$ grupou $\mathrm{U}(1)$ s jadrom $\{\pm 1\}$. Odvoďte z toho, že projektívna priamka $\mathbb{R}P^1$ rovnako ako grupa $\mathrm{SO}(2) \cong \mathrm{U}(1)$ majú obe tvar kružnice, teda o. i. nie sú jednoducho súvislé.

(b) Pre spojité homomorfizmy maticových grúp zadefinujte pojem n-násobného nakrytia. Pre každé $n \geq 1$ nájdite n-násobne nakrývajúci homomorfizmus $\mathrm{U}(1) \to \mathrm{U}(1)$ a určte jeho jadro.

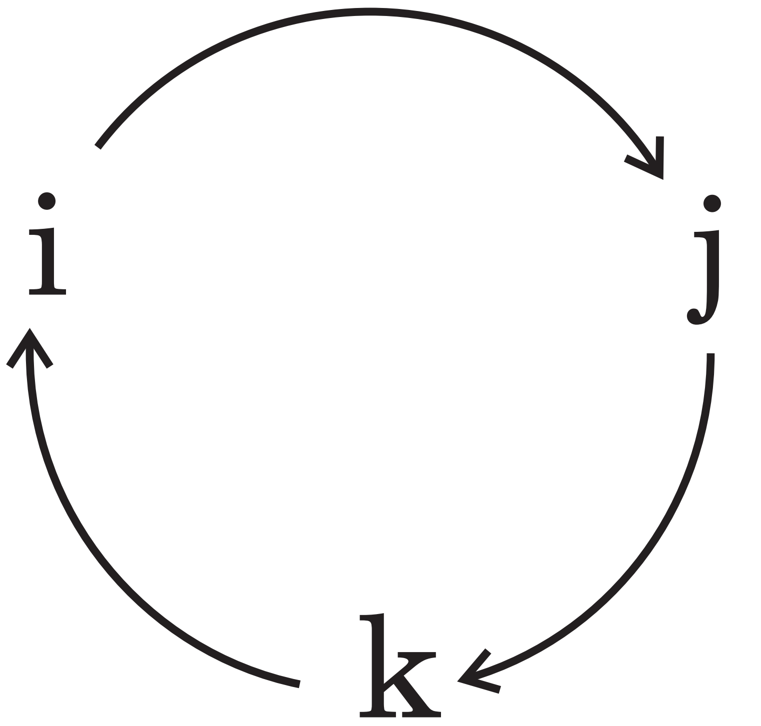
(c) Vysvetlite, čo rozumiete tvrdením, že homomorfizmus $\alpha \mapsto \boldsymbol{R}_\alpha\colon (\mathbb{R}, +) \to (\mathrm{SO}(2), \cdot)$ je nekonečnonásobným nakrytím grupy $(\mathrm{SO}(2), \cdot)$ (ktorá má tvar kružnice) grupou $(\mathbb{R}, +)$ (ktorá má tvar priamky, teda o. i. je jednoducho súvislá, i keď nie kompaktná).

30.29. (a) Vyrobte si *Möbiovu pásku* z pretočeného pruhu látky zošitím jeho dvoch protiľahlých okrajov (obrázok 30.1). Uvedomte si, že výsledný útvar má len jednu stranu a jeho okraj pozostáva len z jednej uzavretej krivky – na rozdiel od plášťa valca (zošitý nepretočený pruh), ktorý má dve strany a jeho okraj tvoria dve kružnice.

(b) Pokúste sa vyrobiť „Möbiov torus“ zošitím protiľahlých bodov na okraji Möbiovej pásky. Prečo sa vám nedarí zošívanie dokončiť? V štvorrozmernom priestore by to nebol problém – dostali by ste uzavretú plochu známu pod názvom *Kleinova fľaša.*

(c) Z valcovej plochy možno vyrobiť valec „prišitím“ dvoch kruhových podstáv k jej hraničným kružniciam. Pokúste sa prišiť k Möbiovej páske (ktorá má len jednu uzavretú hraničnú krivku) kus látky v tvare kruhu, ktorého obvod je dvojnásobkom dĺžky pôvodného pruhu. Prečo sa vám ani tentokrát nedarí prišívanie dokončiť? V štvorrozmernom priestore by ste opäť uspeli – uzavretá plocha, ktorú by ste dostali, má tvar projektívnej roviny $\mathbb{R}P^2$. Teda na rozdiel od projektívnej priamky $\mathbb{R}P^1$, ktorá má tvar kružnice $S^{(1)}$, projektívna rovina $\mathbb{R}P^2$ nemá tvar dvojrozmernej sféry $S^{(2)}$ – je to plocha nerealizovateľná v trojrozmernom priestore.

30.30. Dokážte podrobne vetu 30.4.4, t. j. overte izomorfizmus grúp $S^{(3)} = \mathrm{Sp}(1) \cong \mathrm{SU}(2)$ na základe izomorfizmu algebier $\boldsymbol{U}\colon \mathbb{H} \cong \mathcal{U}$ z tvrdení 30.4.1, 30.4.2.



Obr. 30.1. Möbiova páska

30.31. Euklidovský priestor $\mathbb{R}^4$ stotožníme s priestorom kvaterniónov $\mathbb{H}$ a ukážeme si, ako možno priame zhodnosti (t. j. matice) $\boldsymbol{A} \in \mathrm{SO}(4)$ reprezentovať dvojicami kvaterniónov.

(a) Dokážte, že priradenie $(p, q) \mapsto L_p \circ R_{q^*}$ je homomorfizmus grúp $S^{(3)} \times S^{(3)} \to \mathrm{SO}(4)$ s jadrom $\{(1, 1), (-1, -1)\}$.

(b) Dokážte, že uvedený homomorfizmus je surjektívny, teda je to dvojnásobné nakrytie grupy $\mathrm{SO}(4)$ grupou $\mathrm{Spin}(4) \cong \mathrm{Sp}(1) \times \mathrm{Sp}(1) \cong \mathrm{SU}(2) \times \mathrm{SU}(2)$. (*Návod*: Nech $\boldsymbol{A} \in \mathrm{SO}(4)$. Stĺpce matice $\boldsymbol{A}$ označte a, b, c, d a považujte ich za kvaternióny – potom (a, b, c, d) je kladne orientovaná ortonormálna báza v euklidovskom priestore $\mathbb{H} \cong \mathbb{R}^4$. Nech $q \in S^{(3)}$ je jednotkový kvaternión taký, že zhodnosť $\boldsymbol{\Phi}_q = L_q \circ R_{q^*} \colon \mathbb{R}^3 \to \mathbb{R}^3$ pretransformuje kladne orientovanú ortonormálnu bázu $(\mathrm{i}, \mathrm{j}, \mathrm{k})$ do kladne orientovanej ortonormálnej bázy (a^*b, a^*c, a^*d). Dokážte, že $\boldsymbol{A} = L_{aq} \circ R_{q^*}$.)

(c) Odvoďte z (a) a (b), že každá priama zhodnosť v euklidovskom priestore $\mathbb{H} \cong \mathbb{R}^4$ má tvar $x \mapsto pxq$ pre vhodnú dvojicou jednotkových kvaterniónov $p, q \in S^{(3)}$.

30.32. (a) Dokážte, že determinant $\det \boldsymbol{X}$ je kvadratická forma na reálnom vektorovom priestore $\mathcal{H} \subseteq \mathbb{C}^{2\times 2}$ všetkých hermitovských matíc rozmeru 2×2 a určte jej signatúru.

(b) Overte, že polárna forma ku kvadratickej forme $\det \boldsymbol{X}$ má tvar $\langle \boldsymbol{X}, \boldsymbol{Y} \rangle =$

$\frac{1}{2}\big(\operatorname{tr}\boldsymbol{X}\operatorname{tr}\boldsymbol{Y} - \operatorname{tr}(\boldsymbol{X}\boldsymbol{Y})\big)$ pre $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{H}$.

30.33. (a) Overte aj zvyšné tri podmienky pre maticu $\boldsymbol{A} = \boldsymbol{Rh}_{\gamma/2}$ a Pauliho matice zo záverečnej časti dôkazu vety 30.8.1.

(b) Nech $v \in \mathbb{R}$, $|v| < 1$, je také, že $\boldsymbol{L}_v = \boldsymbol{Rh}_\gamma$. Nájdite $u \in \mathbb{R}$, $|u| < 1$, také, že $\boldsymbol{L}_u = \boldsymbol{Rh}_{\gamma/2}$, t. j. $\boldsymbol{L}_u \cdot \boldsymbol{L}_u = \boldsymbol{L}_v$, a vyjadrite ho ako funkciu v (u je „relativistická polovica“ rýchlosti v).

(c) Uvedomte si, že pre rýchlosti, ktoré sú malé v porovnaní s rýchlosťou svetla, naozaj platí $u \approx v/2$. Na druhej strane, pre $v = c$ je $u = c$.

31. Lieove algebry a maticové grupy

Lieove algebry patria k najdôležitejším typom lineárnych algebier. Teória Lieových grúp a Lieových algebier je však značne rozsiahla a pokročilá časť modernej matematiky, pre ktorú je rámec samotnej lineárnej algebry priúzky. Ak sa chceme vyhnúť teórii vektorových polí na hladkých varietách a všeobecným Lieovým grupám, čo by nás predsa len zaviedlo priďaleko od témy nášho kurzu a zväčšilo by rozsah tejto už aj bez toho značne hrubej učebnice nad únosné medze, nezostáva nám nič iné než sa popri niekoľkých základných pojmoch a výsledkoch tejto teórie obmedziť na viac-menej informatívny prehľad Lieových algebier nám už známych maticových grúp a niekoľko málo jednoduchých ukážok ich použitia.

31.1 Lieove algebry

Lieove algebry tesne súvisia s asociatívnymi lineárnymi algebrami. Bilineárnu binárnu operáciu v Lieových algebrách však nenazývame násobením ale *Lieovými zátvorkami* alebo *komutátorom* a obvykle ju značíme hranatými zátvorkami $[\ ,\ ]$.

Lieova algebra $\mathcal{L}$ nad poľom K teda je lineárna K-algebra s operáciou $[\ ,\ ]$, ktorá je popri bilinearite navyše alternujúca, t. j.

$$[\boldsymbol{a},\boldsymbol{a}] = \mathbf{0}$$

pre všetky $\boldsymbol{a} \in \mathcal{L}$, a spĺňa tzv. *Jacobiho identitu*

$$[\boldsymbol{a},[\boldsymbol{b},\boldsymbol{c}]] + [\boldsymbol{b},[\boldsymbol{c},\boldsymbol{a}]] + [\boldsymbol{c},[\boldsymbol{a},\boldsymbol{b}]] = \mathbf{0}$$

pre všetky $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c} \in \mathcal{L}$. Z podmienok bilinearity a alternovania vyplýva antisymetria

$$[\boldsymbol{a},\boldsymbol{b}] = -[\boldsymbol{b},\boldsymbol{a}]$$

– pozri lemu 10.1.1, najmä dôkaz časti (b). Pomocou nej možno Jacobiho identitu prepísať do podoby

$$[\boldsymbol{a},[\boldsymbol{b},\boldsymbol{c}]] = [[\boldsymbol{a},\boldsymbol{b}],\boldsymbol{c}] + [\boldsymbol{b},[\boldsymbol{a},\boldsymbol{c}]],$$

ktorá vlastne hovorí, že pre pevné $\boldsymbol{a} \in \mathcal{L}$ sa lineárne zobrazenie $\mathrm{ad}_{\boldsymbol{a}} = [\boldsymbol{a},\ ]\colon \mathcal{L} \to \mathcal{L}$ správa voči Lieovým zátvorkám podobne ako derivácia voči súčinu. Táto podobnosť lepšie vynikne pri menej symetrickom zápise

$$\mathrm{ad}_{\boldsymbol{a}}[\boldsymbol{b},\boldsymbol{c}] = [\mathrm{ad}_{\boldsymbol{a}}(\boldsymbol{b}),\boldsymbol{c}] + [\boldsymbol{b},\mathrm{ad}_{\boldsymbol{a}}(\boldsymbol{c})].$$

S komutátorom matíc sme sa už stretli v paragrafe 22.7. Skutočne, vektorový priestor $K^{n\times n}$ štvorcových matíc nad ľubovoľným poľom K s operáciou komutátora $[\boldsymbol{A},\boldsymbol{B}]=\boldsymbol{AB}-\boldsymbol{BA}$ tvorí Lieovu algebru. Tento príklad pripúšťa isté zovšeobecnenie.

Komutátorom prvkov $\boldsymbol{a}$, $\boldsymbol{b}$ *asociatívnej algebry* $(\mathcal{A},\cdot)$ nazývame výraz

$$[\boldsymbol{a},\boldsymbol{b}]=\boldsymbol{ab}-\boldsymbol{ba}.$$

Potom algebru $\mathcal{A}^L=\big(\mathcal{A},[\ ,\ ]\big)$ nazývame *Lieovou algebrou asociatívnej algebry* $\mathcal{A}$. Tento názov je oprávnený nasledujúcim tvrdením.

31.1.1. Tvrdenie. *Nech* $(\mathcal{A},\cdot)$ *je asociatívna lineárna algebra nad poľom* K. *Potom* $\mathcal{A}^L$ *je Lieova algebra nad poľom* K.

Dôkaz. Bilinearita a alternovanie komutátora $[\ ,\ ]$ sú splnené triviálne. Vďaka asociatívnosti násobenia v $\mathcal{A}$ možno ľahko overiť, že pre pevné $\boldsymbol{a}\in\mathcal{A}$ sa lineárne zobrazenie $\mathrm{ad}_{\boldsymbol{a}}=[\boldsymbol{a},\ ]\colon\mathcal{A}\to\mathcal{A}$ správa voči súčinu ako derivácia, t. j. pre ľubovoľné $\boldsymbol{b},\boldsymbol{c}\in\mathcal{A}$ platí

$$[\boldsymbol{a},\boldsymbol{bc}]=[\boldsymbol{a},\boldsymbol{b}]\boldsymbol{c}+\boldsymbol{b}[\boldsymbol{a},\boldsymbol{c}].$$

Z toho sa už jednoducho odvodí rovnosť $[\boldsymbol{a},[\boldsymbol{b},\boldsymbol{c}]]=[[\boldsymbol{a},\boldsymbol{b}],\boldsymbol{c}]+[\boldsymbol{b},[\boldsymbol{a},\boldsymbol{c}]]$ a následne Jacobiho identita.

Nasledujúci príklad ukazuje, že nie každá Lieova algebra $\mathcal{L}$ je izomorfná s Lieovou algebrou $\mathcal{A}^L$ nejakej asociatívnej algebry $\mathcal{A}$. Na druhej strane každá Lieova algebra je izomorfná s *podalgebrou* Lieovej algebry $\mathcal{A}^L$ vhodnej asociatívnej algebry $\mathcal{A}$. Dôkaz tohto faktu však prekračuje rámec nášho kurzu. To je tiež dôvod, prečo sa *Lieova algebra* $\mathcal{L}$ nazýva *komutatívna*, ak v nej platí

$$[\boldsymbol{a},\boldsymbol{b}]=\boldsymbol{0}$$

pre všetky $\boldsymbol{a},\boldsymbol{b}\in\mathcal{L}$. Pre Lieovu algebru $\mathcal{A}^L$ je totiž táto podmienka ekvivalentná s komutatívnosťou pôvodnej asociatívnej algebry $\mathcal{A}$. Komutátor $[\boldsymbol{a},\boldsymbol{b}]$ tak v istom zmysle zachytáva „mieru nekomutatívnosti“ súčinu $\boldsymbol{ab}$ v $\mathcal{A}$.

Čitateľ už pozná aj ďalší jednoduchý no dôležitý príklad Lieovej algebry.

31.1.2. Príklad. Trojrozmerný vektorový priestor $\mathbb{R}^3$ nad poľom $\mathbb{R}$ s operáciou vektorového súčinu $\times$ tvorí Lieovu algebru. Overenie Jacobiho identity

$$(\boldsymbol{u}\times(\boldsymbol{v}\times\boldsymbol{w}))+(\boldsymbol{v}\times(\boldsymbol{w}\times\boldsymbol{u}))+(\boldsymbol{w}\times(\boldsymbol{u}\times\boldsymbol{v}))=\boldsymbol{0}$$

pre $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in\mathbb{R}^3$ prenechávame čitateľovi ako súčasť cvičenia 31.8. V cvičení 3.1.8 nájde čitateľ návod ako dokázať, že na $\mathbb{R}^3$ neexistuje asociatívna bilineárna operácia $(\boldsymbol{u},\boldsymbol{v})\mapsto\boldsymbol{uv}$ taká, že $\boldsymbol{u}\times\boldsymbol{v}=\boldsymbol{uv}-\boldsymbol{vu}$ pre všetky $\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^3$. Inak povedané, Lieova algebra $\big(\mathbb{R}^3,\times\big)$ nie je izomorfná so žiadnou algebrou tvaru $\mathcal{A}^L$, kde $(\mathcal{A},\cdot)$ je asociatívna $\mathbb{R}$-algebra.

31.2 Lieova algebra maticovej grupy

Z predošlých dvoch kapitol by čitateľ mal nadobudnúť dojem, že maticové grupy sú vo všeobecnosti dosť zložité objekty, ktoré sa jedna od druhej môžu výrazne líšiť svojou algebraickou, geometrickou a topologickou štruktúrou. Na druhej strane, štruktúra konečnorozmerných vektorových priestorov je pomerne jednoduchá a nám už dobre známa. Je jednoznačne určená jediným celočíselným invariantom – dimenziou (pozri vetu 6.3.4). Prekvapivo veľa informácií o štruktúre maticovej grupy však možno získať z istej Lieovej algebry (čo je vlastne „iba“ konečnorozmerný vektorový priestor doplnený binárnou operáciou komutátora) priradenej tejto grupe. Pôjde o *dotykový priestor* „priložený“ bodom $\mathbf{0}$ k maticovej grupe G v jednotkovom prvku $\boldsymbol{I} \in G$. Kvôli upresneniu tejto definície je však potrebné malé odbočenie do diferenciálnej geometrie.

Nech $M \subseteq \mathbb{R}^n$ je ľubovoľná množina. *Hladkou krivkou* v M budeme nazývať akékoľvek spojité zobrazenie $f\colon J \to M$ netriviálneho intervalu $J \subseteq \mathbb{R}$ do M, ktoré má na celom intervale J spojitú prvú deriváciu f'.[1] Hovoríme, že vektor $\boldsymbol{u} \in \mathbb{R}^n$ je *dotykový vektor* množiny M v bode $\boldsymbol{a} \in M$, ak existuje hladká krivka $f\colon J \to M$ a *vnútorný* bod t_0 intervalu J taký, že $\boldsymbol{a} = f(t_0)$ a $\boldsymbol{u} = f'(t_0)$. Zrejme $\boldsymbol{u} \in \mathbb{R}^n$ je dotykový vektor množiny $M \subseteq \mathbb{R}^n$ v bode $\boldsymbol{a} \in M$ práve vtedy, keď existuje $\varepsilon > 0$ a hladká krivka $f\colon (-\varepsilon, \varepsilon) \to M$ taká, že $\boldsymbol{a} = f(0)$ a $\boldsymbol{u} = f'(0)$.

Každá grupa reálnych matíc je podmnožinou n^2-rozmerného vektorového priestoru $\mathbb{R}^{n\times n}$ pre nejaké $n \in \mathbb{N}$; každú grupu komplexných matíc možno zasa po stotožnení $\mathbb{C} \cong \mathbb{R}^2$ chápať ako podmnožinu vhodného $2n^2$-rozmerného vektorového priestoru $(\mathbb{C}^{n\times n})_{\mathbb{R}} \cong \left(\mathbb{R}^2\right)^{n\times n}$. Množinu všetkých dotykových vektorov maticovej grupy $G \subseteq K^{n\times n}$ (kde K je niektoré z polí $\mathbb{R}$, $\mathbb{C}$) v bode $\boldsymbol{I} \in G$ budeme značiť $\mathcal{L}(G)$.

31.2.1. Veta. *Nech K je niektoré z polí $\mathbb{R}$, $\mathbb{C}$ a $G \subseteq \mathrm{GL}(n, K)$ je maticová grupa. Potom pre ľubovoľné $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{L}(G)$, $c \in \mathbb{R}$, $\boldsymbol{X} \in G$ platí*
(a) $\boldsymbol{A} + \boldsymbol{B}$, $c\boldsymbol{A} \in \mathcal{L}(G)$;
(b) $\boldsymbol{X} \cdot \boldsymbol{A} \cdot \boldsymbol{X}^{-1} \in \mathcal{L}(G)$;
(c) $[\boldsymbol{A}, \boldsymbol{B}] \in \mathcal{L}(G)$.

Dôkaz. Nech $f, g\colon (-\varepsilon, \varepsilon) \to G$ sú hladké krivky také, že $f(0) = g(0) = \boldsymbol{I}$ a $f'(0) = \boldsymbol{A}$, $g'(0) = \boldsymbol{B}$. V (a) budeme navyše predpokladať, že $c \neq 0$ (prípad $c = 0$ je totiž triviálny).

(a) Potom predpismi $t \mapsto f(t) \cdot g(t)$ resp. $t \mapsto f(ct)$ sú určené hladké krivky $f \cdot g\colon (-\varepsilon, \varepsilon) \to G$ resp. $f_c\colon (-\varepsilon/|c|, \varepsilon/|c|) \to G$, pre ktoré zrejme platí

[1] Bežne sa v matematike v definícii hladkej krivky vyžaduje existencia a spojitosť derivácií všetkých rádov. Na naše účely by to však bol zbytočný luxus.

$(f \cdot g)(0) = \boldsymbol{I}$, $f_c(0) = \boldsymbol{I}$. Jednoduchými výpočtami s využitím tvrdenia 22.4.1 dostávame

$$(f \cdot g)'(0) = f'(0) \cdot g(0) + f(0) \cdot g'(0) = \boldsymbol{A} + \boldsymbol{B},$$
$$(f_c)'(0) = cf'(0) = c\boldsymbol{A},$$

teda $\boldsymbol{A} + \boldsymbol{B}$, $c\boldsymbol{A} \in \mathcal{L}(G)$.

(b) Keďže G je grupa, predpisom $t \mapsto \boldsymbol{X} \cdot f(t) \cdot \boldsymbol{X}^{-1}$ je definovaná hladká krivka $\boldsymbol{X} \cdot f \cdot \boldsymbol{X}^{-1} \colon (-\varepsilon, \varepsilon) \to G$, pričom $\big(\boldsymbol{X} \cdot f \cdot \boldsymbol{X}^{-1}\big)(0) = \boldsymbol{I}$. Podľa tvrdenia 22.4.1

$$\big(\boldsymbol{X} \cdot f \cdot \boldsymbol{X}^{-1}\big)'(0) = \boldsymbol{X} \cdot f'(0) \cdot \boldsymbol{X}^{-1} = \boldsymbol{X} \cdot \boldsymbol{A} \cdot \boldsymbol{X}^{-1},$$

takže aj $\boldsymbol{X} \cdot \boldsymbol{A} \cdot \boldsymbol{X}^{-1} \in \mathcal{L}(G)$.

(c) Podľa (b) je predpisom $t \mapsto f(t) \cdot \boldsymbol{B} \cdot f(t)^{-1}$ definovaná funkcia $f \cdot \boldsymbol{B} \cdot f^{-1} \colon (-\varepsilon, \varepsilon) \to \mathcal{L}(G)$; zrejme je to hladká krivka v množine $\mathcal{L}(G) \subseteq K^{n\times n}$. Podľa (a) je $\mathcal{L}(G)$ lineárny podpriestor konečnorozmerného vektorového priestoru $\mathbb{R}^{n\times n}$, prípadne $(\mathbb{C}^{n\times n})_{\mathbb{R}}$. Z definície derivácie $h'(t_0)$ funkcie h v bode t_0 ako limity výrazov tvaru $\big(h(t) - h(t_0)\big)/(t - t_0)$ pre $t \to t_0$ a z uzavretosti lineárnych podpriestorov ako podmnožín v $\mathbb{R}^n$ vyplýva, že i derivácia funkcie $f \cdot \boldsymbol{B} \cdot f^{-1}$ v každom bode intervalu $(-\varepsilon, \varepsilon)$ patrí do podpriestoru $\mathcal{L}(G)$ (pozri cvičenie 22.13). Špeciálne tam patrí matica

$$\begin{aligned}\big(f \cdot \boldsymbol{B} \cdot f^{-1}\big)'(0) &= f'(0) \cdot \boldsymbol{B} \cdot f(0)^{-1} - f(0) \cdot \boldsymbol{B} \cdot f(0)^{-1} \cdot f'(0) \cdot f(0)^{-1}\\ &= \boldsymbol{A} \cdot \boldsymbol{B} - \boldsymbol{B} \cdot \boldsymbol{A} = [\boldsymbol{A}, \boldsymbol{B}].\end{aligned}$$

Pri výpočte sme použili vzorec $\big(f^{-1}\big)'(t_0) = f(t_0)^{-1} \cdot f'(t_0) \cdot f(t_0)^{-1}$ z cvičenia 22.14.

31.2.2. Dôsledok. *Nech K je niektoré z polí $\mathbb{R}$, $\mathbb{C}$ a $G \subseteq \mathrm{GL}(n, K)$ je maticová grupa. Potom dotykový priestor $\mathcal{L}(G)$ s operáciou komutátora je Lieova algebra nad poľom $\mathbb{R}$.*

Podľa (b) vety 31.2.1 je $\mathcal{L}(G)$ navyše uzavretá vzhľadom na konjugáciu prvkami grupy G.

Lieovu algebru $\mathcal{L}(G)$ nazývame *Lieovou algebrou maticovej grupy G*. Pre istotu ešte raz zdôraznime, že bez ohľadu na to, či G je reálna alebo komplexná maticová grupa, jej Lieova algebra $\mathcal{L}(G)$ je vo všeobecnosti len *reálna* Lieova algebra. Podľa okolností je to Lieova podalgebra Lieovej algebry $\mathbb{R}^{n\times n}$ resp. $(\mathbb{C}^{n\times n})_{\mathbb{R}}$ s operáciou komutátora $[\boldsymbol{A}, \boldsymbol{B}] = \boldsymbol{A} \cdot \boldsymbol{B} - \boldsymbol{B} \cdot \boldsymbol{A}$. To samozrejme neznamená, že by Lieova algebra $\mathcal{L}(G)$ nejakej maticovej grupy $G \subseteq \mathrm{GL}(n, \mathbb{C})$ nemohla byť uzavretá aj na komplexné skalárne násobky a tvoriť tak Lieovu algebru nad $\mathbb{C}$ – to je však niečo ako „bonus navyše“.

Len na okraj poznamenajme, že Lieovu algebru možno uvedeným spôsobom priradiť každej *Lieovej grupe*, čo je, veľmi približne povedané, grupa vybavená hladkou topologickou štruktúrou lokálne zhodnou s topológiou nejakého euklidovského priestoru $\mathbb{R}^n$, vzhľadom na ktorú sú grupové operácie súčinu a inverzného prvku spojito diferencovateľné.

Lieova algebra maticovej (či, všeobecnejšie, Lieovej) grupy sa zvykne značiť príslušným malým gotickým písmenom, napr. $\mathcal{L}(G) = \mathfrak{g}$, $\mathcal{L}(H) = \mathfrak{h}$ a pod.

Z definície Lieovej algebry $\mathfrak{g}$ maticovej grupy G vyplýva, že $\mathfrak{g}$ zaznamenáva chovanie prvkov grupy G len v blízkosti jej jednotky $\boldsymbol{I}$; naopak o prvkoch, do ktorých sa z $\boldsymbol{I}$ nemožno dostať po hladkej krivke, takpovediac „nevie nič“. Presnejšie možno tento jav popísať pomocou pojmu *lokálneho izomorfizmu*, ktorý však už presahuje rámec nášho kurzu. Jeden zrejmý dôsledok našej definície však môžeme zaznamenať.

31.2.3. Tvrdenie. *Nech G je reálna alebo komplexná maticová grupa. Potom*

$$\mathcal{L}(G) = \mathcal{L}\big(\Upsilon(G)\big).$$

Inak povedané, Lieova algebra $\mathfrak{g}$ grupy G je totožná s Lieovou algebrou súvislej komponenty $\Upsilon(G)$ jednotkovej matice $\boldsymbol{I} \in G$.

Napr. ak G je *diskrétna maticová grupa*, t. j. existuje $\varepsilon > 0$ také, že jediný prvok $\boldsymbol{A} \in G$, pre ktorý platí $\max_{ij} |a_{ij} - \delta_{ij}| < \varepsilon$, je jednotková matica $\boldsymbol{I}$, tak $\mathfrak{g} = \{\boldsymbol{0}\}$ je triválna Lieova algebra, rovnako ako v prípade jednoprvkovej grupy $G = \{\boldsymbol{I}\}$.

31.3 Exponenciálne zobrazenie

Prvky Lieovej algebry $\mathfrak{g} = \mathcal{L}(G)$ maticovej grupy G možno alternatívne charakterizovať s využitím exponenciály matice. V zhode s príkladom 27.3.11 nazývame *jednoparametrickou podgrupou* maticovej grupy G ľubovoľný spojitý grupový homomorfizmus $\Phi\colon (\mathbb{R}, +) \to (G, \cdot)$, prípadne jeho obraz $\operatorname{Im}\Phi = \{\Phi(t);\ t \in \mathbb{R}\}$. Možno dokázať, že každá jednoparametrická podgrupa $\Phi\colon \mathbb{R} \to G$ je zároveň hladká krivka, ktorá vyhovuje homogénnej autonómnej diferenciálnej rovnici $\Phi'(t) = \boldsymbol{A} \cdot \Phi(t)$ s počiatočnou podmienkou $\Phi(0) = \boldsymbol{I}$, kde $\boldsymbol{A} = \Phi'(0)$. Φ má preto tvar $\Phi(t) = \mathrm{e}^{\boldsymbol{A}t}$. Maticu $\boldsymbol{A}$ potom nazývame *infinitezimálnym generátorom jednoparametrickej grupy Φ*. Ide o zjavnú analógiu s cyklickými grupami, ktoré možno vyjadriť v tvare $\{a^k;\ k \in \mathbb{Z}\}$, kde a je generátor grupy.

31.3.1. Veta. *Nech K je niektoré z polí $\mathbb{R}$, $\mathbb{C}$ a $G \subseteq \mathrm{GL}(n, K)$ je maticová grupa. Potom*

$$\mathcal{L}(G) = \big\{\boldsymbol{A} \in K^{n\times n};\ (\forall\, t \in \mathbb{R})\big(\mathrm{e}^{\boldsymbol{A}t} \in G\big)\big\}.$$

To znamená, že Lieova algebra $\mathfrak{g} = \mathcal{L}(G)$ pozostáva z infinitezimálnych generátorov jednoparametrických podgrúp grupy G.
Dôkaz. Zrejme $f(t) = \mathrm{e}^{\boldsymbol{A}t}$ je hladká krivka, $f(0) = \boldsymbol{I}$, $f'(0) = \boldsymbol{A}$. Preto ak $\mathrm{e}^{\boldsymbol{A}t} \in G$ pre všetky t z nejakého netriviálneho intervalu $(-\varepsilon, \varepsilon)$, tak $\boldsymbol{A} \in \mathfrak{g}$.

Pri dôkaze obrátenej inklúzie využijeme doteraz úspešne obchádzaný predpoklad uzavretosti maticovej grupy G ako podmnožiny priestoru $K^{n\times n}$. Nech teda $\boldsymbol{A} \in K^{n\times n}$ a $f\colon (-\varepsilon, \varepsilon) \to G$ je hladká krivka, pre ktorú platí $f(0) = \boldsymbol{I}$, $f'(0) = \boldsymbol{A}$. Pre každé celé číslo $k \geq 1$ položme $f_k(t) = f(t/k)^k$. Potom $f_k\colon (-k\varepsilon, k\varepsilon) \to G$ je hladká krivka, $f_k(0) = \boldsymbol{I}$ a – keďže matica $f_k(0) = \boldsymbol{I}$ komutuje s každou maticou – podľa dôsledku 22.4.2 je $f_k'(0) = \big[kf(t/k)^{k-1} \cdot (1/k)f'(t/k)\big]_{t=0} = \boldsymbol{A}$.

Dokážeme, že pre každé $t \in \mathbb{R}$ platí

$$\lim_{\substack{k\to\infty \\ |t|<k\varepsilon}} f_k(t) = \mathrm{e}^{\boldsymbol{A}t}\,.$$

Potom z uzavretosti grupy G dostaneme $\mathrm{e}^{\boldsymbol{A}t} \in G$. Keďže funkcia f má spojitú deriváciu na intervale $(-\varepsilon, \varepsilon)$, z podmienok $f(0) = \boldsymbol{I}$, $f'(0) = \boldsymbol{A}$ vyplýva existencia okolia $(-\delta, \delta) \subseteq (-\varepsilon, \varepsilon)$ bodu 0 a spojitej maticovej funkcie $\boldsymbol{A}(t)$ takej, že $\lim_{t\to 0} \boldsymbol{A}(t) = \boldsymbol{A}$ a pre $|t| < \delta$ platí $f(t) = \boldsymbol{I} + \boldsymbol{A}(t)t$. Potom pre dosť veľké k máme $|t| < k\delta$ ako aj

$$f_k(t) = f(t/k)^k = \left(\boldsymbol{I} + \boldsymbol{A}(t/k)\frac{t}{k}\right)^k,$$

v dôsledku čoho

$$\lim_{k\to\infty} f_k(t) = \lim_{k\to\infty}\left(\boldsymbol{I} + \boldsymbol{A}(t/k)\frac{t}{k}\right)^k = \mathrm{e}^{\boldsymbol{A}t}$$

(porovnaj s cvičením 22.20).

Podľa vety 31.3.1 určuje maticová grupa G *exponenciálne zobrazenie* $\exp_G\colon \mathfrak{g} \to G$, pričom samozrejme

$$\exp_G(\boldsymbol{A}) = \exp \boldsymbol{A} = \mathrm{e}^{\boldsymbol{A}}$$

pre $\boldsymbol{A} \in \mathfrak{g}$. Obraz exponeciálneho zobrazenia

$$\mathrm{e}^{\mathfrak{g}} = \operatorname{Im}\exp_G = \big\{\mathrm{e}^{\boldsymbol{A}};\, \boldsymbol{A} \in \mathfrak{g}\big\}$$

je tvorený zjednotením všetkých jednoparametrických podgrúp grupy G a – ako obraz súvislej množiny $\mathfrak{g}$ v spojitom zobrazení – je i sám súvislý. Z toho už okamžite vyplýva nasledujúce pozorovanie.

31.3.2. Tvrdenie. *Obraz* $\mathrm{e}^{\mathfrak{g}}$ *exponenciálneho zobrazenia maticovej grupy* G *nad niektorým z polí* $\mathbb{R}$ *alebo* $\mathbb{C}$ *je súvislá podmnožina jej súvislej komponenty jednotky* $\Upsilon(G)$.

Žiadalo by sa nám dodať, že $\mathrm{e}^{\mathfrak{g}}$ je dokonca podgrupa súvislej komponenty $\Upsilon(G)$. To by však bol unáhlený záver. Hoci obraz $\mathrm{e}^{\mathfrak{g}}$ očividne obsahuje jednotkovú maticu a je uzavretý vzhľadom na maticovú inverziu, vo všeobecnosti nie je uzavretý vzhľadom na súčin matíc. Pre nekomutujúce matice totiž nemusí platiť $\mathrm{e}^{\boldsymbol{A}} \cdot \mathrm{e}^{\boldsymbol{B}} = \mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}$.

Dôkaz toho, že množina $\mathrm{e}^{\mathfrak{g}}$ generuje podgrupu $\Upsilon(G)$, teda, ak G je súvislá, priamo celú grupu G, už presahuje rámec tohto kurzu.

Exponenciálne zobrazenie $\exp_G$ je spojité a na nejakom dosť malom okolí M nulovej matice $\mathbf{0} \in \mathfrak{g}$ (to môže mať napr. tvar $M = \{\boldsymbol{A} \in \mathfrak{g};\ \max|a_{ij}| \leq \varepsilon\}$ pre vhodné $\varepsilon > 0$) aj prosté. Potom však zúženie $\exp_G \restriction M$ je nevyhnutne bijektívne zobrazenie okolia M na nejaké okolie N jednotkovej matice $\boldsymbol{I} \in G$ spojité zároveň so svojim inverzným tzv. *logaritmickým zobrazením* $(\exp_G \restriction M)^{-1} = \ln_G \restriction N$. Čo je však ešte zaujímavejšie, súčin matíc z N je plne určený štruktúrou Lieovej algebry $\mathfrak{g}$. Presnejšie, pre ľubovoľné $\boldsymbol{A}, \boldsymbol{B} \in M$ existuje jednoznačne určená matica $\boldsymbol{C} \in \mathfrak{g}$ taká, že

$$\mathrm{e}^{\boldsymbol{A}} \cdot \mathrm{e}^{\boldsymbol{B}} = \mathrm{e}^{\boldsymbol{C}},$$

ktorú možno vyjadriť ako súčet tzv. *Campbellovho-Bakerovho-Hausdorffovho radu*, pozostávajúceho z rôznym spôsobom do seba vložených komutátorov matíc $\boldsymbol{A}$ a $\boldsymbol{B}$. Na tomto mieste bez dôkazu uvádzame len jeho prvé členy

$$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B} + \frac{1}{2}[\boldsymbol{A}, \boldsymbol{B}] + \frac{1}{12}\big([\boldsymbol{A}, [\boldsymbol{A}, \boldsymbol{B}]] + [[\boldsymbol{A}, \boldsymbol{B}], \boldsymbol{B}]]\big) + \dots.$$

Vyššie členy treba chápať ako postupné „opravy“ k súčtu prvých dvoch členov $\boldsymbol{A} + \boldsymbol{B}$. Pre komutujúce $\boldsymbol{A}$, $\boldsymbol{B}$ dostávame známu formulu $\mathrm{e}^{\boldsymbol{A}} \cdot \mathrm{e}^{\boldsymbol{B}} = \mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}$ z tvrdenia 22.2.2; ak $\boldsymbol{A}$, $\boldsymbol{B}$ komutujú so svojim komutátorom, dostávame rovnosť

$$\mathrm{e}^{\boldsymbol{A}} \cdot \mathrm{e}^{\boldsymbol{B}} = \exp\left(\boldsymbol{A} + \boldsymbol{B} + \frac{1}{2}[\boldsymbol{A}, \boldsymbol{B}]\right) = \exp(\boldsymbol{A} + \boldsymbol{B}) \cdot \exp\frac{1}{2}[\boldsymbol{A}, \boldsymbol{B}]$$

z vety 22.7.3.

Nakoľko každé okolie $N \subseteq \Upsilon(G)$ jednotkovej matice $\boldsymbol{I} \in G$ už generuje súvislú komponentu $\Upsilon(G)$, vyplýva z toho, že štruktúra podgrupy $\Upsilon(G) \subseteq G$ je plne určená štruktúrou Lieovej algebry $\mathfrak{g}$. Špeciálne štruktúru súvislej maticovej grupy G možno plne zrekonštruovať na základe znalosti jej Lieovej algebry $\mathfrak{g}$. Ešte raz však upozorňujeme, že $\mathfrak{g}$ nenesie nijakú informáciu o diskrétnej faktorovej grupe $G/\Upsilon(G)$, teda o tom, ako je G zostavená z jednotlivých tried rozkladu $\Upsilon(G) \cdot \boldsymbol{A} \in G/\Upsilon(G)$.

Na druhej strane, napr. z komutatívnosti Lieovej algebry $\mathfrak{g}$ očividne vyplýva komutatívnosť súvislej komponenty $\Upsilon(G)$ jej maticovej grupy G.

Spomínaná rekonštrukcia je najjednoduchšia v prípade surjektívnosti exponenciálneho zobrazenia $\exp_G \colon \mathfrak{g} \to G$, t. j. ak každá matica $\boldsymbol{X} \in G$ má tvar $\boldsymbol{X} = \mathrm{e}^{\boldsymbol{A}}$ pre nejakú maticu $\boldsymbol{A} \in \mathfrak{g}$, – vtedy hovoríme, že G je *exponenciálna maticová grupa*.

Zrejme každá exponenciálna grupa je súvislá. Na druhej strane, nie každá súvislá maticová grupa je exponenciálna (pozri cvičenie 31.19). Ako uvidíme v nasledujúcom paragrafe, niektoré nám už známe súvislé maticové grupy sú exponenciálne.

31.4 Lieove algebry konkrétnych maticových grúp

Lieova algebra nejakej konkrétnej maticovej grupy, typicky označenej skratkou z veľkých písmen, sa značí rovnakou skratkou z malých písmen, napr. $\mathcal{L}\big(\mathrm{GL}(n,K)\big) = \mathrm{gl}(n,K)$, $\mathcal{L}\big(\mathrm{SO}(n)\big) = \mathrm{so}(n)$, $\mathcal{L}\big(\mathrm{U}(n)\big) = \mathrm{u}(n)$ a pod.

Preskúmajme teraz Lieove algebry niektorých najdôležitejších maticových grúp.

Z definujúcej podmienky $\det \boldsymbol{A} \neq 0$ regularity matice vyplýva, že všeobecné lineárne grupy $\mathrm{GL}(n,\mathbb{R})$, $\mathrm{GL}(n,\mathbb{C})$ sú otvorené podmnožiny v $\mathbb{R}^{n\times n}$ resp. $\mathbb{C}^{n\times n}$. Preto dotykový priestor v každom ich bode splýva s priestorom všetkých matíc nad príslušným poľom. Ak ešte uvážime, že súvislá komponenta jednotky v $\mathrm{GL}(n,\mathbb{R})$ je $\mathrm{GL}^+(n,\mathbb{R})$ (pozri vetu 29.7.7), podľa tvrdenia 31.2.3 z toho vyplýva:

31.4.1. Veta. $\mathrm{gl}(n,\mathbb{R}) = \mathrm{gl}^+(n,\mathbb{R}) = \mathbb{R}^{n\times n}$, $\qquad \mathrm{gl}(n,\mathbb{C}) = \mathbb{C}^{n\times n}$.

31.4.2. Veta. *Nech K je niektoré z polí $\mathbb{R}$, $\mathbb{C}$. Potom*

$$\mathrm{sl}(n,K) = \big\{\boldsymbol{A} \in K^{n\times n};\ \operatorname{tr}\boldsymbol{A} = 0\big\}.$$

Dôkaz. Nech $\boldsymbol{A} \in K^{n\times n}$ je ľubovoľná matica Podľa Liouvilleovej formuly (veta 22.2.6) pre každé $t \in \mathbb{R}$ platí

$$\det \mathrm{e}^{\boldsymbol{A}t} = \mathrm{e}^{\operatorname{tr}\boldsymbol{A}t}\,.$$

Na základe vety 31.3.1 z toho vyplýva, že $\boldsymbol{A} \in \mathrm{sl}(n,K)$ práve vtedy, keď $\mathrm{e}^{\operatorname{tr}\boldsymbol{A}t} = 1$ pre každé $t \in \mathbb{R}$. Táto podmienka je zrejme ekvivalentná s rovnosťou $\operatorname{tr}\boldsymbol{A} = 0$.

Z viet 31.4.1–2 vyplýva, že $\mathrm{gl}(n,\mathbb{C})$ aj $\mathrm{sl}(n,\mathbb{C})$ možno v prípade potreby považovať aj za Lieove algebry nad poľom $\mathbb{C}$.

K popisu Lieových algebier ďalších maticových grúp pristúpime jednotným spôsobom.

31.4.3. Veta. *(a) Nech $G = \mathrm{Stb}\,\boldsymbol{Z}$ je stabilizátor matice $\boldsymbol{Z} \in \mathbb{R}^{n\times n}$ vzhľadom na akciu $\boldsymbol{X} \mapsto \boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{X} \cdot \boldsymbol{A}$ grupy $\mathrm{GL}(n,\mathbb{R})$ na priestore matíc $\mathbb{R}^{n\times n}$. Potom*

$$\mathfrak{g} = \mathcal{L}(G) = \big\{\boldsymbol{A} \in \mathbb{R}^{n\times n};\ \boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{Z} = -\boldsymbol{Z} \cdot \boldsymbol{A}\big\}.$$

(b) Nech $G = \mathrm{Stb}\,\boldsymbol{Z}$ je stabilizátor matice $\boldsymbol{Z} \in \mathbb{C}^{n\times n}$ vzhľadom na akciu $\boldsymbol{X} \mapsto \boldsymbol{A}^* \cdot \boldsymbol{X} \cdot \boldsymbol{A}$ grupy $\mathrm{GL}(n,\mathbb{C})$ na priestore matíc $\mathbb{C}^{n\times n}$. Potom

$$\mathfrak{g} = \mathcal{L}(G) = \big\{\boldsymbol{A} \in \mathbb{C}^{n\times n};\ \boldsymbol{A}^* \cdot \boldsymbol{Z} = -\boldsymbol{Z} \cdot \boldsymbol{A}\big\}.$$

Dôkaz. Dokážeme len reálny prípad (a); komplexný prípad (b) je analogický. Nech $\boldsymbol{A} \in \mathfrak{g}$ a f je hladká krivka v G taká, že $f(0) = \boldsymbol{I}$, $f'(0) = \boldsymbol{A}$. To znamená, že pre každé t z jej definičného oboru platí $f(t)^\mathsf{T} \cdot \boldsymbol{Z} \cdot f(t) = \boldsymbol{Z}$. Derivovaním tejto rovnosti v bode $t = 0$ s použitím tvrdenia 22.4.1 dostaneme

$$f'(0)^\mathsf{T} \cdot \boldsymbol{Z} \cdot f(0) + f(0) \cdot \boldsymbol{Z} \cdot f'(0) = \boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{Z} + \boldsymbol{Z} \cdot \boldsymbol{A} = \boldsymbol{0},$$

teda, $\boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{Z} = -\boldsymbol{Z} \cdot \boldsymbol{A}$.

Naopak, nech $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ spĺňa uvedenú rovnosť. Dokážeme, že potom pre každé $t \in \mathbb{R}$ platí $\mathrm{e}^{\boldsymbol{A}t} \in G$, t. j.

$$\big(\mathrm{e}^{\boldsymbol{A}t}\big)^\mathsf{T} \cdot \boldsymbol{Z} \cdot \mathrm{e}^{\boldsymbol{A}t} = \boldsymbol{Z}.$$

V priestore $\mathbb{R}^{n\times n}$ uvažujme hladkú krivku $h(t) = \big(\mathrm{e}^{\boldsymbol{A}t}\big)^\mathsf{T} \cdot \boldsymbol{Z} \cdot \mathrm{e}^{\boldsymbol{A}t}$ a konštantnú krivku $z(t) = \boldsymbol{Z}$. Priamym výpočtom možno overiť, že obe vyhovujú počiatočnej úlohe $x'(t) = \boldsymbol{A}^\mathsf{T} \cdot x(t) + x(t) \cdot \boldsymbol{A}$, $x(0) = \boldsymbol{Z}$. Potrebný záver $h(t) = z(t) = \boldsymbol{Z}$ pre každé t potom vyplýva z jednoznačnosti riešenia tejto úlohy (pozri vetu 22.5.1 a nasledujúcu poznámku (a)).

Lieove algebry jednotlivých (pseudo)ortogonálnych či (pseudo)unitárnych grúp dostaneme vhodnou voľbou matice $\boldsymbol{Z}$. Vzhľadom na tvrdenie 31.2.3 a vetu 29.7.7 by nás už nemalo prekvapiť, že Lieova algebra nerozlišuje celú grupu a jej súvislú komponentu jednotky (ani podgrupy H, pre ktoré platí $\Upsilon(G) \subseteq H \subseteq G$).

31.4.4. Veta. *Nech k, l, n sú kladné celé čísla, $n = k+l$ a $\boldsymbol{D}_{kl} = \mathrm{diag}(\boldsymbol{I}_k, -\boldsymbol{I}_l)$. Potom*

$$\begin{aligned}
\mathrm{o}(n) = \mathrm{so}(n) &= \big\{\boldsymbol{A} \in \mathbb{R}^{n\times n};\ \boldsymbol{A}^\mathsf{T} = -\boldsymbol{A}\big\},\\
\mathrm{o}(k,l) = \mathrm{so}(k,l) = \mathrm{so}^+(k,l) &= \big\{\boldsymbol{A} \in \mathbb{R}^{n\times n};\ \boldsymbol{A}^\mathsf{T} \cdot \boldsymbol{D}_{kl} = -\boldsymbol{D}_{kl} \cdot \boldsymbol{A}\big\},\\
\mathrm{u}(n) &= \big\{\boldsymbol{A} \in \mathbb{C}^{n\times n};\ \boldsymbol{A}^* = -\boldsymbol{A}\big\},\\
\mathrm{su}(n) = \mathrm{u}(n) \cap \mathrm{sl}(n,\mathbb{C}) &= \big\{\boldsymbol{A} \in \mathbb{C}^{n\times n};\ \boldsymbol{A}^* = -\boldsymbol{A}\ \&\ \mathrm{tr}\,\boldsymbol{A} = 0\big\},\\
\mathrm{u}(k,l) &= \big\{\boldsymbol{A} \in \mathbb{C}^{n\times n};\ \boldsymbol{A}^* \cdot \boldsymbol{D}_{kl} = -\boldsymbol{D}_{kl} \cdot \boldsymbol{A}\big\},\\
\mathrm{su}(k,l) = \mathrm{u}(k,l) \cap \mathrm{sl}(n,\mathbb{C}) &= \big\{\boldsymbol{A} \in \mathbb{C}^{n\times n};\ \boldsymbol{A}^* \cdot \boldsymbol{D}_{kl} = -\boldsymbol{D}_{kl} \cdot \boldsymbol{A}\ \&\ \mathrm{tr}\,\boldsymbol{A} = 0\big\}.
\end{aligned}$$

Všimnite si, že Lieova algebra $\mathrm{o}(n) = \mathrm{so}(n)$ je tvorená všetkými antisymetrickými maticami $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. Takéto matice už automaticky majú nulovú celú diagonálu, teda aj stopu. Podobne, Lieova algebra $\mathrm{u}(n)$ pozostáva zo všetkých antihermitovských matíc $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ – tie však môžu mať nenulovú stopu.

Konečne si treba uvedomiť, že Lieove algebry $\mathrm{u}(n)$, $\mathrm{su}(n)$, $\mathrm{u}(k,l)$, $\mathrm{su}(k,l)$ nie sú uzavreté vzhľadom na komplexné skalárne násobky.

Jedna z charakteristík maticovej (všeobecnejšie Lieovej) grupy G, ktorú vieme určiť zo znalosti jej Lieovej algebry, je dimenzia. Samotná definícia dimenzie topologického priestoru nie je nijako jednoduchou záležitosťou, pre vektorové priestory nad poľom $\mathbb{R}$ však topologická dimenzia splýva s ich dimenziou ako počtom prvkov bázy. Intuitívne by malo byť jasné, že vzhľadom na translačnú symetriu je topológia maticovej grupy „všade rovnaká“, a taktiež, že okolie jednotky v grupe sa od okolia nuly v k nemu priloženom dotykovom priestore, ktoré mu zodpovedá v exponenciálnom zobrazení, líši len prípadným zakrivením, teda ich dimenzie sú rovnaké. To nám umožňuje významne si zjednodušť život tým, že *dimenziu maticovej grupy* priamo definujeme ako dimenziu jej Lieovej algebry. Teda presnejšie, ak G je maticová grupa nad niektorým z polí $\mathbb{R}$, $\mathbb{C}$, tak

$$\dim G = \dim \mathfrak{g} = \dim_{\mathbb{R}} \mathcal{L}(G).$$

Vety 31.4.1–4 majú spolu s uvedenou definíciou nasledujúci dôsledok. Samozrejme, najprv treba určiť dimenzie príslušných Lieových algebier – to prenechávame čitateľovi ako cvičenie.

31.4.5. Veta. *Nech k, l, n sú kladné celé čísla a $n = k + l$. Potom*

$$\begin{aligned}
&\dim \mathrm{GL}(n,\mathbb{R}) = \dim \mathrm{GL}^+(n,\mathbb{R}) = n^2, && \dim \mathrm{GL}(n,\mathbb{C}) = 2n^2,\\
&\dim \mathrm{SL}(n,\mathbb{R}) = n^2 - 1, && \dim \mathrm{SL}(n,\mathbb{C}) = 2(n^2 - 1),\\
&\dim \mathrm{O}(n) = \dim \mathrm{SO}(n) = \frac{1}{2}n(n-1), && \dim \mathrm{U}(n) = n^2,\\
& && \dim \mathrm{SU}(n) = n^2 - 1,\\
&\dim \mathrm{O}(k,l) = \dim \mathrm{SO}(k,l) = \frac{1}{2}n(n-1), && \dim \mathrm{U}(k,l) = n^2,\\
& && \dim \mathrm{SU}(k,l) = n^2 - 1.
\end{aligned}$$

31.4.6. Príklad. Prvky kanonickej bázy $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ v $\mathbb{R}^3$ spĺňajú komutačné vzťahy

$$\boldsymbol{e}_1 \times \boldsymbol{e}_2 = \boldsymbol{e}_3, \qquad \boldsymbol{e}_2 \times \boldsymbol{e}_3 = \boldsymbol{e}_1, \qquad \boldsymbol{e}_3 \times \boldsymbol{e}_1 = \boldsymbol{e}_2.$$

Podobne, jednu z báz Lieovej algebry so(3) očividne tvoria antisymetrické matice

$$\boldsymbol{E}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \boldsymbol{E}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad \boldsymbol{E}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

ktoré vyhovujú analogickým komutačným vzťahom

$$[\boldsymbol{E}_1, \boldsymbol{E}_2] = \boldsymbol{E}_3, \qquad [\boldsymbol{E}_2, \boldsymbol{E}_3] = \boldsymbol{E}_1, \qquad [\boldsymbol{E}_3, \boldsymbol{E}_1] = \boldsymbol{E}_2.$$

Z toho dôvodu je priradením $\boldsymbol{e}_i \mapsto \boldsymbol{E}_i$ pre $i = 1, 2, 3$ definovaný izomorfizmus Lieových algebier $\big(\mathbb{R}^3, \times\big) \cong \big(\mathrm{so}(3), [\ ,\ ]\big)$.

Jednotkové kvaternióny i, j, k tvoria bázu Lieovej algebry $\mathbb{R}^3 = [\mathrm{i}, \mathrm{j}, \mathrm{k}] \subseteq \mathbb{H}$ vektorových kvaterniónov s obvyklým komutátorom $[p, q] = pq - qp$ a spĺňajú komutačné vzťahy

$$[\mathrm{i}, \mathrm{j}] = 2\mathrm{k}, \qquad [\mathrm{j}, \mathrm{k}] = 2\mathrm{i}, \qquad [\mathrm{k}, \mathrm{i}] = 2\mathrm{j}.$$

Keďže zobrazenie $\boldsymbol{U}\colon \mathbb{H} \to \mathbb{C}^{2\times 2}$, definované tesne pred tvrdením 30.4.2, je injektívny homomorfizmus asociatívnych algebier, je zrejmé, že aj antihermitovské matice s nulovou stopou

$$\boldsymbol{U}(\mathrm{i}) = \mathrm{i}\sigma_3 = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}, \quad \boldsymbol{U}(\mathrm{j}) = \mathrm{i}\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \boldsymbol{U}(\mathrm{k}) = \mathrm{i}\sigma_1 = \begin{pmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix},$$

tvoriace bázu Lieovej algebry su(2) (kde σ_1, σ_2, σ_3 sú Pauliho matice z paragrafu 30.8), vyhovujú obdobným komutačným vzťahom

$$[\boldsymbol{U}(\mathrm{i}), \boldsymbol{U}(\mathrm{j})] = 2\boldsymbol{U}(\mathrm{k}), \qquad [\boldsymbol{U}(\mathrm{j}), \boldsymbol{U}(\mathrm{k})] = 2\boldsymbol{U}(\mathrm{i}), \qquad [\boldsymbol{U}(\mathrm{k}), \boldsymbol{U}(\mathrm{i})] = 2\boldsymbol{U}(\mathrm{j}).$$

Zúženie zobrazenia $\boldsymbol{U}$ na podpriestor vektorových kvaterniónov $\mathbb{R}^3 \subseteq \mathbb{H}$ je tak izomorfizmom Lieových algebier $\big(\mathbb{R}^3, [\ ,\ ]\big) \cong \big(\mathrm{su}(2), [\ ,\ ]\big)$.

Nakrývajúci homomorfizmus $\mathrm{SU}(2) \to \mathrm{SO}(3)$ zas dáva tušiť, že aj Lieove algebry so(3) a su(2) sú izomorfné. Priradenie $\boldsymbol{E}_1 \mapsto \boldsymbol{U}(\mathrm{i})$, $\boldsymbol{E}_2 \mapsto \boldsymbol{U}(\mathrm{j})$, $\boldsymbol{E}_3 \mapsto \boldsymbol{U}(\mathrm{k})$, ktoré nás asi napadne ako prvé, však nezachováva komutačné vzťahy. Dokážete ho upraviť tak, aby dávalo izomorfizmus $\mathrm{so}(3) \cong \mathrm{su}(2)$?

Splňme ešte sľub, ktorý sme dali na konci predchádzajúceho paragrafu.

31.4.7. Veta. *Pre každé $n \geq 1$ maticové grupy $\mathrm{U}(n)$, $\mathrm{SU}(n)$, $\mathrm{SO}(n)$, ako aj vlastná ortochrónna Lorentzova grupa $\Lambda_+^{\uparrow}(1)$ sú exponenciálne.*

Dôkaz. Exponenciálnosť grúp $\mathrm{U}(n)$ a $\mathrm{SU}(n)$ je vlastne len iným vyjadrením druhej časti vety 23.4.9 v spojení s vetou 31.4.4. Dôkaz exponenciálnosti grúp $\mathrm{SO}(n)$ a $\Lambda_+^{\uparrow}(1)$ prenechávame ako cvičenie čitateľovi.

31.5 Jednoparametrické podgrupy pseudoortogonálnej grupy

Stojí za trochu námahy domyslieť do podrobností, ako vyzerajú matice z Lieových algebier pseudoortogonálnych a pseudounitárnych grúp. Algebra $\mathrm{o}(k,l) = \mathrm{so}(k,l) = \mathrm{so}^+(k,l)$ je tvorená maticami blokového tvaru

$$\begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{B} \\ \boldsymbol{B}^\mathsf{T} & \boldsymbol{A}_1 \end{pmatrix},$$

kde $\boldsymbol{A}_0 \in \mathbb{R}^{k\times k}$, $\boldsymbol{A}_1 \in \mathbb{R}^{l\times l}$ sú antisymetrické matice a $\boldsymbol{B} \in \mathbb{R}^{k\times l}$. Podobne, algebra $\mathrm{u}(k,l)$ je tvorená maticami v blokovom tvare

$$\begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{B} \\ \boldsymbol{B}^* & \boldsymbol{A}_1 \end{pmatrix},$$

kde $\boldsymbol{A}_0 \in \mathbb{C}^{k\times k}$, $\boldsymbol{A}_1 \in \mathbb{C}^{l\times l}$ sú antihermitovské matice a $\boldsymbol{B} \in \mathbb{C}^{k\times l}$. Teda v oboch prípadoch príslušné matice obsahujú diagonálne umiestnenú antisymetrickú resp. antihermitovskú časť $\mathrm{diag}(\boldsymbol{A}_1, \boldsymbol{A}_2)$, priečne doplnenú symetrickou časťou zloženou z blokov $\boldsymbol{B}$, $\boldsymbol{B}^\mathsf{T}$ resp. hermitovskou časťou zloženou z blokov $\boldsymbol{B}$, $\boldsymbol{B}^*$. Matice z algebry $\mathrm{o}(k,l) = \mathrm{so}(k,l)$ majú nulovú diagonálu, a tým skôr stopu, rovnako ako matice z algebry $\mathrm{o}(k+l) = \mathrm{so}(k+l)$. Matica $\boldsymbol{A} \in \mathrm{u}(k,l)$ patrí do algebry $\mathrm{su}(k,l)$ práve vtedy, keď má nulovú stopu, t. j. práve vtedy, keď $\mathrm{tr}\,\boldsymbol{A}_0 = -\mathrm{tr}\,\boldsymbol{A}_1$, čo je opäť analogické vzťahu medzi algebrami $\mathrm{u}(k+l)$ a $\mathrm{su}(k+l)$.

Znalosť Lieových algebier $\mathrm{o}(k,l)$, $\mathrm{u}(k,l)$ a $\mathrm{su}(k,l)$ umožňuje nájsť ich viacmenej kanonické bázy a popísať jednoparametrické podgrupy k nim prislúchajúcich maticových grúp, ktorých infinitezimálne generátory tvoria prvky týchto báz. To nám dáva možnosť aspoň trochu nahliadnuť do štruktúry týchto grúp. Bližšie si to predvedieme na príklade pseudoortogonálnej grupy $\mathrm{O}(k,l)$.

31.5.1. Príklad. Lieova algebra $\mathrm{o}(k,l)$ má bázu pozostávajúcu z generátorov dvoch typov: Jednak je to $\binom{k}{2} + \binom{l}{2}$ antisymetrických matíc $\boldsymbol{E}_{sr} - \boldsymbol{E}_{rs}$, kde $1 \le r < s \le k$ alebo $k+1 \le r < s \le k+l$, jednak kl symetrických matíc $\boldsymbol{E}_{rs} + \boldsymbol{E}_{sr}$, kde $1 \le k \le r$, $k+1 \le s \le k+l$.

Matica $\boldsymbol{E}_{sr} - \boldsymbol{E}_{rs}$ má na mieste (r,s) prvok -1, na mieste (s,r) prvok 1 a všade inde nuly. Kľúčom k ňou generovanej jednoparametrickej grupe je jednoparametrická grupa

$$\mathrm{e}^{-\mathrm{i}t\sigma_2} = \exp\begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} = \boldsymbol{R}_t.$$

Jednoparametrická grupa $\exp\bigl[(\boldsymbol{E}_{sr} - \boldsymbol{E}_{rs})t\bigr]$ má preto na miestach (r,r) a (s,s) hodnotu $\cos t$, na miestach (r,s) resp. (s,r) hodnoty $-\sin t$ resp. $\sin t$, na zvyšných miestach diagonály 1 a všade inde 0. Je to teda rotácia o uhol t v dvojrozmernom kladne alebo záporne definitnom invariantnom podpriestore $[\boldsymbol{e}_r, \boldsymbol{e}_s] \subseteq \mathbb{R}^{(k,l)}$.

Matica $\boldsymbol{E}_{rs} + \boldsymbol{E}_{sr}$ má na miestach (r, s) a (s, r) prvok 1 a všade inde nuly. V najjednoduchšom prípade, keď $k = l = 1$, generuje jednoparametrickú grupu

$$\mathrm{e}^{t\sigma_1} = \exp\begin{pmatrix} 0 & t \\ t & 0 \end{pmatrix} = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix} = \boldsymbol{Rh}_t.$$

Jednoparametrická grupa $\exp\bigl[(\boldsymbol{E}_{rs} + \boldsymbol{E}_{sr})t\bigr]$ má preto na miestach (r, r) a (s, s) hodnotu $\cosh t$, na miestach (r, s) a (s, r) hodnotu $\sinh t$, na zvyšných miestach diagonály 1 a všade inde 0. Ide teda o hyperbolickú rotáciu, čiže *boost*, o hyperbolický uhol t v dvojrozmernom indefinitnom invariantnom podpriestore $[\boldsymbol{e}_r, \boldsymbol{e}_s] \subseteq \mathbb{R}^{(k,l)}$.

Všetky matice z vlastnej špeciálnej pseudoortogonálnej grupy $\mathrm{SO}^+(k, l)$ možno teda vyjadriť v tvare súčinu konečného počtu matíc uvedených dvoch typov.

Cvičenia

31.1. (a) Vyjadrite podmienku alternovania resp. antisymetrie komutátora v Lieovej algebre v reči štruktúrnych konštánt.

(b) Vyjadrite pomocou štruktúrnych konštánt Jacobiho identitu.

31.2. (a) Dokážte, že Lieova algebra nad poľom charakteristiky $\neq 2$ je komutatívna práve vtedy, keď je komutatívna v pôvodnom zmysle tohto slova, t. j. platí v nej $[\boldsymbol{a}, \boldsymbol{b}] = [\boldsymbol{b}, \boldsymbol{a}]$.

(b) Ako je to s Lieovými algebrami nad poľom charakteristiky 2? Je ich veľa, vzniknú napr. z každej asociatívnej algebry definíciou komutátora $[x, y] = xy - yx = xy + yx$ a sú nevyhnutne komutatívne v obyčajnom zmysle, ale nie ako Lieove algebry. Uveďte konkrétne príklady.

31.3. Nájdite (až na izomorfizmus) všetky Lieove algebry nad poľom $\mathbb{R}$ so základným vektorovým priestorom $\mathbb{R}$ resp. $\mathbb{R}^2$. (V oboch prípadoch je len jedna a to komutatívna, t. j. $[\boldsymbol{a}, \boldsymbol{b}] = \boldsymbol{0}$ pre všetky $\boldsymbol{a}, \boldsymbol{b}$.)

31.4. *Heisenbergova algebra* je Lieova algebra so základným priestorom $\mathbb{R}^3$, pre bázické vektory platí $[\boldsymbol{e}_1, \boldsymbol{e}_2] = \boldsymbol{e}_3$, $[\boldsymbol{e}_1, \boldsymbol{e}_3] = [\boldsymbol{e}_2, \boldsymbol{e}_3] = \boldsymbol{0}$. Dá sa realizovať striktne hornými trojuholníkovými maticami $\boldsymbol{A} \in \mathbb{R}^{3\times 3}$ (t. j. $a_{ij} = 0$ pre $i \geq j$) – nájdite príslušný izomorfizmus.

31.5. (a) Pre každý z vektorov $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3 \in \mathbb{R}^3$, napíšte maticu $\boldsymbol{E}_i \in \mathbb{R}^{3\times 3}$ lineárneho operátora $\boldsymbol{x} \mapsto \boldsymbol{e}_i \times \boldsymbol{x}$ na priestore $\mathbb{R}^3$ vzhľadom na štandardnú bázu $\boldsymbol{\varepsilon} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$.

(b) Predpis $\boldsymbol{e}_i \mapsto \boldsymbol{E}_i$ určuje izomorfizmus $\mathbb{R}$-algebier $\bigl(\mathbb{R}^3, \times\bigr) \cong \bigl(\mathrm{so}(3), [\ ,\ ]\bigr)$. Dokážte a porovnajte s príkladom 31.4.6. Odvoďte z toho Jacobiho identitu pre vektorový súčin.

(c) Nájdite obraz všeobecného vektora $(a, b, c)^{\mathsf{T}} \in \mathbb{R}^3$ v uvedenom izomorfizme.

31.6. Skúste riešiť rovnakú úlohu ako v cvičení 31.3 pre vektorový priestor $\mathbb{R}^3$. (Malo by ich byť 9; hovorí sa tomu *Bianchiho klasifikácia*.)

31.7. (a) Nájdite explicitný tvar pre komutátor $[p,q] = pq - qp$ kvaterniónov $p,q \in \mathbb{H}$.

(b) Dokážte, že predpisom $\boldsymbol{e}_1 \mapsto \mathrm{i}/2$, $\boldsymbol{e}_2 \mapsto \mathrm{j}/2$ $\boldsymbol{e}_3 \mapsto \mathrm{k}/2$ je definovaný injektívny homomorfizmus Lieových algebier $\big(\mathbb{R}^3, \times\big) \to \big(\mathbb{H}, [\ ,\ ]\big)$.

(c) Nájdite explicitný tvar čo najprirodzenejšieho izomorfizmu Lieových algebier $\mathrm{so}(3) \cong \mathrm{su}(2)$.

31.8. Dokážeme, že Lieova algebra $(\mathbb{R}^3, \times)$ nie je izomorfná so žiadnou Lieovou algebrou, ktorá vznikne ako $\mathcal{A}^L$ z nejakej asociatívnej algebry $\mathcal{A}$. Predpokladajme, že $(\boldsymbol{u},\boldsymbol{v}) \mapsto \boldsymbol{uv}$ je asociatívna bilineárna operácia na $\mathbb{R}^3$ taká, že $\big(\mathbb{R}^3, \times\big) = \big(\mathbb{R}^3, \cdot\big)^L$, t. j. $\boldsymbol{u} \times \boldsymbol{v} = \boldsymbol{uv} - \boldsymbol{vu}$ pre všetky $\boldsymbol{u},\boldsymbol{v} \in \mathbb{R}^3$.

(a) Dokážte, že pre každé $\boldsymbol{u} \in \mathbb{R}^3$ platí $\boldsymbol{u}^2 \times \boldsymbol{u} = \boldsymbol{0}$. Odvoďte z toho, že vektory $\boldsymbol{u}$ a $\boldsymbol{u}^2 = \boldsymbol{uu}$ sú lineárne závislé.

(b) Pre $i = 1,2,3$ označte a_i reálne čísla také, že $\boldsymbol{e}_i^2 = a_i\boldsymbol{e}_i$. Dokážte, že pre $i \neq j$ platí $a_i\boldsymbol{e}_j = \boldsymbol{e}_i\boldsymbol{e}_j + \boldsymbol{e}_j\boldsymbol{e}_i = a_j\boldsymbol{e}_i$. Odvoďte z toho, že pre ľubovoľné $\boldsymbol{u},\boldsymbol{v} \in \mathbb{R}^3$ potom nutne platí $\boldsymbol{vu} = -\boldsymbol{uv}$, t. j. bilineárna operácia $\boldsymbol{uv}$ je antisymetrická.

(c) Dokážte, že potom $\boldsymbol{u} \times \boldsymbol{v} = 2\boldsymbol{uv}$, v dôsledku čoho by aj vektorový súčin musel byť asociatívny. Nájdite príklad dosvedčujúci, že nie je.

(d) Nájdite asociatívnu algebru $\mathcal{A} = \big(\mathbb{R}^4, \cdot\big)$, pre ktorú existuje injektívny homomorfizmus Lieových algebier $(\mathbb{R}^3, \times) \to \mathcal{A}^L$.

31.9. Nech G je maticová grupa. Nájdite chybu v nasledujúcom „dôkaze" uzavretosti Lieovej algebry $\mathcal{L}(G)$ vzhľadom na súčet matíc: Nech $\boldsymbol{A},\boldsymbol{B} \in \mathcal{L}(G)$. Potom podľa vety 31.3.1 platí $\mathrm{e}^{\boldsymbol{A}t}, \mathrm{e}^{\boldsymbol{B}t} \in G$ pre všetky $t \in \mathbb{R}$. Preto tiež $\mathrm{e}^{(\boldsymbol{A}+\boldsymbol{B})t} = \mathrm{e}^{\boldsymbol{A}t} \cdot \mathrm{e}^{\boldsymbol{B}t} \in G$ pre všetky $t \in \mathbb{R}$, a z tej istej vety vyplýva $\boldsymbol{A} + \boldsymbol{B} \in \mathcal{L}(G)$.

31.10. (a) Na základe kanonického izomorfizmu z vety 29.2.2 stotožníme afinné rozšírenie G^{af} maticovej grupy $G \subseteq \mathrm{GL}(n,K)$ nad poľom $K = \mathbb{R}$ resp. $K = \mathbb{C}$ s grupou matíc tvaru $\left(\begin{smallmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \boldsymbol{0} & 1 \end{smallmatrix}\right)$, kde $\boldsymbol{A} \in G$, $\boldsymbol{u} \in K^n$. Dokážte, že základný priestor Lieovej algebry $\mathcal{L}\big(G^{\mathrm{af}}\big) \subseteq K^{(n+1)\times(n+1)}$ pozostáva zo všetkých matíc tvaru $\left(\begin{smallmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \boldsymbol{0} & 0 \end{smallmatrix}\right)$, kde $\boldsymbol{A} \in \mathcal{L}(G)$, $\boldsymbol{u} \in K^n$.

(b) Nech K je ľubovoľné pole. Nájdite explicitný tvar komutátora matíc $\left(\begin{smallmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \boldsymbol{0} & 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} \boldsymbol{B} & \boldsymbol{v} \\ \boldsymbol{0} & 0 \end{smallmatrix}\right)$, kde $\boldsymbol{A},\boldsymbol{B} \in K^{n\times n}$, $\boldsymbol{u},\boldsymbol{v} \in K^n$. Dokážte, že všetky matice tvaru $\left(\begin{smallmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \boldsymbol{0} & 0 \end{smallmatrix}\right)$, kde $\boldsymbol{A} \in K^{n\times n}$, $\boldsymbol{u} \in K^n$, tvoria podalgebru Lieovej algebry $K^{(n+1)\times(n+1)}$.

(c) Nech $\mathcal{A} \subseteq K^{n\times n}$ je Lieova algebra nad nejakým podpoľom F poľa K (s komutátorom $[\boldsymbol{A},\boldsymbol{B}] = \boldsymbol{AB} - \boldsymbol{BA}$). *Afinným rozšírením Lieovej algebry* $\mathcal{A}$ nazývame algebru $\mathcal{A}^{\mathrm{af}} = \mathcal{A} \times K^n$ so štruktúrou vektorového priestoru nad F definovanou po zložkách a s komutátorom $\big[(\boldsymbol{A},\boldsymbol{u}),(\boldsymbol{B},\boldsymbol{v})\big] = \big([\boldsymbol{A},\boldsymbol{B}], \boldsymbol{u} + \boldsymbol{Av} - \boldsymbol{v} - \boldsymbol{Bu}\big)$. Dokážte, že $\mathcal{A}^{\mathrm{af}}$ je Lieova algebra nad poľom F izomorfná s podalgebrou Lieovej algebry $K^{(n+1)\times(n+1)}$ tvorenou všetkými maticami tvaru $\left(\begin{smallmatrix} \boldsymbol{A} & \boldsymbol{u} \\ \boldsymbol{0} & 0 \end{smallmatrix}\right)$, kde $\boldsymbol{A} \in \mathcal{A}$, $\boldsymbol{u} \in K^n$.

(d) Za podmienok ako v časti (a) popíšte prirodzený izomorfizmus Lieových algebier $\mathcal{L}\big(G^{\mathrm{af}}\big) \cong \mathcal{L}(G)^{\mathrm{af}}$. Taktiež overte, že pri reprezentácii afinného rošírenia G^{af} polopriamym súčinom $G \ltimes K^n$ platí priamo rovnosť $\mathcal{L}\big(G^{\mathrm{af}}\big) = \mathcal{L}(G)^{\mathrm{af}}$.

(e) Vypočítajte dimenzie afinných rozšírení Lieových algebier z viet 31.4.1, 31.4.2 a 31.4.4 a maticových grúp z vety 31.4.5.

31.11. Vypočítajte deriváciu funkcie h z dôkazu vety 31.4.3 a dokážte, že h vyhovuje počiatočnej úlohe $h' = \boldsymbol{A}^{\mathsf{T}} \cdot h + h \cdot \boldsymbol{A}$, $h(0) = \boldsymbol{Z}$.

31.12. Nech $G, H \subseteq \mathrm{GL}(n, K)$, kde K je niektoré z polí $\mathbb{R}$, $\mathbb{C}$, sú maticové grupy. Dokážte rovnosť $\mathcal{L}(G \cap H) = \mathcal{L}(G) \cap \mathcal{L}(H)$. Zdôvodnite tým tvar Lieových algebier $\mathrm{su}(n)$ a $\mathrm{su}(k, l)$ z vety 31.4.4.

31.13. Vysvetlite prečo dimenzie Lieových algebier $\mathrm{o}(k, l) = \mathrm{so}(k, l)$, resp. $\mathrm{u}(k, l)$, $\mathrm{su}(k, l)$ nezávisia priamo na k, l ale iba na ich súčte $n = k + l$.
(*Návod*: $\boldsymbol{A} \mapsto \boldsymbol{A} \cdot \boldsymbol{D}_{kl}$ je izomorfizmus vektorových priestorov (nie nutne Lieových algebier) nad $\mathbb{R}$ $\mathrm{so}(n) \to \mathrm{so}(k, l)$ resp. $\mathrm{u}(n) \to \mathrm{u}(k, l)$ resp. $\mathrm{su}(n) \to \mathrm{su}(k, l)$.)

31.14. Dokážte vetu 31.4.5 výpočtom dimenzií Lieových algebier príslušných maticových grúp. (S použitím predchádzajúceho cvičenia minimalizujte počet nevyhnutných výpočtov.)

31.15. Dokážte, že z komutatívnosti Lieovej algebry $\mathfrak{g}$ vyplýva komutatívnosť súvislej komponenty $\Upsilon(G)$ jej maticovej grupy G.

31.16. (a) Dokážte, že pre maticu $\boldsymbol{E} = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, ktorá tvorí bázu Lieovej algebry $\mathrm{so}(2)$, naozaj platí $\mathrm{e}^{\boldsymbol{E}t} = \boldsymbol{R}_t$ pre ľubovoľné $t \in \mathbb{R}$. Odvoďte z toho, že grupa $\mathrm{SO}(2)$ je exponenciálna.

(b) Pomocou (a) dokážte, že pre matice $\boldsymbol{E}_1, \boldsymbol{E}_2, \boldsymbol{E}_3$ z príkladu 31.4.6, tvoriace bázu Lieovej algebry $\mathrm{so}(3)$, platí $\mathrm{e}^{\boldsymbol{E}_i t} = \boldsymbol{R}_t^{\boldsymbol{e}_i}$ pre $i = 1, 2, 3$ a ľubovoľné $t \in \mathbb{R}$. S využitím tvrdenia 22.1.2, dôsledku 23.5.4 a vety 31.2.1 (b) z toho odvoďte, že aj $\mathrm{SO}(3)$ je exponenciálna grupa.

(c) Podobne ako v (b) dokážte, že aj grupy $\mathrm{SO}(n)$ pre $n > 3$ sú exponenciálne.

31.17. (a) Odvoďte explicitný tvar jednoparametrických podgrúp $\mathrm{e}^{\boldsymbol{U}(\mathrm{i})t} = \begin{pmatrix} \mathrm{e}^{\mathrm{i}t} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}t} \end{pmatrix}$, $\mathrm{e}^{\boldsymbol{U}(\mathrm{j})t} = \left(\begin{smallmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{smallmatrix}\right)$, $\mathrm{e}^{\boldsymbol{U}(\mathrm{k})t} = \left(\begin{smallmatrix} \cos t & \mathrm{i}\sin t \\ \mathrm{i}\sin t & \cos t \end{smallmatrix}\right)$ grupy $\mathrm{SU}(2)$, zodpovedajúcich infinitezimálnym generátorom $\boldsymbol{U}(\mathrm{i}) = \mathrm{i}\sigma_3$, $\boldsymbol{U}(\mathrm{j}) = \mathrm{i}\sigma_2$, $\boldsymbol{U}(k) = \mathrm{i}\sigma_1$, ktoré tvoria bázu jej Lieovej algebry $\mathrm{su}(2)$.

(b) S použitím vety 17.4.3 dokážte, že ľubovoľnú maticu $\boldsymbol{A} = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \in \mathrm{SU}(2)$ možno vyjadriť v tvare

$$\begin{aligned} \boldsymbol{A} &= \exp\left(\frac{\varphi}{2}\,\mathrm{i}\sigma_3\right) \cdot \exp\left(\frac{\chi}{2}\,\mathrm{i}\sigma_1\right) \cdot \exp\left(\frac{\psi}{2}\,\mathrm{i}\sigma_3\right) \\ &= \begin{pmatrix} \cos\frac{\chi}{2} \exp\frac{\mathrm{i}(\varphi+\psi)}{2} & \mathrm{i}\sin\frac{\chi}{2} \exp\frac{\mathrm{i}(\varphi-\psi)}{2} \\ \mathrm{i}\sin\frac{\chi}{2} \exp\frac{\mathrm{i}(\psi-\varphi)}{2} & \cos\frac{\chi}{2} \exp\frac{-\mathrm{i}(\varphi+\psi)}{2} \end{pmatrix}, \end{aligned}$$

kde $0 \le \varphi < 2\pi$, $0 \le \chi < \pi$, $-2\pi < \psi \le 2\pi$ sú parametre, ktoré opäť nazývame *Eulerove uhly*, – porovnajte s paragrafom 23.6.

(c) Dokážte, že obrazom matice $\exp\left(\frac{\alpha}{2}\,\mathrm{i}\sigma_p\right)$, kde $p = 1, 2, 3$, v nakrývajúcom homomorfizme $\mathrm{SU}(2) \to \mathrm{SO}(3)$ je matica otočenia $\boldsymbol{R}_\alpha^{\boldsymbol{e}_p}$.

31.18. Pre dané $\vartheta \in \mathbb{R}$ nájdite maticu $\boldsymbol{A} \in \mathrm{so}(1, 1)$ takú, že $\mathrm{e}^{\boldsymbol{A}} = \boldsymbol{Rh}_\vartheta$. Odvoďte z toho, že $\Lambda_+^{\uparrow}(1)$ je exponenciálna grupa izomorfná s $(\mathbb{R}, +)$.

31.19. Dokážeme, že hoci grupa $\mathrm{SL}(2, \mathbb{R})$ je súvislá (veta 29.7.6 (c)), exponenciálne zobrazenie $\mathrm{sl}(2, \mathbb{R}) \to \mathrm{SL}(2, \mathbb{R})$ nie je surjektívne. Uvažujme ľubovoľnú maticu $\boldsymbol{A} = \left(\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right) \in \mathrm{sl}(2, \mathbb{R})$ a označme $\delta = -\det \boldsymbol{A} = a^2 + bc$, $\rho = \sqrt{|\delta|}$. Overte nasledujúce rovnosti:

(a) $\boldsymbol{A}^2 = \delta \boldsymbol{I}_2$;

(b) $e^{\boldsymbol{A}t} = \begin{pmatrix} \cosh(\rho t)+(a/\rho)\sinh(\rho t) & (b/\rho)\sinh(\rho t) \\ (c/\rho)\sinh(\rho t) & \cosh(\rho t)-(a/\rho)\sinh(\rho t) \end{pmatrix}$, ak $\delta > 0$;

(c) $e^{\boldsymbol{A}t} = \begin{pmatrix} 1+at & bt \\ ct & 1-at \end{pmatrix}$, ak $\delta = 0$;

(d) $e^{\boldsymbol{A}t} = \begin{pmatrix} \cos(\rho t)+(a/\rho)\sin(\rho t) & (b/\rho)\sin(\rho t) \\ (c/\rho)\sin(\rho t) & \cos(\rho t)-(a/\rho)\sin(\rho t) \end{pmatrix}$, ak $\delta < 0$.

(e) V každom z uvedených troch prípadov vypočítajte stopu matice $e^{\boldsymbol{A}t}$ a ohraničte hodnoty, aké môže nadobúdať. Presvedčte sa, že pre ľubovoľnú maticu $\boldsymbol{A} \in \mathrm{sl}(2, \mathbb{R})$ platí $\operatorname{tr} e^{\boldsymbol{A}t} \geq -2$.

(f) Nájdite aspoň jednu maticu $\boldsymbol{X} \in \mathrm{SL}(2, \mathbb{R})$ takú, že $\operatorname{tr} \boldsymbol{X} < -2$. Táto matica nemôže mať tvar $\boldsymbol{X} = e^{\boldsymbol{A}t}$ pre žiadnu maticu $\boldsymbol{A} \in \mathrm{sl}(2, \mathbb{R})$. Ukážte, že existuje nekonečne mnoho takých matíc $\boldsymbol{X}$.

Časť V

Multilineárna algebra

32. Úvod do tenzorového počtu

Tenzory sú z matematického hľadiska jednoducho multilineárne formy. Vo fyzike sa tenzory vyskytujú najmä ako hodnoty veličín popisujúcich vlastnosti anizotropných prostredí a dejov, ktoré v nich prebiehajú.

Významnejšie príklady využitia tenzorov presahujú rámec lineárnej algebry, a tým aj nášho kurzu. Z matematických disciplín, v ktorých hrajú tenzory významnú úlohu, treba spomenúť aspoň diferenciálnu geometriu, prostredníctvom ktorej ďalej vstupujú napr. do teoretickej mechaniky či všeobecnej teórie relativity, a reperezentácie grúp, s rozsiahlymi aplikáciámi napr. v kryštalografii či časticovej fyzike. Bez tenzorov sa nezaobídeme ani pri štúdiu elektromagnetizmu či zložitejších systémov v kvantovej mechanike. Vo väčšine týchto aplikácií však nejde o „konštantné“ ale o „premenné“ tenzory, čiže o *tenzorové polia.*

Zoči-voči spomínaným faktom sú ciele záverečných dvoch kapitol tejto knihy pomerne skromné. Mali by čitateľovi sprostredkovať istú „tenzorovú gramotnosť“, ktorá by mu umožnila ďalej študovať oné dôležitejšie, no zložitejšie aplikácie. Naše príklady využitia tenzorov budú preto pomerne jednoduché. K nim patrí aj nasledujúci úvodný motivačný príklad.

32.0 Tenzor napätia

Ako sme spomínali ešte v úvode ku kapitole 1, skaláry (presnejšie prvky poľa $\mathbb{R}$) sú hodnoty fyzikálnych veličín určených jedine svojou veľkosťou (prípadne aj znamienkom). Vektory zodpovedajú veličinám, ktoré sú okrem veľkosti určené navyše smerom a orientáciou. Ukazuje sa však, že popri nich je užitočné zaviesť a študovať aj veličiny, ktoré zachytávajú vzťahy medzi dvoma či dokonca viacerými vektorovými veličinami. Takýmto veličinám hovoríme *tenzory.* Príkladom tenzorovej fyzikálnej veličiny je *tenzor napätia*, ktorý popisuje silové pôsobenie v rôznych smeroch v danom prostredí. Napätie sa po latinsky povie *tensio*, čo prešlo do slovenčiny v podobe odborného termínu *tenzia.* Tenzor napätia, teda doslova *tenzor tenzie*, je tak prototypom všetkých tenzorových veličín.

Predstavme si nejaké spojité prostredie (napr. pevné alebo kvapalné teleso) v trojrozmernom euklidovskom priestore $\mathbb{R}^3$ vystavené silovému pôsobeniu, ktoré sa v ňom šíri. Také pôsobenie môže spočívať v tlaku resp. ťahu v rôznych smeroch, ako aj v torzii. Do nejakého bodu tohto prostredia umiestnime myslenú plôšku $\Delta\boldsymbol{S} = \boldsymbol{n}\Delta S$, orientovanú v smere jednotkového vektora normály $\boldsymbol{n}$, dosť malú na to, aby sme veľkosť sily $\Delta\boldsymbol{F}$, ktorá na ňu pôsobí zo strany protiľahlej vektoru $\boldsymbol{n}$, mohli považovať za priamoúmernú veľkosti ΔS tejto plôšky. Keby aj sila $\Delta\boldsymbol{F}$ mala smer normály $\boldsymbol{n}$, mohli by sme písať $\Delta\boldsymbol{F} = p\,\Delta\boldsymbol{S}$ pre nejaký koeficient úmernosti $p = p(\boldsymbol{n}) \in \mathbb{R}$, závislý od normálového vektora $\boldsymbol{n}$. V prostredí spĺňajúcom *Pascalov zákon*, ako je tomu v *ideálnych kvapalinách*, by koeficient úmernosti p bol navyše konštantný, čiže mal rovnakú hodnotu pre všetky smery normál $\boldsymbol{n}$, zodpovedajúcu *tlaku* v danom bode.

Smery vektorov $\Delta\boldsymbol{S}$ a $\Delta\boldsymbol{F}$ sa však môžu líšiť, napr. v dôsledku torzie. Zovšeobecnením podmienky priamej úmernosti je podmienka *lineárnej závislosti* sily $\Delta\boldsymbol{F}$ od orientovanej plôšky $\Delta\boldsymbol{S}$, čiže

$$\Delta\boldsymbol{F} = T(\Delta\boldsymbol{S}) = T(\boldsymbol{n})\Delta S$$

pre nejaký lineárny operátor $T\colon \mathbb{R}^3 \to \mathbb{R}^3$, ktorému hovoríme *tenzor napätia* v danom bode. V limite pre $\Delta S \to 0$ dostávame vyjadrenie vektora $T(\boldsymbol{n})$ v tvare derivácie

$$T(\boldsymbol{n}) = \frac{\mathrm{d}\boldsymbol{F}}{\mathrm{d}S}.$$

Tenzor napätia môžeme stotožniť s jeho maticou

$$T = \big(T(\boldsymbol{e}_1), T(\boldsymbol{e}_2), T(\boldsymbol{e}_3)\big) = \big(T^i_j\big) \in \mathbb{R}^{3\times 3}$$

v kanonickej báze $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ priestoru $\mathbb{R}^3$. Stĺpec $T(\boldsymbol{e}_j)$ matice T predstavuje silu v prepočte na jednotku plošného obsahu pôsobiacu na infinitezimálnu orientovanú plôšku s normálou $\boldsymbol{n} = \boldsymbol{e}_j$ a $T^i_j\boldsymbol{e}_i$ je kolmý priemet vektora $T(\boldsymbol{e}_j)$ do osi $x_i = [\boldsymbol{e}_i]$. Inak povedané,

$$T(\boldsymbol{e}_j) = T^1_j\boldsymbol{e}_1 + T^2_j\boldsymbol{e}_2 + T^3_j\boldsymbol{e}_3$$

je rozklad vektora $T(\boldsymbol{e}_j)$ do zložiek v smeroch jednotlivých súradných osí. V prípade platnosti Pascalovho zákona má tenzor napätia diagonálnu maticu $T = p\boldsymbol{I}_3$.

Dospievame teda k záveru, že niektorým fyzikálnym veličinám môžu zodpovedať aj zložitejšie matematické objekty než len skaláry alebo vektory. V našom zatiaľ jedinom príklade to bolo lineárne zobrazenie, no vo všeobecnosti to môžu byť multilineárne zobrazenia. Takýmto objektom budeme hovoriť *tenzory*.

Skôr než pristúpime k samotnému matematickému výkladu tenzorov, treba ešte upozorniť na istú nejednoznačnosť v terminológii, ktorá by občas mohla viesť k nedorozumeniam. V našom príklade sme každému bodu nejakej množiny $M \subseteq \mathbb{R}^3$ predstavujúcej skúmané teleso priradili tenzor napätia v tomto bode. Napätie v danom prostredí je tak popísané (typicky spojitou či dokonca spojito diferencovateľnou) funkciou $T\colon M \to \mathbb{R}^{3\times 3}$, ktorá každému bodu $\boldsymbol{x} \in M$ priradí tenzor (maticu) $T(\boldsymbol{x}) = \big(T^i_j(\boldsymbol{x})\big)$ zodpovedajúci napätiu v bode $\boldsymbol{x}$. Ak chceme zachytiť aj vývoj napätia v čase, musíme uvažovať funkciu $T\colon M \times I \to \mathbb{R}^{3\times 3}$, kde $I \subseteq \mathbb{R}$ predstavuje nejaký časový interval.

Často – či už v diferenciálnej geometrii, fyzike alebo v technických aplikáciách – sa pod tenzorom rozumie práve takáto funkcia. Metodicky správnejšie je však nazývať ju *tenzorovým poľom* (v prvom prípade *statickým*, v druhom *časovo premenným*) a názov *tenzor* používať len pre jej hodnoty $T(\boldsymbol{x})$ resp. $T(\boldsymbol{x}, t)$. V našom výklade sa zameriame iba na jednotlivé tenzory, a nie na tenzorové polia. Vo viacerých príkladoch však bude aspoň v pozadí prítomná možnosť priradiť tenzor príslušného typu všetkým bodom nejakej množiny $M \subseteq \mathbb{R}^n$, teda chápať „naše“ tenzory ako hodnoty istého tenzorového poľa.

32.1 Dualita

V tejto i v nasledujúcej kapitole označuje K ľubovoľné pevne zvolené pole a pod vektorovým priestorom rozumieme vždy vektorový priestor nad poľom K. Pripomeňme z paragrafu 6.5, že všetky lineárne zobrazenia z vektorového priestoru V do vektorového priestoru W tvoria vektorový priestor, ktorý značíme $\mathcal{L}(V, W)$. Vektorový priestor $\mathcal{L}(V, K)$ značíme V^* a nazývame *duálnym priestorom* alebo len krátko *duálom* vektorového priestoru V. Jeho prvky nazývame *lineárne funkcionály* alebo tiež *lineárne formy* na V.

Konečnorozmerný vektorový priestor V je izomorfný so svojim duálom V^* a taktiež so svojim *druhým duálom* V^{**}. Konkrétny izomorfizmus $V \cong V^*$ závisí na voľbe bázy v priestore V. Na druhej strane existuje *kanonický izomorfizmus* $V \cong V^{**}$, ktorý nám umožňuje stotožniť každý vektor $\boldsymbol{x} \in V$ s lineárnym funkcionálom $\widehat{\boldsymbol{x}}\colon V^* \to K$, kde $\widehat{\boldsymbol{x}}(\eta) = \eta(\boldsymbol{x})$ pre $\eta \in V^*$. V dôsledku toho nebudeme ďalej rozlišovať vektor $\boldsymbol{x} \in V$ a funkcionál $\widehat{\boldsymbol{x}} \in V^{**}$ ani priestor V a jeho druhý duál V^{**}, čiže pod duálom priestoru V^* budeme rozumieť pôvodný priestor V. Zároveň zavádzame symetrické označenie

$$\langle \eta, \boldsymbol{x} \rangle = \eta(\boldsymbol{x}) = \boldsymbol{x}(\eta) = \langle \boldsymbol{x}, \eta \rangle.$$

Takto zavedené lomené zátvorky sú tiež bilineárnou formou $\langle \cdot\,, \cdot \rangle\colon V^* \times V \to K$; hovoríme jej *párovanie* alebo *dualita* medzi V^* a V.

Pozrime sa teraz bližšie na nekanonické izomorfizmy $V \cong V^*$. Každá báza $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ priestoru V určuje lineárny izorfizmus

$$\boldsymbol{x} \mapsto (\boldsymbol{x})_{\boldsymbol{\beta}}\colon V \to \mathbb{R}^n,$$

kde $n = \dim V$. Ak položíme $(\boldsymbol{x})_{\boldsymbol{\beta}} = (x^1, \dots, x^n)^T$ (dôvody náhleho presunu indexov zdola nahor začneme odhaľovať v závere paragrafu), tak každé z priradení $\boldsymbol{v}^i(\boldsymbol{x}) = x^i$ určuje lineárny funkcionál $\boldsymbol{v}^i\colon V \to K$. Pri takomto označení teda môžeme písať

$$(\boldsymbol{x})_{\boldsymbol{\beta}} = \big(\boldsymbol{v}^1(\boldsymbol{x}), \ldots, \boldsymbol{v}^n(\boldsymbol{x})\big)^T,$$
$$\boldsymbol{x} = x^1\boldsymbol{v}_1 + \ldots + x^n\boldsymbol{v}_n = \boldsymbol{v}^1(\boldsymbol{x})\boldsymbol{v}_1 + \ldots + \boldsymbol{v}^n(\boldsymbol{x})\boldsymbol{v}_n.$$

Každý z lineárnych funkcionálov $\boldsymbol{v}^i \colon V \to K$ je jednoznačne určený svojimi hodnotami

$$\langle \boldsymbol{v}^i, \boldsymbol{v}_j \rangle = \delta^i_j$$

na vektoroch $\boldsymbol{v}_j$ bázy $\boldsymbol{\beta}$. Treba si však uvedomiť, že – napriek označeniu – lineárny funkcionál $\boldsymbol{v}^i$ závisí od celej bázy $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$, a nielen od vektora $\boldsymbol{v}_i$.

Lineárne formy $\boldsymbol{v}^1, \ldots, \boldsymbol{v}^n$ tvoria bázu duálneho priestoru V^*; označujeme ju

$$\boldsymbol{\beta}^* = (\boldsymbol{v}^1, \ldots, \boldsymbol{v}^n)$$

a nazývame *duálna báza* k báze $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$. Priradenie $\boldsymbol{v}_j \mapsto \boldsymbol{v}^j$ možno jednoznačne rozšíriť do lineárneho izomorfizmu $V \to V^*$, pri ktorom sa vektor $\boldsymbol{x} = \sum_j x^j \boldsymbol{v}_j$ zobrazí na funkcionál $\sum_j x^j \boldsymbol{v}^j$.

Pre lineárny funkcionál $\eta \in V^*$ a vektor $\boldsymbol{x} = \sum_i x^i \boldsymbol{v}_i \in V$ zrejme platí

$$\eta(\boldsymbol{x}) = \sum_i x^i \eta(\boldsymbol{v}_i) = \sum_i \eta(\boldsymbol{v}_i)\boldsymbol{v}^i(\boldsymbol{x}),$$

čiže $\eta = \sum_i \eta(\boldsymbol{v}_i)\boldsymbol{v}^i$. To znamená, že súradnice funkcionálu η v báze $\boldsymbol{\beta}^*$ tvorí stĺpec

$$(\eta)_{\boldsymbol{\beta}^*} = \big(\eta(\boldsymbol{v}_1), \ldots, \eta(\boldsymbol{v}_n)\big)^T = (\eta)^T_{1,\boldsymbol{\beta}},$$

t. j. transponovaná matica k matici $(\eta)_{1,\boldsymbol{\beta}}$ lineárneho zobrazenia $\eta \colon V \to K$ vzhľadom na bázy $\boldsymbol{\beta}$ vo V a 1 v K.

Uvažujme dve bázy $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$, $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ priestoru V. Označme P^i_j prvok na mieste (i, j) matice prechodu $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ z bázy $\boldsymbol{\beta}$ do bázy $\boldsymbol{\alpha}$. S použitím práve zavedeného označenia môžeme písať

$$P^i_j = \langle \boldsymbol{u}^i, \boldsymbol{v}_j \rangle.$$

Súradnice $(\boldsymbol{x})_{\boldsymbol{\alpha}} = \big(\boldsymbol{u}^1(\boldsymbol{x}), \ldots, \boldsymbol{u}^n(\boldsymbol{x})\big)^T$, $(\boldsymbol{x})_{\boldsymbol{\beta}} = \big(\boldsymbol{v}^1(\boldsymbol{x}), \ldots, \boldsymbol{v}^n(\boldsymbol{x})\big)^T$ vektora $\boldsymbol{x} \in V$ v bázach $\boldsymbol{\alpha}$ resp. $\boldsymbol{\beta}$ sa transformujú podľa vzťahu $(\boldsymbol{x})_{\boldsymbol{\alpha}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot (\boldsymbol{x})_{\boldsymbol{\beta}}$ (pozri paragraf 7.5), čo rozmenené na drobné znamená

$$\boldsymbol{u}^i(\boldsymbol{x}) = \sum_j P^i_j \boldsymbol{v}^j(\boldsymbol{x}),$$

čiže $\boldsymbol{u}^i = \sum_j P^i_j \boldsymbol{v}^j$ pre $i \leq n$, a v súhrne $\boldsymbol{\alpha}^* = \boldsymbol{\beta}^* \cdot \boldsymbol{P}^T_{\boldsymbol{\alpha},\boldsymbol{\beta}}$. Pre maticu prechodu medzi duálnymi bázami $\boldsymbol{\alpha}^*$, $\boldsymbol{\beta}^*$ z toho vyplýva

$$\boldsymbol{P}_{\boldsymbol{\beta}^*,\boldsymbol{\alpha}^*} = \boldsymbol{P}^T_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \big(\boldsymbol{P}^{-1}_{\boldsymbol{\beta},\boldsymbol{\alpha}}\big)^T.$$

Súradnice lineárnej formy $\eta \in V^*$ vzhľadom na duálne bázy $\boldsymbol{\alpha}^* = (\boldsymbol{u}^1, \dots, \boldsymbol{u}^n)$, $\boldsymbol{\beta}^* = (\boldsymbol{v}^1, \dots, \boldsymbol{v}_n)$ sa preto budú transformovať inverzne voči transformácií súradníc $(\boldsymbol{x})_{\boldsymbol{\alpha}}$, $(\boldsymbol{x})_{\boldsymbol{\beta}}$:

$$(\eta)_{\boldsymbol{\beta}^*} = \boldsymbol{P}_{\boldsymbol{\beta}^*,\boldsymbol{\alpha}^*} \cdot (\eta)_{\boldsymbol{\alpha}^*} = \boldsymbol{P}^T_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot (\eta)_{\boldsymbol{\alpha}^*},$$

teda po rozpise do jednotlivých zložiek

$$\boldsymbol{v}_i(\eta) = \sum_j P_i^j \boldsymbol{u}_j(\eta),$$

t. j. $\boldsymbol{v}_i = \sum_j P_i^j \boldsymbol{u}_j$, čo je v zhode s transformačným vzťahom $\boldsymbol{\beta} = \boldsymbol{\alpha} \cdot \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ ešte z paragrafu 7.5.

Znovu si však uvedomme, že zatiaľ čo matica $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \left(P_j^i\right)$ sprostredkúva prechod od súradníc $(\boldsymbol{x})_{\boldsymbol{\beta}}$ k súradniciam $(\boldsymbol{x})_{\boldsymbol{\alpha}}$, k nej transponovaná matica $\boldsymbol{P}^T_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \left(P_i^j\right) = \boldsymbol{P}_{\boldsymbol{\beta}^*,\boldsymbol{\alpha}^*}$ zabezpečuje prechod v opačnom smere, totiž od súradníc $(\eta)_{\boldsymbol{\alpha}^*}$ k súradniciam $(\eta)_{\boldsymbol{\beta}^*}$. Ak označíme maticu prechodu $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \boldsymbol{P}$, máme $\boldsymbol{\beta} = \boldsymbol{\alpha} \cdot \boldsymbol{P}$, $\boldsymbol{\beta}^* = (\boldsymbol{\alpha} \cdot \boldsymbol{P})^*$, a rozdielny charakter transformačných rovníc pre súradnice vektorov $\boldsymbol{x} \in V$ resp. funkcionálov $\eta \in V^*$ sa prejaví výrazne:

$$(\boldsymbol{x})_{\boldsymbol{\alpha}\cdot\boldsymbol{P}} = \boldsymbol{P}^{-1} \cdot (\boldsymbol{x})_{\boldsymbol{\alpha}} \qquad \text{resp.} \qquad (\eta)_{(\boldsymbol{\alpha}\cdot\boldsymbol{P})^*} = \boldsymbol{P}^T \cdot (\eta)_{\boldsymbol{\alpha}^*}.$$

Hovoríme, že súradnice vektorov $\boldsymbol{x} \in V$ sa transformujú *kontravariatne* (t. j. inverzne voči transformácii báz), zatiaľ čo súradnice lineárnych foriem $\eta \in V^*$ sa transformujú *kovariantne* (t. j. súhlasne s transformáciou báz). Prvky pôvodného priestoru V preto nazývame *kontravariantné vektory* alebo len krátko *vektory*, kým prvky jeho duálu V^* nazývame *kovariantné vektory*, stručne len *kovektory*. Je to však terminológia „zaujatého pozorovateľa“, ktorý celú vec posudzuje z hľadiska báz v pôvodnom priestore V.

V tejto súvislosti prijmeme nasledujúcu dohodu, ktorú sme vlastne už začali uplatňovať: súradnice vektorov budeme číslovať hornými a súradnice kovektorov dolnými indexmi; naopak, usporiadané n-tice vektorov budeme číslovať dolnými a usporiadané n-tice kovektorov hornými indexmi. Význam tejto dohody vyjde najavo až trochu neskôr v súvislosti s tzv. *Einsteinovou sumačnou konvenciou.*

Dualita umožňuje zaviesť aj akýsi vzťah *ortogonálnosti medzi vektormi a kovektormi*, ktorý do veľkej miery pripomína ortogonálnosť v (pseudo)euklidovských resp. konečnorozmerných unitárnych priestoroch. Hovoríme, že vektor $\boldsymbol{x} \in V$ a kovektor $\eta \in V^*$ sú *ortogonálne*, označenie $\boldsymbol{x} \perp \eta$, prípadne $\eta \perp \boldsymbol{x}$, ak

$$\langle \boldsymbol{x}, \eta \rangle = 0.$$

Pre množiny $X \subseteq V$, $\Omega \subseteq V^*$ definujeme ich *ortokomplementy*

$$X^\perp = \left\{\eta \in V^*;\ (\forall \boldsymbol{x} \in X)(\boldsymbol{x} \perp \eta)\right\}$$

resp.

$$\Omega^{\perp} = \big\{\boldsymbol{x} \in V;\ (\forall\,\eta \in \Omega)(\boldsymbol{x} \perp \eta)\big\}.$$

Pre naše účely postačí, ak zaznamenáme nasledujúce jednoduché tvrdenie, ktorého dôkaz prenechávame čitateľovi.

32.1.1. Tvrdenie. *Nech V je konečnorozmerný vektorový priestor a $\Omega \subseteq V^*$ je ľubovoľná množina. Potom $\Omega^{\perp} \subseteq V$ je lineárny podpriestor a pre jeho dimenziu platí*

$$\dim \Omega^{\perp} = \dim V - \dim[\Omega].$$

Špeciálne Ω generuje V^* práve vtedy, keď $\Omega^{\perp} = \{\mathbf{0}\}$.

32.2 Tenzorový súčin

Jestvuje viacero ekvivalentných spôsobov ako definovať tenzorový súčin vektorových priestorov. Z nich asi najmenej abstraktný a pedagogicky najstráviteľnejší začína veľmi jednoduchou a konkrétnou definíciou tenzorového súčinu funkcií, ktorá je základom konštrukcie tenzorového súčinu vektorových priestorov funkcií. Ako špeciálny prípad dostaneme tenzorový súčin duálnych priestorov k daným vektorovým priestorom. Trik s dualitou, pri ktorom pôvodný konečnorozmerný priestor V stotožníme s jeho druhým duálom V^{**}, prípadne nekonečnorozmerný priestor V iba vnoríme do V^{**}, nám napokon umožní zaviesť tenzorový súčin pre akékoľvek vektorové priestory nad daným poľom. My sa vydáme práve touto cestou, hoci zďaleka nie je najkratšia možná.

Nech X, Y sú množiny a K je pole. *Tenzorovým súčinom funkcií* $f\colon X \to K$, $g\colon Y \to K$ nazývame funkciu $f \otimes g\colon X \times Y \to K$ danú predpisom

$$(f \otimes g)(x, y) = f(x)g(y)$$

pre $x \in X$, $y \in Y$.

Funkcie $h\colon X \times Y \to K$ tvaru $h = f \otimes g$, kde $f\colon X \to K$, $g\colon Y \to K$, nazývame *rozložiteľné*. Uvedomme si, že zďaleka nie každá funkcia $h \in K^{X\times Y}$ je rozložiteľná, t. j. pripúšťa *separáciu premenných* $h(x, y) = f(x)g(y)$ pre nejaké funkcie $f \in K^X$, $g \in K^Y$.

Tenzorovým súčinom lineárnych podpriestorov $U \subseteq K^X$, $V \subseteq K^Y$ nazývame lineárny podpriestor $U \otimes V$ vektorového priestoru $K^{X\times Y}$ generovaný všetkými funkciami $f \otimes g$, kde $f \in U$, $g \in V$. Teda tenzorový súčin $U \otimes V$ je tvorený všetkými súčtami tvaru $(f_1 \otimes g_1) + \ldots + (f_n \otimes g_n)$, kde $f_1, \ldots, f_n \in U$, $g_1, \ldots, g_n \in V$. Tento fakt sa zakladá na pozorovaní, že pre $a \in K$ a $f \in K^X$, $g \in K^Y$ platí

$$a(f \otimes g) = (af) \otimes g = f \otimes (ag).$$

Navyše pre $f, f_1, f_2 \in K^X$, $g, g_1, g_2 \in K^Y$ máme

$$f \otimes (g_1 + g_2) = (f \otimes g_1) + (f \otimes g_2), \qquad (f_1 + f_2) \otimes g = (f_1 \otimes g) + (f_2 \otimes g).$$

Takže môžeme zhrnúť:

32.2.1. Tvrdenie. *Operácia tenzorového súčinu funkcií $(f, g) \mapsto f \otimes g$ je bilineárne zobrazenie $K^X \times K^Y \to K^{X \times Y}$.*

Pre funkcie $f\colon X \to K$, $g\colon Y \to K$, $h\colon Z \to K$ a prvky $x \in X$, $y \in Y$, $z \in Z$ zrejme platí

$$(f \otimes g)(x, y) = f(x)g(y) = (g \otimes f)(y, x),$$
$$(f \otimes (g \otimes h)(x, (y, z)) = f(x)g(y)h(z) = ((f \otimes g) \otimes h)((x, y), z).$$

Pre ľubovoľné lineárne podpriestory $U \subseteq K^X$, $V \subseteq K^Y$, $W \subseteq K^Z$ tak dostávame prirodzené izomorfizmy

$$U \otimes V \cong V \otimes U,$$
$$U \otimes (V \otimes W) \cong (U \otimes V) \otimes W,$$

ktoré interpretujeme ako *komutatívnosť* a *asociatívnosť tenzorového súčinu* priestorov. K nim sa druží ešte *distributívnosť*: pre lineárne podpriestory $U_1, U_2 \subseteq K^X$, $V \subseteq K^Y$ bilineárnosť tenzorového súčinu funkcií má za následok priamo rovnosť lineárnych podpriestorov

$$(U_1 + U_2) \otimes V = (U_1 \otimes V) + (U_2 \otimes V),$$

pričom súčet vpravo je priamy práve vtedy, keď $U_1 + U_2$ je priamy súčet. Keďže priamy súčin $U_1 \times U_2$ je izomorfný s priamym súčtom svojich lineárnych podpriestorov $U_1 \times \{\mathbf{0}\} \cong U_1$ a $\{\mathbf{0}\} \times U_2 \cong U_2$, vyplýva z toho aj distributivita tenzorového súčinu voči priamemu súčinu:

$$(U_1 \times U_2) \otimes V \cong (U_1 \otimes V) \times (U_2 \otimes V).$$

Najdôležitejšia z uvedených vlastností je asociatívnosť, ktorá nám umožňuje vynechávať zátvorky vo viacnásobných tenzorových súčinoch a stotožniť tenzorový súčin $V_1 \otimes \ldots \otimes V_n$ (s ľubovoľným rozmiestnením zátvoriek) lineárnych podpriestorov $V_i \subseteq K^{X_i}$ s lineárnym podpriestorom v $K^{X_1 \times \ldots \times X_n}$ generovaným všetkými funkciami tvaru $f_1 \otimes \ldots \otimes f_n$, kde $f_i \in V_i$.

32.2.2. Tvrdenie. *Ak $U \subseteq K^X$ a $V \subseteq K^Y$ sú konečnorozmerné lineárne podpriestory s bázami $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$ resp. $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$, tak systém mn vektorov*

$$\boldsymbol{\alpha} \otimes \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{u}_1 \otimes \boldsymbol{v}_1, & \boldsymbol{u}_1 \otimes \boldsymbol{v}_2, & \ldots & \boldsymbol{u}_1 \otimes \boldsymbol{v}_n, \\ \boldsymbol{u}_2 \otimes \boldsymbol{v}_1, & \boldsymbol{u}_2 \otimes \boldsymbol{v}_2, & \ldots & \boldsymbol{u}_2 \otimes \boldsymbol{v}_n, \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{u}_m \otimes \boldsymbol{v}_1, & \boldsymbol{u}_m \otimes \boldsymbol{v}_2, & \ldots & \boldsymbol{u}_m \otimes \boldsymbol{v}_n \end{pmatrix}$$

tvorí bázu tenzorového súčinu $U \otimes V$. Preto $\dim(U \otimes V) = \dim U \cdot \dim V = mn$.

Dôkaz. To, že uvedené vektory generujú tenzorový súčin $U \otimes V$, je triviálne. Dokážeme, že sú lineárne nezávislé. Nech $c^{ij} \in K$ sú skaláry také, že

$$\sum_{i,j} c^{ij} (\boldsymbol{u}_i \otimes \boldsymbol{v}_j) = \mathbf{0}.$$

Uvedomme si, že $\boldsymbol{u}_i \colon X \to K$, $\boldsymbol{v}_j \colon Y \to K$ sú funkcie, takže uvedená rovnosť znamená, že pre všetky $x \in X$, $y \in Y$ platí

$$\sum_{i,j} c^{ij} \boldsymbol{u}_i(x) \boldsymbol{v}_j(y) = \sum_i \boldsymbol{u}_i(x) \left(\sum_j c^{ij} \boldsymbol{v}_j(y) \right) = 0.$$

Zvoľne pevne nejaké $y \in Y$. Z platnosti poslednej rovnosti pre každé $x \in X$ vyplýva

$$\sum_i \left(\sum_j c^{ij} \boldsymbol{v}_j(y) \right) \boldsymbol{u}_i = \mathbf{0},$$

a na základe lineárnej nezávislosti vektorov $\boldsymbol{u}_1, \dots, \boldsymbol{u}_m$

$$\sum_j c^{ij} \boldsymbol{v}_j(y) = 0$$

pre každé i. Nakoľko $y \in Y$ sme zvolili ľubovoľne, vyplýva z toho

$$\sum_j c^{ij} \boldsymbol{v}_j = \mathbf{0}.$$

Keďže aj vektory $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ sú lineárne nezávislé, konečne dostávame $c^{ij} = 0$ pre všetky i, j.

Uvažujme teraz ľubovoľné vektorové priestory U a V nad poľom K. Ich duály U^*, V^* tvoria lineárne podpriestory priestorov funkcií K^U resp. K^V. Teda tenzorový súčin $U^* \otimes V^*$ je dobre definovaný.

32.2.3. Tvrdenie. *Pre ľubovoľné lineárne funkcionály $\eta \in U^*$, $\vartheta \in V^*$ je ich tenzorový súčin $\eta \otimes \vartheta \colon U \times V \to K$ bilineárna forma na vektorových priestoroch U, V. Ak U a V sú konečnorozmerné, tak ich tenzorový súčin $U^* \otimes V^*$ splýva s vektorovým priestorom $\mathcal{L}_2(U, V, K)$ všetkých bilineárnych foriem $U \times V \to K$.*

Dôkaz. Prvé tvrdenie možno overiť priamym výpočtom, ktorý vynecháme. Zostáva dokázať, že každú bilineárnu formu $F\colon U \times V \to K$ na konečnorozmerných priestoroch U, V možno vyjadriť v tvare súčtu

$$F = (\eta^1 \otimes \vartheta^1) + \ldots + (\eta^k \otimes \vartheta^k),$$

pre vhodné lineárne formy $\eta^i \in U^*$, $\vartheta^i \in V^*$.

Nech $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$, $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ sú bázy priestorov U resp. V a $\boldsymbol{\alpha}^* = (\boldsymbol{u}^1, \ldots, \boldsymbol{u}^m)$, $\boldsymbol{\beta}^* = (\boldsymbol{v}^1, \ldots, \boldsymbol{v}^n)$ sú k nim duálne bázy. Položme $A_{ij} = F(\boldsymbol{u}_i, \boldsymbol{v}_j)$; potom $\boldsymbol{A} = (A_{ij})_{m\times n}$ je matica formy F vzhľadom na bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ a pre všetky $\boldsymbol{x} \in U$, $\boldsymbol{y} \in V$ platí

$$F(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i,j} A_{ij} \boldsymbol{u}^i(\boldsymbol{x}) \boldsymbol{v}^j(\boldsymbol{y}) = \sum_{i,j} A_{ij} (\boldsymbol{u}^i \otimes \boldsymbol{v}^j)(\boldsymbol{x}, \boldsymbol{y})$$

(pozri paragraf 11.1). To znamená, že

$$F = \sum_{i,j} A_{ij} (\boldsymbol{u}^i \otimes \boldsymbol{v}^j).$$

Tým sme dokázali ešte jeden výsledok.

32.2.4. Tvrdenie. *Nech U, V sú konečnorozmerné vektorové priestory s bázami $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$ resp. $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ a $F\colon U \times V \to K$ je bilineárna forma. Potom súradnice formy $F \in U^* \times V^*$ vzhľadom na tenzorový súčin duálnych báz $\boldsymbol{\alpha}^* \otimes \boldsymbol{\beta}^*$ tvoria prvky $A_{ij} = F(\boldsymbol{u}_i, \boldsymbol{v}_j)$ matice $[F]_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ formy F vzhľadom na bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$.*

Zatiaľ máme definovaný tenzorový súčin pre vektorové priestory funkcií, pod ktoré spadajú aj duály ľubovoľných vektorových priestorov. Chýba nám však definícia tenzorového súčinu pre abstraktné vektorové priestory. Našťastie môžeme každý konečnorozmerný vektorový priestor V stotožniť s jeho druhým duálom V^{**}, prípadne nekonečnorozmerný priestor V vnoriť do V^{**}. V oboch prípadoch sa teda na V môžeme dívať ako na lineárny podpriestor priestoru funkcií K^{V^*}.

Nech U a V sú vektorové priestory nad poľom K. Podľa predchádzajúcej definície *tenzorový súčin vektorov* $\boldsymbol{u} \in U$, $\boldsymbol{v} \in V$ je bilineárna forma $\boldsymbol{u} \otimes \boldsymbol{v}\colon U^* \times V^* \to K$ daná predpisom

$$(\boldsymbol{u} \otimes \boldsymbol{v})(\eta, \vartheta) = \boldsymbol{u}(\eta)\boldsymbol{v}(\vartheta) = \eta(\boldsymbol{u})\vartheta(\boldsymbol{v}) = (\eta \otimes \vartheta)(\boldsymbol{u}, \boldsymbol{v})$$

pre $\eta \in V^*$, $\vartheta \in U^*$. *Tenzorový súčin $U \otimes V$ vektorových priestorov U* a *V* je lineárny podpriestor priestoru $K^{U^* \times V^*}$ generovaný všetkými bilineárnymi formami $\boldsymbol{u} \otimes \boldsymbol{v}$, kde $\boldsymbol{u} \in U$, $\boldsymbol{v} \in V$.

Pre vektorové priestory funkcií $U \subseteq K^X$, $V \subseteq K^Y$ však teraz máme dve definície tenzorového súčinu: pôvodnú definíciu a novú definíciu, v ktorej vystupujú ako vektorové priestory funkcií $U \subseteq K^{U^*}$, $V \subseteq K^{V^*}$. Patrilo by sa ukázať, že „oba tenzorové súčiny“ $U \otimes V$ sú izomorfné. Na chvíľu teda budeme pre ten neskôr definovaný používať znak $\boxtimes$. Kľúčovým momentom celého dôkazu bude pozorovanie, že dosadenie ľubovoľného prvku $x \in X$ do funkcie $f \in U$, t. j. priradenie $f \mapsto f(x)$, definuje lineárnu formu $\sigma^x\colon U \to K$, pomocou ktorej možno množinu X stotožniť s množinou $\{\sigma^x;\ x \in X\} \subseteq U^*$. Podobne možno Y stotožniť s množinou $\{\sigma^y;\ y \in Y\} \subseteq V^*$.

Najprv si však treba vyjasniť, kedy sa dva prvky $(f_1 \otimes g_1) + \ldots + (f_n \otimes g_n)$, $(f_1' \otimes g_1') + \ldots + (f_m' \otimes g_m')$ pôvodného tenzorového súčinu $K^X \otimes K^Y$ rovnajú. To platí vtedy a len vtedy, keď ich rovnosť možno (v konečnom počte krokov) odvodiť z vlastností bilinearity operácie tenzorového súčinu. Presnejšie to vystihuje nasledujúce tvrdenie, podľa ktorého o rovnosti takých výrazov rozhodujú lineárne resp. bilineárne formy. Zrejme stačí vedieť, kedy platí $(f_1 \otimes g_1) + \ldots + (f_n \otimes g_n) = \mathbf{0}$. Vyslovíme celý rad podmienok ekvivalentných s touto rovnosťou.

32.2.5. Tvrdenie. *Nech $U \subseteq K^X$, $V \subseteq K^Y$ sú lineárne podpriestory, $f_1, \ldots, f_n \in U$, $g_1, \ldots, g_n \in V$ sú ľubovoľné funkcie a $U_0 = [f_1, \ldots, f_n] \subseteq U$, $V_0 = [g_1, \ldots, g_n] \subseteq V$ sú ich lineárne obaly. Potom nasledujúce podmienky sú ekvivalentné:*

(i) $(f_1 \otimes g_1) + \ldots + (f_n \otimes g_n) = \mathbf{0}$;

(ii) $\eta(f_1)\vartheta(g_1) + \ldots + \eta(f_n)\vartheta(g_n) = 0$
pre všetky lineárne formy $\eta \in U_0^*$, $\vartheta \in V_0^*$;

(iii) $\eta(f_1)\vartheta(g_1) + \ldots + \eta(f_n)\vartheta(g_n) = 0$
pre všetky lineárne formy $\eta \in U^*$, $\vartheta \in V^*$;

(iv) $F(f_1, g_1) + \ldots + F(f_n, g_n) = 0$
pre všetky bilineárne formy $F\colon U_0 \times V_0 \to K$;

(v) $F(f_1, g_1) + \ldots + F(f_n, g_n) = 0$
pre všetky bilineárne formy $F\colon U \times V \to K$;

(vi) $H(f_1, g_1) + \ldots + H(f_n, g_n) = \mathbf{0}$
pre ľubovoľný vektorový priestor W a bilineárne zobrazenie $H\colon U_0 \times V_0 \to W$;

(vii) $H(f_1, g_1) + \ldots + H(f_n, g_n) = \mathbf{0}$
pre ľubovoľný vektorový priestor W a bilineárne zobrazenie $H\colon U \times V \to W$.

Dôkaz vykonáme podľa schémy (i) ⇒ (ii) ⇒ (iv) ⇒ (vi) ⇒ (vii) ⇒ (v) ⇒ (iii) ⇒ (i).

(i) ⇒ (ii) Pre každé $x \in X$ je predpisom $f \mapsto f(x)$ definovaný lineárny funkcionál $\sigma^x\colon U_0 \to K$. Potom pre množinu $\Omega = \{\sigma^x;\ x \in X\} \subseteq U_0^*$ zrejme platí $\Omega^{\perp} = \{\mathbf{0}\}$. Nakoľko U_0 konečnorozmerný, podľa záverečnej časti tvrdenia 32.1.1 to znamená, že Ω generuje jeho duál U_0^*. Keďže i ten je konečnorozmerný, existuje konečná množina $X_0 \subseteq X$ taká, že už podmnožina $\{\sigma^x;\, x \in X_0\} \subseteq \Omega$ generuje U_0^*. Podobne možno nájsť konečnú množinu $Y_0 \subseteq Y$ takú, že lineárne formy $\sigma^y\colon V_0 \to K$, $y \in Y_0$, generujú duál V_0^*.

Podľa (i), $f_1(x)g_1(y) + \ldots + f_n(x)g_n(y) = 0$ pre všetky $x \in X$, $y \in Y$. Zvoľme ľubovoľné $\eta \in U_0^*$, $\vartheta \in V_0^*$. Z vykonaných úvah vyplýva existencia skalárov $a_x, b_y \in K$, pre $x \in X_0$, $y \in Y_0$, takých že $\eta = \sum_{x\in X_0} a_x\sigma^x$ a $\vartheta = \sum_{y\in Y_0} b_y\sigma^y$. Potom

$$\begin{aligned}
\eta(f_1)\vartheta(g_1) + \ldots + \eta(f_n)\vartheta(g_n) &= \sum_{i=1}^{n}\Bigg(\sum_{x\in X_0} a_x\sigma^x(f_i)\Bigg)\Bigg(\sum_{y\in Y_0} b_y\sigma^y(g_i)\Bigg)\\
&= \sum_{i=1}^{n}\Bigg(\sum_{x\in X_0} a_x f_i(x)\Bigg)\Bigg(\sum_{y\in Y_0} b_y g_i(y)\Bigg)\\
&= \sum_{x\in X_0}\sum_{y\in Y_0} a_x b_y \sum_{i=1}^{n} f_i(x)g_i(y) = 0.
\end{aligned}$$

(ii) $\Rightarrow$ (iv) vyplýva z tvrdenia 32.2.3, podľa ktorého každá bilineárna forma na konečnorozmerných priestoroch U_0, V_0 je súčtom foriem tvaru $\eta \otimes \vartheta$, kde $\eta \in U_0^*$, $\vartheta \in V_0^*$.

(iv) $\Rightarrow$ (vi) Nech platí (iv) a predpokladajme, že pre nejaké bilineárne zobrazenie $H\colon U_0 \times V_0 \to W$ je

$$\boldsymbol{w} = H(f_1, g_1) + \ldots + H(f_n, g_n) \neq \mathbf{0}.$$

Potom existuje lineárny funkcionál $\varphi\colon W \to K$ taký, že $\varphi(\boldsymbol{w}) \neq 0$ (pozri cvičenie 6.15). To je však v spore s (iv), lebo predpisom $F(f, g) = \varphi H(f, g)$ je definovaná bilineárna forma $\varphi \circ H\colon U_0 \times V_0 \to K$.

(vi) $\Rightarrow$ (vii) je triviálne, lebo pre bilineárne zobrazenie $H\colon U \times V \to W$ jeho zúženie $H \restriction (U_0 \times V_0)\colon U_0 \times V_0 \to W$ je opäť bilineárne zobrazenie.

Implikácie (vii) $\Rightarrow$ (v) a (v) $\Rightarrow$ (iii) sú tak isto triviálne.

(iii) $\Rightarrow$ (i) vyplýva z pozorovania, že pre $x \in X$, $y \in Y$ sú priradeniami $f \mapsto f(x)$, $g \mapsto g(y)$ definované lineárne formy $\sigma^x\colon U \to K$ resp. $\sigma^y\colon V \to K$. Potom

$$(f_1 \otimes g_1)(x, y) + \ldots + (f_n \otimes g_n)(x, y) = \sigma^x(f_1)\sigma^y(g_1) + \ldots + \sigma^x(f_n)\sigma^y(g_n) = 0$$

pre všetky $x \in X$, $y \in Y$.

Ešte poznamenajme, že ekvivalenciu podmienok (i)–(v), prípadne aj (vi) a (vii) pre konečnorozmerný vektorový priestor W, možno dokázať aj bez predpokladu existencie Hamelových báz z cvičenia 6.15.

32.2.6. Dôsledok. *Nech $U \subseteq K^X$, $V \subseteq K^Y$ sú vektorové priestory funkcií nad poľom K. Potom priradenie $f \otimes g \mapsto f \boxtimes g$ určuje prirodzený lineárny izomorfizmus ich tenzorových súčinov $U \otimes V \to U \boxtimes V$.*

Dôkaz. Vďaka ekvivalencii podmienok (i) a (iii) predchádzajúceho tvrdenia je priradením $f \otimes g \mapsto f \boxtimes g$ dobre definované injektívne lineárne zobrazenie $U \otimes V \to U \boxtimes V$. Kedže $U \otimes V$ je generovaný prvkami tvaru $f \otimes g$ a $U \boxtimes V$ je generovaný prvkami tvaru $f \boxtimes g$, je to zároveň surjektívne zobrazenie.

Odteraz teda nebudeme rozlišovať medzi „konkrétnym" tenzorovým súčinom lineárnych podpriestorov funkcií $U \subseteq K^X$, $V \subseteq K^Y$ a ich „abstraktným" tenzorovým súčinom, pri ktorom ich chápeme ako lineárne podpriestory $U \subseteq K^{U^*}$, $V \subseteq K^{V^*}$. Oba značíme rovnako $U \otimes V$ a používame tú reprezentáciu, ktorá sa v danej chvíli lepšie hodí. Ak bude treba zdôrazniť úlohu poľa K, budeme používať označenie $U \otimes_K V$.

Najdôležitejšie vlastnosti tenzorového súčinu „abstraktných" vektorových priestorov sú zhrnuté v nasledujúcej vete.

32.2.7. Veta. *Nech U, V W sú vektorové priestory nad poľom K. Potom*

(a) operácia tenzorového súčinu $(\boldsymbol{u}, \boldsymbol{v}) \mapsto \boldsymbol{u} \otimes \boldsymbol{v}$ je bilineárne zobrazenie $U \times V \to U \otimes V$;

(b) ku každému bilineárnemu zobrazeniu $H\colon U \times V \to W$ existuje práve jedno lineárne *zobrazenie $H'\colon U \otimes V \to W$ také, že*

$$H'(\boldsymbol{u} \otimes \boldsymbol{v}) = H(\boldsymbol{u}, \boldsymbol{v})$$

pre všetky $\boldsymbol{u} \in U$, $\boldsymbol{v} \in V$;

(c) priradenie $H \mapsto H'$ je lineárny izomorfizmus $\mathcal{L}_2(U, V, W) \to \mathcal{L}(U \otimes V, W)$;

(d) ak U, V sú konečnorozmerné, tak tenzorový súčin $U \otimes V$ splýva s vektorovým priestorom $\mathcal{L}_2(U^*, V^*, K)$ všetkých bilineárnych foriem $F\colon U^* \times V^* \to K$ a pre jeho dimenziu platí

$$\dim(U \otimes V) = \dim U \cdot \dim V.$$

Dôkaz. (a) je obsahom tvrdenia 32.2.1.

(b) Nech $H\colon U \times V \to W$ je bilineárne zobrazenie. Keďže $U \otimes V$ je generovaný prvkami tvaru $\boldsymbol{u} \otimes \boldsymbol{v}$, každé lineárne zobrazenie $U \otimes V \to W$ je jednoznačne určené svojimi hodnotami na nich. Teda k danému H existuje najviac jedno H', ktoré spĺňa uvedenú podmienku. Presvedčíme sa, že také H' naozaj existuje.

Problém by mohli robiť jedine prvky $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n \in U$, $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n \in V$ také, že $\sum_i (\boldsymbol{u}_i \otimes \boldsymbol{v}_i) = \boldsymbol{0}$, ale $\sum_i H(\boldsymbol{u}_i, \boldsymbol{v}_i) \neq \boldsymbol{0}$. To však odporuje implikácii (i) $\Rightarrow$ (vii) z tvrdenia 32.2.5.

(c) Injektívnosť aj surjektívnosť zobrazenia $H \mapsto H'$ je zrejmá. Linearitu možno jednoducho overiť priamym výpočtom.

(d) je dôsledkom tvrdení 32.2.2 a 32.2.3.

Tvrdenia (a) a (b) hovoria, že priradenie $(\boldsymbol{u}, \boldsymbol{v}) \mapsto \boldsymbol{u} \otimes \boldsymbol{v}$ je *univerzálne bilineárne zobrazenie* na priamom súčine vektorových priestorov $U \times V$: akékoľvek zobrazenie $H\colon U \times V \to W$ do tretieho vektorového priestoru W je totiž bilineárne práve vtedy, keď ho možno rozložiť na kompozíciu $H(\boldsymbol{u}, \boldsymbol{v}) = H'(\boldsymbol{u} \otimes \boldsymbol{v})$ tenzorového súčinu $\otimes\colon U \times V \to U \otimes V$ a *lineárneho* zobrazenia $H'\colon U \otimes V \to W$. Vďaka (d) (porovnaj tiež s tvrdením 32.2.3) môžeme *tenzorový súčin* $U \otimes V$ akýchkoľvek *vektorových priestorov* U, V nad poľom K stručne definovať ako vektorový priestor všetkých bilineárnych foriem $U^* \times V^* \to K$ na duálnych priestoroch U^*, V^*. Toto je naozaj jedna z obvyklých definícií bežne používaných v matematickej literatúre. Význam tvrdenia (c) sa naplno prejaví až v paragrafe 32.4. Jeden jeho dôsledok však môžeme zaznamenať už teraz.

32.2.8. Dôsledok. *Nech U a V sú konečnorozmerné vektorové priestory. Potom duál $(U \otimes V)^*$ tenzorového súčinu $U \otimes V$ je kanonicky izomorfný s tenzorovým súčinom $U^* \otimes V^*$ duálnych priestorov U^* a V^*.*

Dôkaz. Podľa 32.2.7 (c) možno každú bilineárnu formu $F\colon U \times V \to K$ stotožniť s lineárnym funkcionálom $F'\colon U \otimes V \to K$. Vďaka tvrdeniu 32.2.3 tak máme

$$U^* \otimes V^* = \mathcal{L}_2(U, V, K) \cong \mathcal{L}(U \otimes V, K) = (U \otimes V)^*.$$

V prípade, že K je niektoré z polí $\mathbb{R}$, $\mathbb{C}$, nekonečnorozmerné vektorové priestory U, V sú často vybavené vhodnou topológiou (napr. pochádzajúcou od nejakej normy). Ich duály U^*, V^*, ktoré sa obvykle definujú ako priestory všetkých *spojitých* lineárnych funkcionálov na príslušnom priestore, sú opäť vybavené prirodzenou topológiou. Tenzorovým súčinom $U \otimes V$ sa v takom prípade väčšinou rozumie priestor všetkých *spojitých* bilineárnych foriem $U^* \times V^* \to K$. [1] Zvyčajne neplatí, že by množina $\{\boldsymbol{u} \otimes \boldsymbol{v};\ \boldsymbol{u} \in U, \boldsymbol{v} \in V\}$, (v algebraickom zmysle) generovala takto definovaný tenzorový súčin $U \otimes V$, no možno aspoň dokázať, že jej lineárny obal je *hustý* v $U \otimes V$ v nejakej „prirodzenej“ topológii na priestore $U \otimes V$, čiže každú bilineárnu formu $F \in U \otimes V$ možno ľubovoľne presne aproximovať formami tvaru konečných

[1] Aj konečnorozmerné vektorové priestory nad $\mathbb{R}$ alebo $\mathbb{C}$ majú svoju prirodzenú topológiu totožnú s obvyklou topológiou v $\mathbb{R}^n$ resp. $\mathbb{C}^n$, kde n je dimenzia príslušného priestoru. Nakoľko však všetky lineárne aj bilineárne formy na nich sú automaticky spojité, je požiadavka spojitosti bilineárnych foriem pri definícii tenzorového súčinu konečnorozmerných vektorových priestorov nadbytočná.

súčtov $\sum_{i=1}^{m}(\boldsymbol{u}_i \otimes \boldsymbol{v}_i)$. My sa však budeme naďalej držať jednotnej čisto algebraickej definície tenzorového súčinu pre akékoľvek vektorové priestory.

32.2.9. Príklad. Tenzorový súčin $K[x] \otimes K[y]$ priestorov polynómov jednej premennej x resp. y v dôsledku rovnosti $x^m \otimes y^n = x^m y^n$ splýva s priestorom $K[x, y]$ polynómov dvoch (komutujúcich) premenných x a y.

32.2.10. Príklad. Na druhej strane tenzorový súčin $\mathcal{C}(\mathbb{R}) \otimes \mathcal{C}(\mathbb{R})$ dvoch exemplárov vektorového priestoru všetkých spojitých funkcií $\mathbb{R} \to \mathbb{R}$ s priestorom $\mathcal{C}(\mathbb{R} \times \mathbb{R})$ všetkých spojitých funkcií $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$ nesplýva, no je v ňom hustý v tom zmysle, že pre každú spojitú funkciu $h \in \mathcal{C}(\mathbb{R} \times \mathbb{R})$, $\varepsilon > 0$ a ľubovoľné intervaly $\langle a, b\rangle$, $\langle c, d\rangle$ existujú funkcie $f_1, g_1, \dots, f_k, g_k \in \mathcal{C}(\mathbb{R})$ také, že pre všetky $x \in \langle a, b\rangle$, $y \in \langle c, d\rangle$ platí [2]

$$|f_1(x)g_1(y) + \ldots + f_k(x)g_k(y) - h(x, y)| < \varepsilon.$$

To sa využíva pri riešení diferenciálnych rovníc. V tom najjednoduchšom prípade možno pri dostatočne presnej aproximácii funkcie

$$h(x, t) \approx f_1(x)g_1(t) + \ldots + f_k(x)g_k(t)$$

na dostatočne veľkom dvojrozmernom intervale $\langle a, b\rangle \times \langle c, d\rangle$ aproximovať riešenie diferenciálnej rovnice $x'(t) = h(x(t), t)$ s ľubovoľnou presnosťou a na ľubovoľne veľkom intervale súčtom

$$x(t) \approx x_1(t) + \ldots + x_k(t)$$

riešení dielčích rovníc $x_i'(t) = f_i(x_i(t))g_i(t)$. Tie sa riešia ľahko zapamätateľnou metódou *separácie premenných*:

$$\begin{aligned}
\frac{\mathrm{d}x}{\mathrm{d}t} &= f(x)g(t), \\
\frac{\mathrm{d}x}{f(x)} &= g(t)\,\mathrm{d}t, \\
\int \frac{\mathrm{d}x}{f(x)} &= \int g(t)\,\mathrm{d}t, \\
\Phi(x) &= G(t) + C, \\
x(t) &= \Phi^{-1}(G(t) + C),
\end{aligned}$$

kde G je primitívna funkcia k funkcii g, Φ je primitívna funkcia k funkcii $1/f$, Φ^{-1} je k nej inverzná funkcia a C je integračná konštanta, ktorú možno stanoviť na základe prípadnej počiatočnej podmienky $x(t_0) = x_0$. Detaily však v konkrétnych prípadoch môžu byť trochu zložitejšie, než prezrádza náš zbežný náčrt.

[2] Ide o hustotu vzhľadom na tzv. *topológiu rovnomernej konvergencie na kompaktoch*, nazývanú tiež *kompaktno-otvorená topológia*.

32.3 Zdvih poľa skalárov

Pomocou tenzorových súčinov možno jednoducho popísať konštrukciu nazývanú *zdvihom poľa skalárov*, ktorou z vektorového priestoru nad poľom K dostaneme vektorový priestor nad jeho rozšírením L. Špeciálnym prípadom tejto konštrukcie je komplexifikácia reálneho vektorového priestoru z paragrafu 19.4.

Nech pole L je rozšírením poľa K a V je vektorový priestor nad K. Potom aj L môžeme považovať za vektorový priestor nad poľom K a vytvoriť tenzorový súčin $L \otimes_K V$. Ten je, samozrejme, vektorovým priestorom nad K, no jednoducho z neho možno vyrobiť vektorový priestor nad L. Stačí zadefinovať násobenie rozložiteľných prvkov $c \otimes \boldsymbol{x} \in L \otimes V$ skalármi $b \in L$ rovnosťou

$$b(c \otimes \boldsymbol{x}) = bc \otimes \boldsymbol{x}.$$

Výsledný vektorový priestor nad L značíme tiež $V^L = L \otimes_K V$ a nazývame *zdvihom poľa skalárov vektorového priestoru* V z K na L. na pole L.

Pre dimenzie priestoru $L \otimes_K V$ platí

$$\dim_K(L \otimes_K V) = [L : K] \dim_K V, \qquad \dim_L(L \otimes_K V) = \dim_K V.$$

Prípad nekonečného rozšírenia je triválny. Ak $[L : K] = m$ a $\boldsymbol{\alpha} = (u_1 = 1, u_2, \dots, u_m)$ je báza L ako vektorového priestoru nad K, tak $L \otimes_K V$ je tvorený všetkými súčtami tvaru

$$(u_1 \otimes \boldsymbol{x}_1) + (u_2 \otimes \boldsymbol{x}_2) + \dots + (u_m \otimes \boldsymbol{x}_m),$$

kde $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_m \in V$, čo možno zjednodušene zapísať ako lineárnu kombináciu

$$\boldsymbol{x}_1 + u_2\boldsymbol{x}_2 + \dots + u_m\boldsymbol{x}_m,$$

pričom pre dva takéto výrazy platí

$$\boldsymbol{x}_1 + u_2\boldsymbol{x}_2 + \dots + u_m\boldsymbol{x}_m = \boldsymbol{y}_1 + u_2\boldsymbol{y}_2 + \dots + u_m\boldsymbol{y}_m \\ \Leftrightarrow\ \boldsymbol{x}_1 = \boldsymbol{y}_1 \ \&\ \dots\ \&\ \boldsymbol{x}_m = \boldsymbol{y}_m$$

a ich súčet je definovaný zrejmým spôsobom po zložkách. Pre $K = \mathbb{R}$, $L = \mathbb{C}$ s bázou $(1, \mathrm{i})$ v tom čitateľ určite spozná reprezentáciu vektorov komplexifikácie $V^{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} V$ v tvare $\boldsymbol{x} + \mathrm{i}\boldsymbol{y}$, kde $\boldsymbol{x}, \boldsymbol{y} \in V$.

Ak $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ je báza V, tak $\boldsymbol{\beta}$ je zároveň bázou V^L ako vektorového priestoru nad poľom L; jeho bázu ako vektorového priestoru nad K podľa tvrdenia 32.2.2 tvoria všetky vektory $u_i \otimes \boldsymbol{v}_j = u_i\boldsymbol{v}_j$ systému $\boldsymbol{\alpha} \otimes \boldsymbol{\beta}$.

Pole L možno zároveň chápať ako asociatívnu lineárnu algebru nad poľom K (pozri paragraf 30.1). Označme $\gamma_{ij}^k \in K$ štruktúrne konštanty algebry

L vzhľadom na bázu $\boldsymbol{\alpha}$. To znamená, že pre všetky $i, j \leq m$ platí

$$u_i u_j = \sum_{k=1}^{m} \gamma_{ij}^k u_k.$$

Vyjadrenie násobku všeobecného skalára $a_1 + a_2 u_2 + \ldots + a_m u_m \in L$, $a_i \in K$, a všeobecného vektora $\boldsymbol{x}_1 + u_2 \boldsymbol{x}_2 + \ldots + u_m \boldsymbol{x}_m \in V^L$, $\boldsymbol{x}_j \in V$, pomocou štruktúrnych konštánt možno na prvý pohľad vyzerá trochu zložito:

$$\begin{aligned}(a_1 + a_2 u_2 + \ldots + a_m u_m)(\boldsymbol{x}_1 + u_2 \boldsymbol{x}_2 + \ldots + u_m \boldsymbol{x}_m) &= \sum_{i,j} a_i u_i u_j \boldsymbol{x}_j \\ &= \sum_{i,j} \Bigg(\sum_k \gamma_{ij}^k u_k \Bigg) a_i \boldsymbol{x}_j = \sum_k u_k \sum_{i,j} \gamma_{ij}^k a_i \boldsymbol{x}_j.\end{aligned}$$

Čitateľ by si mal preto samostatne overiť, že pre komplexifikáciu reálneho vektorového priestoru dáva táto formula presne to, čo sa od nej očakáva.

32.4 Priestory (multi)lineárnych zobrazení

Pomocou konštrukcií duálu a tenzorového súčinu možno jednotným spôsobom popísať vektorové priestory lineárnych a multilineárnych zobrazení.

Najprv si pripomeňme označenie $\mathcal{L}_n(V_1, \ldots, V_n, W)$ vektorového priestoru všetkých n-lineárnych zobrazení $V_1 \times \ldots \times V_n \to W$, ktoré sme zaviedli v cvičení 10.2. Teda špeciálne $\mathcal{L}_1(V, W) = \mathcal{L}(V, W)$ a $\mathcal{L}_2(U, V, W)$ je priestor všetkých bilineárnych zobrazení $U \times V \to W$. Pre úplnosť ešte kladieme $\mathcal{L}_0(W) = W$. Multilineárne zobrazenia $V_1 \times \ldots \times V_n \to K$ nazývame tiež *multilineárnymi formami.*

Uvedomme si, že na funkciu dvoch premenných $F\colon X \times Y \to Z$ sa možno dívať aj ako na funkciu $F : X \to Z^Y$, ktorá predpisom $F_x(y) = F(x, y)$ priradí prvku $x \in X$ funkciu $F_x\colon Y \to Z$. Naopak, každú funkciu $G\colon X \to Z^Y$ možno chápať aj ako funkciu dvoch premenných $G\colon X \times Y \to Z$ takú, že $G(x, y) = G(x)(y)$ pre $x \in X$, $y \in Y$. Tým sú popísané dve prirodzené navzájom inverzné bijekcie medzi množinami funkcií $Z^{X \times Y}$ a $\left(Z^Y\right)^X$, prostredníctvom ktorých ich možno priamo stotožniť.

Niečo podobné funguje aj pre lineárne a multilineárne zobrazenia medzi vektorovými priestormi. Každé *bilineárne zobrazenie* $H\colon U \times V \to W$ určuje priradením $\boldsymbol{u} \mapsto H_{\boldsymbol{u}} = H(\boldsymbol{u}, \cdot)$ *lineárne zobrazenie* $H\colon U \to \mathcal{L}(V, W)$, ktoré umožňuje prirodzene stotožniť priestory $\mathcal{L}_2(U, V, W)$ a $\mathcal{L}(U, \mathcal{L}(V, W))$.

Konečnorozmerné vektorové priestory V a W možno navyše prirodzene stotožniť s ich druhými duálmi V^{**} resp. W^{**}, čím dostávame ďalšie kanonické izomorfizmy

$$\begin{aligned}\mathcal{L}(V, W) &\cong \mathcal{L}(V, W^{**}) = \mathcal{L}(V, \mathcal{L}(W^*, K)) \\ &\cong \mathcal{L}_2(V, W^*, K) \cong \mathcal{L}_2(V^{**}, W^*, K) \cong V^* \otimes W.\end{aligned}$$

Tento výsledok stojí za zaznamenanie.

32.4.1. Tvrdenie. *Nech V, W sú konečnorozmerné vektorové priestory. Potom vektorový priestor $\mathcal{L}(V,W)$ všetkých lineárnych zobrazení $V \to W$ je kanonicky izomorfný s tenzorovým súčinom $V^* \otimes W$.*

Kanonický izomorfizmus $\mathcal{L}(V,W) \to V^* \otimes W$ možno popísať aj priamo. Lineárne zobrazenie $\varphi\colon V \to W$ sa v ňom zobrazí na (rovnako značenú) bilineárnu formu $\varphi\colon V \times W^* \to K$ (t. j. prvok priestoru $V^* \otimes W$) danú predpisom

$$\varphi(\boldsymbol{x}, \vartheta) = \vartheta(\varphi\boldsymbol{x}) = (\vartheta \circ \varphi)(\boldsymbol{x})$$

pre $\boldsymbol{x} \in V$, $\vartheta \in W^*$.

Nech $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$, $\boldsymbol{\gamma} = (\boldsymbol{w}_1, \dots, \boldsymbol{w}_k)$ sú bázy priestorov V resp. W a $\boldsymbol{\beta}^* = (\boldsymbol{v}^1, \dots, \boldsymbol{v}^n)$, $\boldsymbol{\gamma}^* = (\boldsymbol{w}^1, \dots, \boldsymbol{w}^k)$ sú k nim duálne bázy. Potom súradnice bilineárnej formy φ vzhľadom na bázu $\boldsymbol{\beta}^* \otimes \boldsymbol{\gamma} = (\boldsymbol{v}^j \otimes \boldsymbol{w}_i)$ priestoru $V^* \otimes W$ tvoria podľa tvrdenia 32.2.4 skaláry

$$A^i_j = \varphi(\boldsymbol{v}_j, \boldsymbol{w}^i) = \boldsymbol{w}^i(\varphi\boldsymbol{v}_j),$$

čiže prvky matice $\boldsymbol{A} = (A^i_j) = (\varphi)_{\boldsymbol{\gamma},\boldsymbol{\beta}}$ lineárneho zobrazenia $\varphi\colon V \to W$. Samotné φ teda možno vyjadriť ako lineárnu kombináciu

$$\varphi = \sum_{i,j} A^i_j (\boldsymbol{v}^j \otimes \boldsymbol{w}_i).$$

Podľa (b) a (c) vety 32.2.7 stotožňujeme bilineárne zobrazenie $H\colon U \times V \to W$ s lineárnym zobrazením $H'\colon U \otimes V \to W$ a následne aj vektorové priestory $\mathcal{L}_2(U,V,W)$ a $\mathcal{L}(U \otimes V, W)$. V konečnorozmernom prípade je ten druhý podľa tvrdenia 32.4.1 kanonicky izomorfný s tenzorovým súčinom $(U \otimes V)^* \otimes W$, a podľa dôsledku 32.2.8 tiež s tenzorovým súčinom $U^* \otimes V^* \otimes W$. Situáciu možno indukciou zovšeobecniť na ľubovoľný konečný počet činiteľov.

32.4.2. Veta. *Nech $V_1, \dots, V_m$ a W sú konečnorozmerné vektorové priestory. Potom priestor $\mathcal{L}_m(V_1, \dots, V_m, W)$ všetkých m-lineárnych zobrazení $V_1 \times \dots \times V_m \to W$ je kanonicky izomorfný s tenzorovým súčinom $V_1^* \otimes \dots \otimes V_m^* \otimes W$.*

Explicitne stotožňujeme m-lineárne zobrazenie $H\colon V_1 \times \dots \times V_m \to W$ s lineárnym funkcionálom $H\colon V_1 \otimes \dots \otimes V_m \otimes W^* \to K$ (t. j. prvkom duálu $(V_1 \otimes \dots \otimes V_m \otimes W^*)^*$, ako aj s $(m+1)$-lineárnou formou $H\colon V_1 \times \dots \times V_m \times W^* \to K$ (t. j. prvkom tenzorového súčinu $V_1^* \otimes \dots \otimes V_m^* \otimes W$) takými, že

$$\begin{aligned} H(\boldsymbol{x}_1 \otimes \dots \otimes \boldsymbol{x}_m \otimes \vartheta) &= H(\boldsymbol{x}_1, \dots, \boldsymbol{x}_m, \vartheta) \\ &= \vartheta H(\boldsymbol{x}_1, \dots, \boldsymbol{x}_m) = (\vartheta \circ H)(\boldsymbol{x}_1, \dots, \boldsymbol{x}_m) \end{aligned}$$

pre $\boldsymbol{x}_1 \in V_1, \ldots, \boldsymbol{x}_m \in V_m$, $\vartheta \in W^*$.

Vo všeobecnosti môžeme uvažovať tenzorové súčiny tvaru

$$V_1 \otimes \ldots \otimes V_p \otimes V_{p+1}^* \otimes \ldots \otimes V_{p+q}^*$$

vektorových priestorov V_i resp. ich duálov V_i^*. My sa však ďalej obmedzíme na prípad, keď všetky priestory V_i sú inštanciami jedného a toho istého konečnorozmerného vektorového priestoru V.

32.5 Tenzory nad vektorovým priestorom

Nad každým konečnorozmerným vektorovým priestorom V možno pomocou konštrukcií duálu a tenzorového súčinu vybudovať rozvetvenú štruktúru ďalších vektorových priestorov.

Ľubovoľnú $(p+q)$-lineárnu formu

$$T\colon \underbrace{V^* \times \ldots \times V^*}_{p} \times \underbrace{V \times \ldots \times V}_{q} \to K,$$

nazývame *tenzorom typu* $\binom{p}{q}$ *nad vektorovým priestorom* V. Skaláry, t. j. prvky poľa K, považujeme za tenzory typu $\binom{0}{0}$. Číslo $p+q$ nazývame *stupňom tenzora* T. [3] Tenzor typu $\binom{p}{q}$ je tak prvkom tenzorového súčinu

$$\mathcal{T}_q^p(V) = \underbrace{V \otimes \ldots \otimes V}_{p} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{q}$$

p exemplárov vektorového priestoru V a q exemplárov jeho duálu V^*. $\mathcal{T}_q^p(V)$ nazývame *priestorom tenzorov typu* $\binom{p}{q}$ nad vektorovým priestorom V. Špeciálne pre $p = q = 0$ kladieme $\mathcal{T}_0^0(V) = K$.

Hovoríme, že $T \in \mathcal{T}_q^p(V)$ je *rozložiteľný tenzor*, ak má tvar

$$T = \boldsymbol{x}_1 \otimes \ldots \otimes \boldsymbol{x}_p \otimes \eta_1 \otimes \ldots \otimes \eta_q$$

pre nejaké vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p \in V$ a kovektory $\eta_1, \ldots, \eta_q \in V^*$; ostatné *tenzory* sa nazývajú *nerozložiteľné*. Priestor tenzorov $\mathcal{T}_q^p(V)$ teda pozostáva z lineárnych kombinácií rozložiteľných tenzorov.

Poznamenajme, že v prípade vektorových priestorov nad poľom komplexných čísel je namieste zaviesť a študovať aj tenzory typu $\binom{p,r}{q,s}$, čiže formy

$$T\colon \underbrace{V^* \times \ldots \times V^*}_{p} \times \underbrace{V \times \ldots \times V}_{q} \times \underbrace{\bar{V}^* \times \ldots \times \bar{V}^*}_{r} \times \underbrace{\bar{V} \times \ldots \times \bar{V}}_{s} \to \mathbb{C}$$

[3] Značenie ani terminológia nie sú v tejto záležitosti jednotné. To, čo sme práve nazvali tenzorom typu $\binom{p}{q}$, sa niekde nazýva *tenzorom typu* (p, q) a inde zasa *tenzorom typu* (q, p). Tenzor stupňa m sa zvykne tiež nazývať *m-valentný tenzor*, prípadne *tenzor rádu* alebo *rangu m*.

lineárne v prvých $p + q$ argumentoch a semilineárne v ďalších $r + s$ argumentoch. Vektorový priestor takýchto foriem možno opísať ako tenzorový súčin

$$\begin{aligned}\mathcal{T}^{p,r}_{q,s}(V) &= \mathcal{T}^p_q(V) \otimes \mathcal{T}^r_s(\overline{V})\\ &= \underbrace{V \otimes \ldots \otimes V}_{p} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{q} \otimes \underbrace{\overline{V} \otimes \ldots \otimes \overline{V}}_{r} \otimes \underbrace{\overline{V}^* \otimes \ldots \otimes \overline{V}^*}_{s},\end{aligned}$$

kde $\overline{V}$ označuje vektorový priestor V, v ktorom bola pôvodná operácia skalárneho násobku nahradená operáciou násobku konjugovaným skalárom $(c, \boldsymbol{x}) \mapsto \bar{c}\boldsymbol{x}$. Napríklad každá poldruhalineárna forma $F\colon V \times V \to \mathbb{C}$ je ako tenzor prvkom tenzorového súčinu $\mathcal{T}^{0,0}_{1,1}(V) = V^* \otimes \overline{V}^*$. Priestormi tenzorov tohto druhu sa však podrobnejšie zaoberať nebudeme.

Podľa vety 32.4.2 je priestor $\mathcal{T}^p_q(V)$ kanonicky izomorfný s ďalšími priestormi (multi)lineárnych zobrazení

$$\begin{aligned}\mathcal{T}^p_q(V) &= \mathcal{T}^p_0(V) \otimes \mathcal{T}^q_0(V^*) = \mathcal{T}^p_0(V) \otimes \mathcal{T}^0_q(V) = \mathcal{L}_{p+q}(\underbrace{V^*, \ldots, V^*}_{p}, \underbrace{V, \ldots, V}_{q}, K)\\ &\cong \mathcal{L}(\underbrace{V^* \otimes \ldots \otimes V^*}_{p} \otimes \underbrace{V \otimes \ldots \otimes V}_{q}, K) = \mathcal{T}^q_p(V^*) = \mathcal{T}^q_p(V)^*\\ &\cong \mathcal{L}(\underbrace{V \otimes \ldots \otimes V}_{q}, \underbrace{V \otimes \ldots \otimes V}_{})_p = \mathcal{L}\big(\mathcal{T}^q_0(V), \mathcal{T}^p_0(V)\big).\end{aligned}$$

Tenzor $T \in \mathcal{T}^p_q(V)$ preto môžeme podľa potreby považovať tak za lineárny funkcionál

$$T\colon \underbrace{V^* \otimes \ldots \otimes V^*}_{p} \otimes \underbrace{V \otimes \ldots \otimes V}_{q} \to K,$$

čiže $T\colon \mathcal{T}^q_p(V) \to K$, ako aj za lineárne zobrazenie

$$T\colon \underbrace{V \otimes \ldots \otimes V}_{q} \to \underbrace{V \otimes \ldots \otimes V}_{p},$$

t. j. $T\colon \mathcal{T}^q_0(V) \to \mathcal{T}^p_0(V)$.

32.5.1. Veta. *Nech* $\dim V = n$. *Potom* $\dim \mathcal{T}^p_q(V) = n^{p+q}$.

Dôkaz vety je vedľajším produktom nasledujúcej úvahy.

Ak $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ je nejaká báza priestoru V a $\boldsymbol{\alpha}^* = (\boldsymbol{u}^1, \ldots, \boldsymbol{u}^n)$ je k nej duálna báza, tak bázu priestoru $\mathcal{T}^p_q(V)$ tvoria v dôsledku tvrdenia 32.2.2 tenzory

$$\boldsymbol{u}_{i_1} \otimes \ldots \boldsymbol{u}_{i_p} \otimes \boldsymbol{u}^{j_1} \otimes \ldots \boldsymbol{u}^{j_q},$$

kde indexy $i_1, \ldots, i_p$, $j_1, \ldots, j_q$ nadobúdajú hodnoty od 1 do n. Každý tenzor $T \in \mathcal{T}_q^p(V)$ tak možno vyjadriť ako lineárnu kombináciu

$$T = \sum_{\substack{i_1,\ldots,i_p \\ j_1,\ldots,j_q}} A^{i_1\ldots i_p}_{j_1\ldots j_q} \big(\boldsymbol{u}_{i_1} \otimes \ldots \otimes \boldsymbol{u}_{i_p} \otimes \boldsymbol{u}^{j_1} \otimes \ldots \otimes \boldsymbol{u}^{j_q}\big)$$

pre jednoznačne určené koeficienty

$$A^{i_1\ldots i_p}_{j_1\ldots j_q} = T\big(\boldsymbol{u}^{i_1}, \ldots, \boldsymbol{u}^{i_p}, \boldsymbol{u}_{j_1}, \ldots, \boldsymbol{u}_{j_q}\big) \in K.$$

Horné, tzv. *kontravariantné indexy* $i_1, \ldots, i_p$ zodpovedajú v tejto lineárnej kombinácii vektorom $\boldsymbol{u}_{i_1}, \ldots, \boldsymbol{u}_{i_p} \in V$ a dolné, tzv. *kovariantné indexy* $j_1, \ldots, j_q$ zasa kovektorom $\boldsymbol{u}^{j_1}, \ldots, \boldsymbol{u}^{j_q} \in V^*$.

Pre kovektory $\eta^1, \ldots, \eta^p \in V^*$ a vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q \in V$ so súradnicami

$$\eta_i^k = \langle \eta^k, \boldsymbol{u}_i \rangle \qquad \text{resp.} \qquad x_l^j = \langle \boldsymbol{x}_l, \boldsymbol{u}^j \rangle \qquad (k \le p,\ l \le q,\ i, j \le n)$$

vzhľadom na bázu $\boldsymbol{\alpha}$ potom platí

$$T\big(\eta^1, \ldots, \eta^p, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_q\big) = \sum_{\substack{i_1,\ldots,i_p \\ j_1,\ldots,j_q}} A^{i_1\ldots i_p}_{j_1\ldots j_q} \eta^1_{i_1} \ldots \eta^p_{i_p} x^{j_1}_1 \ldots x^{j_q}_q.$$

Systém koeficientov $\boldsymbol{A} = \big(A^{i_1\ldots i_p}_{j_1,\ldots j_q}\big)$, kde každý z indexov i_k, j_l nadobúda hodnoty od 1 do n, hrá podobnú úlohu ako matica lineárnej transformácie $V \to V$ alebo bilineárnej formy $V \times V \to K$; nazývame ho *súradnicami tenzora* T vzhľadom na bázu $\boldsymbol{\alpha}$ a značíme

$$\boldsymbol{A} = \big(A^{i_1\ldots i_p}_{j_1\ldots j_q}\big) = (T)_{\boldsymbol{\alpha}}.$$

$\boldsymbol{A}$ je naozaj niečo ako „$(p+q)$-rozmerná matica“ s p hornými a q dolnými indexmi. Znázorniť si takúto „maticu“ prehľadne ako tabuľku však pre tenzory stupňa > 2 už nie je ďalej možné – také niečo by si vyžadovalo schopnosť čítať z „$(p+q)$-rozmerného papiera,“ čo je úloha, ktorú by mohli zvládnuť leda ak obyvatelia $(p+q+1)$-rozmerného priestoru, kde by list takého „papiera“ predstavoval časť nadroviny.

Z toho dôvodu nebudeme ďalej viazať reprezentáciu súradníc tenzorov na „geometriu listu papiera“, inak povedané, nebudeme si ich predstavovať ako tabuľky. Potrebnú informáciu budeme fixovať polohou indexov a ich poradím. Napr. obdobou operácie transponovania matice bude zámena poradia dvoch indexov rovnakého typu. Miesto rozlišovania vektorov-stĺpcov a vektorov-riadkov budeme rozlišovať vektory, ktorých súradnice budú mať indexy hore, a kovektory, ktorých súradnice budú mať indexy dole. Naopak, systémy vektorov budeme číslovať dolnými a systémy kovektorov hornými indexmi. Túto dohodu sme už medzičasom začali požívať. Kvôli sprehľadneniu zápisov však prijmeme aj tzv. *Einsteinovu konvenciu*, ktorá nám umožní vynechávať znaky sumácie $\sum$.

32.5.2. Einsteinova konvencia o implicitnej sumácii. Vo svete tenzorov sa výrazy „správne“ sčitujú len cez tzv. sumačné indexy, t. j. indexy, ktoré sa v nich vyskytujú raz v hornej a raz v dolnej polohe. Sumácia cez takéto indexy sa nemusí explicitne vyznačovať znakom $\sum$. Výsledok sumácie potom závisí iba od zvyšných (t. j. „nesumačných“) tzv. voľných indexov.

Zrozumiteľnosť Einsteinovej sumačnej konvencie závisí na kontexte. Bez toho nemusí byť zrejmé či napr. výraz A^i_i označuje jeden všeobecný diagonálny prvok matice $\boldsymbol{A} = \big(A^i_j\big)$ alebo jej stopu $\operatorname{tr}\boldsymbol{A} = \sum_i A^i_i$. Celá vec sa stane jasnejšou po niekoľkých známych príkladoch, ktoré si teraz prepíšeme pomocou tejto konvencie. Keďže našim cieľom nie je túto konvenciu systematicky využívať, ale naučiť sa jej rozumieť a používať ju, budeme väčšinou uvádzať oba zápisy – s explicitnou aj s implicitnou sumáciou.

32.5.3. Príklad. (a) Vektory sú tenzory typu $\binom{1}{0}$. Vektor $\boldsymbol{x} \in V$ so súradnicami $(\boldsymbol{x})_{\boldsymbol{\alpha}} = (x^j)$ vzhľadom na bázu $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ má tvar

$$\boldsymbol{x} = \sum_j x^j \boldsymbol{u}_j = x^j \boldsymbol{u}_j.$$

(b) Kovektory sú tenzory typu $\binom{0}{1}$. Kovektor $\eta \in V$ so súradnicami $(\eta)_{\boldsymbol{\alpha}} = (\eta_i)$ možno zapísať aj ako lineárnu kombináciu

$$\eta = \sum_i \eta_i \boldsymbol{u}^i = \eta_i \boldsymbol{u}^i.$$

(c) Samotné párovanie $\langle \cdot\,, \cdot \rangle$ je tenzor typu $\binom{1}{1}$, ktorý zodpovedá identickému lineárnemu zobrazeniu $\mathrm{id}_V \colon V \to V$ so súradnicami (maticou) $(\mathrm{id}_V)_{\boldsymbol{\alpha}} = (\delta^i_j)$. Potom

$$\mathrm{id}_V = \sum_{i,j} \delta^i_j \big(\boldsymbol{u}_i \otimes \boldsymbol{u}^j\big) = \sum_i \boldsymbol{u}_i \otimes \boldsymbol{u}^i = \delta^i_j \big(\boldsymbol{u}_i \otimes \boldsymbol{u}^j\big) = \boldsymbol{u}_i \otimes \boldsymbol{u}^i$$

a

$$\langle \eta, \boldsymbol{x} \rangle = \sum_{i,j} \delta^i_j \eta_i x^j = \sum_i \eta_i x^i = \delta^i_j \eta_i x^j = \eta_i x^i.$$

(d) Lineárna transformácia $\varphi \colon V \to V$ je rovnako tenzor typu $\binom{1}{1}$. Jej súradnice (t. j. matica) v báze $\boldsymbol{\alpha}$ sú $\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha}} = (A^i_j)$, kde $A^i_j = \langle \boldsymbol{u}^i, \varphi \boldsymbol{u}_j \rangle$. Potom

$$\varphi = \sum_{i,j} A^i_j (\boldsymbol{u}_i \otimes \boldsymbol{u}^j) = A^i_j (\boldsymbol{u}_i \otimes \boldsymbol{u}^j)$$

a
$$\varphi(\eta,\boldsymbol{x}) = \eta(\varphi\boldsymbol{x}) = \langle\eta,\varphi\boldsymbol{x}\rangle = \sum_{i,j} A^i_j\eta_i x^j = A^i_j\eta_i x^j.$$

Špeciálne pre $\eta = \boldsymbol{u}^i$ máme
$$\varphi(\boldsymbol{u}^i,\boldsymbol{x}) = \boldsymbol{u}^i(\varphi\boldsymbol{x}) = \langle\boldsymbol{u}^i,\varphi\boldsymbol{x}\rangle = \sum_j A^i_j x^j = A^i_j x^j.$$

(e) Bilineárna forma $F\colon V\times V\to K$ je tenzor typu $\binom{0}{2}$. Jej súradnice (matica) v báze $\boldsymbol{\alpha}$ je $\boldsymbol{A} = (F)_{\boldsymbol{\alpha}} = (A_{ij})$, kde $A_{ij} = F(\boldsymbol{u}_i,\boldsymbol{u}_j)$. Potom
$$F = \sum_{i,j} A_{ij}(\boldsymbol{u}^i\otimes\boldsymbol{u}^j) = A_{ij}(\boldsymbol{u}^i\otimes\boldsymbol{u}^j)$$
a
$$F(\boldsymbol{x},\boldsymbol{y}) = \sum_{i,j} A_{ij}x^i y^j = A_{ij}x^i y^j$$
pre vektory $\boldsymbol{x} = x^i\boldsymbol{u}_i$, $\boldsymbol{y} = y^j\boldsymbol{u}_j$.

(f) Konečne všeobecný tenzor $T\in\mathcal{T}^p_q(V)$ so súradnicami $\boldsymbol{A} = (T)_{\boldsymbol{\alpha}} = \big(A^{i_1\dots i_p}_{j_1\dots j_q}\big)$, kde
$$A^{i_1\dots i_p}_{j_1\dots j_q} = T\big(\boldsymbol{u}^{i_1},\dots,\boldsymbol{u}^{i_p},\boldsymbol{u}_{j_1},\dots,\boldsymbol{u}_{j_q}\big),$$
má tvar
$$\begin{aligned} T &= \sum_{\substack{i_1,\dots,i_p\\ j_1,\dots,j_q}} A^{i_1\dots i_p}_{j_1\dots j_q}\big(\boldsymbol{u}_{i_1}\otimes\dots\otimes\boldsymbol{u}_{i_p}\otimes\boldsymbol{u}^{j_1}\otimes\dots\otimes\boldsymbol{u}^{j_q}\big)\\ &= A^{i_1\dots i_p}_{j_1\dots j_q}\big(\boldsymbol{u}_{i_1}\otimes\dots\otimes\boldsymbol{u}_{i_p}\otimes\boldsymbol{u}^{j_1}\otimes\dots\otimes\boldsymbol{u}^{j_q}\big).\end{aligned}$$
Potom pre $\eta^1,\dots,\eta^p\in V^*$, $\boldsymbol{x}_1,\dots,\boldsymbol{x}_q\in V$ máme
$$\begin{aligned} T\big(\eta^1,\dots,\eta^p,\boldsymbol{x}_1,\dots,\boldsymbol{x}_q\big) &= \sum_{\substack{i_1,\dots,i_p\\ j_1,\dots,j_q}} A^{i_1\dots i_p}_{j_1\dots j_q}\eta^1_{i_1}\dots\eta^p_{i_p}x^{j_1}_1\dots x^{j_q}_q\\ &= A^{i_1\dots i_p}_{j_1\dots j_q}\eta^1_{i_1}\dots\eta^p_{i_p}x^{j_1}_1\dots x^{j_q}_q.\end{aligned}$$

Pozrime sa teraz bližšie, ako sa transformujú súradnice tenzora pri zmene bázy. Najprv preskúmame jednoduchšie špeciálne príklady vektorov, kovektorov, bilineárnych foriem a lineárnych zobrazení.

32.5.4. Príklad. Nech $\boldsymbol{\alpha} = (\boldsymbol{u}_1,\dots,\boldsymbol{u}_n)$, $\boldsymbol{\beta} = (\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ sú dve bázy priestoru V a $\boldsymbol{\alpha}^* = (\boldsymbol{u}^1,\dots,\boldsymbol{u}^n)$ resp. $\boldsymbol{\beta}^* = (\boldsymbol{v}^1,\dots,\boldsymbol{v}^n)$ sú k nim duálne bázy. Označme $\boldsymbol{P} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \big(P^i_j\big)$ a $\boldsymbol{Q} = \boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\alpha}} = \big(Q^i_j\big)$ matice prechodu medzi nimi. Potom $\boldsymbol{Q} = \boldsymbol{P}^{-1}$.

(a) Náš doterajší zápis $(\boldsymbol{x})_{\boldsymbol{\alpha}} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot (\boldsymbol{x})_{\boldsymbol{\beta}}$ transformačnej rovnice súradníc vektora $\boldsymbol{x} \in V$ nadobúda s využitím Einsteinovej sumačnej konvencie podobu

$$\boldsymbol{u}^i(\boldsymbol{x}) = \sum_j P^i_j \boldsymbol{v}^j(\boldsymbol{x}) = P^i_j \boldsymbol{v}^j(\boldsymbol{x}).$$

V opačnom smere prebieha zámena súradníc podľa rovnice

$$\boldsymbol{v}^i(\boldsymbol{x}) = \sum_j Q^i_j \boldsymbol{u}^j(\boldsymbol{x}) = Q^i_j \boldsymbol{u}^j(\boldsymbol{x}).$$

(b) Pre kovektor $\eta \in V^*$ sa pôvodná transformačná rovnica $(\eta)_{\boldsymbol{\beta}^*} = \boldsymbol{P}^{\mathsf{T}}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot (\eta)_{\boldsymbol{\alpha}^*}$ zapíše takto:

$$\boldsymbol{v}_i(\eta) = \sum_i P^j_i \boldsymbol{u}_j(\eta) = P^j_i \boldsymbol{u}_j(\eta).$$

(c) Transformačný vzťah $\boldsymbol{B} = \boldsymbol{P}^{\mathsf{T}}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \cdot \boldsymbol{A} \cdot \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ medzi maticami $\boldsymbol{A} = [F]_{\boldsymbol{\alpha}} = (A_{ij})$, $\boldsymbol{B} = [F]_{\boldsymbol{\beta}} = (B_{kl})$ bilineárnej formy $F\colon V \times V \to K$ vzhľadom na bázy $\boldsymbol{\alpha}$ resp. $\boldsymbol{\beta}$ prejde do podoby

$$B_{kl} = \sum_{i,j} P^i_k A_{ij} P^j_l = A_{ij} P^i_k P^j_l.$$

(d) Konečne transformačný vzťah $\boldsymbol{B} = \boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\alpha}} \cdot \boldsymbol{A} \cdot \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ medzi maticami $\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha}} = \left(A^i_j\right)$, $\boldsymbol{B} = (\varphi)_{\boldsymbol{\beta}} = \left(B^k_l\right)$ lineárnej transformácie $\varphi\colon V \to V$ vzhľadom na bázy $\boldsymbol{\alpha}$ resp. $\boldsymbol{\beta}$ má tvar

$$B^k_l = \sum_{i,j} Q^k_i A^i_j P^j_l = A^i_j Q^k_i P^j_l.$$

Obzvlášť v poslednom príklade (d) pekne vidno ako sa pri prechode od súradníc v báze $\boldsymbol{\alpha}$ k súradniciam v báze $\boldsymbol{\beta}$ dvojica i, k kontravariantných (horných) indexov transformuje prostredníctvom prvku Q^k_i matice $\boldsymbol{Q} = \boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\alpha}} = \boldsymbol{P}^{-1}$, teda kontravariantne, kým dvojica j, l kovariantných (dolných) indexov sa transformuje prostredníctvom prvku P^j_l matice $\boldsymbol{P} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$, teda kovariantne. Vo všetkých príkladoch nám párovanie horných a dolných sumačných indexov a požiadavka zhody voľných indexov na oboch stranách rovností poskytujú účinný kontrolný prostriedok. S trochou cviku možno transformačné rovnice písať tak povediac mechanicky. Niečo podobné platí aj pre transformáciu súradníc všeobecného tenzora.

32.5.5. Veta. *Nech $T \in \mathcal{T}^p_q(V)$ je tenzor typu $\binom{p}{q}$ nad n-rozmerným vektorovým priestorom V, $\boldsymbol{A} = (T)_{\boldsymbol{\alpha}} = \left(A^{i_1\dots i_p}_{j_1\dots j_q}\right)$, $\boldsymbol{B} = (T)_{\boldsymbol{\beta}} = \left(B^{k_1\dots k_p}_{l_1\dots l_q}\right)$ sú jeho* *súradnice v bázach* $\boldsymbol{\alpha}$ *resp.* $\boldsymbol{\beta}$ *a* $\boldsymbol{P} = \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \left(P^i_j\right)$ *a* $\boldsymbol{Q} = \boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\alpha}} = \left(Q^i_j\right)$ *sú matice prechodu medzi nimi. Potom*

$$B^{k_1\dots k_p}_{l_1\dots l_q} = \sum_{\substack{i_1,\dots,i_p\\ j_1,\dots,j_q}} Q^{k_1}_{i_1}\dots Q^{k_p}_{i_p} A^{i_1\dots i_p}_{j_1\dots j_q} P^{j_1}_{l_1}\dots P^{j_q}_{l_q} = A^{i_1\dots i_p}_{j_1\dots j_q} Q^{k_1}_{i_1}\dots Q^{k_p}_{i_p} P^{j_1}_{l_1}\dots P^{j_q}_{l_q}.$$

Dôkaz prenechávame ako cvičenie čitateľovi.

Tradične sa tenzory vo fyzikálnej, technickej, no i v (najmä staršej) matematickej literatúre chápu ako systémy skalárov $T = \left(T^{i_1\dots i_p}_{j_1\dots j_q}\right)$ s rôznym počtom horných a dolných indexov. Tenzor je tak stotožňovaný s jeho súradnicami v nejakej, podľa možnosti kanonickej, báze daného vektorového priestoru. Vzápätí je však potrebné dodať, že jeden taký systém skalárov ešte netvorí tenzor – tenzor je totiž až celý systém $\left(T_{\boldsymbol{\alpha}}\right)$ takýchto indexovaných systémov skalárov, jeden pre každú bázu $\boldsymbol{\alpha}$ priestoru V, ktoré sa (pri označeni $T_{\boldsymbol{\alpha}} = (T)_{\boldsymbol{\alpha}} = \boldsymbol{A}$, $T_{\boldsymbol{\beta}} = (T)_{\boldsymbol{\beta}} = \boldsymbol{B}$) transformujú podľa rovnice z vety 32.5.5. Za všetkými jednotlivými súradnicovými reprezentáciami tenzora sa však skrýva jednoznačne určený matematický objekt, ktorý ich drží pohromade. Tým je práve multilineárna forma $T\colon (V^*)^p \times V^q \to K$. Z toho dôvodu dávame v našom kurze prednosť *invariantnej*, t. j. na jednotlivých bázach nezávislej reprezentácii tenzorov.

32.5.6. Príklad. Násobenie $M(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{xy}$ v lineárnej algebre $\mathcal{A}$ nad poľom K je bilineárne zobrazenie $M\colon \mathcal{A}\times\mathcal{A}\to\mathcal{A}$. To znamená, že $M \in \mathcal{T}^1_2(\mathcal{A})$ je tenzor typu $\binom{1}{2}$ nad vektorovým priestorom $\mathcal{A}$. Systém štruktúrnych konštánt algebry $\mathcal{A}$ vzhľadom na bázu $\boldsymbol{\alpha}$ potom nie je nič iné ako súradnice tenzora M v tejto báze, čiže

$$\{\mathcal{A}\}_{\boldsymbol{\alpha}} = \left(a^k_{ij}\right) = (M)_{\boldsymbol{\alpha}}.$$

Ak

$$\{\mathcal{A}\}_{\boldsymbol{\beta}} = \left(b^k_{ij}\right) = (M)_{\boldsymbol{\beta}}$$

je systém štruktúrnych konštánt algebry $\mathcal{A}$ vzhľadom na druhú bázu $\boldsymbol{\beta}$, tak

$$b^k_{ij} = \sum_{r,s,t} Q^k_r a^r_{st} P^s_i P^t_j = Q^k_r a^r_{st} P^s_i P^t_j$$

kde $\boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \left(P^i_j\right)$, $\boldsymbol{P}_{\boldsymbol{\beta},\boldsymbol{\alpha}} = \left(Q^i_j\right)$ sú príslušné matice prechodu.

Naopak, každý tenzor $M \in \mathcal{T}^1_2(V)$ určuje na vektorovom priestore V štruktúru lineárnej algebry. To znamená, že priestor tenzorov $\mathcal{T}^1_2(V)$ možno chápať ako *vektorový priestor všetkých lineárnych algebier s podkladovým vektorovým priestorom* V.

32.6 Operácie na tenzoroch

Na priestoroch tenzorov $\mathcal{T}^p_q(V)$ nad konečnorozmerným vektorovým priestorom V možno zaviesť niekoľko ďalších algebraických operácií. O ich súradnicovej podobe sa niekedy zvykne mierne ironicky hovoriť ako o *indexovej gymnastike.*

32.6.1. Súčin tenzorov. *Súčin $S \otimes T$ tenzorov $S \in \mathcal{T}^p_q(V)$, $T \in \mathcal{T}^r_s(V)$*je tenzor $S \otimes T \in \mathcal{T}^{p+r}_{q+s}(V)$, t. j. multilineárna forma daná predpisom

$$(S \otimes T)(\eta^1, \dots, \eta^p, \vartheta^1, \dots, \vartheta^r, \boldsymbol{x}_1, \dots, \boldsymbol{x}_q, \boldsymbol{y}_1, \dots, \boldsymbol{y}_s) \\ = S(\eta^1, \dots, \eta^p, \boldsymbol{x}_1, \dots, \boldsymbol{x}_q)\, T(\vartheta^1, \dots, \vartheta^r, \boldsymbol{y}_1, \dots, \boldsymbol{y}_s)$$

pre $\eta^1, \dots, \eta^p, \vartheta^1, \dots, \vartheta^r \in V^*$, $\boldsymbol{x}_1, \dots, \boldsymbol{x}_q, \boldsymbol{y}_1, \dots, \boldsymbol{y}_s \in V$. Pre každú štvoricu prirodzených čísel p, q, r, s je tak určená bilineárna binárna operácia

$$\otimes\colon \mathcal{T}^p_q(V) \times \mathcal{T}^r_s(V) \to \mathcal{T}^{p+r}_{q+s}(V),$$

indukujúca lineárny izomorfizmus vektorových priestorov $\mathcal{T}^p_q(V) \otimes \mathcal{T}^r_s(V) \cong \mathcal{T}^{p+r}_{q+s}(V)$.

Ak $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ je nejaká báza priestoru V, a $\boldsymbol{A} = \big(A^{i_1 \dots i_p}_{j_1 \dots j_q}\big) = (S)_{\boldsymbol{\beta}}$, $\boldsymbol{B} = \big(B^{k_1 \dots k_r}_{l_1 \dots l_s}\big) = (T)_{\boldsymbol{\beta}}$ sú súradnice tenzorov S resp. T vzhľadom na ňu, tak pre súradnice $\boldsymbol{C} = \big(C^{i_1 \dots i_p k_1 \dots k_r}_{j_1 \dots j_q l_1 \dots l_s}\big) = (S \otimes T)_{\boldsymbol{\beta}}$ tenzora $S \otimes T$ v báze $\boldsymbol{\beta}$ platí

$$C^{i_1 \dots i_p\, k_1 \dots k_r}_{j_1 \dots j_q\, l_1 \dots l_s} = A^{i_1 \dots i_p}_{j_1 \dots j_q}\, B^{k_1 \dots k_r}_{l_1 \dots l_s}.$$

Píšeme tiež $\boldsymbol{C} = \boldsymbol{A} \otimes \boldsymbol{B}$ a $(p + q + r + s)$-rozmernú maticu $\boldsymbol{C}$ nazývame *tenzorovým súčinom matíc* $\boldsymbol{A}$ a $\boldsymbol{B}$.

Z priestorov tenzorov $\mathcal{T}^p_q(V)$ jednotlivých typov vytvoríme ich priamy súčet

$$\mathcal{T}(V) = \bigoplus_{p,q \in \mathbb{N}} \mathcal{T}^p_q(V)$$

(pozri cvičenie 6.19). Dielčie tenzorové súčiny $\otimes\colon \mathcal{T}^p_q(V) \times \mathcal{T}^r_s(V) \to \mathcal{T}^{p+r}_{q+s}(V)$ možno potom jednoznačne rozšíriť do bilineárnej binárnej operácie

$$\otimes\colon \mathcal{T}(V) \times \mathcal{T}(V) \to \mathcal{T}(V).$$

Výsledný tenzorový súčin $\otimes$ určuje na nekonečnorozmernom priestore $\mathcal{T}(V)$ štruktúru asociatívnej lineárnej algebry s jednotkou $1 \in K = \mathcal{T}^0_0(V)$, nazývanej *algebrou tenzorov* nad vektorovým priestorom V. Algebra $\big(\mathcal{T}(V), \otimes\big)$ je prirodzene *graduovaná* do podpriestorov $\mathcal{T}^p_q(V)$ (pozri paragraf 30.2). Podpriestor kontravariantných tenzorov $\mathcal{T}_0(V) = \bigoplus_{p \in \mathbb{N}} \mathcal{T}^p_0(V)$ a podpriestor kovariantných tenzorov $\mathcal{T}^0(V) = \bigoplus_{q \in \mathbb{N}} \mathcal{T}^0_q(V)$ tvoria podalgebry tenzorovej algebry $\mathcal{T}(V)$, ktoré sú samy tiež prirodzene graduované. Kým však graduácie podalgebier kontra- resp. kovariantných tenzorov sú indexované prirodzenými číslami, graduácia celej tenzorovej algebry je indexovaná dvojicami čísel $p, q \in \mathbb{N}$.

32.6.2. Kontrakcie alebo zovšeobecnené stopy tenzora. Sumácia cez vybraný horný a dolný index určuje lineárne zobrazenie $\mathcal{T}_q^p(V) \to \mathcal{T}_{q-1}^{p-1}(V)$ nazývané *kontrakciou* alebo *zúžením* alebo tiež *zovšeobecnenou stopou*. Názornejšie je opísať toto zobrazenie najprv v jeho súradnicovej podobe.

Nech $n = \dim V$ a $p, q \geq 1$, $k \leq p$, $l \leq q$. Ak $\boldsymbol{A} = \big(A^{i_1 \dots i_p}_{j_1 \dots j_q}\big)$ je matica nad poľom K s p hornými a q dolnými indexmi nadobúdajúcimi hodnoty od 1 do n, tak maticu $\boldsymbol{B}$ s $p-1$ hornými a $q-1$ dolnými indexmi s jednotlivými zložkami

$$B^{i_1 \dots i_{p-1}}_{j_1 \dots j_{q-1}} = \sum_{i=1}^{n} A^{i_1 \dots i_{k-1}\, i\, i_k \dots i_{p-1}}_{j_1 \dots j_{l-1}\, i\, j_l \dots j_{q-1}} = A^{i_1 \dots i_{k-1}\, i\, i_k \dots i_{p-1}}_{j_1 \dots j_{l-1}\, i\, j_l \dots j_{q-1}}$$

nazývame *kontrakciou*, prípadne *zúžením* alebo *zovšeobecnenou stopou matice* $\boldsymbol{A}$ podľa k-teho horného a l-tého dolného indexu; píšeme $\boldsymbol{B} = \operatorname{tr}_l^k \boldsymbol{A}$. (Uvedená formula si vyžaduje miernu modifikáciu, ak sa niektorý z indexov i_k, j_l nachádza v krajnej polohe – to však prenechávame čitateľovi.)

Zrejme pre dvojrozmernú maticu $\boldsymbol{A} = (A^i_j)$ kontrakcia $\operatorname{tr}_1^1 \boldsymbol{A} = \sum_i A^i_i = A^i_i$ splýva s jej „obyčajnou" stopou $\operatorname{tr} \boldsymbol{A}$.

Obvyklý súčin dvojrozmerných matíc $\boldsymbol{A} = \big(A^i_j\big)$, $\boldsymbol{B} = \big(B^j_k\big)$ možno vyjadriť ako kontrakciu ich tenzorového súčinu

$$\boldsymbol{A} \cdot \boldsymbol{B} = \operatorname{tr}_1^2(\boldsymbol{A} \otimes \boldsymbol{B}).$$

Podobne možno pomocou kontrakcií a tenzorového súčinu vyjadriť hodnotu bilineárnej formy $F(\boldsymbol{x}, \boldsymbol{y}) = A_{ij} x^i y^j$ na K^n:

$$F(\boldsymbol{x}, \boldsymbol{y}) = \operatorname{tr}_{12}^{12}(\boldsymbol{A} \otimes \boldsymbol{x} \otimes \boldsymbol{y}) = \big(\operatorname{tr}_1^1 \circ \operatorname{tr}_2^2\big)\big(\boldsymbol{A} \otimes (\boldsymbol{x} \otimes \boldsymbol{y})\big).$$

Presvedčte sa o jednom i druhom rozpisom do jednotlivých zložiek.

Ak $\boldsymbol{A} = (T)_{\boldsymbol{\beta}}$ je matica súradníc tenzora $T \in \mathcal{T}_q^p(V)$ v báze $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ priestoru V, tak jej kontrakcia $\operatorname{tr}_l^k \boldsymbol{A}$ tvorí súradnice tenzora $\operatorname{tr}_l^k T \in \mathcal{T}_{q-1}^{p-1}(V)$ v tej istej báze. Tento tenzor ako multilineárna forma funguje podľa predpisu

$$\begin{aligned}\big(\operatorname{tr}_l^k T\big)&\big(\eta^1, \dots, \eta^{p-1}, \boldsymbol{x}_1, \dots, \boldsymbol{x}_{q-1}\big)\\ &= \sum_{i=1}^{n} T\big(\eta^1, \dots, \eta^{k-1}, \boldsymbol{v}^i, \eta^k, \dots, \eta^{p-1}, \boldsymbol{x}_1, \dots, \boldsymbol{x}_{l-1}, \boldsymbol{v}_i, \boldsymbol{x}_l, \dots, \boldsymbol{x}_{q-1}\big)\end{aligned}$$

pre $\eta^1, \dots, \eta^{p-1} \in V^*$, $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{q-1} \in V$. Presvedčte sa o tom samostatne a dokážte, že multilineárna forma $\operatorname{tr}_l^k T \in \mathcal{T}_{q-1}^{p-1}(V)$ nezávisí na báze $\boldsymbol{\beta}$.

32.6.3. Permutácie indexov. Ľubovoľná dvojica permutácií ϱ množiny $\{1,\dots,p\}$ a σ množiny $\{1,\dots,q\}$ indukuje lineárny izomorfizmus $T \mapsto T^\varrho_\sigma\colon \mathcal{T}^p_q(V) \to \mathcal{T}^p_q(V)$, daný predpisom

$$T^\varrho_\sigma\bigl(\eta^1,\dots,\eta^p,\boldsymbol{x}_1,\dots,\boldsymbol{x}_q\bigr) = T\bigl(\eta^{\varrho(1)},\dots,\eta^{\varrho(p)},\boldsymbol{x}_{\sigma(1)},\dots,\boldsymbol{x}_{\sigma(q)}\bigr)$$

pre $T \in \mathcal{T}^p_q(V)$ a $\eta^1,\dots,\eta^p \in V^*$, $\boldsymbol{x}_1,\dots,\boldsymbol{x}_q \in V$. Inak povedané, poradie kovektorových argumentov η^k poprehadzujeme permutáciou ϱ a poradie vektorových argumentov $\boldsymbol{x}_l$ permutáciou σ. Znak identickej permutácie ι (či už na množine $\{1,\dots,p\}$ alebo $\{1,\dots,q\}$) nepíšeme, teda $T_\sigma = T^\iota_\sigma$ a $T^\varrho = T^\varrho_\iota$.

Ak matica $\boldsymbol{A} = \bigl(A^{i_1\dots i_p}_{j_1\dots j_q}\bigr)$ tvorí súradnice tenzora $T \in \mathcal{T}^p_q(V)$ v báze $\boldsymbol{\beta}$ priestoru V, tak zložky matice $\boldsymbol{A}^\varrho_\sigma = \boldsymbol{B} = \bigl(B^{i_1\dots i_p}_{j_1\dots j_q}\bigr)$ súradníc tenzora T^ϱ_σ v tejto báze sú dané rovnosťou

$$B^{i_1\dots i_p}_{j_1\dots j_q} = A^{i_{\varrho(1)}\dots i_{\varrho(p)}}_{j_{\sigma(1)}\dots j_{\sigma(q)}}.$$

To znamená, že poradie horných indexov je poprehadzované permutáciou ϱ a poradie dolných indexov permutáciou σ.

32.7 Tenzory v (pseudo)euklidovských priestoroch

Predpokladajme, že V je konečnorozmerný vektorový priestor nad poľom $\mathbb{R}$ s bázou $\boldsymbol{\beta} = (\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ a $G\colon V\times V \to \mathbb{R}$ je symetrická regulárna bilineárna forma na V s Gramovou maticou $[G]_{\boldsymbol{\beta}} = \boldsymbol{g} = (g_{ij})$, kde $g_{ij} = G(\boldsymbol{v}_i,\boldsymbol{v}_j)$. Teda V je euklidovský alebo pseudoeuklidovský priestor a G je skalárny resp. pseudoskalárny súčin na V (navyše bez ujmy na všeobecnosti môžeme predpokladať, že G nie je záporne definitná). Formu G obvykle nazývame *metrický* prípadne *pseudometrický tenzor*.

Hoci prípad poľa $\mathbb{R}$ je najdôležitejší, väčšina úvah z tohto paragrafu si zachováva platnosť v konečnorozmerných vektorových priestoroch nad ľubovoľným poľom K vybavených symetrickou regulárnou bilineárnou formou $G\colon V\times V \to K$. V prostredí takýchto priestorov možno indexovú gymnastiku obohatiť o ďalší súbor cvikov. Každý vektor $\boldsymbol{x} \in V$ totiž určuje lineárnu formu (kovektor) $\downarrow\!\boldsymbol{x} \in V^*$ predpisom $(\downarrow\!\boldsymbol{x})(\boldsymbol{y}) = G(\boldsymbol{x},\boldsymbol{y})$ pre $\boldsymbol{y} \in V$. Naopak, každý kovektor $\eta \in V^*$ má tvar $\eta = \downarrow\boldsymbol{x}$ pre j ednoznačne určený vektor $\boldsymbol{x} = \uparrow\eta$. Priestor V tak možno stotožniť s jeho duálom V^* prostredníctvom dvojice navzájom inverzných lineárnych izomorfizmov $\boldsymbol{x} \mapsto \downarrow\!\boldsymbol{x}$, $\eta \mapsto \uparrow\!\eta$.

V súradniciach funguje izomorfizmus $\boldsymbol{x} \mapsto \downarrow\!\boldsymbol{x}$ ako *spustenie indexu* (čo je aj dôvod pre jeho označenie šípkou $\downarrow$): Ak $(\boldsymbol{x})_{\boldsymbol{\beta}} = (x^i)$, tak $(\downarrow\!\boldsymbol{x})_{\boldsymbol{\beta}} = (x_j)$, kde

$$x_j = \sum_i g_{ij}x^i = g_{ij}x^i.$$

V opačnom smere ide o *zdvih indexu*

$$x^i = \sum_j g^{ij} x_j = g^{ij} x_j,$$

kde $g_{ij} g^{jk} = \delta_i^k$, čiže $\big(g^{jk}\big) = \boldsymbol{g}^{-1}$ je inverzná matica k matici $\boldsymbol{g}$. Systémy skalárov (x^i), (x_j) sa niekedy nazývajú *kontravariantné* resp. *kovariantné súradnice* (toho istého kontravariantného) *vektora* $\boldsymbol{x}$.

Rovnako možno popísať aj súradnicovú podobu izomorfizmu $\eta \mapsto \uparrow \eta$, ktorá sa opäť prejaví *zdvihom indexu*: Ak $(\eta)_{\boldsymbol{\beta}} = (\eta_j)$, tak $(\uparrow \eta)_{\boldsymbol{\beta}} = (\eta^i)$, kde

$$\eta^i = \sum_j g^{ij} \eta_j = g^{ij} \eta_j.$$

V opačnom smere ide o *spustenie indexu*

$$\eta_j = \sum_i g_{ij} \eta^i = g_{ij} \eta^i.$$

Systémy skalárov (η_j), (η^i) sa zvyknú nazývať *kovariantné* resp. *kontravariantné súradnice kovektora* η.

Tým sa otvára možnosť voľného prechodu oboma smermi medzi vektormi a kovektormi. Ak (V, G) je euklidovský priestor a báza $\boldsymbol{\beta}$ je navyše ortonormálna, t. j. $g_{ij} = \delta_{ij}$, tak aj $g^{jk} = \delta^{jk}$ a $x_j = x^j$, $\eta^i = \eta_i$. Potreba rozlišovať medzi vektormi a kovektormi ako aj medzi hornými a dolnými indexmi sa potom celkom vytráca. V pseudoeuklidovskom priestore však (v dôsledku indefinitnosti formy G) pre žiadnu bázu $\boldsymbol{\beta}$ neplatí $[G]_{\boldsymbol{\beta}} = \boldsymbol{I}_n$, preto $(g_{ij}) \neq (\delta_{ij})$ a súradnice (x^i), (x_i) či (η_j), (η^j) sa pre niektoré (ko)vektory nevyhnutne líšia.

Operácie spustenia a zdvihu indexov možno z vektorov resp. kovektorov rozšíriť aj na všeobecné tenzory. V (pseudo)euklidovskom priestore (V, G) máme za predpokladu $p \geq 1$ pre každú dvojicu $k \leq p$, $l \leq q + 1$ lineárny izomorfizmus tenzorových priestorov

$$\downarrow_l^k \colon \mathcal{T}_q^p(V) \to \mathcal{T}_{q+1}^{p-1}(V),$$

nazývaný *spustenie k-teho horného indexu na l-té miesto*. Podobne, za predpokladu $q \geq 1$ máme pre každú dvojicu $l \leq q$, $k \leq p+1$ lineárny izomorfizmus tenzorových priestorov

$$\uparrow_l^k \colon \mathcal{T}_q^p(V) \to \mathcal{T}_{q-1}^{p+1}(V),$$

nazývaný *zdvih l-teho dolného indexu na k-té miesto*. V invariantnom zápise fungujú tieto zobrazenia podľa predpisov

$$\begin{aligned}
&\big(\downarrow_l^k T\big)\big(\eta^1, \ldots, \eta^{p-1}, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{q+1}\big) \\
&\qquad = T\big(\eta^1, \ldots, \eta^{k-1}, \downarrow \boldsymbol{x}_l, \eta^k, \ldots, \eta^{p-1}, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{l-1}, \boldsymbol{x}_{l+1}, \ldots, \boldsymbol{x}_{q+1}\big)
\end{aligned}$$

pre $\eta^1, \dots, \eta^{p-1} \in V^*$, $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{q+1} \in V$, resp.

$$\begin{aligned}&\big({\uparrow}_l^k\, T\big)\big(\eta^1, \dots, \eta^{p+1}, \boldsymbol{x}_1, \dots, \boldsymbol{x}_{q-1}\big)\\&\qquad = T\big(\eta^1, \dots, \eta^{k-1}, \eta^{k+1}, \dots, \eta^{p+1}, \boldsymbol{x}_1, \dots, \boldsymbol{x}_{l-1}, {\uparrow}\,\eta^k, \boldsymbol{x}_l, \dots, \boldsymbol{x}_{q-1}\big)\end{aligned}$$

pre $\eta^1, \dots, \eta^{p+1} \in V^*$, $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{q-1} \in V$. [4](Uvedené formuly si opäť vyžadujú miernu modifikáciu pre krajnú polohu niektorého z indexov i_k, j_l.)

V súradniciach tenzora vidno spúšťanie a dvíhanie indexov jasnejšie: Ak $(T)_{\boldsymbol{\beta}} = \big(A^{i_1\dots i_p}_{j_1\dots j_q}\big)$, tak $\big({\downarrow}_l^k\, T\big)_{\boldsymbol{\beta}} = \big(B^{i_1\dots i_{p-1}}_{j_1\dots j_{q+1}}\big)$ a $\big({\uparrow}_l^k\, T\big)_{\boldsymbol{\beta}} = \big(C^{i_1\dots i_{p+1}}_{j_1\dots j_{q-1}}\big)$, kde

$$\begin{aligned}B^{i_1\dots i_{p-1}}_{j_1\dots j_{q+1}} &= \sum_i g_{ij_l} A^{i_1\dots i_{k-1} i i_k \dots i_{p-1}}_{j_1\dots j_{l-1}\ j_{l+1}\dots j_{q+1}} = g_{ij_l} A^{i_1\dots i_{k-1} i i_k \dots i_{p-1}}_{j_1\dots j_{l-1}\ j_{l+1}\dots j_{q+1}},\\C^{i_1\dots i_{p+1}}_{j_1\dots j_{q-1}} &= \sum_j g^{i_k j} A^{i_1\dots i_{k-1}\ i_{k+1}\dots i_{p+1}}_{j_1\dots j_{l-1} j j_l \dots j_{q-1}} = g^{i_k j} A^{i_1\dots i_{k-1}\ i_{k+1}\dots i_{p+1}}_{j_1\dots j_{l-1} j j_l \dots j_{q-1}}.\end{aligned}$$

Samozrejme, spúšťať alebo dvíhať môžeme aj viacero indexov, v krajnom prípade všetky. Označenia ako ${\downarrow}_{l\,s}^{kr}$, ${\uparrow}_{1\dots q}^{k_1\dots k_q}$ a pod. hovoria samy za seba. Napríklad spustením oboch horných indexov v tenzore T typu $\binom{2}{0}$ so súradnicami $(T)_{\boldsymbol{\beta}} = \big(A^{ij}\big)$ dostaneme tenzor ${\downarrow}_{12}^{12}\, T$ typu $\binom{0}{2}$ so súradnicami $\big({\downarrow}_{12}^{12}\, T\big)_{\boldsymbol{\beta}} = \big(A_{kl}\big)$, kde

$$A_{kl} = \sum_{ij} g_{ik} g_{jl} A^{ij} = g_{ik} g_{jl} A^{ij}.$$

Ďalšie podrobnosti ponechávame na samostatné premyslenie čitateľovi.

Aj samotná forma G je tenzor typu $\binom{0}{2}$. Zdvihom oboch jej indexov dostaneme tenzor ${\uparrow}_{12}^{12}\, G$ typu $\binom{2}{0}$. Vďaka symetrii matice $\boldsymbol{g}$ a rovnosti $\boldsymbol{g}^{-1} = \boldsymbol{g}^{-1} \cdot \boldsymbol{g} \cdot \boldsymbol{g}^{-1}$ vidíme, že

$$g^{ij} = \sum_{kl} g^{ik} g_{kl} g^{lj} = g^{ik} g^{jl} g_{kl},$$

teda súradnice $\big({\uparrow}_{12}^{12}\, G\big)_{\boldsymbol{\beta}}$ naozaj tvorí inverzná matica $\boldsymbol{g}^{-1} = \big(g^{ij}\big)$ k matici $\boldsymbol{g}$.

V dôsledku komutatívnosti tenzorového súčinu nehrá miesto, kam sa nejaký index spúšťa alebo dvíha, až takú dôležitú úlohu. V konkrétnych prípadoch sa preto ani nezvykne explicitne vyznačovať; implicitne sa predpokladá,

[4]Značenie ${\downarrow}_l^k\, T$, ${\uparrow}_l^k\, T$ spustenia a zdvihu indexov nie je celkom obvyklé, no na rozdiel od bežne používaného značenia (či skôr neznačenia lebo poloznačenia) je názorné a jednoznačné, a preto vhodné na potreby výuky. Pri konkrétnych výpočtoch by bolo zrejme zbytočne pedantské a ťažkopádne. Ak je z kontextu jasné, ktorý index spúšťame resp. dvíhame a na ktoré miesto, môžeme použiť obvyklejšie „hudobné“ značenie ${\downarrow}\, T = T^\flat$, resp. ${\uparrow}\, T = T^\sharp$.

že index spúšťame alebo dvíhame „na to isté miesto" pod resp. nad ním. Kvôli tomu sa indexy tenzorov často zvyknú písať tak, aby žiaden pár pozostávajúci z jedného horného a jedného dolného indexu neležal v jednej zvislej línii. Príslušný index sa potom spustí/zdvihne na voľné miesto pod/nad ním.

V diferenciálnej geometrii napr. *Riemannov tenzor krivosti* (presnejšie jeho súradnice) zapisujeme v tvare $R^i{}_{klm}$. Spustením horného indexu získame tenzor $R_{jklm} = g_{ij}R^i{}_{klm}$. Keby sme hodlali spustiť horný index napr. na druhé miesto, pôvodný tenzor by sme zapísali v tvare $R_k{}^i{}_{lm}$ a po spustení by sme dostali tenzor $R_{kjlm} = g_{ij}R_k{}^i{}_{lm}$.

Ak $\boldsymbol{\alpha}$ je iná báza priestoru V, tak spúšťanie a dvíhanie indexov v súradnicich tenzorov vzhľadom na bázu $\boldsymbol{\alpha}$ funguje rovnako ako v súradnicich vzhľadom na bázu $\boldsymbol{\beta}$, len s tým rozdielom, že úlohu matíc $\boldsymbol{g} = (g_{ij}) = [G]_{\boldsymbol{\beta}}$, $\boldsymbol{g}^{-1} = (g^{ij})$ prevezmú matice $\boldsymbol{f} = (f_{ij}) = [G]_{\boldsymbol{\alpha}}$, $\boldsymbol{f}^{-1} = (f^{ij})$. Býva zvykom obmedziť sa len na ortonormálne bázy $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ pre ktoré platí $[G]_{\boldsymbol{\alpha}} = [G]_{\boldsymbol{\beta}} = \operatorname{diag}(\boldsymbol{I}_r, -\boldsymbol{I}_s)$, kde (r, s) je signatúra formy G. Vo všetkých takých bázach spúšťanie aj dvíhanie indexov funguje s rovnakými koeficientmi $g_{ij} = g^{ij}$. Čitateľ si hádam ešte spomenie, že v Minkowského časopriestore (kapitola 16) sme takýmto bázam hovorili *inerciálne*.

32.7.1. Príklad. Nech $M \subseteq \mathbb{R}^n$ je otvorená množina, $f\colon M \to \mathbb{R}$ je (spojito) diferencovateľná reálna funkcia n premenných a $\boldsymbol{a} \in M$ je ľubovoľný bod. V diferenciálnom počte funkcií viac premenných chápeme *gradient funkcie f* v bode $\boldsymbol{a}$ primárne ako vektor

$$\operatorname{grad} f(\boldsymbol{a}) = \nabla f(\boldsymbol{a}) = \left(\frac{\partial f(\boldsymbol{a})}{\partial x^1}, \dots, \frac{\partial f(\boldsymbol{a})}{\partial x^n}\right) = \sum_{i=1}^n \frac{\partial f(\boldsymbol{a})}{\partial x^i}\,\boldsymbol{e}_i = \frac{\partial f(\boldsymbol{a})}{\partial x^i}\,\boldsymbol{e}_i$$

vychádzajúci z bodu $\boldsymbol{a}$ a ukazujúci v smere, v ktorom sa hodnota funkcie f mení najrýchlejšie, s dĺžkou priamo úmernou rýchlosti tejto zmeny. *Derivácia funkcie f* v bode $\boldsymbol{a}$ *v smere vektora* $\boldsymbol{v} \in \mathbb{R}^n$ je definovaná ako limita

$$\nabla_{\boldsymbol{v}} f(\boldsymbol{a}) = \lim_{t\to 0} \frac{f(\boldsymbol{a} + t\boldsymbol{v}) - f(\boldsymbol{a})}{t}.$$

Špeciálne pre vektory $\boldsymbol{e}_i$ kanonickej ortonormálnej bázy smerové derivácie splývajú s parciálnymi deriváciami:

$$\nabla_{\boldsymbol{e}_i} f(\boldsymbol{a}) = \frac{\partial f(\boldsymbol{a})}{\partial x^i}.$$

Pre všeobecný vektor $\boldsymbol{v} \in \mathbb{R}^n$ platí

$$\nabla_{\boldsymbol{v}} f(\boldsymbol{a}) = \nabla f(\boldsymbol{a}) \cdot \boldsymbol{v} = \sum_{i=1}^n \frac{\partial f(\boldsymbol{a})}{\partial x^i}\, v^i = \frac{\partial f(\boldsymbol{a})}{\partial x^i}\, v^i,$$

kde $\cdot$ označuje štandardný skalárny súčin v $\mathbb{R}^n$. Na gradient funkcie f v bode $\boldsymbol{a}$ sa tak môžeme dívať aj ako na kovektor, t. j. lineárnu formu $\nabla f(\boldsymbol{a})\colon \mathbb{R}^n \to \mathbb{R}$, ktorá vektoru $\boldsymbol{v} \in \mathbb{R}^n$ priradí skalár

$$\nabla f(\boldsymbol{a})(\boldsymbol{v}) = \nabla_{\boldsymbol{v}} f(\boldsymbol{a}) = \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{a})}{\partial x^i}\, v^i = \frac{\partial f(\boldsymbol{a})}{\partial x^i}\, v^i.$$

Výraz pre totálny diferenciál funkcie f v bode $\boldsymbol{a}$

$$\mathrm{d}f(\boldsymbol{a}, \boldsymbol{x}) = \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{a})}{\partial x^i}\,(x^i - a^i) = \frac{\partial f(\boldsymbol{a})}{\partial x^i}\,(x^i - a^i) = \nabla f(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a})$$

je tak aplikáciou gradientu $\nabla f(\boldsymbol{a})$ ako lineárneho funkcionálu na vektor $\boldsymbol{x} - \boldsymbol{a}$.

V „plochých" priestoroch, ako sú napr. otvorené množiny M v euklidovskom priestore $\mathbb{R}^n$, je rozdiel medzi gradientom ako vektorom a gradientom ako kovektorom viac-menej formálny a nie príliš zreteľný. V zakrivenom priestore však prechod medzi oboma rolami gradientu sprostredkúva Riemannov metrický tenzor $g_{ij}(\boldsymbol{a})$ spojite závislý na bodoch $\boldsymbol{a} \in M$, ktorý určuje jeho geometriu (a vo všeobecnosti nesplýva s Kroneckerovým symbolom δ_{ij}). Gradient je tu primárne lineárna forma na dotykovom priestore v danom bode; zdvihom indexu z neho dostaneme vektor

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{n} g^{ij}(\boldsymbol{a})\, \frac{\partial f(\boldsymbol{a})}{\partial x^i} \right) \boldsymbol{e}_j = g^{ij}(\boldsymbol{a})\, \frac{\partial f(\boldsymbol{a})}{\partial x^i}\, \boldsymbol{e}_j.$$

Ak prestaneme fixovať bod $\boldsymbol{a} \in M$ a gradient budeme chápať ako funkciu $M \to \mathbb{R}^n$, stane sa z neho (ko)vektorové pole, ako je to obvyklé v diferenciálnej geometrii.

Na záver ešte poznamenajme, že spúšťať a dvíhať indexy možno aj bez predpokladu symetrie formy G – dôležitá je len jej regularita. Tak je tomu napr. v tzv. *symplektickej geometrii*, ktorá je zadaná pomocou *antisymetrickej* bilineárnej formy G a má o. i. značné využitie v hamiltonovskej mechanike. Pri absencii symetrie by však bolo treba viaceré formuly tohto paragrafu mierne upraviť.

Cvičenia

32.1. Dokážte tvrdenie 32.1.1.

32.2. Nech X, Y sú konečné množiny a K je pole.
(a) Dokážte rovnosť vektorových priestorov funkcií $K^X \otimes K^Y = K^{X \times Y}$ jednak porovnaním ich dimenzií, jednak explicitným vyjadrením každej funkcie $h\colon X \times Y$ v tvare $h = (f_1 \otimes g_1) + \ldots + (f_k \otimes g_k)$ pre $f_i \in K^X$, $g_i \in K^Y$. Odhadnite počet

potrebných sčítancov k v závislosti na počtoch $m = \# X$, $n = \# Y$.

(b) Ak aj pole K je konečné, pričom $\# K = q$, porovnajte počet prvkov tenzorového súčinu $K^X \otimes K^Y$ a priameho súčinu $K^X \times K^Y$.

(c) Aká je pravdepodobnosť, že náhodne vybraná funkcia $h \in K^{X \times Y}$ je rozložiteľná?

32.3. (a) Nech $\varphi\colon V \to U$ je lineárne zobrazenie. Dokážte, že predpisom $\varphi^*(\eta) = \eta \circ \varphi$ je definované lineárne zobrazenie $\varphi^*\colon U^* \to V^*$.

(b) Dokážte, že pre identické zobrazenie $\mathrm{id}_V\colon V \to V$ je $\mathrm{id}_V^* = \mathrm{id}_{V^*}\colon V^* \to V^*$ taktiež identické zobrazenie.

(c) Pre lineárne zobrazenia $\psi\colon W \to V$, $\varphi\colon V \to U$ platí $(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$. Nakreslite si komutatívny diagram a dokážte.
(V jazyku *teórie kategórií* hovoríme, že priradenia $V \mapsto V^*$, $\varphi \mapsto \varphi^*$ určujú *kontravariantný funktor* na kategórii vektorových priestorov nad poľom K a lineárnych zobrazení.)

32.4. (a) Nech $\eta\colon U \to W$, $\vartheta\colon V \to W$ sú lineárne zobrazenia. Overte, že ich tenzorový súčin $\eta \otimes \vartheta\colon U \times V \to W$ je naozaj bilineárne zobrazenie.

(b) Nech U je vektorový podpriestor priestoru funkcií K^X. Overte, že dosadením prvku $x \in X$ do funkcie $f \in U$, t. j. predpisom $\sigma^x(f) = f(x)$, je definovaná lineárna forma $\sigma^x\colon U \to K$.

(c) Dokážte, že priradením $x \mapsto \sigma^x$ je definované prosté zobrazenie $\sigma\colon X \to U^*$.

32.5. Nech V je vektorový priestor nad poľom K. Nájdite prirodzené izomorfizmy vektorových priestorov

(a) $K \otimes V \cong V$; (b) $K^n \otimes V \cong V^n$ pre $n \geq 0$;

(c) $K^m \otimes K^n \cong K^{m \times n}$; (d) $K^{m \times n} \otimes K^{r \times s} \cong K^{mr \times ns}$.

32.6. *Kroneckerovým súčinom vektorov* $\boldsymbol{x} \in K^m$, $\boldsymbol{y} \in K^n$ nazývame maticu $\boldsymbol{x} \otimes \boldsymbol{y} = (x_i y_j) \in K^{m \times n}$. Kroneckerov súčin $\boldsymbol{x} \otimes \boldsymbol{y}$ niekedy nazývame priamo tenzorovým súčinom vektorov $\boldsymbol{x}$, $\boldsymbol{y}$.

(a) Dokážte, že pre hodnosť matice $\boldsymbol{x} \otimes \boldsymbol{y}$ platí $h(\boldsymbol{x} \otimes \boldsymbol{y}) \leq 1$, pričom $h(\boldsymbol{x} \otimes \boldsymbol{y}) = 1 \;\Leftrightarrow\; \boldsymbol{x} \neq \boldsymbol{0}$ & $\boldsymbol{y} \neq \boldsymbol{0}$.

(b) *Matica* $\boldsymbol{A} \in K^{m \times n}$ sa nazýva *rozložiteľná*, ak $\boldsymbol{A} = \boldsymbol{x} \otimes \boldsymbol{y}$ pre nejaké vektory $\boldsymbol{x} \in K^m$, $\boldsymbol{y} \in K^n$. Dokážte, že matica $\boldsymbol{A} \in K^{m \times n}$ je rozložiteľná práve vtedy, keď je rozložiteľná ako tenzor $\boldsymbol{A} \in K^m \otimes K^n$, a to je ekvivalentné s podmienkou $h(\boldsymbol{A}) \leq 1$.

(c) Dokážte, že hodnosť matice $\boldsymbol{A} \in K^{m \times n}$ splýva s najmenším prirodzeným číslom k, pre ktoré sa $\boldsymbol{A}$ dá vyjadriť ako súčet $\boldsymbol{A} = (\boldsymbol{x}_1 \otimes \boldsymbol{y}_1) + \ldots + (\boldsymbol{x}_k \otimes \boldsymbol{y}_k)$ k nerozložiteľných matíc $\boldsymbol{x}_i \otimes \boldsymbol{y}_i$.

32.7. *Kroneckerovým súčinom matíc* $\boldsymbol{A} \in K^{m \times n}$, $\boldsymbol{B} \in K^{r \times s}$ nazývame maticu

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \ldots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \ldots & a_{2n}\boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & \ldots & a_{mn}\boldsymbol{B} \end{pmatrix} \in K^{mr \times ns}.$$

Kroneckerov súčin $\boldsymbol{A} \otimes \boldsymbol{B}$ niekedy nazývame priamo tenzorovým súčinom matíc

$\boldsymbol{A}$, $\boldsymbol{B}$. Dokážte nasledujúce rovnosti:
(a) $\boldsymbol{I}_m \otimes \boldsymbol{I}_n = \boldsymbol{I}_{mn}$;

(b) $(\boldsymbol{A}_1 \otimes \boldsymbol{B}_1) \cdot (\boldsymbol{A}_2 \otimes \boldsymbol{B}_2) = \boldsymbol{A}_1\boldsymbol{A}_2 \otimes \boldsymbol{B}_1\boldsymbol{B}_2$ pre $\boldsymbol{A}_1, \boldsymbol{A}_2 \in K^{m\times n}$, $\boldsymbol{B}_1, \boldsymbol{B}_2 \in K^{r\times s}$;

(c) $(\boldsymbol{A} \otimes \boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} \otimes \boldsymbol{B}^{-1}$, ak $\boldsymbol{A} \in K^{n\times n}$, $\boldsymbol{B} \in K^{s\times s}$ sú regulárne;

(d) pre hodnosť matice $\boldsymbol{A} \otimes \boldsymbol{B}$ platí $h(\boldsymbol{A} \otimes \boldsymbol{B}) = h(\boldsymbol{A})\, h(\boldsymbol{B})$;

(e) $\operatorname{tr}(\boldsymbol{A} \otimes \boldsymbol{B}) = \operatorname{tr} \boldsymbol{A} \operatorname{tr} \boldsymbol{B}$ a $\det(\boldsymbol{A} \otimes \boldsymbol{B}) = (\det \boldsymbol{A})^n (\det \boldsymbol{B})^s$ pre $\boldsymbol{A} \in K^{n\times n}$, $\boldsymbol{B} \in K^{s\times s}$.

32.8. Dokážte, že komplexifikácia $V^{\mathbb{C}}$ reálneho vektorového priestoru V tak, ako bola definovaná v paragrafe 19.4, je izomorfná s tenzorovým súčinom $\mathbb{C} \otimes_{\mathbb{R}} V$.

32.9. Predpokladajme, že pole L je rozšírením poľa K a V je vektorový priestor nad K.
(a) Za prepokladu, že L je konečné rozšírenie stupňa $[L : K] = k$, podrobne dokážte, že prvky vektorového priestoru $V^L = L \otimes_K V$ možno jednoznačne až na poradie sčítancov reprezentovať v tvare $(b_1 \otimes \boldsymbol{x}_1) + (b_2 \otimes \boldsymbol{x}_2) + \ldots + (b_k \otimes \boldsymbol{x}_k)$, kde $b_1, \ldots, b_k$ je báza L ako vektorového priestoru nad poľom K a $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in V$.

(b) Dokážte, že $\boldsymbol{x} \mapsto 1 \otimes \boldsymbol{x}$ je injektívne K-lineárne zobrazenie $V \to V^L$.

(c) Dokážte, že rovnosť $\dim_K(L \otimes_K V) = [L : K] \dim_K V$ platí, aj keď L je nekonečné rozšírenie poľa K.

(d) Dokážte rovnosť $\dim_L V^L = \dim_K V$.

(e) Nech U, V sú vektorové priestory nad poľom K, s bázami $\boldsymbol{\alpha}$ resp. $\boldsymbol{\beta}$ a $\varphi\colon V \to U$ je K-lineárne zobrazenie s maticou $\boldsymbol{A} = (\varphi)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$. Dokážte, že predpisom $\varphi^L(a \otimes \boldsymbol{x}) = a \otimes \varphi(\boldsymbol{x})$ je jednoznačne určené L-lineárne zobrazenie $\varphi^L\colon V^L \to U^L$ s tou istou maticou $\left(\varphi^L\right)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \boldsymbol{A}$ vzhľadom na dané bázy.

32.10. Nech $V_1, \ldots, V_p, W_1, \ldots, W_q$ sú konečnorozmerné vektorové priestory nad poľom K.
(a) Podrobne popíšte prirodzené izomorfizmy medzi každou dvojicou z nasledujúcich vektorových priestorov:
$V_1^* \otimes \ldots \otimes V_p^* \otimes W_1 \otimes \ldots \otimes W_q$, $\mathcal{L}_{p+q}(V_1, \ldots, V_p, W_1^*, \ldots, W_q^*, K)$,
$\mathcal{L}(V_1 \otimes \ldots \otimes V_p, W_1 \otimes \ldots \otimes W_q)$.

(b) Aká je dimenzia uvedených priestorov, ak $\dim V_i = n_i$, $\dim W_j = m_j$?

32.11. Preverte výpočty v jednotlivých častiach príkladu 32.5.4 a na ich základe dokážte vetu 32.5.5.

32.12. Nech $\boldsymbol{A} = \left(A^i_j\right)$, $\boldsymbol{B} = \left(B^j_k\right)$ sú matice rozmeru $n \times n$ nad poľom K a $\boldsymbol{x} = (x^1, \ldots, x^n)$, $\boldsymbol{y} = (y^1, \ldots, y^n)$ sú vektory z K^n. Dokážte nasledujúce rovnosti:
(a) $\boldsymbol{A} \cdot \boldsymbol{B} = \operatorname{tr}^2_1(\boldsymbol{A} \otimes \boldsymbol{B})$;

(b) $A_{ij}x^iy^j = (\operatorname{tr}^1_1 \circ \operatorname{tr}^2_2)\big(\boldsymbol{A} \otimes (\boldsymbol{x} \otimes \boldsymbol{y})\big)$.

32.13. Dokážte formulu pre kontrakciu $\operatorname{tr}^k_l T$ tenzora T z konca odseku 32.6.2 a overte, že $\operatorname{tr}^k_l T$ ako multilineárna forma nezávisí na báze $\boldsymbol{\beta}$.

32.14. Nech V je konečnorozmerný vektorový priestor nad poľom K. Dokážte nasledujúce dve tvrdenia:
(a) Predpisom $(T, \varrho, \sigma) \mapsto T^\varrho_\sigma$ je definovaná pravá akcia grupy $\mathcal{S}_p \times \mathcal{S}_q$ na priestore tenzorov $\mathcal{T}^p_q(V)$.

(b) Pre pevné $\varrho \in \mathcal{S}_p$, $\sigma \in \mathcal{S}_q$ je $T \mapsto T^\varrho_\sigma$ lineárny izomorfizmus $\mathcal{T}^p_q(V) \to \mathcal{T}^p_q(V)$.

(c) Ako vyzerajú pevné body uvedenej akcie? Osobitne sa zamerajte na prípady

$p = 0$ resp. $q = 0$.

32.15. Nech $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ je báza vektorového priestoru V a $\boldsymbol{\beta}^* = (\boldsymbol{v}^1, \dots, \boldsymbol{v}^n)$ je k nej duálna báza. V častiach (d), (e) navyše predokladáme, že $G\colon V \times V \to K$ je symetrická regulárna bilineárna forma s Gramovou maticou $[G]_{\boldsymbol{\beta}} = (g_{ij})$. Overte, že operácie tenzorového súčinu, kontrakcie, permutácie indexov a spúšťania a dvíhania indexov na bázických prvkoch fungujú podľa nasledujúcich formúl:

(a) $\big(\boldsymbol{v}_{i_1} \otimes \dots \otimes \boldsymbol{v}_{i_p} \otimes \boldsymbol{v}^{j_1} \otimes \dots \otimes \boldsymbol{v}^{j_q}\big) \otimes \big(\boldsymbol{v}_{k_1} \otimes \dots \otimes \boldsymbol{v}_{k_r} \otimes \boldsymbol{v}^{l_1} \otimes \dots \otimes \boldsymbol{v}^{l_s}\big)$
$= \boldsymbol{v}_{i_1} \otimes \dots \otimes \boldsymbol{v}_{i_p} \otimes \boldsymbol{v}_{k_1} \otimes \dots \otimes \boldsymbol{v}_{k_r} \otimes \boldsymbol{v}^{j_1} \otimes \dots \otimes \boldsymbol{v}^{j_q} \otimes \boldsymbol{v}^{l_1} \otimes \dots \otimes \boldsymbol{v}^{l_s}$;

(b) $\operatorname{tr}^k_l\big(\boldsymbol{v}_{i_1} \otimes \dots \otimes \boldsymbol{v}_{i_p} \otimes \boldsymbol{v}^{j_1} \otimes \dots \otimes \boldsymbol{v}^{j_q}\big)$
$= \big\langle \boldsymbol{v}_{i_k}, \boldsymbol{v}^{j_l}\big\rangle \big(\boldsymbol{v}_{i_1} \otimes \dots \otimes \boldsymbol{v}_{i_{k-1}} \otimes \boldsymbol{v}_{i_{k+1}} \otimes \dots \otimes \boldsymbol{v}_{i_p}$
$\otimes \boldsymbol{v}^{j_1} \otimes \dots \otimes \boldsymbol{v}^{j_{l-1}} \otimes \boldsymbol{v}^{j_{l+1}} \otimes \dots \otimes \boldsymbol{v}^{j_q}\big)$;

(c) $\big(\boldsymbol{v}_{i_1} \otimes \dots \otimes \boldsymbol{v}_{i_p} \otimes \boldsymbol{v}^{j_1} \otimes \dots \otimes \boldsymbol{v}^{j_q}\big)^{\varrho}_{\sigma} = \boldsymbol{v}_{i_{\varrho^{-1}(1)}} \otimes \dots \otimes \boldsymbol{v}_{i_{\varrho^{-1}(p)}} \otimes \boldsymbol{v}^{j_{\sigma^{-1}(1)}} \otimes \dots \otimes \boldsymbol{v}^{j_{\sigma^{-1}(q)}}$ pre $\varrho \in \mathcal{S}_p$, $\sigma \in \mathcal{S}_q$;

(d) $\downarrow^k_l \big(\boldsymbol{v}_{i_1} \otimes \dots \otimes \boldsymbol{v}_{i_p} \otimes \boldsymbol{v}^{j_1} \otimes \dots \otimes \boldsymbol{v}^{j_q}\big)$
$= \sum_j g_{i_k j} \big(\boldsymbol{v}_{i_1} \otimes \dots \otimes \boldsymbol{v}_{i_{k-1}} \otimes \boldsymbol{v}_{i_{k+1}} \otimes \dots \otimes \boldsymbol{v}_{i_p}$
$\otimes \boldsymbol{v}^{j_1} \otimes \dots \otimes \boldsymbol{v}^{j_{l-1}} \otimes \boldsymbol{v}^{j} \otimes \boldsymbol{v}^{j_l} \otimes \dots \otimes \boldsymbol{v}^{j_q}\big)$;

(e) $\uparrow^k_l \big(\boldsymbol{v}_{i_1} \otimes \dots \otimes \boldsymbol{v}_{i_p} \otimes \boldsymbol{v}^{j_1} \otimes \dots \otimes \boldsymbol{v}^{j_q}\big)$
$= \sum_i g^{i j_l} \big(\boldsymbol{v}_{i_1} \otimes \dots \otimes \boldsymbol{v}_{i_{k-1}} \otimes \boldsymbol{v}_{i} \otimes \boldsymbol{v}_{i_k} \otimes \dots \otimes \boldsymbol{v}_{i_p}$
$\otimes \boldsymbol{v}^{j_1} \otimes \dots \otimes \boldsymbol{v}^{j_{l-1}} \otimes \boldsymbol{v}^{j_{l+1}} \otimes \dots \otimes \boldsymbol{v}^{j_q}\big)$.

V časti (c) si všimnite zámenu permutácií ϱ, σ inverznými permutáciami ϱ^{-1}, σ^{-1} a ich presun v indexoch zhora nadol resp. zdola nahor. V častiach (d), (e) zasa stojí za povšimnutie, že na bázických vektoroch sa spúšťanie indexu prejavi jeho *zdvihom* a zdvih indexu jeho *spustením*.

32.16. (a) Overte, že lineárny izomorfizmus $\boldsymbol{x} \mapsto \downarrow \boldsymbol{x}$ na (pseudo)eklidovskom priestore (V, G) môžeme vyjadriť pomocou operácií tenzorového súčinu a kontrakcie $\downarrow \boldsymbol{x} = \operatorname{tr}^1_1(G \otimes \boldsymbol{x})$.

(b) Vyjadrite aj všeobecné formuly spúšťania a dvíhania indexov pomocou tenzorového súčinu a zúženia.

32.17. Nech $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ je báza vektorového priestoru V nad poľom $\mathbb{R}$ a $\boldsymbol{\beta}^* = (\boldsymbol{v}^1, \dots, \boldsymbol{v}^n)$ je k nej duálna báza. Nech ďalej G je (pseudo)skalárny súčin na V. Za akých podmienok platí $\downarrow \boldsymbol{v}_i = \boldsymbol{v}^i$ resp. $\uparrow \boldsymbol{v}^j = \boldsymbol{v}_j$ pre všetky $i, j \le n$?

32.18. Popíšte ako „do kríža“ funguje v priestoroch tenzorov nad konečnorozmerným vektorovým priestorom V nad poľom $\mathbb{C}$ s regulárnou kososymetrickou poldruhalineárnou formou G spúšťanie indexov $\mathcal{T}^{p,r}_{q,s}(V) \to \mathcal{T}^{p-1,r}_{q,s+1}(V)$ resp. $\mathcal{T}^{p,r}_{q,s}(V) \to \mathcal{T}^{p,r-1}_{q+1,s}(V)$ a dvíhanie indexov $\mathcal{T}^{p,r}_{q,s}(V) \to \mathcal{T}^{p,r+1}_{q-1,s}(V)$ resp. $\mathcal{T}^{p,r}_{q,s}(V) \to \mathcal{T}^{p+1,r}_{q,s-1}(V)$. Ktoré z týchto zobrazení sú lineárne a ktoré semilineárne izomorfizmy?

33. Symetrické a alternujúce tenzory

Mnoho tenzorov (dalo by sa povedať, že ich väčšina), s ktorými sa stretávame v geometrii či vo fyzike, vykazuje istú symetriu voči zámene poradia argumentov resp. indexov. Napr. kvadratické formy reprezentujeme pomocou symetrických bilineárnych foriem (t. j. symetrických kovariantných tenzorov stupňa 2). Riemannov tenzor krivosti ${R^i}_{jkl}$ je zasa antisymetrický v poslednej dvojici indexov a po spustení indexu i aj v prvej dvojici, čiže ${R^i}_{jkl} = -{R^i}_{jlk}$ a $R_{ijkl} = -R_{jikl}$. [1] Symetrický je aj tenzor napätia (po zdvihu alebo spustení niektorého indexu) a tenzor momentu zotrvačnosti, zatiaľ čo tenzor elektromagnetického poľa je antisymetrický. Antisymetrické multilineárne formy hrajú dôležitú úlohu v diferenciálnej geometrii a algebraickej topológii. Algebry symetrických resp. antisymetrických tenzorov sa objavujú v kvantovej mechanike ako stavové priestory systémov identických bozónov resp. fermiónov (t. j. častíc s celočíselným resp. poloceločíselným spinom).

Symetria prípadne antisymetria tenzora $T \in \mathcal{T}^p_0(V)$ umožňuje značnú redukciu počtu parametrov potrebných na jeho popis. Formálne sa to prejavuje tým, že dimenzie podpriestorov takýchto tenzorov sú podstatne menšie než dimenzia priestoru $\mathcal{T}^p_0(V)$ (porovnaj vety 33.1.3 a 33.4.3 s vetou 32.5.1).

Podrobné a ucelené spracovanie problematiky symetrických a alternujúcich (antisymetrických) tenzorov, zahŕňajúce aj ich najdôležitejšie aplikácie, by podstatne presiahlo rozumný rozsah jednej na to vyhradenej kapitoly. A tak sa voľky-nevoľky obmedzíme len na výklad základných pojmov a súvislostí, doplnený niekoľkými ukážkami aplikácií mimo rámca lineárnej algebry.

33.1 Symetrické tenzory

Prototypom symetrických tenzorov nad vektorovým priestorom V sú symetrické bilineárne formy $F\colon V \times V \to K$, ktoré tvoria lineárny podpriestor priestoru tenzorov $\mathcal{T}^0_2(V) \cong \mathcal{T}^2_0(V^*)$.

Hovoríme, že *p-lineárne zobrazenie* $H\colon V^p \to W$ je *symetrické*, ak zámenou poradia ľubovoľných dvoch argumentov $\boldsymbol{x}_i, \boldsymbol{x}_j \in V$ sa hodnota výrazu $H(\boldsymbol{x}_1, \dots, \boldsymbol{x}_p) \in W$ nezmení. Pre $W = K$ dostávame pojem *symetrického kovariantného tenzora* nad priestorom V. Zrejme kovariantný tenzor $T \in \mathcal{T}^0_p(V)$ stupňa p (t. j. typu $\binom{0}{p}$) je symetrický práve vtedy, keď pre ľubovoľnú permutáciu $\sigma \in \mathcal{S}_p$ platí $T_\sigma = T$. V reči súradníc (vzhľadom na ľubovoľnú

[1] Riemannov tenzor R_{ijkl} je navyše symetrický voči zámene poradia dvojíc indexov (i,j) a (k,l), t. j. $R_{ijkl} = R_{klij}$.

bázu $\boldsymbol{\beta}$ priestoru V) to znamená, že platí $T_{j_1\dots j_p} = T_{j_{\sigma(1)}\dots j_{\sigma(p)}}$ pre všetky $j_1, \dots, j_p \leq n$. Podobne možno definovať a charakterizovať aj pojem *symetrického kontravariantného tenzora* $T \in \mathcal{T}_0^p(V)$ stupňa p (t. j. typu $\binom{p}{0}$).

Symetrické tenzory tvoria v oboch prípadoch lineárne podpriestory priestorov tenzorov $\mathcal{T}_p^0(V)$ resp. $\mathcal{T}_0^p(V)$; značíme ich $\mathrm{Sym}_p(V)$ resp. $\mathrm{Sym}^p(V)$. [2] Vzhľadom na kanonické izomorfizmy $\mathcal{T}_p^0(V) \cong \mathcal{T}_0^p(V^*)$ a $\mathrm{Sym}_p(V) \cong \mathrm{Sym}^p(V^*)$ pre začiatok stačí zaoberať sa priestormi kontravariantných symetrických tenzorov $\mathrm{Sym}^p(V) \subseteq \mathcal{T}_0^p(V)$.

Všetky skaláry $a \in K$ a vektory $\boldsymbol{x} \in V$ budeme považovať za symetrické tenzory (rozmyslite si, že je to aj formálne v zhode s definíciou symetrického tenzora). Inak povedané, $\mathrm{Sym}^0(V) = K$ a $\mathrm{Sym}^1(V) = V$.

Naším cieľom je vybaviť *priestor všetkých symetrických tenzorov*

$$\mathrm{Sym}(V) = \bigoplus_{p\in\mathbb{N}} \mathrm{Sym}^p(V)$$

prirodzenou štruktúrou asociatívnej graduovanej lineárnej algebry. Bežný tenzorový súčin symetrických tenzorov však vo všeobecnosti nie je symetrický tenzor. Napr. vektory $\boldsymbol{x}, \boldsymbol{y} \in V$ sú symetrické tenzory, ale ich súčin $\boldsymbol{x} \otimes \boldsymbol{y}$ nemusí byť symetrický. Pre $\eta, \vartheta \in V^*$ totiž platí

$$(\boldsymbol{x} \otimes \boldsymbol{y})(\eta, \vartheta) = \boldsymbol{x}(\eta)\boldsymbol{y}(\vartheta) = \boldsymbol{y}(\vartheta)\boldsymbol{x}(\eta) = (\boldsymbol{y} \otimes \boldsymbol{x})(\vartheta, \eta).$$

Teda symetria tenzora $\boldsymbol{x} \otimes \boldsymbol{y}$, t. j. rovnosť $(\boldsymbol{x} \otimes \boldsymbol{y})(\eta, \vartheta) = (\boldsymbol{x} \otimes \boldsymbol{y})(\vartheta, \eta)$ pre všetky $\eta, \vartheta \in V^*$, je ekvivalentná s podmienkou $\boldsymbol{x} \otimes \boldsymbol{y} = \boldsymbol{y} \otimes \boldsymbol{x}$. Ak sú ale $\boldsymbol{x}$, $\boldsymbol{y}$ lineárne nezávislé vektory, tak $\boldsymbol{x} \otimes \boldsymbol{y}$ a $\boldsymbol{y} \otimes \boldsymbol{x}$ sú rôzne bilineárne formy na priestore V^*. Ako však vyplýva z našej úvahy, bilineárna operácia súčinu, ktorá zo symetrických tenzorov vyrába opäť symetrické tenzory, musí byť aspoň na vektoroch $\boldsymbol{x}, \boldsymbol{y} \in V$ komutatívna.

Ku konštrukcii algebry symetrických tenzorov preto pristúpime z inej strany. Ak si zvolíme v priestore V nejakú bázu $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$, tak komutatívny a asociatívny súčin ľubovoľných p vektorov $\boldsymbol{v}_{i_1}, \dots, \boldsymbol{v}_{i_p}$ tejto bázy je jednoznačne určený len počtom výskytov $k_1, \dots, k_n$ každého z vektorov $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ v uvedenom zozname. Tento súčin preto zapíšeme v tvare

$$\boldsymbol{v}_{i_1} \dots \boldsymbol{v}_{i_p} = \boldsymbol{v}_1^{k_1} \dots \boldsymbol{v}_n^{k_n}.$$

Všetky formálne lineárne kombinácie

$$\sum a_{k_1\dots k_n} \boldsymbol{v}_1^{k_1} \dots \boldsymbol{v}_n^{k_n}$$

[2] Používa sa tiež označenie $\mathcal{S}_p(V)$ resp. $\mathcal{S}^p(V)$, ktoré by však v našom prípade kolidovalo s označením grupy permutácií $\mathcal{S}_p$.

len s konečným počtom nenulových koeficientov $a_{k_1\dots k_n} \in K$ tvoria nekonečnorozmerný vektorový priestor $K[\boldsymbol{v}_1, \dots, \boldsymbol{v}_n]$ a pravidlo pre súčin výrazov

$$\boldsymbol{v}_1^{k_1} \dots \boldsymbol{v}_n^{k_n} \cdot \boldsymbol{v}_1^{l_1} \dots \boldsymbol{v}_n^{l_n} = \boldsymbol{v}_1^{k_1+l_1} \dots \boldsymbol{v}_n^{k_n+l_n}$$

spolu s podmienkou bilinearity jednoznačne určuje na $K[\boldsymbol{v}_1, \dots, \boldsymbol{v}_n]$ komutatívnu a asociatívnu binárnu operáciu, ktorú nazývame *symetrickým tenzorovým súčinom*. Takto získanú algebru možno prirodzene stotožniť s graduovanou algebrou polynómov $K[x_1, \dots, x_n]$ z príkladu 30.2.1, resp. chápať ju priamo ako algebru polynómov v premenných $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ nad poľom K. Presnejšie, priradenie $x_i \mapsto \boldsymbol{v}_i$ pre $i \le n$ určuje izomorfizmus algebier

$$f(x_1, \dots, x_n) \mapsto f(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) \colon K[x_1, \dots, x_n] \to K[\boldsymbol{v}_1, \dots, \boldsymbol{v}_n]$$

a zároveň indukuje prirodzenú graduáciu algebry $K[\boldsymbol{v}_1, \dots, \boldsymbol{v}_n]$, v ktorej prvky podpriestoru $K[\boldsymbol{v}_1, \dots, \boldsymbol{v}_n]_p$ zodpovedajú homogénnym polynómom p-teho stupňa. Graduovanú algebru $K[\boldsymbol{v}_1, \dots, \boldsymbol{v}_n]$ nazývame *symetrickou algebrou vektorového priestoru* V a značíme ju $\Sigma(V)$. Špeciálne každý vektor $\boldsymbol{x} = x^i \boldsymbol{v}_i \in V = \mathcal{T}_0^1(V)$ môžeme chápať ako prvok vrstvy $\Sigma^1(V) = V$. Priradenie

$$\boldsymbol{v}_{i_1} \otimes \dots \otimes \boldsymbol{v}_{i_p} \mapsto \boldsymbol{v}_1^{k_1} \dots \boldsymbol{v}_n^{k_n},$$

kde $k_i = \#\{r \le p;\ i = i_r\}$ je počet výskytov indexu i v zozname $i_1, \dots, i_p$, určuje surjektívny homomorfizmus graduovaných algebier $\mathcal{T}_0(V) \to \Sigma(V)$, nazývaný *symetrická kanonická projekcia*. Čitateľ by si mal rozmyslieť, že samotná algebra $\Sigma(V)$ ani jej graduácia na vrstvy $\Sigma^p(V) = K[\boldsymbol{v}_1, \dots, \boldsymbol{v}_n]_p$, na rozdiel od homomorfizmu $\mathcal{T}_0(V) \to \Sigma(V)$, nezávisia od výberu bázy $\boldsymbol{\beta}$ (pozri cvičenie 33.2 (a)).

Vďaka spomínanému surjektívnemu homomorfizmu sa na symetrickú algebru $\Sigma(V)$ možno dívať ako na *faktorovú algebru* graduovanej tenzorovej algebry $\mathcal{T}_0(V)$, ktorá vznikne stotožnením všetkých tenzorových súčinov $\boldsymbol{x}_1 \otimes \dots \otimes \boldsymbol{x}_p \in \mathcal{T}_0^p(V)$ líšiacich sa len poradím činiteľov. Tento popis symetrickej algebry $\Sigma(V)$ je už invariantný, teda samotná algebra $\Sigma(V)$ ani jej graduácia $\Sigma^p(V) = K[\boldsymbol{v}_1, \dots, \boldsymbol{v}_n]_p$, na rozdiel od izomorfizmu $\Sigma(V) \cong K[x_1, \dots, x_n]$, nezávisia od výberu bázy $\boldsymbol{\beta}$ (pozri cvičenie 33.2 (b)).

Navzájom rôznych jednočlenov $x_1^{k_1} \dots x_n^{k_n}$ stupňa p je práve toľko, koľkými spôsobmi možno číslo p rozložiť na súčet n celých nezáporných sčítancov, t. j. toľko, koľko je p-prvkových kombinácií s opakovaním z n prvkov (pozri cvičenie 30.10). Keďže tieto polynómy tvoria bázu priestoru $K[x_1, \dots, x_n]_p \cong \Sigma^p(V)$, môžeme tak určiť dimenziu priestoru p-homogénnych polynómov $\Sigma^p(V)$.

33.1.1. Tvrdenie. *Symetrická algebra $\Sigma(V)$ n-rozmerného vektorového priestoru V je komutatívna a asociatívna nekonečnorozmerná graduovaná K-algebra s jednotkou $1 \in K$. Pre dimenzie jej jednotlivých vrstiev platí*

$$\dim \Sigma^p(V) = \binom{n+p-1}{p} = \frac{n(n+1)\dots(n+p-1)}{p!}\,.$$

Symetrický tenzorový súčin $\boldsymbol{xy}$ vektorov $\boldsymbol{x}, \boldsymbol{y} \in V$ sa niekedy tiež značí $\boldsymbol{x} \underset{\text{sym}}{\otimes} \boldsymbol{y}$ alebo $\boldsymbol{x} \vee \boldsymbol{y}$. Priestor $\Sigma^p(V)$ potom značíme aj ako symetrický tenzorový súčin p exemplárov priestoru V, t. j.

$$\Sigma^p(V) = \underbrace{V \underset{\text{sym}}{\otimes} \ldots \underset{\text{sym}}{\otimes} V}_{p} = \underbrace{V \vee \ldots \vee V}_{p},$$

a nazývame *p-tou symetrickou mocninou* priestoru V.

Symetrický tenzorový súčin sa spomedzi symetrických bilineárnych zobrazení vydeľuje podobnou *univerzálnou vlastnosťou* ako bežný tenzorový súčin spomedzi všetkých bilineárnych zobrazení (porovnaj s vetou 32.2.7).

33.1.2. Veta. *Nech V, W sú vektorové priestory nad poľom K, pričom V je konečnorozmerný a $p \in \mathbb{N}$. Potom*

(a) *operácia symetrického tenzorového súčinu $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p) \mapsto \boldsymbol{x}_1 \ldots \boldsymbol{x}_p$ je symetrické p-lineárne zobrazenie $V^p \to \Sigma^p(V)$;*

(b) *ku každému symetrickému p-lineárnemu zobrazeniu $H \colon V^p \to W$ existuje práve jedno* lineárne *zobrazenie $\widetilde{H} \colon \Sigma^p(V) \to W$ také, že*

$$\widetilde{H}(\boldsymbol{x}_1 \ldots \boldsymbol{x}_p) = H(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p)$$

pre všetky $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p \in V$;

(c) *priradenie $H \mapsto \widetilde{H}$ je lineárny izomorfizmus priestoru všetkých symetrických p-lineárnych zobrazení $V^p \to W$ na priestor $\mathcal{L}\big(\Sigma^p(V), W\big)$.*

Dôkaz prenechávame ako cvičenie čitateľovi.

Symetrická mocnina $\Sigma^p(V)$ je podmienkami (a), (b) určená jednoznačne až na lineárny izomorfizmus. Zostáva nám vyjasniť, ako súvisia jednotlivé priestory symetrických tenzorov $\mathrm{Sym}^p(V) \subseteq \mathcal{T}_0^p(V)$ s priestormi $\Sigma^p(V)$ homogénnych polynómov stupňa p a celý nekonečnorozmerný priestor $\mathrm{Sym}(V) \subseteq \mathcal{T}_0(V)$ s graduovanou symetrickou algebrou $\Sigma(V)$. Kľúčom k odpovedi na túto otázku sú body (b) a (c) vety 33.1.2.

Každý symetrický tenzor $T \in \mathrm{Sym}^p(V) \subseteq \mathcal{T}_0^p(V)$ je symetrická p-lineárna forma $T \colon (V^*)^p \to K$. Táto podľa (b) určuje lineárny funkcionál $\widetilde{T} \colon \Sigma^p(V^*) \to K$ taký, že pre ľubovoľné $k_1 + \ldots + k_n = p$ platí

$$\widetilde{T}\big((\boldsymbol{v}^1)^{k_1} \ldots (\boldsymbol{v}^n)^{k_n}\big) = T(\underbrace{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^1}_{k_1}, \ldots, \underbrace{\boldsymbol{v}^n, \ldots, \boldsymbol{v}^n}_{k_n}),$$

kde $\boldsymbol{\beta}^* = (\boldsymbol{v}^1, \ldots, \boldsymbol{v}^n)$ je duálna báza k báze $\boldsymbol{\beta}$ (vo výrazoch $(\boldsymbol{v}^i)^{k_i}$ treba rozlišovať *horný index* i a *exponent* k_i). V dôsledku symetrie tenzora T

výraz na pravej strane rovnosti závisí len od počtu výskytov jednotlivých kovektorov $\boldsymbol{v}^j$ a nie od ich poradia. Podľa (c) je priradenie $T \mapsto \widetilde{T}$ lineárny izomorfizmus vektorového priestoru $\mathrm{Sym}^p(V)$ na vektorový priestor $\mathcal{L}(\Sigma^p(V^*), K) = \Sigma^p(V^*)^*$. Lineárny funkcionál $\widetilde{T} \in \Sigma^p(V^*)^*$ môžeme prirodzene stotožniť s polynómom

$$\widetilde{T} = \sum_{k_1+\ldots+k_n=p} \widetilde{T}\big((\boldsymbol{v}^1)^{k_1} \ldots (\boldsymbol{v}^n)^{k_n}\big)\, \boldsymbol{v}_1^{k_1} \ldots \boldsymbol{v}_n^{k_n} \in \Sigma^p(V)$$

a priradenie $T \mapsto \widetilde{T}$ chápať ako lineárny izomorfizmus $\mathrm{Sym}^p(V) \cong \Sigma^p(V)$.[3] Vzor polynómu $f \in \Sigma^p(V)$ v tomto zobrazení, teda vlastne inverzné zobrazenie k zobrazeniu $T \mapsto \widetilde{T}$, budeme značiť f^+. Symetrický tenzor, ktorý sa v tomto izomorfizme zobrazí na jednočlen $\boldsymbol{v}_1^{k_1} \ldots \boldsymbol{v}_n^{k_n}$ stupňa p, má tvar súčtu

$$\big(\boldsymbol{v}_1^{k_1} \ldots \boldsymbol{v}_n^{k_n}\big)^+ = \sum_{i_1,\ldots,i_p} \boldsymbol{v}_{i_1} \otimes \ldots \otimes \boldsymbol{v}_{i_p}$$

cez všetky usporiadané p-tice $(i_1, \ldots, i_p)$ indexov z množiny $\{1, \ldots, n\}$, ktoré obsahujú každý z indexov $i \le n$ práve k_i-krát. Počet týchto sčítancov udáva tzv. *polynomický koeficient*

$$\binom{p}{k_1, \ldots, k_n} = \frac{p!}{k_1! \ldots k_n!}\,.$$

Násobenie v priestore symetrických tenzorov $\mathrm{Sym}(V) = \bigoplus_{p\in\mathbb{N}} \mathrm{Sym}^p(V)$ zavedieme tak, aby priradenie $T \mapsto \widetilde{T}$ bolo homomorfizmom, a teda izomorfizmom graduovaných algebier $\mathrm{Sym}(V) \cong \Sigma(V)$. Pre $S \in \mathrm{Sym}^p(V)$, $T \in \mathrm{Sym}^q(V)$ preto položíme

$$S \cdot T = \big(\widetilde{S} \cdot \widetilde{T}\big)^+.$$

Potom zrejme $\widetilde{ST} = \widetilde{S} \cdot \widetilde{T}$. Takže môžeme zhrnúť:

33.1.3. Veta. *Symetrické (kontravariantné) tenzory nad n-rozmerným vektorovým priestorom V tvoria nekonečnorozmernú komutatívnu a asociatívnu graduovanú algebru $\big(\mathrm{Sym}(V), \cdot\big)$ s jednotkou $1 \in K$, izomorfnú so symetrickou algebrou $\big(\Sigma(V), \cdot\big)$ priestoru V. Pre dimenzie jej jednotlivých vrstiev platí*

$$\dim \mathrm{Sym}^p(V) = \binom{n+p-1}{p}\,.$$

[3] Tým sme len tak medzi rečou dokázali očakávaný izomorfizmus $\Sigma^p(V)^* \cong \Sigma^p(V^*)$, ktorý je dokonca kanonický. Chyba krásy uvedeného popisu, ktorý závisí od bázy $\boldsymbol{\beta}$, sa totiž dá odstrániť (pozri cvičenie 33.2 (c)).

Ešte raz výslovne upozorňujeme, že – napriek inklúzii $\mathrm{Sym}(V) \subseteq \mathcal{T}_0(V)$ – symetrické tenzory *netvoria podalgebru* graduovanej algebry všetkých kontravariantných tenzorov nad V. Nad poľami nekonečnej charakteristiky však existuje možnosť ako rozšíriť symetrický tenzorový súčin na celý priestor $\mathcal{T}_0(V)$ a urobiť z neho komutatívnu a asociatívnu graduovanú algebru, v ktorej symetrické tenzory *tvoria graduovanú podalgebru* $\mathrm{Sym}(V) \subseteq \mathcal{T}_0(V)$ izomorfnú so symetrickou graduovanou algebrou $\Sigma(V)$. Na ten účel stačí definovať súčiny bázických tenzorov $\boldsymbol{v}_{i_1} \otimes \ldots \otimes \boldsymbol{v}_{i_p} \in \mathcal{T}_0^p(V)$, $\boldsymbol{v}_{j_1} \otimes \ldots \otimes \boldsymbol{v}_{j_q} \in \mathcal{T}_0^q(V)$. Ak označíme $k_i = \#\{r \leq p;\ i_r = i\}$, $l_j = \#\{s \leq q;\ j_s = j\}$, môžeme písať

$$\begin{aligned}\left(\boldsymbol{v}_1^{k_1} \ldots \boldsymbol{v}_n^{k_n}\right)^+ \cdot \left(\boldsymbol{v}_1^{l_1} \ldots \boldsymbol{v}_n^{l_n}\right)^+ &= \left(\boldsymbol{v}_1^{k_1+l_1} \ldots \boldsymbol{v}_n^{k_n+l_n}\right)^+ \\ &= \sum_{m_1,\ldots,m_{p+q}} \boldsymbol{v}_{m_1} \otimes \ldots \otimes \boldsymbol{v}_{m_{p+q}},\end{aligned}$$

kde posledný súčet beží cez všetky $(p+q)$-tice $(m_1, \ldots, m_{p+q})$, v ktorých sa každý z indexov $m \leq n$ vyskytuje práve $(k_m + l_m)$-krát. Túto hodnotu rovnomerne rozdelíme medzi $\binom{p}{k_1,\ldots,k_n}\binom{q}{l_1,\ldots,l_n}$ sčítancov tvoriacich výraz vľavo, t. j. položíme

$$(\boldsymbol{v}_{i_1} \otimes \ldots \otimes \boldsymbol{v}_{i_p}) \cdot (\boldsymbol{v}_{j_1} \otimes \ldots \otimes \boldsymbol{v}_{j_q}) = \frac{k_1! \ldots k_n!}{p!}\,\frac{l_1! \ldots l_n!}{q!} \sum_{m_1,\ldots,m_{p+q}} \boldsymbol{v}_{m_1} \otimes \ldots \otimes \boldsymbol{v}_{m_{p+q}}\,.$$

V nasledujúcom paragrafe uvidíme, že nad poľami nekonečnej charakteristiky jestvuje ešte jedna prirodzená možnosť ako urobiť z priestoru $\mathcal{T}_0(V)$ komutatívnu a asociatívnu graduovanú algebru s graduovanou podalgebrou $\mathrm{Sym}(V) \subseteq \mathcal{T}_0(V)$, izomorfnou so symetrickou graduovanou algebrou $\Sigma(V)$. Tento nový symetrický súčin sa však od práve definovaného symetrického súčinu bude líšiť len istými faktormi kombinatorického charakteru pred výrazmi $\sum_{m_1,\ldots,m_{p+q}} \boldsymbol{v}_{m_1} \otimes \ldots \otimes \boldsymbol{v}_{m_{p+q}}$ (pozri cvičenie 33.5).

33.2 Symetrické tenzory a p-homogénne formy

Súvis medzi (symetrickými) bilineárnymi formami a kvadratickými formami, ktorý sme študovali v kapitole 11, má svoju analógiu vo vzťahu (symetrických) multilineárnych foriem, t. j. kovariantných tenzorov stupňa p a p-homogénnych foriem. Pripomeňme si z príkladu 30.2.1, časť (b), že $K[V]_p$ označuje lineárny podpriestor algebry $K[V]$ všetkých polynomických funkcií $V \to K$, tvorený práve p-homogénnymi funkciami. Ak si zafixujeme nejakú bázu $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ priestoru V a k nej duálnu bázu $\boldsymbol{\beta}^* = (\boldsymbol{v}^1, \ldots, \boldsymbol{v}^n)$, tak priradenie $x_i \mapsto \boldsymbol{v}^i$ možno jednoznačne rozšíriť do surjektívneho homomorfizmu K-algebier $K[x_1, \ldots, x_n] \to K[V]$, pri ktorom sa podpriestor p-homogénnych polynómov $K[x_1, \ldots, x_n]_p$ zobrazí na podpriestor p-homogénnych polynomických funkcií (foriem) $K[V]_p$.

Každý kovariantný tenzor $T \in \mathcal{T}_p^0(V)$, t. j. p-lineárna forma $T\colon V^p \to K$ určuje predpisom
$$T^\Delta(\boldsymbol{x}) = T(\boldsymbol{x}, \dots, \boldsymbol{x})$$
p-homogénnu formu $T^\Delta\colon V \to K$. Naopak, každá p-homogénna forma $f \in K[V]_p$ má zrejme tvar $f(\boldsymbol{x}) = T^\Delta(\boldsymbol{x}) = T(\boldsymbol{x}, \dots, \boldsymbol{x})$ pre viacero rôznych tenzorov $T \in \mathcal{T}_p^0(V)$. Predpisom $T \mapsto T^\Delta$ je tak definované surjektívne (a okrem prípadu $n = 1$ neinjektívne) lineárne zobrazenie $\mathcal{T}_p^0(V) \to K[V]_p$. Navyše, pre $S \in \mathcal{T}_p^0(V)$, $T \in \mathcal{T}_q^0(V)$ zrejme platí
$$(S \otimes T)^\Delta = S^\Delta \cdot T^\Delta \in K[V]_{p+q}.$$

Vieme, že v špeciálnom prípade $p = 2$ a za predpokladu $\operatorname{char} K \neq 2$ má každá kvadratická forma $f\colon V \to K$ tvar $f(\boldsymbol{x}) = T^\Delta(\boldsymbol{x}) = T(\boldsymbol{x}, \boldsymbol{x})$ pre jednoznačne určený *symetrický* tenzor $T \in \mathcal{T}_2^0(V)$ (pozri paragraf 11.2). Jeden z tenzorov $T \in \mathcal{T}_p^0(V)$, ktorý po stotožnení svojich argumentov dáva p-homogénnu formu
$$f(\boldsymbol{x}) = \sum_{k_1+\ldots+k_n=p} a_{k_1\ldots k_n} x_1^{k_1} \ldots x_n^{k_n},$$
kde $x_i = \boldsymbol{v}^i(\boldsymbol{x})$ sú súradnice vektora $\boldsymbol{x} \in V$ v báze $\boldsymbol{\beta}$,[4] je daný predpisom
$$T(\boldsymbol{v}_{i_1}, \dots, \boldsymbol{v}_{i_p}) = \begin{cases} a_{k_1\ldots k_n}, & \text{ak } (i_1, \dots, i_p) = (\underbrace{1, \dots, 1}_{k_1}, \dots, \underbrace{n, \dots, n}_{k_n}), \\ 0 & \text{inak.} \end{cases}$$
To napr. pre $p = 2$ zodpovedá reprezentácii kvadratickej formy $q(\boldsymbol{x}) = \sum_{i\le j} a_{ij} x_i x_j$ bilineárnou formou $F(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i\le j} a_{ij} x_i y_j$ s hornou trojuholníkovou maticou (a_{ij}). Teda tento tenzor má vo všeobecnosti do symetrie ďaleko. Na druhej strane symetrický tenzor $T = f^+$, ktorý zodpovedá polynómu $f(x_1, \dots, x_n)$ v lineárnom izomorfizme
$$T \mapsto \widetilde{T}\colon \operatorname{Sym}_p(V) \to \Sigma_p(V) \cong K[x_1, \dots, x_n]_p,$$
dáva po stotožnení svojich argumentov polynomickú funkciu
$$T^\Delta(\boldsymbol{x}) = f^+(\boldsymbol{x}, \dots, \boldsymbol{x}) = \sum_{k_1,\ldots,k_n} \binom{p}{k_1, \dots, k_n} a_{k_1\ldots k_n} x_1^{k_1} \ldots x_n^{k_n},$$
ktorá okrem prípadu, že f je priamo lineárnou kombináciou p-tych mocnín x_i^p, nesplýva s $f(\boldsymbol{x})$.

[4]Všimnite si, že sme práve porušili dohodu o značení súradníc vektorov hornými indexmi. Pri označení $(x^j)^{k_j}$ by sa nám však plietli horné indexy a exponenty.

Symetrickú p-lineárnu formu $T\colon V^p \to K$, ktorá po stotožnení svojich argumentov dáva p-homogénnu polynomickú funkciu f, čiže platí pre ňu $T^\Delta = f$, nazývame *polárnou formou funkcie* f a značíme ju $\operatorname{pol} f$. Jej existencia a jednoznačnosť je však viazaná na možnosť deliť v poli K kladnými celými číslami $\leq p$ (a ich súčinmi).

Predpokladajme, že $\operatorname{char} K > p$. Keďže $\operatorname{char} K$ je ∞ alebo prvočíslo – v tom druhom prípade je počet prvkov poľa K mocninou jeho charakteristiky, tým skôr platí $\# K > p$. Podľa príkladu 30.2.1 z toho vyplýva, že priradenie $x_i \mapsto \boldsymbol{v}^i$ určuje lineárny izomorfizmus $K[x_1, \dots, x_n]_p \to K[V]_p$, prostredníctvom ktorého môžeme stotožniť p-homogénne formy $f : V \to K$ s p-homogénnymi polynómami $f(x_1, \dots, x_n)$. Navyše, žiaden z polynomických koeficientov $\binom{p}{k_1,\dots,k_n}$ nie je násobkom čísla $\operatorname{char} K$. V takom prípade vieme polárnu formu funkcie $f \in K[V]_p$ určiť len zo znalosti jej koeficientov. Vďaka linearite stačí poznať polárne formy jednočlenov p-teho stupňa $x_1^{k_1} \dots x_n^{k_n}$ – tie majú zrejme tvar

$$\begin{aligned} \operatorname{pol}\bigl(x_1^{k_1} \dots x_n^{k_n}\bigr) &= \binom{p}{k_1, \dots, k_n}^{-1} \bigl(x_1^{k_1} \dots x_n^{k_n}\bigr)^+ \\ &= \frac{k_1! \dots k_n!}{p!} \sum_{i_1,\dots,i_p} \bigl(\boldsymbol{v}^{i_1} \otimes \dots \otimes \boldsymbol{v}^{i_p}\bigr), \end{aligned}$$

pričom sčítame cez všetky usporiadané p-tice $(i_1, \dots, i_p)$ indexov z množiny $\{1, \dots, n\}$, ktoré obsahujú každý z indexov $i \leq n$ práve k_i-krát.

Posledný výraz preskúmame v trochu všeobecnejšom kontexte. *Symetrizáciou tenzora* $T \in \mathcal{T}_p^0(V)$ nazývame tenzor

$$\operatorname{sym} T = \frac{1}{p!} \sum_{\sigma \in \mathcal{S}_p} T_\sigma.$$

Predpisom $T \mapsto \operatorname{sym} T$ je definované lineárne zobrazenie $\operatorname{sym}\colon \mathcal{T}_p^0(V) \to \mathcal{T}_p^0(V)$, ktoré nazývame *symetrizáciou.* Tenzor $T \in \mathcal{T}_p^0(V)$ je zrejme symetrický práve vtedy, keď $T = \operatorname{sym} T$. Navyše pre *ľubovoľný* tenzor $T \in \mathcal{T}_p^0(V)$ a $\boldsymbol{x} \in V$ očividne platí

$$(\operatorname{sym} T)(\boldsymbol{x}, \dots, \boldsymbol{x}) = T(\boldsymbol{x}, \dots, \boldsymbol{x}),$$

teda symetrizácia tenzora T a samotný tenzor indukujú tú istú p-homogénnu formu $T^\Delta = (\operatorname{sym} T)^\Delta \in K[V]_p$.

Pre každý z tenzorov $\boldsymbol{v}^{j_1} \otimes \dots \otimes \boldsymbol{v}^{j_p}$, v ktorom sa každý z činiteľov $\boldsymbol{v}^j$ vyskytuje práve k_j-krát, zrejme platí

$$\operatorname{sym}\bigl(\boldsymbol{v}^{j_1} \otimes \dots \otimes \boldsymbol{v}^{j_p}\bigr) = \frac{k_1! \dots k_n!}{p!} \sum_{i_1,\dots,i_p} \bigl(\boldsymbol{v}^{i_1} \otimes \dots \otimes \boldsymbol{v}^{i_p}\bigr),$$

kde súčet beží cez všetky usporiadané p-tice indexov $(i_1, \ldots, i_p)$, ktoré majú túto vlastnosť. Inak povedané výraz

$$\binom{p}{k_1, \ldots, k_n} \operatorname{sym}\bigl(\boldsymbol{v}^{j_1} \otimes \ldots \otimes \boldsymbol{v}^{j_p}\bigr)^{\sim} = (\boldsymbol{v}^1)^{k_1} \ldots (\boldsymbol{v}^n)^{k_n}$$

je symetrická kanonická projekcia bázického tenzora $\boldsymbol{v}^{j_1} \otimes \ldots \otimes \boldsymbol{v}^{j_p}$. Spojením oboch rovností dostávame

$$\operatorname{pol}\bigl(x_1^{k_1} \ldots x_n^{k_n}\bigr) = \operatorname{sym}\bigl(\boldsymbol{v}^{j_1} \otimes \ldots \otimes \boldsymbol{v}^{j_p}\bigr),$$

kde $(j_1, \ldots, j_p)$ je ľubovoľná usporiadaná p-tica indexov, ktorá obsahuje každý index $j \leq n$ práve k_j-krát.

Ak $p + q < \operatorname{char} K$, tak môžeme definovať ešte jeden súčin symetrických kovariantných tenzorov $S \in \operatorname{Sym}_p(V)$, $T \in \operatorname{Sym}_q(V)$ formulou

$$S \odot T = \operatorname{sym}(S \otimes T) = \frac{1}{(p+q)!} \sum_{\sigma \in \mathcal{S}_{p+q}} (S \otimes T)_\sigma,$$

ktorá však dáva zmysel pre akékoľvek tenzory $S \in \mathcal{T}_p^0(V)$, $T \in \mathcal{T}_q^0(V)$ a stále platí $S \odot T \in \operatorname{Sym}_{p+q}(V)$. Navyše oba tenzory $S \odot T = \operatorname{sym}(S \otimes T)$ a $S \otimes T$ určujú tú istú $(p+q)$-homogénnu formu

$$(S \odot T)^\Delta = (S \otimes T)^\Delta = S^\Delta \cdot T^\Delta,$$

takže lineárne zobrazenie $T \mapsto T^\Delta$ je homomorfizmom operácie symetrického tenzorového súčinu $\odot$ a zároveň aj operácie bežného tenzorového súčinu $\otimes$ na súčin polynómov v $K[V]$. Ak si uvedomíme, že po stotožnení kanonicky izomorfných algebier priradenie $T \mapsto T^\Delta$ nie je vlastne nič iné ako symetrická kanonická projekcia $\mathcal{T}_0(V^*) \to K[x_1, \ldots, x_n] \cong \Sigma(V^*)$, nemalo by nás to prekvapiť. Z jednoznačnosti polárnej formy a zo symetrie tenzorov $S \odot T$, $\operatorname{sym}(S \otimes T)$ ďalej vyplýva

$$S \odot T = \operatorname{sym}(S \otimes T) = \operatorname{pol}\bigl(S^\Delta \cdot T^\Delta\bigr) = \operatorname{sym} S \odot \operatorname{sym} T.$$

Za predpokladu, že pole K má nekonečnú charakteristiku, definuje súčin $\odot$ na podkladovom priestore symetrických kovariantných tenzorov

$$\operatorname{Sym}(V^*) = \bigoplus_{p \in \mathbb{N}} \operatorname{Sym}^p(V^*) = \bigoplus_{p \in \mathbb{N}} \operatorname{Sym}_p(V)$$

štruktúru graduovanej komutatívnej a asociatívnej lineárnej algebry. I keď rovnosť $S \cdot T = S \odot T$ pre symetrické tenzory vo všeobecnosti neplatí, vytvárajú oba symetrické tenzorové súčiny izomorfné graduované algebry. Pre každú jednotlivú vrstvu máme totiž lineárne izomorfizmy

$$\operatorname{Sym}_p(V) \cong K[x_1, \ldots, x_n]_p \cong K[V]_p,$$

ktoré dohromady dávajú izomorfizmy príslušných graduovaných algebier

$$\big(\mathrm{Sym}(V^*),\cdot\,\big) \cong \big(K[x_1,\dots,x_n],\cdot\,\big) \cong \big(K[V],\cdot\,\big) \cong \big(\mathrm{Sym}(V^*),\odot\big).$$

Symetrizácia $T \mapsto \operatorname{sym} T$ je navyše homomorfizmom každej z graduovaných algebier $\big(\mathcal{T}^0(V),\otimes\big)$, $\big(\mathcal{T}^0(V),\odot\big)$ na graduovanú algebru $\big(\mathrm{Sym}(V^*),\cdot\big)$.

Zatiaľ čo súčin $\cdot$ je univerzálny v tom zmysle, že aspoň na symetrických tenzoroch nezávisí na charakteristike poľa K, súčin $\odot$ funguje čo len na podpriestore symetrických tenzorov iba nad poľami nekonečnej charakteristiky, priamo sa však zhoduje s násobením v súradnicovom okruhu polynomických funkcií $K[V]$. To je hlavný dôvod, prečo sa nad poľom charakteristiky ∞ niekedy dáva prednosť symetrickému tenzorovému súčinu $S \odot T$ pred „univerzálnym“ symetrickým tenzorovým súčinom $S{\cdot}T$. Označenie $S{\cdot}T$, prípadne $S \underset{\mathrm{sym}}{\otimes} T$ sa potom zvykne používať pre „uprednostňovaný“ symetrický tenzorový súčin.

Ešte poznamenajme, že symetrický tenzorový súčin $S \odot T = \operatorname{sym}(S \otimes T)$ možno samozrejme definovať aj pre kontravariantné tenzory $S \in \mathcal{T}_0^p(V)$, $T \in \mathcal{T}_0^q(V)$. Úvahy vykonané v paragrafe 33.2 nás teraz oprávňujú vysloviť nasledujúci záver.

33.2.1. Veta. *Nech V je konečnorozmerný vektorový priestor nad poľom K charakteristiky ∞. Potom súčin $\odot$ vytvára na priestore kontravariantných tenzorov $\mathcal{T}_0(V)$ štruktúru komutatívnej a asociatívnej graduovanej algebry $\big(\mathcal{T}_0(V),\odot\big)$ s jednotkou $1 \in K$, v ktorej symetrické tenzory tvoria graduovanú podalgebru $\mathrm{Sym}(V)$ a symetrická kanonická projekcia $T \mapsto T^{\Delta}$ je homomorfizmus každej z graduovaných algebier $\big(\mathcal{T}_0(V),\otimes\big)$, $\big(\mathcal{T}_0(V),\odot\big)$ na symetrickú graduovanú algebru $\big(\Sigma(V),\cdot\big)$. Jej zúženie na podalgebru symetrických tenzorov je izomorfizmus graduovaných algebier $\big(\mathrm{Sym}(V),\odot\big) \cong \big(\Sigma(V),\cdot\big)$.*

33.3 Totálne derivácie vyšších rádov

V prípade, že V je vektorový priestor nad poľom $\mathbb{R}$, možno pre p-homogénne formy $f\colon V \to \mathbb{R}$ očividným spôsobom definovať vlastnosti kladnej resp. zápornej definitnosti či semidefinitnosti, ako aj vlastnosť indefinitnosti, rovnako ako pre kvadratické formy v paragrafe 12.2. Zrejme každá homogénna funkcia *nepárneho* stupňa p, ktorá nie je identicky rovná nule, je automaticky indefinitná. Hovoriť o kladnej resp. zápornej (semi)definitnosti má preto praktický zmysel len pre homogénne formy *párneho* stupňa p. Ukážeme si, ako možno pomocou týchto pojmov aplikovaných na parciálne derivácie vyšších rádov rozhodnúť o existencii a charaktere extrémov funkcie n-premenných aj v niektorých prípadoch, keď nám to len znalosť prvých a druhých parciálnych derivácií nedovoľuje.

Nech $M \subseteq \mathbb{R}^n$ je otvorená množina a $f\colon M \to \mathbb{R}$ je funkcia. *Totálnou p-tou deriváciou* prípadne *totálnou deriváciou p-teho rádu* funkcie f v bode v bode $\boldsymbol{a} \in M$ nazývame p-rozmernú maticu

$$f^{(p)}(\boldsymbol{a}) = \left(\frac{\partial^p f(\boldsymbol{a})}{\partial x_{i_1} \dots \partial x_{i_p}} \right),$$

tvorenú parciálnymi deriváciami $\partial^p f(\boldsymbol{a})/\partial x_{i_1} \dots \partial x_{i_p}$, za predpokladu, že všetky jej zložky existujú a sú konečné. Pre $p = 1$ dostávame gradient a pre $p = 2$ Hesseho maticu funkcie f (pozri paragraf 12.3 a príklad 32.7.1).

Na totálnu p-tu deriváciu sa možno dívať ako na kovariantný tenzor $f^{(p)}(\boldsymbol{a}) \in \mathcal{T}_p^0(\mathbb{R}^n)$. Ak navyše všetky parciálne derivácie funkcie f až do stupňa p sú spojité v nejakom okolí bodu $\boldsymbol{a}$, tak podľa Clairautovej-Schwartzovej vety $f^{(p)}(\boldsymbol{a})$ je symetrický tenzor. Ním indukovaná p-homogénna forma priradí vektoru $\boldsymbol{u} = \boldsymbol{x} - \boldsymbol{a} \in \mathbb{R}^n$, kde $\boldsymbol{x} \in M$, výraz

$$\mathrm{d}^p f(\boldsymbol{a}, \boldsymbol{x}) = \left(f^{(p)}(\boldsymbol{a})\right)^{\Delta}(\boldsymbol{u}) = \frac{\partial^p f(\boldsymbol{a})}{\partial x_{i_1} \dots \partial x_{i_p}} \left(x_{i_1} - a_{i_1}\right) \dots \left(x_{i_p} - a_{i_p}\right)$$

(používame Einsteinovu sumačnú konvenciu), ktorý nazývame *p-ty totálny diferenciál* alebo *totálny diferenciál p-teho rádu* funkcie f v bode $\boldsymbol{a}$.

Hodnotu funkcie f, ktorá má v bode $\boldsymbol{a}$ spojité parciálne derivácie až do rádu p vrátane, možno v okolí tohto bodu vyjadriť pomocou Taylorovho rozvoja

$$f(\boldsymbol{x}) = f(\boldsymbol{a}) + \mathrm{d}f(\boldsymbol{a}, \boldsymbol{x}) + \frac{1}{2!}\,\mathrm{d}^2(\boldsymbol{a}, \boldsymbol{x}) + \dots + \frac{1}{p!}\,\mathrm{d}^p(\boldsymbol{a}, \boldsymbol{x}) + \Theta(\boldsymbol{x})(\boldsymbol{x} - \boldsymbol{a}),$$

kde Θ je spojité zobrazenie, ktoré každému bodu $\boldsymbol{x}$ z onoho okolia bodu $\boldsymbol{a}$ priradí p-homogénnu formu $\Theta(\boldsymbol{x})\colon \mathbb{R}^n \to \mathbb{R}$ a $\Theta(\boldsymbol{a})$ je forma identicky rovná nule. Z uvedeného vyjadrenia možno podobnou úvahou ako v paragrafe 12.3 odvodiť nasledujúce závery:

33.3.1. Veta. *Nech $M \subseteq \mathbb{R}^n$ je otvorená množina, $\boldsymbol{a} \in M$ a $f\colon M \to \mathbb{R}$ je funkcia, ktorá má v bode $\boldsymbol{a}$ spojité všetky parciálne derivácie až do rádu p vrátane. Predpokladajme, že v bode $\boldsymbol{a}$ sú všetky parciálne derivácie rádov 1 až $p - 1$ nulové, t. j. $f^{(r)}(\boldsymbol{a}) = \mathbf{0}$ pre $1 \leq r < p$, a $f^{(p)}(\boldsymbol{a}) \neq \mathbf{0}$. Potom*
(a) ak forma $f^{(p)}(\boldsymbol{a})$ je kladne definitná (čo môže nastať len pre párne p), tak funkcia f má v bode $\boldsymbol{a}$ ostré lokálne minimum;
(b) ak forma $f^{(p)}(\boldsymbol{a})$ je záporne definitná (čo opäť môže nastať len pre párne p), tak funkcia f má v bode $\boldsymbol{a}$ ostré lokálne maximum;
(c) ak forma $f^{(p)}(\boldsymbol{a})$ je indefinitná (čo nastane vždy, keď p je nepárne), tak funkcia f nemá v bode $\boldsymbol{a}$ extrém; ak p je párne, tak f má v bode $\boldsymbol{a}$ sedlo.

Pre istotu pripomíname, že funkcia $f\colon M \to \mathbb{R}$ má v bode $\boldsymbol{a} \in M$ *sedlo*, ak existujú dva lineárne nezávislé vektory $\boldsymbol{u}, \boldsymbol{v}$ a kladné číslo ε také, že $\{\boldsymbol{a} + s\boldsymbol{u} + t\boldsymbol{v};\ s^2 + t^2 \le \varepsilon^2\} \subseteq M$, a funkcie $f_{\boldsymbol{a}}^{\boldsymbol{u}}, f_{\boldsymbol{a}}^{\boldsymbol{v}}\colon (-\varepsilon, \varepsilon) \to \mathbb{R}$, kde $f_{\boldsymbol{a}}^{\boldsymbol{u}}(t) = f(\boldsymbol{a} + t\boldsymbol{u})$, $f_{\boldsymbol{a}}^{\boldsymbol{v}}(t) = f(\boldsymbol{a} + t\boldsymbol{v})$ majú v bode $t = 0$ ostré lokálne maximum resp. minimum. Zrejme v prípade (c) sa sedlo v bode $\boldsymbol{a} \in M$ môže (no nemusí) vyskytnúť aj pre nepárne p.

Ak nenastane žiaden z prípadov (a), (b), (c), t.j. ak p je párne a forma $f^{(p)}(\boldsymbol{a})$ je kladne alebo záporne semidefinitná, ale nie definitná, tak o prítomnosti ani charaktere prípadných extrémov funkcie f v bode $\boldsymbol{a}$ nemožno rozhodnúť len na základe znalosti p-tej totálnej derivácie $f^{(p)}(\boldsymbol{a})$. Niekedy nám v takom prípade môže pomôcť „výlet do okolia“ bodu $\boldsymbol{a}$ (pozri cvičenie 33.6).

Bohužiaľ, na rozdiel od kvadratických foriem, pre p-homogénne formy párneho stupňa ≥ 4 nemáme k dispozícii jednoduchú a efektívnu metódu na určenie ich definitnosti. A tak má veta 33.3.1 väčšmi teoretický než praktický význam.

33.4 Alternujúce tenzory

Nech V, W sú vektorové priestory nad poľom K a $H\colon V^p \to W$ je *p-lineárne zobrazenie*. Hovoríme, že H je *alternujúce zobrazenie*, ak $H(\boldsymbol{x}_1, \dots, \boldsymbol{x}_p) = \boldsymbol{0}$ pre všetky $\boldsymbol{x}_1, \dots, \boldsymbol{x}_p \in V$, len čo pre nejaké $i < j \le n$ platí $\boldsymbol{x}_i = \boldsymbol{x}_j$. *Zobrazenie* H sa nazýva *antisymetrické*, ak zámenou poradia ľubovoľných argumentov $\boldsymbol{x}_i, \boldsymbol{x}_j \in V$, kde $i \ne j$, sa výraz $H(\boldsymbol{x}_1, \dots, \boldsymbol{x}_p) \in W$ zmení na opačný.

Úplne rovnako ako v dôkaze lemy 10.1.1 sa možno presvedčiť, že každé alternujúce multilineárne zobrazenie je antisymetrické. Ak charakteristika poľa K nie je 2, tak aj naopak každé antisymetrické multilineárne zobrazenie je očividne alternujúce. Ak však $\operatorname{char} K = 2$, tak podmienky antisymetrie a symetrie sú zrejme ekvivalentné.

Ako špeciálny prípad pre $W = K$ dostávame pojmy *alternujúceho* resp. *antisymetrického kovariantného tenzora* nad priestorom V. Zrejme tenzor $T \in \mathcal{T}_p^0(V)$ je antisymetrický práve vtedy, keď pre ľubovoľnú permutáciu $\sigma \in \mathcal{S}_p$ platí $T_\sigma = (\operatorname{sgn} \sigma)\, T$, kde $\operatorname{sgn} \sigma = (-1)^{|\sigma|}$ je *znamienko permutácie* σ (pozri paragraf 0.5). Podobne (zámenou vektorového priestoru V jeho duálom V^*) možno definovať aj pojmy *alternujúceho* resp. *antisymetrického kontravariantného tenzora* $T \in \mathcal{T}_0^p(V)$.

Nad poľami charakteristiky $\ne 2$, špeciálne nad poľami nekonečnej charakteristiky sa pojmy *alternujúci* a *antisymetrický tenzor* môžu používať ako synonymá. Nakoľko však nehodláme zo svojich úvah zbytočne vylučovať ani polia charakteristiky 2, radšej budeme dôsledne hovoriť o alternujúcich tenzoroch.

Kontravariantné alternujúce tenzory $T \in \mathcal{T}_0^p(V)$ tvoria lineárny podpriestor priestoru tenzorov $\mathcal{T}_0^p(V)$, ktorý značíme $\mathrm{Alt}^p(V)$. Podobne tvoria lineárny podpriestor $\mathrm{Alt}_p(V) \subseteq \mathcal{T}_p^0(V)$ kovariantné alternujúce tenzory. Vzhľadom na kanonický izomorfizmus $\mathrm{Alt}_p(V) \cong \mathrm{Alt}^p(V^*)$ sa stačí zaoberať kontravariantnými alternujúcimi tenzormi. Kvôli úplnosti (no v plnom súlade s definíciou) kladieme $\mathrm{Alt}^0(V) = K$, $\mathrm{Alt}^1(V) = V$, čiže skaláry $a \in K$ a vektory $\boldsymbol{x} \in V$ považujeme za alternujúce tenzory.

Jeden typ alternujúceho tenzora na vektorovom priestore $V = K^n$ všetkých usporiadaných n-tíc, zapisovaných ako stĺpce, už dôverne poznáme. Je to determinant chápaný ako zobrazenie $D = \det\colon K^{n\times n} \to K$, t. j. ako alternujúci kovariantný tenzor $D \in \mathrm{Alt}_n(K^n)$. Čoskoro uvidíme, že všetky alternujúce tenzory sú v podstate „iba“ determinanty.

Ďalej si všimnime, že pre $p > n = \dim V$ neexistujú nenulové alternujúce tenzory $T \in \mathcal{T}_0^p(V)$, teda $\mathrm{Alt}^p(V) = \{\mathbf{0}\}$. Z toho dôvodu je *priestor alternujúcich tenzorov* nad vektorovým priestorom V, čiže priamy súčet

$$\mathrm{Alt}(V) = \bigoplus_{p\in\mathbb{N}} \mathrm{Alt}^p(V) = \bigoplus_{p=0}^{n} \mathrm{Alt}^p(V),$$

konečnorozmerný. Našim cieľom je vybaviť ho štruktúrou asociatívnej graduovanej K-algebry. Obyčajný tenzorový súčin $\boldsymbol{x} \otimes \boldsymbol{y}$ vektorov $\boldsymbol{x}, \boldsymbol{y} \in V$ (čo sú alternujúce tenzory) však vo všeobecnosti alternujúci nie je. Pre ľubovoľné $\eta \in V^*$ totiž platí

$$(\boldsymbol{x} \otimes \boldsymbol{y})(\eta, \eta) = \boldsymbol{x}(\eta)\, \boldsymbol{y}(\eta) = \eta(\boldsymbol{x})\, \eta(\boldsymbol{y}) = (\eta \otimes \eta)(\boldsymbol{x}, \boldsymbol{y}).$$

Z toho vidno, že ak zároveň chceme, aby naša konštrukcia mala všeobecný charakter, t. j. dala sa rovnakým spôsobom aplikovať na ľubovoľný (konečnorozmerný) vektorový priestor, o. i. aj na duál V^* pôvodného priestoru V, je potrebné, aby aj samotný súčin v algebre $\mathrm{Alt}(V)$ bol na vektoroch $\boldsymbol{x}, \boldsymbol{y} \in V$ alternujúcim bilineárnym zobrazením.

Ku konštrukcii tejto algebry preto pristúpime z druhej strany, podobne ako sme si počínali v prípade symetrických tenzorov v paragrafe 33.1. V priestore V si zvolíme ľubovoľnú bázu $\boldsymbol{\beta} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$. Vopred sa dohodneme, že hľadaný súčin budeme značiť znakom $\wedge$ a nazývať *klinovým* alebo častejšie *vonkajším súčinom*. Alternujúci a asociatívny súčin $\boldsymbol{v}_{i_1} \wedge \cdots \wedge \boldsymbol{v}_{i_p}$ ľubovoľných p vektorov $\boldsymbol{v}_{i_1}, \ldots, \boldsymbol{v}_{i_p}$ tejto bázy je nenulový, len ak sú všetky indexy $i_1, \ldots, i_p$ navzájom rôzne. Pre dva takéto súčiny $\boldsymbol{v}_{i_1} \wedge \ldots \wedge \boldsymbol{v}_{i_p}$, $\boldsymbol{v}_{i_{\sigma(1)}} \wedge \ldots \wedge \boldsymbol{v}_{i_{\sigma(p)}}$, líšiace sa len poradím činiteľov, t. j. permutáciou $\sigma \in \mathcal{S}_p$, v dôsledku antisymetrie nevyhnutne platí

$$\boldsymbol{v}_{i_{\sigma(1)}} \wedge \ldots \wedge \boldsymbol{v}_{i_{\sigma(p)}} = (\operatorname{sgn}\sigma)(\boldsymbol{v}_{i_1} \wedge \ldots \wedge \boldsymbol{v}_{i_p}).$$

Každý taký súčin možno preto jednoznačne zapísať v tvare $\pm \boldsymbol{v}_{j_1} \wedge \ldots \wedge \boldsymbol{v}_{j_p}$, kde $1 \leq j_1 < \ldots < j_p \leq n$. Všetky formálne lineárne kombinácie

$$\sum_{j_1 < \ldots < j_p} a_{j_1 \ldots j_p} (\boldsymbol{v}_{j_1} \wedge \ldots \wedge \boldsymbol{v}_{j_p})$$

s koeficientmi $a_{j_1 \ldots j_p} \in K$ tvoria konečnorozmerný vektorový priestor $\Lambda^p(V)$ nad poľom K, ktorý nazývame *p-tou vonkajšou mocninou vektorového priestoru V*. Jej prvkom hovoríme *p-vektory* nad priestorom V. Zrejme $\Lambda^p(V) = \{\mathbf{0}\}$ pre $p > n$.

Pravidlo pre súčin p-vektora $\boldsymbol{v}_{i_1} \wedge \ldots \wedge \boldsymbol{v}_{i_p} \in \Lambda^p(V)$ a q-vektora $\boldsymbol{v}_{j_1} \wedge \cdots \wedge \boldsymbol{v}_{j_q} \in \Lambda^q(V)$

$$(\boldsymbol{v}_{i_1} \wedge \ldots \wedge \boldsymbol{v}_{i_p}) \wedge (\boldsymbol{v}_{j_1} \wedge \cdots \wedge \boldsymbol{v}_{j_q}) = \begin{cases} \mathbf{0}, \quad \text{ak } \{i_1, \ldots, i_p\} \cap \{j_1, \ldots, j_q\} \neq \emptyset, \\ \boldsymbol{v}_{i_1} \wedge \ldots \wedge \boldsymbol{v}_{i_p} \wedge \boldsymbol{v}_{j_1} \wedge \cdots \wedge \boldsymbol{v}_{j_q}, \\ \qquad \text{ak } \{i_1, \ldots, i_p\} \cap \{j_1, \ldots, j_q\} = \emptyset, \end{cases}$$

je vynútené požiadavkami alternovania a asociatívnosti, a spolu s podmienkou bilinearity jednoznačne určuje na priamom súčte

$$\Lambda(V) = \bigoplus_{p \in \mathbb{N}} \Lambda^p(V) = \bigoplus_{p=0}^{n} \Lambda^p(V)$$

štruktúru graduovanej algebry s vrstvami $\Lambda^p(V)$. Graduovanú algebru $(\Lambda(V), \wedge)$ nazývame *vonkajšou* alebo tiež *Grassmannovou algebrou vektorového priestoru V*. Priradením

$$\boldsymbol{v}_{i_1} \otimes \ldots \otimes \boldsymbol{v}_{i_p} \mapsto \boldsymbol{v}_{i_1} \wedge \cdots \wedge \boldsymbol{v}_{i_p}$$

je určený surjektívny homomorfizmus graduovaných algebier $\mathcal{T}_0(V) \to \Lambda(V)$, nazývaný *alternujúca kanonická projekcia*. Čitateľ by si mal rozmysliet, že samotná algebra $\Lambda(V)$ ani jej graduácia na vrstvy $\Lambda^p(V)$, na rozdiel od homomorfizmu $\mathcal{T}_0(V) \to \Lambda(V)$, nezávisia od výberu bázy $\boldsymbol{\beta}$ (pozri cvičenie 33.8 (a)).

Na vonkajšiu algebru $\Lambda(V)$ sa vďaka tomuto surjektívnemu homomorfizmu možno pozerať ako na *faktorovú algebru* graduovanej tenzorovej algebry $\mathcal{T}_0(V)$, ktorá vznikne stotožnením všetkých tenzorových súčinov $\boldsymbol{x}_1 \otimes \ldots \otimes \boldsymbol{x}_p \in \mathcal{T}_0^p(V)$, v ktorých sú vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p \in V$ lineárne závislé, s nulovým tenzorom $\mathbf{0} \in \mathcal{T}_0^p(V)$. Tento popis algebry $\Lambda(V)$ je už invariantný, t. j. nezávisí na báze $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ (pozri cvičenie 33.8 (b), (c)).

Násobenie vo vonkajšej algebre $\Lambda(V)$ je navyše *kosokomutatívne* alebo *graduovane antikomutatívne*, t. j. pre $A \in \Lambda^p(V)$, $B \in \Lambda^q(V)$ platí

$$B \wedge A = (-1)^{pq}(A \wedge B).$$

Platnosť tejto podmienky pre p-vektor $A = \boldsymbol{v}_{i_1} \wedge \ldots \wedge \boldsymbol{v}_{i_p}$ a q-vektor $B = \boldsymbol{v}_{j_1} \wedge \ldots \wedge \boldsymbol{v}_{j_q}$ možno nahliadnuť okamžite: stačí si uvedomiť, že na to, aby sme v súčine $A \wedge B$ dostali všetkých q činiteľov $\boldsymbol{v}_{j_l}$ dopredu, musíme postupne každý z nich vymeniť s každým z p činiteľov $\boldsymbol{v}_{i_k}$, čo má za následok pq zmien znamienka. Všeobecný prípad vyplýva zo špeciálneho na základe multilinearity klinového súčinu.

Všetkých súčinov $\boldsymbol{v}_{i_1} \wedge \cdots \wedge \boldsymbol{v}_{i_p}$ stupňa p, kde $1 \leq i_1 < \ldots < i_p \leq n$, je práve toľko, koľko je p-prvkových podmnožín množiny $\{1, \ldots, n\}$. Keďže tieto súčiny tvoria bázu priestoru $\Lambda^p(V)$, môžeme tak určiť dimenziu jednotlivých vrstiev $\Lambda^p(V)$ ako aj celej vonkajšej algebry $\Lambda(V)$.

33.4.1. Tvrdenie. *Vonkajšia algebra $\big(\Lambda(V), \wedge\big)$ n-rozmerného vektorového priestoru V je kosokomutatívna a asociatívna graduovaná K-algebra s jednotkou $1 \in K$. Pre dimenzie jej jednotlivých vrstiev resp. celej algebry platí*

$$\dim \Lambda^p(V) = \binom{n}{p} = \frac{n(n-1)\ldots(n-p+1)}{p!},$$
$$\dim \Lambda(V) = \sum_{p=0}^{n} \binom{n}{p} = 2^n.$$

Vonkajšiu p-tu mocninu vektorového priestoru V niekedy značíme aj ako klinový súčin

$$\Lambda^p(V) = \underbrace{V \wedge \ldots \wedge V}_{p}$$

p exemplárov vektorového priestoru V. Vonkajší súčin sa spomedzi alternujúcich multilineárnych zobrazení vydeľuje analogickou *univerzálnou vlastnosťou* ako bežný tenzorový súčin spomedzi všetkých bilineárnych zobrazení resp. symetrický tenzorový súčin spomedzi všetkých symetrických multilineárnych zobrazení (porovnaj s vetami 32.2.7 a 33.1.2).

33.4.2. Veta. *Nech V, W sú vektorové priestory nad poľom K, pričom V je konečnorozmerný a $p \in \mathbb{N}$. Potom*

(a) *operácia vonkajšieho súčinu $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p) \mapsto \boldsymbol{x}_1 \wedge \ldots \wedge \boldsymbol{x}_p$ je alternujúce p-lineárne zobrazenie $V^p \to \Lambda^p(V)$;*

(b) *ku každému alternujúcemu p-lineárnemu zobrazeniu $H \colon V^p \to W$ existuje práve jedno* lineárne *zobrazenie $\widehat{H} \colon \Lambda^p(V) \to W$ také, že*

$$\widehat{H}(\boldsymbol{x}_1 \wedge \ldots \wedge \boldsymbol{x}_p) = H(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p)$$

pre všetky $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p \in V$;

(c) *priradenie $H \mapsto \widehat{H}$ je lineárny izomorfizmus priestoru všetkých alternujúcich p-lineárnych zobrazení $V^p \to W$ na priestor $\mathcal{L}\big(\Lambda^p(V), W\big)$.*

Dôkaz opäť prenechávame ako cvičenie čitateľovi.

Ľahko možno nahliadnuť, že podmienky (a), (b) určujú priestor $\Lambda^p(V)$ jednoznačne až na izomorfizmus. Podmienky (b), (c) sú kľúčom k objasneniu súvisu jednotlivých priestorov alternujúcich tenzorov $\mathrm{Alt}^p(V) \subseteq \mathcal{T}_0^p(V)$ s vonkajšími mocninami $\Lambda^p(V)$ ako aj celého priestoru alternujúcich tenzorov $\mathrm{Alt}(V) \subseteq \mathcal{T}_0(V)$ s graduovanou vonkajšou algebrou $\Lambda(V)$.

Každý alternujúci tenzor $T \in \mathrm{Alt}^p(V) \subseteq \mathcal{T}_0^p(V)$ je alternujúca p-lineárna forma $T\colon (V^*)^p \to K$. Preto podľa (b) určuje lineárny funkcionál $\widehat{T}\colon \Lambda^p(V^*) \to K$ taký, že pre ľubovoľné $1 \le i_1 < \ldots < i_p \le n$ platí

$$\widehat{T}\big(\boldsymbol{v}^{i_1} \wedge \cdots \wedge \boldsymbol{v}^{i_p}\big) = T\big(\boldsymbol{v}^{i_1}, \ldots, \boldsymbol{v}^{i_p}\big),$$

kde $\boldsymbol{\beta}^* = (\boldsymbol{v}^1, \ldots, \boldsymbol{v}^n)$ je duálna báza k báze $\boldsymbol{\beta}$. Podľa (c) je priradenie $T \mapsto \widehat{T}$ lineárny izomorfizmus vektorového priestoru $\mathrm{Alt}^p(V)$ na vektorový priestor $\mathcal{L}(\Lambda^p(V^*), K) = \Lambda^p(V^*)^*$. Lineárny funkcionál $\widehat{T} \in \Lambda^p(V^*)^*$ možno prirodzene stotožniť s výrazom

$$\begin{aligned}\widehat{T} &= \sum_{i_1<\ldots<i_p} \widehat{T}\big(\boldsymbol{v}^{i_1} \wedge \cdots \wedge \boldsymbol{v}^{i_p}\big)(\boldsymbol{v}_{i_1} \wedge \cdots \wedge \boldsymbol{v}_{i_p})\\ &= \sum_{i_1<\ldots<i_p} T\big(\boldsymbol{v}^{i_1}, \ldots, \boldsymbol{v}^{i_p}\big)(\boldsymbol{v}_{i_1} \wedge \cdots \wedge \boldsymbol{v}_{i_p}) \in \Lambda^p(V)\end{aligned}$$

a priradenie $T \mapsto \widehat{T}$ chápať ako lineárny izomorfizmus $\mathrm{Alt}^p(V) \cong \Lambda^p(V)$.[5] Inverzné zobrazenie k tomuto izomorfizmu budeme značiť $A \mapsto A^{\pm}$. Špeciálne, alternujúci tenzor, ktorý sa zobrazí na p-vektor $\boldsymbol{v}_{i_1} \wedge \cdots \wedge \boldsymbol{v}_{i_p}$, je súčtom $p!$ sčítancov

$$(\boldsymbol{v}_{i_1} \wedge \cdots \wedge \boldsymbol{v}_{i_p})^{\pm} = \sum_{\varrho \in \mathcal{S}_p} (\operatorname{sgn} \varrho)(\boldsymbol{v}_{i_{\varrho(1)}} \otimes \ldots \otimes \boldsymbol{v}_{i_{\varrho(p)}}).$$

Súčin na priestore alternujúcich tenzorov opäť zadefinujeme tak, aby priradenie $T \mapsto \widehat{T}$ bolo homomorfizmom, teda izomorfizmom takto vzniknutej graduovanej algebry $\big(\mathrm{Alt}(V), \wedge\big)$ na graduovanú vonkajšiu algebru $\big(\Lambda(V), \wedge\big)$. Preto pre $S \in \mathrm{Alt}^p(V)$, $T \in \mathrm{Alt}^q(V)$ položíme

$$S \wedge T = \big(\widehat{S} \wedge \widehat{T}\big)^{\pm}$$

a podmienka homomorfnosti $\big(S \wedge T\big)^{\widehat{}} = \widehat{S} \wedge \widehat{T}$ bude automaticky splnená.

[5] Tým sme mimochodom dokázali ďalší očakávaný izomorfizmus $\Lambda^p(V)^* \cong \Lambda^p(V^*)$, ktorý je opäť kanonický, t. j. – napriek uvedenému popisu – nezávisí od bázy $\boldsymbol{\beta}$ (pozri paragraf 33.5 a cvičenie 33.8 (d)).

33.4.3. Veta. *Alternujúce (kontravariantné) tenzory nad n-rozmerným vektorovým priestorom V tvoria kosokomutatívnu a asociatívnu graduovanú algebru $\big(\mathrm{Alt}(V), \wedge\big)$ s jednotkou $1 \in K$, izomorfnú s vonkajšou algebrou $\big(\Lambda(V), \wedge\big)$ priestoru V. Pre dimenzie jej jednotlivých vrstiev a celej algebry platí*

$$\dim \mathrm{Alt}^p(V) = \binom{n}{p}, \qquad \textit{resp.} \qquad \dim \mathrm{Alt}(V) = 2^n.$$

Po stotožnení vonkajšej algebry $\Lambda(V^*)$ s algebrou $\mathrm{Alt}(V^*)$ sa alternujúce kovariantné tenzory $T \in \mathrm{Alt}(V^*)$ zvyknú nazývať tiež *formy* na priestore V; tenzorom $T \in \mathrm{Alt}^p(V^*) = \mathrm{Alt}_p(V)$ hovoríme *vonkajšie p-formy*.

Nasledujúci príklad konečne odhaľuje konkrétnu podobu alternujúcich tenzorov.

33.4.4. Príklad. Nech $V = K^n$ je priestor všetkých usporiadaných n-tíc skalárov z poľa K, zapisovaných ako stĺpce, s kanonickou bázou $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ a duálnou bázou $(\boldsymbol{e}^1, \dots, \boldsymbol{e}^n)$. Usporiadanú p-ticu vektorov $(\boldsymbol{a}_1, \dots, \boldsymbol{a}_p) \in V^p$ stotožníme ako obvykle s maticou $\boldsymbol{A} = \big(\boldsymbol{e}^i(\boldsymbol{a}_j)\big) = (a^i_j) \in K^{n\times p}$. Podľa definície vonkajšieho súčinu máme

$$\begin{aligned}
\big(\boldsymbol{e}^{i_1} \wedge \ldots \wedge \boldsymbol{e}^{i_p}\big)(\boldsymbol{A}) &= \sum_{\sigma \in \mathcal{S}_p} (\operatorname{sgn}\sigma)\big(\boldsymbol{e}^{i_{\sigma(1)}} \otimes \ldots \otimes \boldsymbol{e}^{i_{\sigma(p)}}\big)(\boldsymbol{a}_1, \dots, \boldsymbol{a}_p)\\
&= \sum_{\sigma \in \mathcal{S}_p} (\operatorname{sgn}\sigma)\big(\boldsymbol{e}^{i_1} \otimes \ldots \otimes \boldsymbol{e}^{i_p}\big)(\boldsymbol{a}_{\sigma(1)}, \ldots \boldsymbol{a}_{\sigma(p)})\\
&= \sum_{\sigma \in \mathcal{S}_p} (-1)^{|\sigma|}\, a^{i_1}_{\sigma(1)} \ldots a^{i_p}_{\sigma(p)}\\
&= \det \boldsymbol{A}(i_1, \dots, i_p).
\end{aligned}$$

Inak povedané, $\big(\boldsymbol{e}^{i_1} \wedge \ldots \wedge \boldsymbol{e}^{i_p}\big)(\boldsymbol{A})$ je proste *determinant* matice $\boldsymbol{A}(i_1, \dots, i_p) \in K^{p\times p}$ tvorenej postupne i_1-tým až i_p-tým riadkom matice $\boldsymbol{A}$.

Špeciálne pre $p = n$, $\boldsymbol{A} \in K^{n\times n}$ dostávame

$$\big(\boldsymbol{e}^1 \wedge \ldots \wedge \boldsymbol{e}^n\big)(\boldsymbol{A}) = \det \boldsymbol{A}.$$

Keďže $\dim \mathrm{Alt}_n(V) = \binom{n}{n} = 1$, determinant, chápaný ako alternujúca n-lineárna forma $\det\ = \boldsymbol{e}^1 \wedge \ldots \wedge \boldsymbol{e}^n \colon K^{n\times n} \to K$, je bázou priestoru $\mathrm{Alt}_n(V)$ (porovnaj s vetou 10.3.1 a cvičením 10.3).

Nasledujúce tvrdenie je len priamočiarym zovšeobecnením uvedeného príkladu. Jeho dôkaz preto prenechávame čitateľovi.

33.4.5. Tvrdenie. *Nech $\boldsymbol{x}_1, \dots, \boldsymbol{x}_p \in V$, $\eta^1, \dots, \eta^p \in V^*$. Potom*

$$(\eta^1 \wedge \ldots \wedge \eta^p)(\boldsymbol{x}_1, \dots, \boldsymbol{x}_p) = (\boldsymbol{x}_1 \wedge \ldots \wedge \boldsymbol{x}_p)(\eta^1, \dots, \eta^p) = \det\big(\langle \eta^i, \boldsymbol{x}_j\rangle\big).$$

Ani tentokrát, napriek inklúzii $\mathrm{Alt}(V) \subseteq \mathcal{T}_0(V)$, *nie je* $\mathrm{Alt}(V)$ *podalgebrou* graduovanej algebry $\mathcal{T}_0(V)$ všetkých kontravariantných tenzorov nad V. Ak však pole K má nekonečnú charakteristiku, tak – podobne ako v prípade symetrických tenzorov – možno definíciu klinového súčinu rozšíriť na celý priestor kontravariantných tenzorov $\mathcal{T}_0(V)$.

Ak znovu použijeme „metódu rovnomerného rozdelenia", ľahko nahliadneme, že klinový súčin bázických vektorov $\boldsymbol{v}_{i_1} \otimes \ldots \otimes \boldsymbol{v}_{i_p} \in \mathcal{T}_0^p(V)$, $\boldsymbol{v}_{j_1} \otimes \ldots \otimes \boldsymbol{v}_{j_q} \in \mathcal{T}_0^q(V)$ treba zadefinovať formulou

$$(\boldsymbol{v}_{i_1} \otimes \ldots \otimes \boldsymbol{v}_{i_p}) \wedge (\boldsymbol{v}_{j_1} \otimes \ldots \otimes \boldsymbol{v}_{j_q}) = \frac{1}{p!\,q!}(\boldsymbol{v}_{i_1} \wedge \ldots \wedge \boldsymbol{v}_{i_p} \wedge \boldsymbol{v}_{j_1} \wedge \ldots \wedge \boldsymbol{v}_{j_q})^{\pm}.$$

Podrobnejšie preberieme aj druhý variant alternujúceho tenzorového súčinu, ktorému niektorí autori dávajú prednosť. Predpokladajme zatiaľ len, že $\operatorname{char} K > p$. *Alternáciou*, prípadne tiež *antisymetrizáciou tenzora* $T \in \mathcal{T}_0^p(V)$ nazývame výraz

$$\operatorname{alt} T = \frac{1}{p!} \sum_{\sigma \in \mathcal{S}_p} (\operatorname{sgn} \sigma)\, T^{\sigma}.$$

Predpisom $T \mapsto \operatorname{alt} T$ je zrejme definované lineárne zobrazenie $\operatorname{alt}\colon \mathcal{T}_0^p(V) \to \mathcal{T}_0^p(V)$, pričom tenzor $T \in \mathcal{T}_0^p(V)$ je alternujúci práve vtedy, keď $T = \operatorname{alt} T$. Navyše výraz $p!\,(\operatorname{alt} T)^{\wedge} \in \Lambda^p(V)$ je zrejme alternujúca kanonická projekcia tenzora $T \in \mathcal{T}_0^p(V)$.

Ak $p + q < \operatorname{char} K$, tak súčin alternujúcich tenzorov $S \in \mathrm{Alt}^p(V)$, $T \in \mathrm{Alt}^q(V)$ možno definovať formulou

$$S \mathbin{\bar\wedge} T = \operatorname{alt}(S \otimes T) = \frac{1}{(p+q)!} \sum_{\sigma \in \mathcal{S}_{p+q}} (\operatorname{sgn} \sigma)(S \otimes T)^{\sigma}.$$

Túto formulu však možno aplikovať na ľubovoľné tenzory $S \in \mathcal{T}_0^p(V)$, $T \in \mathcal{T}_0^q(V)$, pričom stále platí $S \mathbin{\bar\wedge} T \in \mathrm{Alt}^{p+q}(V)$ a taktiež

$$S \mathbin{\bar\wedge} T = \operatorname{alt}(S \otimes T) = \operatorname{alt} S \mathbin{\bar\wedge} \operatorname{alt} T.$$

Inak povedané, alternácia $T \mapsto \operatorname{alt} T$ je homomorfizmom každej z graduovaných algebier $\big(\mathcal{T}_0(V), \otimes\big)$, $\big(\mathcal{T}_0(V), \bar\wedge\,\big)$ na graduovanú algebru $\big(\mathrm{Alt}(V), \bar\wedge\,\big)$.

Tento „nový" klinový súčin $\bar\wedge$ súvisí s „pôvodným" klinovým súčinom $\wedge$ jednoduchým vzťahom: pre ľubovoľný p-vektor $S \in \mathrm{Alt}^p(V)$ a q-vektor $T \in \mathrm{Alt}^q(V)$ platí

$$S \mathbin{\bar\wedge} T = \frac{p!\,q!}{(p+q)!}(S \wedge T)$$

(overte si samostatne ako cvičenie). Z tejto rovnosti vyplýva izomorfizmus všetkých troch graduovaných algebier $\big(\mathrm{Alt}(V), \bar\wedge\,\big)$, $\big(\mathrm{Alt}(V), \wedge\big)$, $\big(\Lambda(V), \wedge\big)$. Takže, podobne ako vo vete 33.2.1, môžeme zhrnúť:

33.4.6. Veta. *Nech V je konečnorozmerný vektorový priestor nad poľom charakteristiky ∞. Potom súčin $\bar\wedge$ vytvára na priestore kontravariantných tenzorov $\mathcal{T}_0(V)$ štruktúru kosokomutatívnej asociatívnej graduovanej algebry $\big(\mathcal{T}_0(V), \bar\wedge\,\big)$ s jednotkou $1 \in K$, v ktorej alternujúce tenzory tvoria graduovanú podalgebru $\mathrm{Alt}(V)$ a kanonická alternujúca projekcia $T \mapsto p!\,(\mathrm{alt}\,T)\hat{}$ pre $T \in \mathcal{T}_0^p(V)$ je homomorfizmus každej z graduovaných algebier $\big(\mathcal{T}_0(V), \otimes\big)$, $\big(\mathcal{T}_0(V), \bar\wedge\,\big)$ na graduovanú algebru $\big(\Lambda(V), \wedge\big)$. Jej zúženie na podalgebru alternujúcich tenzorov je izomorfizmom graduovaných algebier $\big(\mathrm{Alt}(V), \bar\wedge\,\big) \cong \big(\Lambda(V), \wedge\big)$.*

Z uvedených dôvodov sa aj súčin $S\,\bar\wedge\,T$ štandardne značí $S \wedge T$. Vzhľadom na reprezentáciu alternujúcich tenzorov prvkami vonkajšej algebry $\Lambda(V)$ však voľba konkrétneho alternujúceho súčinu v $\mathcal{T}_0(V)$ resp. v $\mathrm{Alt}(V)$ nehrá až takú významnú úlohu.

33.5 Dualita vo vonkajšej algebre

V celom paragrafe predpokladáme, že V je n-rozmerný vektorový priestor nad poľom K, $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ je nejaká jeho báza a $\boldsymbol{\beta}^* = (\boldsymbol{v}^1, \dots, \boldsymbol{v}^n)$ je k nej duálna báza.

33.5.1. Dualita a vektorový súčin. Vektorové priestory alternujúcich tenzorov $\Lambda^p(V)$, $\Lambda^{n-p}(V)$, $\Lambda_p(V)$, $\Lambda_{n-p}(V)$ majú rovnakú dimenziu $\binom{n}{p} = \binom{n}{n-p}$, sú preto lineárne izomorfné. Pri bližšom pohľade sa ukazuje, že tieto izomorfizmy majú charakter duality.

Predpis $\big(\eta^1, \dots, \eta^p, \boldsymbol{x}_1, \dots, \boldsymbol{x}_p\big) \mapsto \det\big(\langle \eta^i, \boldsymbol{x}_j \rangle\big)$ určuje multilineárne zobrazenie $(V^*)^p \times V^p \to K$, alternujúce v prvých a rovnako aj v druhých p argumentoch. Preto podľa vety 33.4.2 (b) (ktorú treba použiť dvakrát) indukuje bilineárnu formu $F\colon \Lambda^p(V^*) \times \Lambda^p(V) \to K$ takú, že

$$F\big(\eta^1 \wedge \ldots \wedge \eta^p, \boldsymbol{x}_1 \wedge \ldots \wedge \boldsymbol{x}_p\big) = \det\big(\langle \eta^i, \boldsymbol{x}_j \rangle\big)$$

pre $\boldsymbol{x}_1, \dots, \boldsymbol{x}_p \in V$, $\eta^1, \dots, \eta^p \in V^*$. Ľahko možno nahliadnuť, že forma F je regulárna, preto predpisom $\Phi(\eta^1, \dots, \eta^p)(\boldsymbol{x}_1, \dots, \boldsymbol{x}_p) = \det\big(\langle \eta^i, \boldsymbol{x}_j \rangle\big)$ je podľa vety 11.1.7 definovaný kanonický lineárny izomorfizmus $\Phi\colon \Lambda_p(V) = \Lambda^p(V^*) \cong \Lambda^p(V)^*$. Rovnako možno dospieť ku kanonickému izomorfizmu $\Lambda^p(V) = \Lambda_p(V^*) \cong \Lambda_p(V)^*$.

Taktiež klinový súčin medzi vrstvami $\Lambda^p(V)$ a $\Lambda^{n-p}(V)$ vo vonkajšej algebre $\Lambda(V)$ je bilineárne zobrazenie $\wedge\colon \Lambda^p(V) \times \Lambda^{n-p}(V) \to \Lambda^n(V)$. Uvedomme si, že priestor $\Lambda^n(V)$ je jednorozmerný s bázou (napr.) $\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_n$. Pre $T \in \Lambda^n(V)$ označme $B(T)$ ten jediný skalár $b \in K$, pre ktorý platí

$T = b(\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_n)$. Potom B môžeme považovať aj za regulárnu bilineárnu formu $B\colon \Lambda^p(V) \times \Lambda^{n-p}(V) \to K$ takú, že $B(T_1, T_2) = B(T_1 \wedge T_2)$ pre $T_1 \in \Lambda^p(V)$, $T_2 \in \Lambda^{n-p}(V)$. Z nej dostávame lineárny izomorfizmus $\Lambda^p(V) \cong \Lambda^{n-p}(V)^*$.

Uvedené priradenie je však popísané pomocou bázy $\boldsymbol{\beta}$. Vyjasníme, v akom zmysle ho možno považovať za kanonické. S inou bázou $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \ldots, \boldsymbol{v}_n)$ by sme pre každé $T \in \Lambda^n(V)$ mali $T = b(\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_n) = a(\boldsymbol{u}_1 \wedge \ldots \wedge \boldsymbol{u}_n)$, kde $a = b \det \boldsymbol{P}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$. A tak sú forma B aj ňou indukovaný izomorfizmus určené jednoznačne až na multiplikatívnu konštantu.

Obraz p-vektora $\boldsymbol{x}_1 \wedge \ldots \wedge \boldsymbol{x}_p$ v izomorfizme $\Lambda^p(V) \cong \Lambda^{n-p}(V)^*$ je teda lineárny funkcionál $\Lambda^{n-p}(V) \to K$, ktorý značíme $\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_p$, prípadne $\boldsymbol{x}_1 \vee \ldots \vee \boldsymbol{x}_p$ a nazývame *vektorový* alebo tiež *duálny vonkajší súčin* vektorov $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p \in V$. Jeho pôsobenie na $(n-p)$-vektor $\boldsymbol{y}_1 \wedge \ldots \wedge \boldsymbol{y}_{n-p} \in \Lambda^{n-p}(V)$ možno pomocou súradníc v báze $\boldsymbol{\beta}$ opísať názorne ako determinant

$$(\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_p)(\boldsymbol{y}_1 \wedge \ldots \wedge \boldsymbol{y}_{n-p}) = \det\big((\boldsymbol{x}_1)_{\boldsymbol{\beta}}, \ldots, (\boldsymbol{x}_p)_{\boldsymbol{\beta}}, (\boldsymbol{y}_1)_{\boldsymbol{\beta}}, \ldots, (\boldsymbol{y}_{n-p})_{\boldsymbol{\beta}}\big).$$

Inak povedané,

$$\boldsymbol{x}_1 \wedge \ldots \wedge \boldsymbol{x}_p \wedge \boldsymbol{y}_1 \wedge \ldots \wedge \boldsymbol{y}_{n-p} = (\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_p)(\boldsymbol{y}_1 \wedge \ldots \wedge \boldsymbol{y}_{n-p})(\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_n).$$

Ak vezmeme do úvahy aj skôr spomínaný kanonický izomorfizmus $\Lambda^{n-p}(V)^* \cong \Lambda_{n-p}(V) = \Lambda^{n-p}(V^*)$ dostávame ďalší „takmer kanonický“ izomorfizmus

$$\Lambda^p(V) \cong \Lambda^{n-p}(V^*),$$

takže vektorový súčin $\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_p$ môžeme chápať aj ako prvok vonkajšej mocniny $\Lambda^{n-p}(V^*)$.

V euklidovskom priestore V by sme si za $\boldsymbol{\beta}$ samozrejme zvolili nejakú ortonormálnu bázu, ktorou by sme zadali orientáciu priestoru V (pozri paragraf 15.2). Potom pre každú ortonormálnu bázu $\boldsymbol{\alpha}$ zhodne orientovanú s bázou $\boldsymbol{\beta}$ by sme dostali tú istú bilineárnu formu B a ten istý lineárny izomorfizmus $\Lambda^p(V) \cong \Lambda^{n-p}(V)^*$ ako pre bázu $\boldsymbol{\beta}$. Navyše po stotožnení priestoru V s jeho duálom V^* prostredníctvom skalárneho súčinu dostávame izomorfizmy

$$\Lambda^p(V) \cong \Lambda^{n-p}(V^*) \cong \Lambda^{n-p}(V),$$

kanonické z hľadiska všetkých kladne orientovaných ortonormálnych báz. V tomto prípade teda môžeme vektorový súčin $\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_p$ konečne považovať za $(n-p)$-vektor z vonkajšej mocniny $\Lambda^{n-p}(V)$. Zrejme pre $p = n-1$ splýva $\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_{n-1}$ s vektorovým súčinom z paragrafu 15.4.

33.5.2. Vnútorný súčin a Hodgeov operátor. Dosadením pevných vektorov $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p \in V$ na prvých p-miest alternujúcej $(p+q)$-formy $F \in \mathrm{Alt}_{p+q}(V)$ získame q-formu $\boldsymbol{i}_{\boldsymbol{x}_1 \ldots \boldsymbol{x}_p}(F) \in \mathrm{Alt}_q(V)$ takú, že

$$\boldsymbol{i}_{\boldsymbol{x}_1 \ldots \boldsymbol{x}_p}(F)(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_q) = F(\boldsymbol{x}_1 \ldots \boldsymbol{x}_p, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_q)$$

pre $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_q \in V$. Priradením $\big(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p, F\big) \mapsto \boldsymbol{i}_{\boldsymbol{x}_1 \ldots \boldsymbol{x}_p}(F)$ je dané multilineárne zobrazenie $V^p \times \mathrm{Alt}_{p+q}(V) \to \mathrm{Alt}_q(V)$ alternujúce v prvých p-zložkách. To podľa viet 32.2.7 (b) a 33.4.2 (b) indukuje lineárne zobrazenie $\Lambda^p(V) \otimes \mathrm{Alt}_{p+q}(V) \to \mathrm{Alt}_q(V)$ nazývané *vnútorný súčin*. Teda pre $T \in \Lambda_p(V)$, $F \in \mathrm{Alt}_{p+q}(V)$, $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_q \in V$ platí

$$\boldsymbol{i}_T(F)(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_q) = \widehat{F}(T \wedge \boldsymbol{y}_1 \wedge \ldots \wedge \boldsymbol{y}_q).$$

Rovnako funguje aj vnútorný súčin $\Lambda_p(V) \times \mathrm{Alt}^{p+q}(V) \to \mathrm{Alt}^q(V)$.

Pre $q = 0$ nám bilineárne forma $(T, F) \mapsto \boldsymbol{i}_T(F) \colon \Lambda^p(V) \times \mathrm{Alt}_p(V) \to \mathrm{Alt}_0(V) = K$ dáva iný pohľad na dualitu

$$\mathrm{Alt}_p(V) = \mathrm{Alt}^p(V^*) \cong \Lambda^p(V)^*.$$

Pre $q = n - p$ je zasa $\mathrm{Alt}_n(V) \cong K$, takže $\Lambda^p(V) \otimes \mathrm{Alt}_n(V) \cong \Lambda^p(V)$ a predpisom

$$*T = \boldsymbol{i}_T(\boldsymbol{v}^1 \wedge \ldots \wedge \boldsymbol{v}^n)$$

je definovaný izomorfizmus $* : \Lambda^p(V) \cong \mathrm{Alt}_{n-p}(V)$ nazývaný *Hodgeov operátor* alebo *Hodgeova dualita*. Obvyklú podobu Hodgeovho operátora $* : \Lambda^p(V) \cong \Lambda_{n-p}(V)$ dostávame po stotožnení $\mathrm{Alt}_{n-p}(V) \cong \Lambda_{n-p}(V)$. V euklidovskom priestore V môžeme ďalej stotožniť $V \cong V^*$ a $\Lambda_{n-p}(V) = \Lambda^{n-p}(V^*) \cong \Lambda^{n-p}(V)$, takže po voľbe nejakej ortonormálnej bázy $\boldsymbol{\beta}$ funguje Hodgeova dualita $* : \Lambda^p(V) \cong \Lambda^{n-p}(V)$ kanonicky vzhľadom na všetky s ňou súhlasne orientované ortonormálne bázy. Nie je ťažké nahliadnuť, že Hodgeova dualita splýva s dualitami $\Lambda^p(V) \cong \Lambda_{n-p}(V)$, resp. $\Lambda^p(V) \cong \Lambda^{n-p}(V)$ z odstavca 33.5.1. Presnejšie, pre $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p \in V$ platí

$$\boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_p = *(\boldsymbol{x}_1 \wedge \ldots \wedge \boldsymbol{x}_p).$$

33.6 Symetrické a alternujúce tenzory v kvantovej mechanike

Jeden zo základných princípov kvantovej mechaniky postuluje, že stavový priestor W kvantovo-mechanického systému $\mathbf{S}$, ktorý je spojením kvantovomechanických systémov $\mathbf{S}_1, \ldots, \mathbf{S}_p$ so stavovými priestormi $V_1, \ldots, V_p$, je „vhodný“ lineárny podpriestor tenzorového súčinu $V_1 \otimes \ldots \otimes V_p = V_1 \otimes_{\mathbb{C}} \ldots \otimes_{\mathbb{C}}$

V_p.[6] Intuitívne, rozložiteľný vektor $\psi_1 \otimes \ldots \otimes \psi_p$ reprezentuje stav systému $\mathbf{S}$, zodpovedajúci situácii, keď každý z jeho podsystémov $\mathbf{S}_j$ je v stave $\psi_j \in V_j$. Nerozložiteľné stavové vektory však takúto intuitívnu reprezentáciu nemajú, v dôsledku čoho úvahy o tom, aké stavy podsystémov $\mathbf{S}_j$ reprezentuje všeobecný stavový vektor $\sum_k a_k(\psi_{1k} \otimes \ldots \otimes \psi_{pk}) \in W$, ako aj samotná predstava o rozdelení systému $\mathbf{S}$ na pôvodné podsystémy, strácajú jasný zmysel.

Dôležité špeciálne typy kvantovomechanických systémov vznikajú spojením podsystémov pozostávajúcich z identických častíc. Stavy každého z takýchto podsystémov (to znamená stavy každej jednotlivej častice) možno reprezentovať vektormi z toho istého priestoru $V_j = V$. Na celý systém sa potom vzťahuje *princíp nerozlíšiteľnosti* rovnakých častíc. Podľa neho v takomto systéme nemožno rozlíšiť jednotlivé častice, ani určiť stavy, v ktorých sa každá z nich nachádza. Napr. v systéme pozostávajúcom z dvoch rovnakých častíc nemožno rozlíšiť celkový stav, keď je prvá častica v stave $\varphi \in V$ a druhá v stave $\psi \in V$, od celkového stavu, keď je prvá v stave ψ a druhá v stave φ (presnejšie, takáto formulácia problému vôbec nemá zmysel). Pri matematickom modelovaní sa to prejaví tak, že stavový priestor systému je buď podpriestor symetrických tenzorov $\mathrm{Sym}^p(V) \subseteq \mathcal{T}_0^p(V)$ alebo podpriestor alternujúcich (antisymetrických) tenzorov $\mathrm{Alt}^p(V) \subseteq \mathcal{T}_0^p(V)$, kde p je počet identických častíc. Ktorý z oboch prípadov nastane, závisí na spine príslušného druhu častíc.

Systému p rovnakých častíc s celočíselným spinom (vrátane častíc bez spinu, čiže so spinom 0) zodpovedá priestor symetrických tenzorov $\mathrm{Sym}^p(V) \cong \Sigma^p(V)$. Ich chovanie podlieha tzv. *Boseho-Einsteinovej štatistike* a hovoríme im *bozóny*. Medzi bozóny patria napr. fotóny, mezóny a α-častice (jadrá hélia).

Systému p rovnakých častíc s poloceločíselným spinom zodpovedá priestor alternujúcich tenzorov $\mathrm{Alt}^p(V) \cong \Lambda^p(V)$. Ich chovanie sa riadi tzv. *Fermiho-Diracovou štatistikou* a hovoríme im *fermióny*. Medzi fermióny patria napr. elektróny a protóny (a ich antičastice – pozitróny a antiprotóny), ako aj neutróny.

Ak $(\psi_1, \ldots, \psi_n)$ je nejaká ortonormálna báza stavového priestoru V, tak stav systému p identických bozónov „k_1 častíc v stave ψ_1, k_2 častíc v stave ψ_2, …, k_n častíc v stave ψ_n“ fyzici zvyčajne zapisujú ako symetrický tenzor

$$|k_1, \ldots, k_n\rangle = \binom{p}{k_1, \ldots, k_n}^{-1/2} \psi_1^{k_1} \ldots \psi_n^{k_n} \in \Sigma^p(V),$$

kde $k_1, \ldots, k_n$ sú nezáporné celé čísla také, že $k_1 + \ldots + k_n = p$. Podobne, stav systému p identických fermiónov „po jednej častici v stavoch $\psi_{j_1}, \ldots, \psi_{j_p}$,“

[6] V typickom prípade, keď aspoň dva z Hilbertových priestorov $V_1, \ldots, V_p$ sú nekonečnorozmerné, je potrebné nahradiť tenzorový súčin $V_1 \otimes \ldots \otimes V_p$ jeho *zúplnením*. Túto technickú komplikáciu však nebudeme brať do úvahy.

kde $1 \leq j_1 < \ldots < j_p \leq n$, sa zvykne zapisovať ako alternujúci tenzor

$$|k_1, \ldots, k_n\rangle = (p!)^{-1/2}(\psi_{j_1} \wedge \ldots \wedge \psi_{j_p}) \in \Lambda^p(V),$$

kde pre $j \leq n$ platí $k_j = 1$, ak $j \in \{j_1, \ldots, j_p\}$, a $k_j = 0$ v opačnom prípade. Čísla $k_1, \ldots, k_n$ zakaždým udávajú *stupne zaplnenia* jednotlivých bázických stavov $\psi_1, \ldots, \psi_n$. Normujúce koeficienty $\binom{p}{k_1,\ldots,k_n}^{-1/2}$ resp. $(p!)^{-1/2}$ sa môžu líšiť podľa voľby konkrétnej reprezentácie symetrického súčinu v $\mathrm{Sym}^p(V)$ resp. vonkajšieho súčinu v $\mathrm{Alt}^p(V)$.

Symetria stavového tenzora $|k_1, \ldots, k_n\rangle \in \Sigma^p(V)$ znamená, že dva „mysliteľné stavy“ $\psi_{i_1} \otimes \ldots \otimes \psi_{i_p}$, $\psi_{j_1} \otimes \ldots \otimes \psi_{j_p}$ systému p identických bozónov, líšiace sa len poradím činiteľov ψ_j sú nerozlíšiteľné, teda nemôžu vystupovať oddelene, ale iba ako súčasť toho istého symetrického tenzora $\psi_{i_1} \ldots \psi_{i_p} = \psi_{j_1} \ldots \psi_{j_p} = \psi_1^{k_1} \ldots \psi_n^{k_n}$.

Z toho, že stavový tenzor $|k_1, \ldots, k_n\rangle \in \Lambda^p(V)$ systému p identických fermiónov je alternujúci, vyplýva obmedzujúca podmienka $k_j \in \{0, 1\}$ na stupne zaplnenia jednotlivých bázických stavov. Ide o matematický prejav tzv. *Pauliho vylučovacieho princípu*, podľa ktorého v systéme identických fermiónov sa žiadne dve častice nemôžu nachádzať v rovnakom stave. Antisymetria takýchto tenzorov je už čisto matematickým dôsledkom ich alternovania a multilinearity.

Keďže v priebehu vývoja kvantovomechanického systému môžu častice vznikať i zanikať, je potrebné matematicky zachytiť aj systémy s premenným počtom identických bozónov resp. fermiónov. Prvému prípadu zodpovedá symetrická algebra $\Sigma(V) = \bigoplus_{p\in\mathbb{N}} \Sigma^p(V)$, druhému vonkajšia algebra $\Lambda(V) = \bigoplus_{p\in\mathbb{N}} \Lambda^p(V)$. Stojí za povšimnutie, že v typickom prípade nekonečnorozmerného Hilbertovho priestoru V sú netriválne všetky priame sčítance $\Lambda^p(V)$, $p \in \mathbb{N}$. Naopak, ak V má konečný rozmer n, tak z Pauliho vylučovacieho princípu vyplýva, že takýto systém môže obsahovať nanajvýš n častíc. Taktiež si treba uvedomiť, že v oboch prípadoch tenzor $|0, \ldots, 0, \ldots\rangle$ pozostávajúci zo samých núl (či už ich je konečne alebo nekonečne mnoho) nepredstavuje nulový tenzor $\mathbf{0}$ (ktorému napokon nezodpovedá žiaden stav systému), ale nenulový skalár z priestoru $\Sigma^0(V) = \Lambda^0(V) = \mathbb{C}$. Dohodou stanovujeme $|0, \ldots, 0, \ldots\rangle = 1$.

Uvedený tvar stavových tenzorov umožňuje veľmi názorný zápis napr. *anihilačných* a k nim združených *kreačných operátorov.*

Jednotlivé *operátory anihilácie* resp. *kreácie* $A_j, A_j^* : \Sigma(V) \to \Sigma(V)$ bozónu v stave ψ_j vyzerajú takto:

$$A_j|k_1, \ldots, k_j, \ldots, k_p\rangle = \sqrt{k_j}\,|k_1, \ldots, k_j \dot{-} 1, \ldots, k_p\rangle,$$
$$A_j^*|k_1, \ldots, k_j, \ldots, k_p\rangle = \sqrt{k_j + 1}\,|k_1, \ldots, k_j + 1, \ldots, k_p\rangle,$$

kde $k \dot{-} 1 = \max\{0,\, k-1\}$. Jednoduché overenie komutačných vzťahov

$$[A_i, A_j] = [A_i^*, A_j^*] = \mathbf{0}, \qquad\qquad [A_i, A_j^*] = \delta_{ij}\,\mathrm{id}$$

prenechávame čitateľovi.

Pre *operátory anihilácie* resp. *kreácie* $A_j, A_j^* : \Lambda(V) \to \Lambda(V)$ fermiónu v stave ψ_j máme vyjadrenia

$$\begin{aligned} A_j|k_1, \dots, k_j, \dots, k_p\rangle &= (-1)^{k_1+\ldots+k_{j-1}} \sqrt{k_j}\,|k_1, \dots, k_j \dot{-} 1, \dots, k_p\rangle, \\ A_j^*|k_1, \dots, k_j, \dots, k_p\rangle &= (-1)^{k_1+\ldots+k_{j-1}} \sqrt{k_j \oplus 1}\,|k_1, \dots, k_j \oplus 1, \dots, k_p\rangle, \end{aligned}$$

kde $\oplus$ označuje sčítanie modulo 2. Znamienkový faktor $(-1)^{k_1+\ldots+k_{j-1}}$ má konvenčný charakter a mohol by sa voliť aj inak. Ľahko možno overiť, že ľubovoľné dva kreačné resp. anihilačné operátory antikomutujú, t. j. $A_iA_j = -A_jA_i$ a $A_i^*A_j^* = -A_j^*A_i^*$. Pre kompozície anihilačného a kreačného operátora však platí $A_iA_j^* + A_j^*A_i = \delta_{ij}\,\mathrm{id}$. Ak si teda označíme

$$\{A, B\} = AB + BA$$

antikomutátor operátorov A, B, dostávame tzv. *antikomutačné vzťahy*

$$\{A_i, A_j\} = \{A_i^*, A_j^*\} = \mathbf{0}, \qquad\qquad \{A_i, A_j^*\} = \delta_{ij}\,\mathrm{id}\,.$$

Špeciálne $A_j^*A_j^* = \mathbf{0}$, teda v zhode s Pauliho vylučovacím princípom nie je možné kreovať dva identické fermióny v tom istom stave ψ_j.

Význam kreačných a anihilačných operátorov spočíva hlavne v tom, že – podobne ako sme to videli v paragrafe 26.10 pri kvantovom harmonickom oscilátore – s ich pomocou možno vyjadriť hamiltonián systému aj viacero ďalších hermitovských operátorov zodpovedajúcich pozorovateľným fyzikálnym veličinám, popisujúcim jeho vlastnosti. Napr. hermitovský operátor $A_j^*A_j$ popisuje v oboch prípadoch počet častíc v stave ψ_j a operátor $\sum_{j=1}^n A_j^*A_j$ udáva počet všetkých častíc v systéme.

Cvičenia

33.1. (a) Dokážte vetu 33.1.2.

(b) Sformulujte, čo presne znamená, že priestor $\Sigma^p(V)$ je podmienkami (a), (b) tejto vety určený jednoznačne až na izomorfizmus, a dokážte to.

33.2. (a) Nech $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$, $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ sú dve bázy vektorového priestoru V. Dokážte, že priradenie $\boldsymbol{u}_i \mapsto \boldsymbol{v}_i$ možno jednoznačne rozšíriť do izomorfizmu graduovaných algebier $K[\boldsymbol{u}_1, \dots, \boldsymbol{u}_n] \cong K[\boldsymbol{v}_1, \dots, \boldsymbol{v}_n]$. Odvoďte z toho, že symetrická algebra $\Sigma(V)$ nezávisí na báze priestoru V.

(b) Označme J_{sym} najmenší ideál graduovanej algebry $\mathcal{T}_0(V)$, ktorý obsahuje všetky prvky tvaru $T - T^\sigma$, kde $T \in \mathcal{T}_0^p(V)$ a $\sigma \in \mathcal{S}_p$ pre $p \in \mathbb{N}$ (pozri cvičenie 30.6). Dokážte, že graduovaná symetrická algebra $\Sigma(V)$ je izomorfná s faktorovou algebrou $\mathcal{T}_0(V)/J_{\mathrm{sym}}$, čo je invariantný (t. j. na báze priestoru V nezávislý) popis symetrickej algebry $\Sigma(V)$.

(c) Nájdite kanonický (t. j. na báze priestoru V nezávislý) lineárny izomorfizmus $\Sigma^p(V)^* \cong \Sigma^p(V^*)$.

33.3. (a) Dokážte že počet všetkých usporiadaných p-tic $(i_1, \dots, i_p)$ prvkov z množiny $\{1, \dots, n\}$, ktoré obsahujú každý z indexov $i \le n$ práve k_i-krát udáva *polynomický koeficient* $\binom{p}{k_1,\dots,k_n} = \frac{p!}{k_1! \dots k_n!}$.

(b) Dokážte, že tento počet je rovnaký ako počet rozkladov množiny $\{1, \dots, p\}$ na n disjunktných nie nevyhnutne neprázdnych podmnožín, ktoré majú postupne $k_1, \dots, k_n$ prvkov.

(c) Dokážte tzv. *polynomickú vetu* (pre $p = 2$ ide samozrejme o binomickú vetu): Pre prvky $a_1, \dots, a_n \in K$ platí

$$(a_1 + \dots + a_n)^p = \sum \binom{p}{k_1, \dots, k_n} a_1^{k_1} \dots a_n^{k_n},$$

kde súčet prebieha cez všetky usporiadané n-tice nezáporných celých čísel $k_1, \dots, k_n$ takých, že $k_1 + \dots + k_n = p$.

(d) Odvoďte z toho identitu $\sum \binom{p}{k_1,\dots,k_n} = n^p$ (pri rovnakej sumácii ako v (c)).

33.4. Nech $\operatorname{char} K > p$ a $f : K^n \to K$ je p-homogénna forma daná predpisom

$$f(\boldsymbol{x}) = \sum_{k_1 + \dots + k_n = p} a_{k_1 \dots k_n} x_1^{k_1} \dots x_n^{k_n}.$$

Označme pol $f = T = \sum_{i_1 \dots i_p} T_{i_1 \dots i_p} \left(\boldsymbol{e}^{i_1} \otimes \dots \otimes \boldsymbol{e}^{i_p}\right)$. Nájdite explicitné vyjadrenie zložiek $T_{i_1 \dots i_p}$ symetrického tenzora T pomocou koeficientov $a_{k_1 \dots k_n}$.

33.5. Odvoďte explicitný vzťah

$$(\boldsymbol{v}_{i_1} \otimes \dots \otimes \boldsymbol{v}_{i_p}) \odot (\boldsymbol{v}_{j_1} \otimes \dots \otimes \boldsymbol{v}_{j_q}) = \frac{\binom{p}{k_1,\dots,k_n}\binom{q}{l_1,\dots,l_n}}{\binom{p+q}{k_1+l_1,\dots,k_n+l_n}} (\boldsymbol{v}_{i_1} \otimes \dots \otimes \boldsymbol{v}_{i_p}) \cdot (\boldsymbol{v}_{j_1} \otimes \dots \otimes \boldsymbol{v}_{j_q})$$

medzi oboma symetrickými súčinmi bázických vektorov $\boldsymbol{v}_{i_1} \otimes \dots \otimes \boldsymbol{v}_{i_p} \in \mathcal{T}_0^p(V)$, $\boldsymbol{v}_{j_1} \otimes \dots \otimes \boldsymbol{v}_{j_q} \in \mathcal{T}_0^q(V)$, kde $k_i = \#\{r \le p;\ i_r = i\}$, $l_j = \#\{s \le q;\ j_s = j\}$.

33.6. Zdôvodnite platnosť nasledujúceho dodatku k vete 33.3.1 (porovnajte s cvičením 12.9). Nech $M \subseteq \mathbb{R}^n$ je otvorená množina, $\boldsymbol{a} \in M$ a $f\colon M \to \mathbb{R}$ je funkcia, ktorá má v bode $\boldsymbol{a}$ spojité všetky parciálne derivácie až do rádu p vrátane. Predpokladajme, že v bode $\boldsymbol{a}$ sú všetky parciálne derivácie rádov 1 až $p-1$ nulové, t. j. $f^{(r)}(\boldsymbol{a}) = \boldsymbol{0}$ pre $1 \le r < p$, pričom p je párne číslo.

(a) Ak forma $f^{(p)}(\boldsymbol{x})$ je kladne (záporne) definitná pre všetky $\boldsymbol{x} \ne \boldsymbol{a}$ z nejakého okolia $N \subseteq M$ bodu $\boldsymbol{a}$, tak f má v bode $\boldsymbol{a}$ ostré lokálne minimum (maximum);

(b) Ak forma $f^{(p)}(\boldsymbol{x})$ je kladne (záporne) semidefinitná pre všetky $\boldsymbol{x} \ne \boldsymbol{a}$ z nejakého okolia $N \subseteq M$ bodu $\boldsymbol{a}$, tak f má v bode $\boldsymbol{a}$ lokálne minimum (maximum).

(c) Ak forma $f^{(p)}(\boldsymbol{x})$ je indefinitná pre všetky $\boldsymbol{x} \ne \boldsymbol{a}$ z nejakého okolia $N \subseteq M$ bodu $\boldsymbol{a}$, tak f má v bode $\boldsymbol{a}$ sedlo.

33.7. (a) Dokážte vetu 33.4.2.

(b) Sformulujte, čo presne znamená, že priestor $\Lambda^p(V)$ je podmienkami (a), (b) tejto vety určený jednoznačne až na izomorfizmus, a dokážte to.

33.8. Kvôli zdôrazneniu úlohy bázy $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ pri konštrukcii vonkajšej algebry vektorového priestoru V budeme miesto $\Lambda(V)$ na chvíľu písať $\Lambda(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$.

(a) Nech $\boldsymbol{\alpha} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$, $\boldsymbol{\beta} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ sú dve bázy vektorového priestoru V. Dokážte, že priradenie $\boldsymbol{u}_i \mapsto \boldsymbol{v}_i$ možno jednoznačne rozšíriť do izomorfizmu graduovaných algebier $\Lambda(\boldsymbol{u}_1, \dots, \dots \boldsymbol{u}_n) \cong \Lambda(\boldsymbol{v}_1, \dots, \dots \boldsymbol{v}_n)$. Odvoďte z toho, že symetrická algebra $\Lambda(V)$ nezávisí na báze priestoru V.

(b) Označme J_{alt} najmenší ideál graduovanej algebry $\mathcal{T}_0(V)$, ktorý obsahuje všetky prvky tvaru $T-(\operatorname{sgn}\sigma)T^\sigma$, kde $T \in \mathcal{T}_0^p(V)$ a $\sigma \in \mathcal{S}_p$ pre $p \in \mathbb{N}$ (pozri cvičenie 30.6). Dokážte, že graduovaná vonkajšia algebra $\Lambda(V)$ je izomorfná s faktorovou algebrou $\mathcal{T}_0(V)/J_{\mathrm{alt}}$. Odvoďte z toho, že symetrická algebra $\Lambda(V)$ nezávisí na báze priestoru V.

(c) Dokážte, že vektory $\boldsymbol{x}_1, \dots, \boldsymbol{x}_p \in V$ sú lineárne závislé práve vtedy, keď $\boldsymbol{x}_1 \otimes \dots \otimes \boldsymbol{x}_p \in J_{\mathrm{alt}}$, t. j. práve vtedy, keď $\boldsymbol{x}_1 \wedge \dots \wedge \boldsymbol{x}_p = \mathbf{0}$.

(d) Nájdite kanonický (t. j. na báze priestoru V nezávislý) lineárny izomorfizmus $\Lambda^p(V)^* \cong \Lambda^p(V^*)$.

33.9. Odvoďte explicitnú formulu

$$S \wedge T = (\widehat{S} \wedge \widehat{T})^{\pm} = \sum_{\sigma \in \mathcal{S}^{\mathrm{sh}}_{p,q}} (\operatorname{sgn}\sigma)(S \otimes T)^\sigma,$$

pre vonkajší súčin alternujúcich tenzorov $S \in \operatorname{Alt}^p(V)$, $T \in \operatorname{Alt}^q(V)$, kde množina $\mathcal{S}^{\mathrm{sh}}_{p,q} = \{\sigma \in \mathcal{S}_{p+q};\ \sigma(1) < \dots < \sigma(p)\ \&\ \sigma(p+1) < \dots < \sigma(p+q)\}$ je tvorená všetkými tzv. (p,q)-*premiešanými* (t. j. *shuffle*) permutáciami množiny $\{1, \dots, p+q\}$.

33.10. Za predpokladu $\operatorname{char} K > p+q$ dokážte vzťah medzi oboma vonkajšími súčinmi

$$S \mathbin{\bar\wedge} T = \frac{p!\,q!}{(p+q)!}(S \wedge T)$$

alternujúcich tenzorov $S \in \operatorname{Alt}^p(V)$, $T \in \operatorname{Alt}^q(V)$. Rozšírte jeho platnosť (ak treba, po prípadnej úprave) na ľubovoľné kontravariantné tenzory $S \in \mathcal{T}_0^p(V)$, $T \in \mathcal{T}_0^q(V)$.

33.11. Nech V je vektorový priestor nad poľom K, $\dim V = n$. *Anulátorom p-vektora* $T \in \Lambda^p(V)$ nazývame množinu $\operatorname{An} T = \{\boldsymbol{x} \in V;\ T \wedge \boldsymbol{x} = \mathbf{0}\}$. Dokážte postupne nasledujúce tvrdenia:

(a) Anulátor $\operatorname{An} T$ ľubovoľného p-vektora $T \in \Lambda^p(V)$ je lineárny podpriestor vo V.

(b) Ak $\boldsymbol{a}_1, \dots, \boldsymbol{a}_p \in V$ sú lineárne nezávislé 1-vektory, tak $\operatorname{An}(\boldsymbol{a}_1 \wedge \dots \wedge \boldsymbol{a}_p) = [\boldsymbol{a}_1, \dots, \boldsymbol{a}_p]$.

(c) Pre $T \in \Lambda^p(V)$, $U \in \Lambda^q(V)$ platí $\operatorname{An} T \subseteq \operatorname{An} U$ práve vtedy, keď T *delí* U, t. j. $q \geq p$ a $U = T \wedge S$ pre nejaké $S \in \Lambda^{q-p}(V)$.

(d) Pre nenulový p-vektor $T \in \Lambda^p(V)$ platí $\dim \operatorname{An} T = p$ práve vtedy, keď T je *rozložiteľný p-vektor*, t. j. $T = \boldsymbol{x}_1 \wedge \ldots \wedge \boldsymbol{x}_p$ pre nejaké $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p \in V$.

(e) Nenulový 2-vektor $T \in \Lambda^2(V)$ je rozložiteľný práve vtedy, keď $T \wedge T = \mathbf{0}$.

(f) Každý $(n-1)$-vektor $T \in \Lambda^{n-1}(V)$ je rozložiteľný.

33.12. Nech V je euklidovský priestor so skalárnym súčinom $\langle \cdot, \cdot \rangle$.

(a) Dokážte, že formulou $\langle \boldsymbol{x}_1 \wedge \ldots \wedge \boldsymbol{x}_p, \boldsymbol{y}_1 \wedge \ldots \wedge \boldsymbol{y}_p \rangle = \det\big(\langle \boldsymbol{x}_i, \boldsymbol{y}_j \rangle\big)$ je daný skalárny súčin na vonkajšej mocnine $\Lambda^p(V)$.

(b) Pre ľubovoľné 1-vektory $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p \in V$ platí

$$\|\boldsymbol{x}_1 \wedge \ldots \wedge \boldsymbol{x}_p\| = \big(\det \boldsymbol{G}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p)\big)^{1/2} = \mathrm{vol}_p(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p),$$

kde $\boldsymbol{G}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p)$ je Gramova matica a $\mathrm{vol}_p(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p)$ je p-rozmerný objem rovnobežnostena vytvoreného vektormi $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p \in V$. Dokážte a porovnajte s paragrafom 15.3.

33.13. Pre konečnorozmerné priestory a $p \le \dim V + \dim W$ popíšte kanonické lineárne izomorfizmy

$$\Sigma^p(V \oplus W) \cong \bigoplus_{k=0}^{p} \Sigma^k(V) \otimes \Sigma^{p-k}(W),$$

$$\Lambda^p(V \oplus W) \cong \bigoplus_{k=0}^{p} \Lambda^k(V) \otimes \Lambda^{p-k}(W)$$

pre p-tu symetrickú resp. vonkajšiu mocninu priameho súčtu $V \oplus W$.

33.14. (a) Dokážte, že vnútorný súčin $V \times \mathrm{Alt}(V^*) \to \mathrm{Alt}(V^*)$ sa voči vonkajšiemu súčinu správa ako *graduovaná derivácia*, t. j. pre $\boldsymbol{x} \in V$, $F \in \mathrm{Alt}^p(V^*) = \mathrm{Alt}_p(V)$, $G \in \mathrm{Alt}^q(V^*) = \mathrm{Alt}_q(V)$ platí

$$\boldsymbol{i}_{\boldsymbol{x}}(F \wedge G) = \big(\boldsymbol{i}_{\boldsymbol{x}}(F) \wedge G\big) + (-1)^p \big(F \wedge \boldsymbol{i}_{\boldsymbol{x}}(G)\big).$$

(b) Dokážte, že pre ľubovoľné $\boldsymbol{x}, \boldsymbol{y} \in V$ platí $\boldsymbol{i}_{\boldsymbol{x}} \circ \boldsymbol{i}_{\boldsymbol{x}} = \mathbf{0}$, $\boldsymbol{i}_{\boldsymbol{x}} \circ \boldsymbol{i}_{\boldsymbol{y}} = -(\boldsymbol{i}_{\boldsymbol{y}} \circ \boldsymbol{i}_{\boldsymbol{x}})$.

(c) Vyjadrite obe rovnosti v (b) jednotným spôsobom ako *antikomutačný vzťah* pre operátory $\boldsymbol{i}_{\boldsymbol{x}}$, $\boldsymbol{i}_{\boldsymbol{y}}$.

33.15. Zaveďte vnútorný súčin aj pre symetrické tenzory; sformulujte a overte jeho základné vlastnosti.

33.16. Predpokladajme, že V je euklidovský priestor a každá z vonkajších mocnín $\Lambda^p(V)$ je vybavená skalárnym súčinom tak, ako v cvičení 33.12.

(a) Dokážte, že Hodgeov operátor $*\colon \Lambda^p(V) \to \Lambda^{n-p}(V)$ je izometria.

(b) Overte, že kompozícia $* \circ *$ Hodgeových operátorov $*\colon \Lambda^p(V) \to \Lambda^{n-p}(V)$ a $*\colon \Lambda^{n-p}(V) \to \Lambda^p(V)$ je $(-1)^{p(n-p)}$ násobok identického operátora.

33.17. (a) Overte komutačné resp. antikomutačné vzťahy pre anihilačné a kreačné operátory A_i, A_j^* bozónov resp. fermiónov.

(b) Nájdite vlastné hodnoty a k nim príslušné vlastné podpriestory hermitovských operátorov počtu častíc $A_j^* A_j$ a $\sum_{j=1}^n A_j^* A_j$ pre bozóny, t. j. ako operátorov $\Sigma(V) \to \Sigma(V)$, aj pre fermióny, t. j. ako operátorov $\Lambda(V) \to \Lambda(V)$.

Literatúra

[1] A. V. Archangeľskij, *Konečnomernyje linejnyje prostranstva*, Izdateľstvo Moskovskogo universiteta, Moskva, 1982.

[2] L. Bican, *Lineární algebra*, Matematický seminář, SPN, Praha, 1979.

[3] G. Birkhoff, S. Mac Lane, *Prehľad modernej algebry*, Alfa, Bratislava, 1979.

[4] L. Boček, *Tensorový počet*, Matematický seminář, SPN, Praha, 1978.

[5] M. L. Curtis, *Matrix Groups*, Springer Verlag, New York, 1984.

[6] B. A. Dubrovin, A. T. Fomenko, S. P. Novikov, *Sovremennaja geometrija, Metody i priloženija, časť I*, Nauka, Moskva, 1979, (anglický preklad *Modern Geometry, Methods and Applicatons*, Springer Verlag, New York, 1984).

[7] A. K. Faddejev, J. S. Sominskij, *Zbierka úloh z vyššej algebry* Alfa, Bratislava, 1968.

[8] L. D. Faddejev, O. A. Jakubovskij, *Lekcii po kvantovoj mechanike dľa studentov--matematikov*, Izdateľstvo Leningradskogo universiteta, Leningrad, 1980.

[9] M. Fecko, *Diferenciálna geometria a Lieove grupy pre fyzikov* Iris, Bratislava, 2004.

[10] F. R. Gantmacher, *Teorija matric*, Nauka, Moskva, 1966.

[11] I. M. Geľfand, *Lekciji po linejnoj algebre*, Nauka, Moskva, 1966.

[12] L. I. Golovina, *Linejnaja algebra i nekotoryje jejo priloženija* Nauka, Moskva, 1985.

[13] W. H. Greub, *Multilinear Algebra*, Springer-Verlag, Berlin, 1967.

[14] P. R. Halmos, *Finite-Dimensional Vector Spaces*, Springer Verlag, New York, 1974.

[15] L. Hogben (ed.), *Handbook of Linear Algebra*, Chapman & Hall/CRC, Boca Raton, FL 2007.

[16] V. Chvál, M. Mikola, *Lineárna algebra*, Katolícka univerzita, Ružomberok, 2000.

[17] D. Ilkovič, *Fyzika I, II*, (4. vyd.), Alfa, Bratislava, SNTL, Praha, 1968.

[18] V. A. Iľjin, E. G. Pozňak, *Linejnaja algebra* Nauka, Moskva, 1978.

[19] P. Kaprálik, J. Tvarožek, *Zbierka riešených príkladov z lineárnej algebry a analytickej geometrie*, Alfa, Bratislava, 1987.

[20] T. Katriňák, M. Gavalec, E. Gedeonová, J. Smítal, *Algebra a teoretická aritmetika I*, (3. vydanie), Univerzita Komenského, Bratislava, 1999.

[21] J. Korbaš, *Lineárna algebra a geometria I*, Univerzita Komenského, Bratislava, 2003.

[22] A. I. Kostrikin, Ju. I. Manin, *Linejnaja algebra i geometrija*, Nauka, Moskva, 1986, (anglický preklad *Linear algebra and geometry*, Gordon & Breach, New York, 1989).

[23] J. Kvasnica, *Matematický aparát fyziky*, Academia, Praha, 1989.

[24] D. C. Lay, *Linear Algebra and Its Applications*, Addison-Wesley, Boston, 2003.

[25] O. Litzman, M. Sekanina, *Užití grup ve fyzice*, Academia, Praha, 1982.

[26] S. Mac Lane, G. Birkhoff, *Algebra*, (2. vydanie), Alfa, Bratislava, 1974.

[27] A. I. Maľcev, *Osnovy linejnoj algebry*, Nauka, Moskva, 1975.

[28] R. Merris, *Multilinear Algebra*, Gordon Breach, Amsterdam, 1997.

[29] L. Motl, M. Zahradník, *Pěstujeme lineární algebru*, Univerzita Karlova, vydavatelství Karolinum, Praha, 1995.

[30] M. Neusel, *Invariant Theory*, Student Mathematical Library, American Mathematical Society, Providence, Rhode Island, 2006.

[31] J. Pišút, L. Gomolčák, V. Černý, *Úvod do kvantovej mechaniky*, Alfa, Bratislava, 1983.

[32] J. Slovák, *Lineární algebra*, skriptá PřF MU, Brno, 1995, elektronický učebný text dostupný na sieti, `http://www.math.muni.cz/~slovak`

[33] J. Smítal, E. Gedeonová, *Lineárna algebra*, skriptá MFF UK, Bratislava, 1981.

[34] G. Strang, *Linear Algebra and its Applications*, Thomson Brooks/Cole, 2006.

[35] F. Šik, *Lineární algebra zaměřená na numerickou analýzu*, Masarykova univerzita, Brno, 1998.

[36] K. Tapp, *Matrix Groups for Undergraduates*, Student Mathematical Library, American Mathematical Society, 2005.

[37] V. Votruba, *Základy speciální teorie relativity*, Academia, Praha, 1969.

Register