Diskrétní matematika

doc. Lukáš Vokřínek, PhD.

22. září 2024

Obsah

Úvod

Tady bude úvod.

Lukáš Vokřínek

Sylabus přednášky

Tady bude sylabus.

1 Problémy teorie čísel

Přirozená a celá čísla jsou nejjednodušší matematickou strukturou, zkoumání jejich vlastností však postavilo před generace matematiků celou řadu velice obtížných problémů.

Často jsou to problémy, které je možno snadno formulovat, přesto však dodnes neznáme jejich řešení.

V několika přednáškách se teď budeme zabývat úlohami o celých číslech. Převážně v nich půjde o dělitelnost celých čísel, popřípadě o řešení rovnic v oboru celých nebo přirozených čísel.

God made integers, all else is the work of man. (L. Kronecker)

Příklady problémů teorie čísel

- *problém prvočíselných dvojčat* – rozhodnout, zda existuje nekonečně mnoho prvočísel p takových, že i $p+2$ je prvočíslo;
- *problém existence lichého dokonalého čísla* – tj. čísla jehož součet dělitelů je roven dvojnásobku tohoto čísla;
- *Goldbachova hypotéza* (rozhodnout, zda každé sudé číslo větší než 2 je možno psát jako součet dvou prvočísel);
- *velká Fermatova věta* (Fermat's Last Theorem) – rozhodnout, zda existují kladná celá čísla n, x, y, z tak, že $n > 2$ a platí $x^n + y^n = z^n$; Pierre de Fermat jej formuloval cca 1637, vyřešil Andrew Wiles v roce 1995.

Diofantické rovnice

V kouzelném měšci máme neomezené množství dvoukorun a pětikorun. Jaké částky můžeme zaplatit tak, aby nebylo potřeba vracet?

Ptáme se tedy, pro která přirozená čísla n existují přirozená k, ℓ tak, aby

$$2k + 5\ell = n.$$

Asi se dá vcelku snadno uvěřit, že libovolnou vyšší částku takto zaplatíme, po pravdě jakoukoliv častku s výjimkou 1 Kč a 3 Kč.

S vracením pak zvládneme zaplatit libovolnou částku, tj. každé n lze vyjádřit jako

$$2k + 5\ell = n$$

pro nějaká celá k, ℓ.

Umíme to pro jakékoliv hodnoty mincí? Jak by to dopadlo třeba pro $7k + 11\ell = n$? A jak pro $4k + 6\ell = n$?

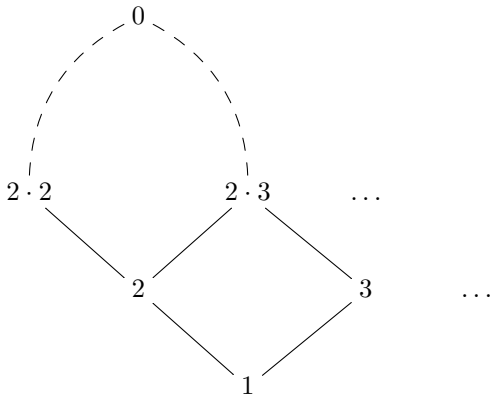
2 Dělitelnost

Základní motivací pro nás bude důkaz věty o rozkladu čísla na součin prvočísel a zejména *jednoznačnost* tohoto rozkladu, která nám například umožní snadno vypsat všechny dělitele daného čísla (vybereme z rozkladu libovolnou kolekci činitelů). Vysvětlíme nyní krátce, jak jednoznačnost rozkladu souvisí s dělitelností: Pro konkrétnost ukážeme, proč nemůže nastat rovnost $2 \cdot 7 = 3 \cdot 5$; pravá strana je totiž součinem lichých čísel a je tedy lichá, zatímco levá strana obsahuje činitel 2 a je tedy sudá. Podobný argument bude fungovat i pro prvočísla různá od 2, pokud dokážeme následující: Součin čísel, která nejsou dělitelná prvočíslem p nemůže být dělitelný p (případně obměnou, pokud je součin čísel dělitelný prvočíslem p, pak musí být dělitelný p některý z činitelů).

Poznamenejme, že v některých číselných oborech jednoznačnost rozkladu neplatí, např. $2 \cdot 3 = (1-\sqrt{-5}) \cdot (1+\sqrt{-5})$ jsou dva různé rozklady čísla 6 v číselném oboru všech těch komplexních čísel, která lze vyjádřit ve tvaru $a + b\sqrt{-5}$ (a opravdu se tam jedná o „prvočísla“). Budeme tedy muset využít nějaké netriviální vlastnosti celých čísel.

Definice. Řekneme, že celé číslo a *dělí* celé číslo b (neboli číslo b je *dělitelné* číslem a, též a je *dělitel* a neboli b je *násobek* a), právě když existuje celé číslo c tak, že platí $a \cdot c = b$. Píšeme pak $a \mid b$.

Snadno se vidí, že dělitelnost je uspořádání na přirozených (tedy nezáporných celých) číslech, tj. splňuje reflexivitu $a \mid a$, tranzitivitu $a \mid b \mid c \implies a \mid c$ a antisymetrii $a \mid b \mid a \implies a = b$. Je vcelku užitečné mít představu o Hasseho diagramu tohoto uspořádání:



V oboru všech celých čísel už dělitelnost antisymetrická není, neboť $-a \mid a$. Pokud v Hasseho diagramu nahradíme každé nenulové číslo a dvojicí $\pm a$, dostaneme velice dobrou představu o dělitelnosti na všech celých číslech. (Obecně relace $\leq$ splňující pouze reflexivitu a tranzitivitu se nazývá předuspořádání a $a \sim b \Leftrightarrow a \leq b \leq a$ zadává relaci ekvivalence, na jejímž rozkladu už je $\leq$ uspořádáním.)

Dále nás budou zajímat algebraické vlastnosti dělitelnosti, které za chvíli o něco přehledněji přepíšeme do vlastností kongruencí:

$$a \mid b \wedge a \mid c \implies a \mid b + c \wedge a \mid b - c$$
$$a \mid b \Longleftrightarrow ac \mid bc \quad (c \neq 0)$$

Příklad. Zjistěte, pro která přirozená čísla n je číslo $n^2 + 1$ dělitelné číslem 3.

Řešení. Uvidí se, že záleží pouze na zbytku n po dělení třemi.

Příklad. Zjistěte, pro která přirozená čísla n je číslo $n^2 + 1$ dělitelné číslem $n + 1$.

Dělení se zbytkem

Věta 1 (o dělení celých čísel se zbytkem)**.** *Pro libovolně zvolená celá čísla a, $m > 0$ existují jednoznačně určená čísla $q \in \mathbb{Z}$, $r \in \{0, 1, \ldots, m-1\}$ tak, že $a = qm + r$.*

Důkaz. Prvně dokážeme pro $a \geq 0$ indukcí: pro $a < m$ zřejmé, pro $a \geq m$ pak rekurzivně s využitím výsledku pro $a - m$ (podíl je potřeba zvětšit o 1, zbytek zůstane stejný). Případ $a < 0$ lze snadno převést na $(-a - 1) \geq 0$. □

Číslo q, resp. r z věty se nazývá *(neúplný) podíl*, resp. *zbytek* při dělení čísla a číslem m se zbytkem. Vhodnost obou názvů je zřejmá, přepíšeme-li rovnost $a = mq + r$ do tvaru

$$\frac{a}{m} = q + \frac{r}{m}, \quad \text{přitom} \quad 0 \leq \frac{r}{m} < 1.$$

Příklad. Dokažte, že jsou-li zbytky po dělení čísel $a, b \in \mathbb{Z}$ číslem $m \in \mathbb{N}$ jedna, je jedna i zbytek po dělení čísla ab číslem m.

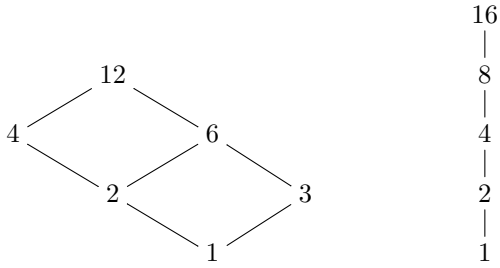
3 Společní dělitelé a společné násobky

Největší společný dělitel (gcd)

Jedním z nejdůležitějších nástrojů výpočetní teorie čísel je výpočet největšího společného dělitele. Protože jde, jak si ukážeme, o relativně rychlou proceduru, je i v moderních algoritmech velmi často využívána.

Definice. Mějme přirozená čísla a, b. Libovolné přirozené číslo m takové, že $m \mid a$, $m \mid b$ se nazývá *společný dělitel* čísel a, b. Společný dělitel čísel a, b, který je dělitelný libovolným společným dělitelem těchto čísel, se nazývá *největší společný dělitel* čísel a, b a značí se (a, b). (Jedná se o infimum vzhledem k dělitelnosti.)

Například $(12, 16) = 4$: společní dělitelé jsou 4, 2, 1, největší z nich (vzhledem k dělitelnosti) je 4.



Poznámka. Přímo z definice plyne, že pro libovolné $a, b \in \mathbb{N}$ platí $(a, b) = (b, a)$, $(a, 1) = 1$, $(a, 0) = a$.

Definici lze snadno rozšířit i na libovolná celá čísla a, b. Budeme požadovat, aby nevjětší společný dělitel byl největší mezi nezápornými společnými děliteli, tím bude jednoznačně určený.

Definice. Mějme přirozená čísla a, b. Libovolné přirozené číslo m takové, že $a \mid m$, $b \mid m$ se nazývá *společný násobek* čísel a, b. Společný násobek čísel a, b, který dělí libovolný společný násobek těchto čísel, se nazývá *nejmenší společný násobek* čísel a, b a značí se $[a, b]$.

Poznámka. Analogicky se definuje i největší společný dělitel a nejmenší společný násobek více než dvou celých čísel a snadno se následně dokáže, že platí

$$(a_1, \ldots, a_n) = ((a_1, \ldots, a_{n-1}), a_n),$$
$$[a_1, \ldots, a_n] = [[a_1, \ldots, a_{n-1}], a_n].$$

Eukleidův algoritmus

Dosud jsme nijak nezdůvodnili, zda pro každou dvojici $a, b \in \mathbb{Z}$ čísla (a, b) a $[a, b]$ vůbec existují. To si lze hezky představit přes rozklad na prvočinitele, ale ten je výpočetně velmi náročný (RSA) a navíc k jeho odvození budeme existenci největšího společného dělitele využívat.

Pokud však existují, jsou určena jednoznačně: Pro každá dvě čísla $m_1, m_2 \in \mathbb{N}_0$ totiž podle definice platí, že pokud $m_1 \mid m_2$ a zároveň $m_2 \mid m_1$, je nutně $m_1 = m_2$. Důkaz existence čísla (a, b) podáme (spolu s algoritmem jeho nalezení) v následující větě.

Věta 2 (Euklidův algoritmus)**.** *Nechť $a_1 \geq a_2$ jsou přirozená čísla. Pro každé $n \geq 3$, pro které $a_{n-1} \neq 0$, označme a_n zbytek po dělení čísla a_{n-2} číslem a_{n-1}. Pak po konečném počtu kroků dostaneme $a_r = 0$ a platí $a_{r-1} = (a_1, a_2)$.*

Důkaz. Jelikož platí $a_3 = a_1 - qa_2$, dostáváme $(a_1, a_2) = (a_3, a_2) = (a_2, a_3)$ a největší společný dělitel tedy zůstává stále stejný, přičemž $(a_{r-1}, a_r) = (a_{r-1}, 0) = a_{r-1}$. □

Algoritmus a důkaz jeho korektnosti demonstrujeme na příkladu:

Příklad. Určete největšího společného dělitele čísel 10175 a 2277.

Vlastnosti gcd

Poznámka. Z definice, z předchozího tvrzení a z toho, že pro libovolná $a, b \in \mathbb{Z}$ platí $(a, b) = (a, -b) = (-a, b) = (-a, -b)$, plyne, že existuje největší společný dělitel libovolných dvou celých čísel.

Věta 3 (Bezoutova)**.** *Pro libovolná celá čísla a_1, a_2 existuje jejich největší společný dělitel (a_1, a_2), přitom existují celá čísla k_1, k_2 tak, že $(a_1, a_2) = k_1 a_1 + k_2 a_2$.*

Důkaz. Při první z metod se největší společný dělitel (a_1, a_2) rekurzivně počítá jako (a_2, a_3) a je vyjadřen jako celočíselná kombinace a_2, a_3 přičemž se následně nahradí $a_3 = a_1 - q a_2$ celočíselnou kombinací a_1, a_2.

Druhá metoda funguje z druhé strany a při počítání $a_1, a_2, a_3, \ldots, a_{r-1}, a_r = 0$ zároveň počítá i vyjádření každého a_n jako celočíselné kombinace a_1, a_2, viz příklad. □

Důsledek. *Pro libovolná celá čísla a_1, a_2 lze jako celočíselné kombinace $n = k_1 a_1 + k_2 a_2$ vyjádřit právě násobky největšího společného dělitele (a_1, a_2).*

Příklad. Výpočet největšího společného dělitele pomocí Euklidova algoritmu je s využitím výpočetní techniky i pro relativně velká čísla poměrně rychlý. V našem příkladu to vyzkoušíme na 2 číslech A, B, z nichž každé je součinem dvou 101-ciferných prvočísel. Všimněme si, že výpočet největšího společného dělitele i takto velkých čísel trval zanedbatelný čas.

```
p = next_prime(5*10^300)
q = next_prime(3*10^300)
r = next_prime(2*10^300)
A = p * q; B = p * r
gcd(A, B)
```

Příklad v systému `SAGE` lze vyzkoušet na `https://cocalc.com/`.

Poznámka. Euklidův algoritmus a Bezoutova věta jsou základními výsledky elementární teorie čísel a tvoří jeden z pilířů algoritmů algebry a teorie čísel.

Nesoudělnost

Definice. Čísla $a, b \in \mathbb{Z}$ se nazývají *nesoudělná*, jestliže platí $(a, b) = 1$. Čísla $a_1, a_2, \ldots, a_n \in \mathbb{Z}$ se nazývají *po dvou nesoudělná*, jestliže pro každé $i \neq j$ platí $(a_i, a_j) = 1$.

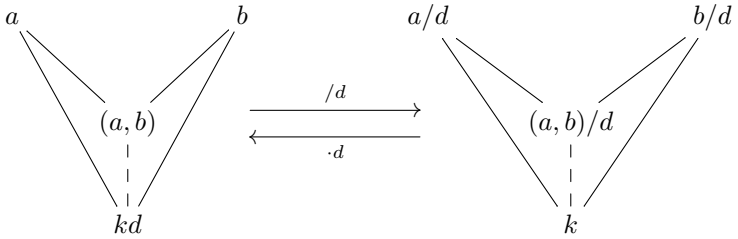
Věta 4. *Pro libovolná přirozená čísla a, b a jejich největšího společného dělitele $(a, b) = d$ jsou čísla a/d, b/d nesoudělná.*

Důkaz. Dokážeme obecněji pro libovolné $d \mid (a, b)$ vztah $(a/d, b/d) = (a, b)/d$, tvrzení je pak speciálním případem. Číslo k je společný dělitel čísel a/d, b/d, právě když

$$k \mid a/d,\ k \mid b/d \Leftrightarrow kd \mid a,\ kd \mid b \Leftrightarrow kd \mid (a, b) \Leftrightarrow k \mid (a, b)/d$$

a největší takové k je tedy $(a, b)/d$. □

Poznámka. Diagramaticky lze předchozí důkaz znázornit takto:



Věta 5. *Pro libovolná přirozená čísla a, b, c platí: jestliže $c \mid ab$, $(c, a) = 1$, pak $c \mid b$.*

Důkaz. Zjevně $c \mid cb$ a podle předpokladu také $c \mid ab$, musí tedy c dělit i jakoukoliv jejich celočíselnou kombinaci; přitom podle Bezoutova lemmatu $kc + la = 1$, takže

$$c \mid (k \cdot cb + l \cdot ab) = (kc + la)b = b. \qquad \square$$

Nejmenší společný násobek

Věta 6. *Pro libovolná přirozená čísla a_1, a_2 existuje jejich nejmenší společný násobek $[a_1, a_2]$ a platí $(a_1, a_2) \cdot [a_1, a_2] = a_1 a_2$.*

Důkaz. Nejlépe se vidí přes rozklad na součin prvočísel. Jinak se vše prvně převede na případ $(a_1, a_2) = 1$ a pak $a_2 \mid x$ implikuje $a_1 a_2 \mid a_1 x$, symetricky $a_1 \mid x$ implikuje $a_1 a_2 \mid a_2 x$ a jejich celočíselnou kombinací pak dostaneme $a_1 a_2 \mid (ka_1 + la_2)x = x$. Jinými slovy každý společný násobek a_1, a_2 je násobkem $a_1 a_2$ a tento je tedy nejmenší společný násobek. $\square$

4 Prvočísla

Prvočíslo je jeden z nejdůležitějších pojmů elementární teorie čísel. Jeho důležitost je dána především větou o jednoznačném rozkladu libovolného přirozeného čísla na součin prvočísel, která je silným a účinným nástrojem při řešení celé řady úloh z teorie čísel.

Definice. Každé přirozené číslo $n \geq 2$ má aspoň dva kladné dělitele: 1 a n. Pokud kromě těchto dvou jiné kladné dělitele nemá, nazývá se *prvočíslo*. V opačném případě hovoříme o *složeném čísle*.

V dalším textu budeme zpravidla prvočíslo značit písmenem p. Nejmenší prvočísla jsou 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, ... (zejména číslo 1 za prvočíslo ani za číslo složené nepovažujeme, je totiž invertibilní, neboli tzv. jednotkou okruhu celých čísel). Prvočísel je, jak brzy dokážeme, nekonečně mnoho, máme ovšem poměrně limitované výpočetní prostředky na zjištění, zda je dané číslo prvočíslem (největší známé prvočíslo $2^{82\,589\,933} - 1$ má pouze 24 862 048 cifer).

Základní věta aritmetiky

Uveďme nyní větu, která udává ekvivalentní podmínku prvočíselnosti a je základní ingrediencí při důkazu jednoznačnosti rozkladu na prvočísla.

Věta 7 (Euklidova o prvočíslech)**.** *Přirozené číslo $p \geq 2$ je prvočíslo, právě když platí: pro každá celá čísla a, b z $p \mid ab$ plyne $p \mid a$ nebo $p \mid b$.*

Důkaz. Jedná se o reformulaci Věty 5 pro $c = p$ prvočíslo s tím, že podmínka $(p, a) = 1$ je ekvivalentní tomu, že $p \nmid a$ (to lze vidět naopak tak, že $(p, a) = p$ je ekvivalentní $p \mid a$). $\square$

Věta 8. *Libovolné přirozené číslo $n \geq 2$ je možné vyjádřit jako součin prvočísel, přičemž je toto vyjádření jediné, nebereme-li v úvahu pořadí činitelů. (Je-li n prvočíslo, pak jde o „součin“ jednoho prvočísla.)*

Důkaz. Existence se dokáže jednoduše metodou rozděl a panuj: Pokud n je prvočíslo, máme hotovo, jinak jej rozdělíme na součin dvou menších čísel, na obě aplikujeme indukci a součiny dáme dohromady.

Jednoznačnost je složitější v tom, že využívá Eukleidovu větu. Předpokládejme $p_1 \cdots p_k = n = q_1 \cdots q_l$. Protože p_k dělí součin vlevo, musí dělit i součin vpravo a podle Eukleidovy věty některý z činitelů q_j; protože se jedná o prvočísla, $p_k = q_j$. Vydělením tímto prvočíslem dostaneme číslo menší než n a jeho dva rozklady, na něž můžeme použít indukci. $\square$

5 Kongruence

Pojem kongruence byl zaveden Gaussem. Ačkoliv je to pojem velice jednoduchý, jeho důležitost a užitečnost v teorii čísel je nedocenitelná; projevuje se zejména ve stručných a přehledných zápisech některých i velmi komplikovaných úvah.

Definice. Jestliže dvě celá čísla a, b mají při dělení přirozeným číslem m týž zbytek, nazývají se a, b *kongruentní modulo* m (též *kongruentní podle modulu* m), což zapisujeme takto:

$$a \equiv b \pmod{m}.$$

V opačném případě řekneme, že a, b nejsou kongruentní modulo m, a píšeme

$$a \not\equiv b \pmod{m}.$$

Lemma. *Pro libovolná $a, b \in \mathbb{Z}$, $m \in \mathbb{N}$ jsou následující podmínky ekvivalentní:*

1. $a \equiv b \pmod{m}$,
2. $a = b + mt$ *pro vhodné* $t \in \mathbb{Z}$,
3. $m \mid a - b$.

Důkaz. Protože se zbytky po dělení m periodicky opakují s periodou m, dvě čísla dávají stejný zbytek po dělení m, právě když se liší o násobek m. □

Základní vlastnosti kongruencí

Kongruence modulo m je *relace ekvivalence*, jejíž třídy budeme nazývat *zbytkové třídy* modulo m; jedná se o řádky v následující tabulce:

$\cdots$	$-m$	0	m	$\cdots$
$\cdots$	$-m+1$	1	$m+1$	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\cdots$	-1	$m-1$	$2m-1$	$\cdots$

Konkrétně tedy kongruence modulo m splňuje následující vlastnosti:

- $a \equiv a \pmod{m}$, tj. kongruence podle modulu m je *reflexivní*,
- $a \equiv b \pmod{m} \Rightarrow b \equiv a \pmod{m}$, tj. kongruence podle modulu m je *symetrická*,
- $a \equiv b \pmod{m}$, $b \equiv c \pmod{m} \Rightarrow a \equiv c \pmod{m}$, tj. kongruence podle modulu m je *tranzitivní*.

Následující triviální vlastnost se hodí při počítání s konkrétními čísly (zejména zápornými, kde zbytek po dělení m není tolik používaný a proto svádí k numerickým chybám, např. $120 \equiv -10 \equiv 3 \pmod{13}$ prvně odečtením 130 a poté přičtením 13):

- K libovolné straně můžeme přičíst libovolný násobek modulu:

$$a \equiv b \pmod{m} \quad \Rightarrow \quad a \equiv b + k \cdot m \pmod{m}.$$

Nyní následují algebraické vlastnosti, které dohromady říkají, že v kongruenci můžeme v libovolném polynomiálním výrazu všechna vystupující čísla nahradit čísly s nimi kongruentními (zejména menšími, aby se výraz zjednodušil); to se ovšem *netýká* exponentů, kterými se budeme zabývat později.

- Kongruence podle téhož modulu můžeme *sčítat*, tedy i *vynásobit týmž číslem*:

$$\begin{array}{l} a_1 \equiv b_1 \pmod m, \\ a_2 \equiv b_2 \pmod m \end{array} \quad \Rightarrow \quad a_1 + a_2 \equiv b_1 + b_2 \pmod m.$$
$$a \equiv b \pmod m \quad \Rightarrow \quad k \cdot a \equiv k \cdot b \pmod m.$$

- Kongruence podle téhož modulu můžeme *násobit*, tedy i *umocnit na totéž číslo.*

$$\begin{array}{l} a_1 \equiv b_1 \pmod m, \\ a_2 \equiv b_2 \pmod m \end{array} \quad \Rightarrow \quad a_1 \cdot a_2 \equiv b_1 \cdot b_2 \pmod m.$$
$$a \equiv b \pmod m \quad \Rightarrow \quad a^k \equiv b^k \pmod m.$$

- Obě strany kongruence můžeme vydělit číslem k, jestliže je *nesoudělné s modulem.*

$$k \cdot a \equiv k \cdot b \pmod m, \quad (k, m) = 1 \quad \Rightarrow \quad a \equiv b \pmod m.$$

Poslední sada vlastností se týká vztahů mezi kongruencemi vzhledem k různým modulům, využijeme je až později, takže se na první čtení dají přeskočit.

- Jestliže $n \mid m$, pak
$$a \equiv b \pmod m \quad \Rightarrow \quad a \equiv b \pmod n.$$
Naopak pokud $a \equiv b \pmod n$, dostáváme $m/n = k$ možných řešení
$$a \equiv b,\ a \equiv b + n,\ \ldots,\ \text{nebo } a \equiv b + (k-1)n \pmod m.$$

- Jestliže $m = [m_1, m_2]$ je nejmenší společný násobek, pak
$$a \equiv b \pmod{m_1},\ a \equiv b \pmod{m_2} \ \Leftrightarrow\ a \equiv b \pmod m.$$

- *Obě strany* kongruence *a modul* lze vynásobit nebo vydělit libovolným číslem
$$a \equiv b \pmod m \quad \Leftrightarrow \quad k \cdot a \equiv k \cdot b \pmod{k \cdot m}.$$

Poznámka. Některé vlastnosti kongruencí jsme již používali, aniž bychom si toho povšimli – například příklad z minulého týdne lze přeformulovat do tvaru

$$a \equiv 1 \pmod m,\ b \equiv 1 \pmod m \quad \Rightarrow \quad ab \equiv 1 \pmod m,$$

což je speciální případ z předhozího tvrzení.

Nejde o náhodu. Libovolné tvrzení používající kongruence můžeme snadno přepsat pomocí dělitelnosti. Užitečnost kongruencí tedy netkví v tom, že bychom pomocí nich mohli řešit úlohy, které bez nich řešit nejsme schopni, ale v tom, že jde o velmi vhodný způsob zápisu. Osvojíme-li si ho, výrazně tím zjednodušíme jak vyjadřování, tak i některé úvahy. Je to typický jev: v matematice hraje vhodná symbolika velmi závažnou úlohu.

Příklad. Nalezněte zbytek po dělení čísla 5^{20} číslem 26.

Příklad. Dokažte, že pro libovolné prvočíslo p a libovolná $a, b \in \mathbb{Z}$ platí

$$a^p + b^p \equiv (a + b)^p \pmod p.$$

Příklad. Najděte "inverzi" k číslu 39 modulo 47, tj. najděte x takové, že $39 \cdot x \equiv 1 \pmod{47}$.

Inverze modulo m

Věta 9. *Je-li a nesoudělné s modulem m, tj. $(a, m) = 1$, pak existuje řešení*

$$a \cdot x \equiv 1 \pmod{m}.$$

Toto řešení značíme $x \equiv a^{-1}$ a nazýváme inverzí *k a modulo m. Jakožto zbytková třída je toto řešení jediné.*

Důkaz. Zobrazení $x \pmod{m} \mapsto a \cdot x \pmod{m}$ na zbytkových třídách je injektivní (to přesně říká vlastnost dělení: $ax \equiv ay \Rightarrow x \equiv y$); protože je zbytkových tříd na obou stranách stejně, totiž m, jedná se o bijekci a jednička $1 \pmod{m}$ má jediný vzor. □

Věta 10. *Nechť $m \in \mathbb{N}$, $a, b \in \mathbb{Z}$. Pokud $(a, m) = 1$, má kongruence*

$$a \cdot x \equiv b \pmod{m}$$

právě jedno řešení $x \pmod{m}$.

V obecném případě označme $d = (a, m)$. Pak má tato kongruence řešení právě tehdy, když $d \mid b$; řešením je pak právě d zbytkových tříd $x \pmod{m}$.

Důkaz. Podle předchozí věty inverze $a^{-1} \pmod{m}$ existuje a vynásobením rovnice

$$a \cdot x \equiv b \pmod{m}$$

touto inverzí dostaneme (jediné) řešení

$$x \equiv a^{-1} \cdot b \pmod{m}.$$

V obecném případě $d \mid m \mid (b - ax)$; protože také $d \mid a$, dostáváme opravdu $d \mid b$. Naopak, pokud $d \mid b$, vydělíme obě strany i modul největším společným dělitelem d a dostaneme při označení $a' = a/d$, $b' = b/d$, $m' = m/d$ ekvivalentní rovnici

$$a' \cdot x \equiv b' \pmod{m'}$$

kde již $(a', m') = 1$ podle Věty 4. Podle první části dostaneme jediné řešení $x \pmod{m'}$, kterému odpovídá právě d řešení $x \pmod{m}$. □

Algoritmus

Začneme s ekvivalentní soustavou dvou kongruencí

$$\begin{aligned} m \cdot x &\equiv 0 \pmod{m} \\ a \cdot x &\equiv b \pmod{m} \end{aligned}$$

a vždy první rovnici systému nahradíme rovnicí vzniklou odečtením vhodného násobku druhé rovnice (tak abychom koeficient m nahradili jeho zbytkem po dělení číslem a), dokud nedostaneme koeficienty d a 0:

$$\begin{aligned} d \cdot x &\equiv c \pmod{m} \\ 0 \cdot x &\equiv k \pmod{m} \end{aligned}$$

Máme dvě možnosti:

- $k \equiv 0$ a soustava, a tedy i původní rovnice, má řešení vzniklé z první rovnice vydělením d, totiž: $x \equiv c/d \pmod{m/d}$;[1]

[1] Zde tvrdíme, že v případě $k \equiv 0$ už bude automaticky $d \mid c$, což je sice pravda, ale dokazovat to nebudeme; v každém příkladu to tak vyjde a není potřeba to tedy v rámci výpočtu používat.

- $k \not\equiv 0$ a soustava, a tedy i původní rovnice, nemá řešení.

Příklad. Řešte $39x \equiv 41 \pmod{47}$

Poznámka. Teoretický, i když ne příliš praktický postup, pro jednoduchost v případě $(a, m) = 1$: z Bezoutovy věty dostaneme $ka + lm = 1$, použijeme

$$a \cdot x \equiv b = (ka + lm)b \equiv kab \pmod{m}$$

a vydělíme a, takže $x \equiv kb \pmod{m}$. (Zbytečně počítáme koeficient l.)

Wilsonova věta

Pomocí věty o řešitelnosti lineárních kongruencí lze dokázat mj. významnou Wilsonovu větu udávající nutnou (i postačující) podmínku prvočíselnosti. Takové podmínky jsou velmi významné ve výpočetní teorii čísel, kdy je třeba efektivně poznat, je-li dané velké číslo prvočíslem. Bohužel dosud není známo, jak rychle vypočítat modulární faktoriál velkého čísla, proto není v praxi Wilsonova věta k tomuto účelu používána.

Věta 11 (Wilsonova)**.** *Přirozené číslo $n > 1$ je prvočíslo, právě když*

$$(n-1)! \equiv -1 \pmod{n}$$

Důkaz. Pokud n není prvočíslo, je některé z čísel $1, \ldots, n-1$ v součinu $(n-1)!$ soudělné s n, proto také celý součin, a nemůže tedy být kongruentní s -1.

Naopak, pokud např. $n = 7$, pak

$$6! = 1 \cdot (2 \cdot 4) \cdot (3 \cdot 5) \cdot 6 \equiv 1 \cdot 1 \cdot 1 \cdot (-1) \equiv -1 \pmod{7},$$

kde do závorek jsme vždy spárovali číslo se svou inverzí (naopak je pak původní číslo inverzí k tomu spárovanému), takže součin každé závorky je 1. Jediná čísla, která nejsou spárována s žádným dalším jsou ta, co jsou spárována sama se sebou, tj. $x^{-1} \equiv x$ neboli $x^2 \equiv 1$. Věta 24, kterou bychom mohli dokázat bez problému nyní, říká, že to jsou právě 1 a -1, proto součin vždy vyjde $\equiv 1 \cdot 1 \cdots 1 \cdot (-1) \equiv -1$. □

6 Soustavy lineárních kongruencí o jedné neznámé

Soustavy lineárních kongruencí

Máme-li soustavu lineárních kongruencí o téže neznámé, můžeme podle předchozí věty rozhodnout o řešitelnosti každé z nich. V případě, kdy aspoň jedna z kongruencí nemá řešení, nemá řešení ani celá soustava. Naopak, jestliže každá z kongruencí řešení má, upravíme ji do tvaru $x \equiv c_i \pmod{m_i}$. Dostaneme tak soustavu kongruencí

$$\begin{aligned} x &\equiv c_1 \pmod{m_1} \\ &\vdots \\ x &\equiv c_k \pmod{m_k} \end{aligned}$$

Zřejmě stačí vyřešit případ $k = 2$, řešení soustavy více kongruencí snadno obdržíme opakovaným řešením soustav dvou kongruencí.

Čínská zbytková věta (CRT)

Ve čtvrtém století se čínský matematik Sun Ze (Sun Tsu) ptal na číslo, které při dělení třemi dává zbytek 2, při dělení pěti zbytek 3 a při dělení sedmi je zbytek opět 2.

Věta 12 (Čínská zbytková věta)**.** *Nechť $m_1, \dots, m_k \in \mathbb{N}$ jsou po dvou nesoudělná, $c_1, \dots, c_k \in \mathbb{Z}$. Pak platí: soustava*

$$\begin{aligned} x &\equiv c_1 \pmod{m_1} \\ &\vdots \\ x &\equiv c_k \pmod{m_k} \end{aligned}$$

má jediné řešení modulo $m_1 \cdot m_2 \cdots m_k$.

Uvědomme si, že jde o docela silné tvrzení (které ve skutečnosti platí v podstatně obecnějších algebraických strukturách), umožňující nám při předepsání libovolných zbytků podle zvolených (po dvou nesoudělných) modulů garantovat, že existuje číslo s těmito předpsanými zbytky.

Důkaz. Budeme se zabývat pouze soustavou dvou kongruencí. Uvažujme zobrazení

$$x \pmod{m_1 m_2} \mapsto (x \pmod{m_1}, x \pmod{m_2}),$$

které zbytkové třídě modulo $m_1 m_2$ přiřadí dvojici odpovídajících zbytkových tříd modulo m_1 a modulo m_2. Toto zobrazení je injektivní (viz vlastnosti kongruencí: $x \equiv y \pmod{m_1}$, $x \equiv y \pmod{m_2} \Rightarrow x \equiv y \pmod{m_1 m_2}$). Na obou stranách přitom máme stejný počet prvků $m_1 m_2$, jedná se tedy o bijekci a dvojice (c_1, c_2) má jediný vzor – tím je zbytková třída $c \pmod{m_1 m_2}$ taková, že $c \equiv c_1 \pmod{m_1}$, $c \equiv c_2 \pmod{m_2}$, tedy řešení soustavy. □

Obecný případ

Věta 13. *Nechť c_1, c_2 jsou celá čísla, m_1, m_2 přirozená. Označme $d = (m_1, m_2)$ a $m = [m_1, m_2]$. Soustava dvou kongruencí*

$$\begin{aligned} x &\equiv c_1 \pmod{m_1} \\ x &\equiv c_2 \pmod{m_2} \end{aligned}$$

v případě $c_1 \not\equiv c_2 \pmod{d}$ nemá řešení. Jestliže naopak $c_1 \equiv c_2 \pmod{d}$, pak existuje celé číslo c tak, že $x \in \mathbb{Z}$ vyhovuje soustavě, právě když vyhovuje kongruenci

$$x \equiv c \pmod{m}.$$

Důkaz. Má-li soustava nějaké řešení $x \in \mathbb{Z}$, platí nutně $x \equiv c_1 \pmod{d}$, $x \equiv c_2 \pmod{d}$, a tedy i $c_1 \equiv c_2 \pmod{d}$. Odtud plyne, že v případě $c_1 \not\equiv c_2 \pmod{d}$ soustava nemůže mít řešení.

Uvažujme opět zobrazení

$$c \pmod{m} \mapsto (c \pmod{m_1}, c \pmod{m_2}),$$

které zbytkové třídě modulo m přiřadí dvojici odpovídajících zbytkových tříd modulo m_1, m_2. Toto zobrazení je opět injektivní. Počítejme dvojice tříd ze zadání, tj. takové, že $c_1 \equiv c_2 \pmod{d}$. Libovolné $c_1 \pmod{m_1}$ určuje $c_2 \pmod{d}$ a to odpovídá právě m_2/d třídám $c_2 \pmod{m_2}$. Dohromady tak je těchto dvojic $m_1 \cdot (m_2/d) = [m_1, m_2] = m$ a zobrazení je opět bijekce (jen jsme potřebovali zmenšit množinu napravo ze všech dvojic na ty "kompatibilní"). Zbytek důkazu je stejný. □

Algoritmus

Prvně obměna na algoritmus pro jednu rovnici: soustavu

$$\begin{aligned} x &\equiv c_1 \pmod{m_1} \\ x &\equiv c_2 \pmod{m_2} \end{aligned}$$

přepíšeme na ekvivalentní

$$\begin{aligned} m_2 \cdot x &\equiv m_2 \cdot c_1 \pmod{m_1 m_2} \\ m_1 \cdot x &\equiv m_1 \cdot c_2 \pmod{m_1 m_2} \end{aligned}$$

a vyřešíme podobně jako předtím postupným odčítáním.

O něco lepší bývá převedení první rovnice na "parametrický" tvar $x = m_1 \cdot t + c_1$, dosazení do druhé rovnice

$$m_1 \cdot t + c_1 \equiv c_2 \pmod{m_2},$$

vyřešení vzhledem k t, dosazení do $x = m_1 \cdot t + c_1$ a převedení na "implicitní" tvar.

Příklad. Řešte systém kongruencí

$$\begin{aligned} x &\equiv 1 \pmod{10} \\ x &\equiv 5 \pmod{18} \\ x &\equiv -4 \pmod{25}. \end{aligned}$$

Řešení. Výsledkem je $x \equiv 221 \pmod{450}$.

Čínskou zbytkovou větu můžeme použít také „v opačném směru".

Příklad. Řešte kongruenci $23\,941x \equiv 915 \pmod{3564}$.

Řešení. Víme, že $x \in \mathbb{Z}$ je řešením dané kongruence, právě když je řešením soustavy

$$\begin{aligned} 23941x &\equiv 915 \pmod{2^2} \\ 23941x &\equiv 915 \pmod{3^4} \\ 23941x &\equiv 915 \pmod{11}. \end{aligned}$$

Vyřešíme-li postupně každou z kongruencí soustavy (což lze paralelně, viz níže), dostaneme ekvivalentní soustavu

$$\begin{aligned} x &\equiv 3 \pmod{4} \\ x &\equiv -3 \pmod{81} \\ x &\equiv -4 \pmod{11}, \end{aligned}$$

odkud snadno postupem pro řešení soustav kongruencí dostaneme $x \equiv -1137 \pmod{3564}$, což je také řešení zadané kongruence.

Modulární reprezentace čísel

Při počítání s velkými čísly je někdy výhodnější než s dekadickým či binárním zápisem čísel pracovat s tzv. *modulární reprezentací* (též *residue number system*), která umožňuje snadnou paralelizaci některých výpočtů s velkými čísly (algebraických, nefunguje např. porovnávání čísel; navíc je potřeba hlídat přetečení). Takový systém je určen k-ticí modulů (obvykle po dvou nesoudělných) a každé číslo menší než jejich součin je pak jednoznačně reprezentováno k-ticí zbytků (jejichž hodnoty nepřevyšují příslušné moduly) – viz např. `http://goo.gl/oM25m`.

Příklad. Pětice modulů $3, 5, 7, 11, 13$ nám umožní jednoznačně reprezentovat čísla menší než 15015 a efektivně provádět (v případě potřeby distribuovaně) běžné aritmetické operace. Vypočtěme např. součin čísel 1234 a 5678, v této modulární soustavě reprezentovaných pěticemi $[1, 4, 2, 2, 12]$ a $[2, 3, 1, 2, 10]$.

Součin provedeme po složkách a dostaneme $[2, 2, 2, 4, 3]$, což na závěr pomocí CRT převedeme zpět na 9662, což je modulo 15015 totéž jako $1234 \cdot 5678$.

7 Prvočísla

PRIMES is in P

Poznámka. Již jsme se zmínili, že je složité o velkých číslech s jistotou rozhodnout, jde-li o prvočíslo (na druhou stranu je o naprosté většině složených čísel snadné prokázat, že jsou skutečně složená). Přesto se v roce 2002 podařilo indickým matematikům (Agrawal, Saxena, Kayal: `http://www.cse.iitk.ac.in/users/manindra/algebra/primality_v6.pdf`) dokázat, že problém prvočíselnosti je možné rozhodnout algoritmem s časovou složitostí polynomiálně závislou na počtu cifer vstupního čísla. Nic podobného se zatím nepodařilo v otázce rozkladu čísla na prvočísla (třebaže se obecně nevěří, že je to možné, exaktní důkaz zatím nebyl podán).

Is FACTOR in P?

Nic podobného se zatím nepodařilo v otázce rozkladu čísla na prvočísla (třebaže se obecně nevěří, že je to možné, exaktní důkaz zatím nebyl podán). Nejrychlejší obecně použitelný faktorizační algoritmus, tzv. *síto v číselném tělese*[2], je sub-exponenciální časové složitosti $O(e^{1.9(\log N)^{1/3}(\log\log N)^{2/3}})$.

Poznámka. Peter Shor v roce 1994 vymyslel algoritmus, který faktorizuje v kubickém čase (tj. $O((\log N)^3)$) na kvantovém počítači. Je k tomu nicméně třeba sestrojit počítače s dostatečným počtem qubits – jak je to obtížné, lze vysledovat z toho, ze v roce 2001 se IBM podařilo pomocí kvantového počítače rozložit číslo 15 a v roce 2012 byl dosažen další faktorizační rekord rozkladem čísla 21.

RSA Challenge

Poznámka. Že je problém rozkladu přirozeného čísla na prvočísla výpočetně složitý, o tom svědčí i (již neplatná) výzva učiněná v roce 1991 firmou RSA Security (viz `http://www.rsasecurity.com/rsalabs/node.asp?id=2093`). Pokud se komukoliv podařilo rozložit čísla označená podle počtu cifer jako RSA-100, ..., RSA-704, RSA-768, ..., RSA-2048, mohl obdržet 1 000,..., 30 000, 50 000, ..., resp. 200 000 dolarů (číslo RSA-100 rozložil v témže roce Arjen Lenstra, číslo RSA-704 bylo rozloženo v roce 2012, některá dosud rozložena nebyla).

Dělitelé znovu

Díky jednoznačnosti rozkladu na prvočísla jsme schopni (se znalostí tohoto rozkladu) snadno odpovědět i na otázky ohledně počtu či součtu dělitelů konkrétního čísla. Stejně snadno dostaneme i (z dřívějška intuitivně známý) postup na výpočet největšího společného dělitele dvou čísel ze znalosti jejich rozkladu na prvočísla.

Důsledek.

- *Každý kladný dělitel čísla* $a = p_1^{n_1}\cdot\dots\cdot p_k^{n_k}$ *je tvaru* $p_1^{m_1}\cdot\dots\cdot p_k^{m_k}$, *kde* $m_1,\dots,m_k \in \mathbb{N}_0$ *a* $m_1 \le n_1$, $m_2 \le n_2$, ..., $m_k \le n_k$.
- *Číslo* a *má tedy právě* $\tau(a) = (n_1+1)(n_2+1)\cdot\dots\cdot(n_k+1)$ *kladných dělitelů, jejichž součet je*
$$\sigma(a) = \frac{p_1^{n_1+1}-1}{p_1-1}\cdots\frac{p_k^{n_k+1}-1}{p_k-1}.$$
- *Jsou-li* $p_1,\dots,p_k$ *navzájem různá prvočísla a* n_1, ..., n_k, m_1, ..., $m_k \in \mathbb{N}_0$ *a označíme-li* $r_i = \min\{n_i, m_i\}$, $s_i = \max\{n_i, m_i\}$ *pro každé* $i = 1, 2, \dots, k$, *platí*
$$(p_1^{n_1}\cdot\dots\cdot p_k^{n_k}, p_1^{m_1}\cdot\dots\cdot p_k^{m_k}) = p_1^{r_1}\cdot\dots\cdot p_k^{r_k},$$
$$[p_1^{n_1}\cdot\dots\cdot p_k^{n_k}, p_1^{m_1}\cdot\dots\cdot p_k^{m_k}] = p_1^{s_1}\cdot\dots\cdot p_k^{s_k}.$$

[2] Pro podrobnosti navštivte `M8190` Algoritmy teorie čísel

Mersenneho prvočísla a dokonalá čísla

S pojmem *součet všech kladných dělitelů čísla a* souvisí pojem tzv. *dokonalého čísla a*, které splňuje podmínku $\sigma(a) = 2a$, resp. slovně: *součet všech kladných dělitelů čísla a menších než a samotné je roven číslu a.*

Takovými čísly jsou např. $6 = 1 + 2 + 3$, $28 = 1 + 2 + 4 + 7 + 14$, 496 a 8128 (jde o všechna dokonalá čísla menší než 10 000).

Poznámka. Lze ukázat, že sudá dokonalá čísla jsou v úzkém vztahu s tzv. *Mersenneho prvočísly.* Platí totiž: *a je sudé dokonalé číslo, právě když je tvaru* $a = 2^{p-1} \cdot (2^p - 1)$*, kde* $2^p - 1$ *je prvočíslo.* (Na cvičeních se dokazuje, že pak p musí být samo prvočíslem.)

Důkaz. Nechť $a = 2^{p-1} \cdot l$ je sudé dokonalé číslo, tedy $p \geq 2$ a l je liché. Pak

$$2^p \cdot l = 2a = \sigma(a) = (2^p - 1) \cdot \sigma(l)$$

(poslední rovnost je speciálním případem multiplikativity funkce σ, viz níže). Protože $2^p - 1$ dělí levou stranu a je nesoudělné s 2^p, musí $2^p - 1 \mid l$, řekněme $l = (2^p - 1) \cdot m$ a rovnici znovu přepíšeme:

$$2^p \cdot (2^p - 1) \cdot m = 2a = \sigma(a) = (2^p - 1) \cdot \sigma((2^p - 1) \cdot m),$$

tedy $2^p \cdot m = \sigma((2^p - 1) \cdot m)$. Protože $(2^p - 1) \cdot m$ má zjevně následující dva dělitele $m < (2^p - 1) \cdot m$, jejichž součet už je $2^p \cdot m$, musí to být právě všichni dělitelé a tedy $m = 1$ a $2^p - 1$ musí být prvočíslo. □

Na druhou stranu popsat lichá dokonalá čísla se dodnes nepodařilo, resp. **dodnes se neví, jestli vůbec nějaké liché dokonalé číslo existuje**.

Hledání velkých prvočísel

Mersenneho prvočísla jsou právě prvočísla tvaru $2^p - 1$. Bez zajímavosti není ani to, že právě Mersenneho prvočísla jsou mezi všemi prvočísly nejlépe „vidět“ – pro Mersenneho čísla existuje poměrně jednoduchý a rychlý postup, jak ověřit, že jde o prvočísla.

Test (Lucas-Lehmerův test)**.** *Definujme posloupnost* $(s_n)_{n=0}^{\infty}$ *rekurzívně předpisem* $s_0 = 4, s_{n+1} = s_n^2 - 2$.

Pak je číslo $M_p = 2^p - 1$ *prvočíslo, právě tehdy, když* M_p *dělí* s_{p-2}.

Proto není náhodou, že největší známá prvočísla jsou obvykle tvaru $2^p - 1$ (viz např. `http://www.utm.edu/research/primes/largest.html`).

Jakkoliv může být hledání největšího známého prvočísla chápáno jako pochybná zábava bez valného praktického užitku[3], jednak posunuje hranice matematického poznání a zdokonaluje použité metody (a často i hardware), jednak může přinést benefit i samotným objevitelům (Electronic Frontier Foundation vypsala odměny EFF Cooperative Computing Awards za nalezení prvočísla majícího alespoň $10^6, 10^7, 10^8$ a 10^9 číslic – odměny 50, resp. 100 tisíc $ za první dvě kategorie byly vyplaceny v letech 2000, resp. 2009 – v obou případech projektu GIMPS – na další odměny si ještě zřejmě nějaký čas počkáme).

Rozložení prvočísel

Nyní se budeme snažit zodpovědět následující otázky:

1. Je prvočísel nekonečně mnoho?
2. Je prvočísel nekonečně mnoho v každé (nebo aspoň některé) aritmetické posloupnosti?
3. Jak jsou prvočísla rozložena mezi přirozenými čísly?

[3] Viz např. titulek iDnes z 6.února 2013: *Největší známé prvočíslo na světě má 17 milionů číslic a je k ničemu*

> There are two facts about the distribution of prime numbers. The first is that, [they are] the most arbitrary and ornery objects studied by mathematicians: they grow like weeds among the natural numbers, seeming to obey no other law than that of chance, and nobody can predict where the next one will sprout. The second fact is even more astonishing, for it states just the opposite: that the prime numbers exhibit stunning regularity, that there are laws governing their behavior, and that they obey these laws with almost military precision.
>
> Don Zagier

Prvočísel je nekonečně mnoho

Věta 14 (Eukleidés)**.** *Mezi přirozenými čísly existuje nekonečně mnoho prvočísel.*

Důkaz. Předpokládejme, že prvočísel je konečně mnoho a označme je $p_1, p_2, \ldots, p_n$. Položme $N = p_1 \cdot p_2 \ldots p_n + 1$. Toto číslo je buď samo prvočíslem nebo je dělitelné nějakým prvočíslem různým od $p_1, \ldots, p_n$ (čísla $p_1, \ldots, p_n$ totiž dělí číslo $N - 1$), což je spor. □

Poznámka. Existuje mnoho variant důkazů nekonečnosti prvočísel z různých oblastí matematiky, uveďme ještě alespoň některá tvrzení, z nichž zároveň získáme alespoň částečnou informaci o rozložení prvočísel mezi přirozenými čísly.

Prvočísel je vcelku hodně

Příklad. Pro celé $n > 2$ existuje mezi čísly n a $n!$ alespoň jedno prvočíslo.

Řešení. Označme p libovolné prvočíslo dělící číslo $n! - 1$ (takové existuje podle Základní věty aritmetiky, protože $n! - 1 > 1$). Kdyby $p \leq n$, muselo by p dělit číslo $n!$ a nedělilo by $n! - 1$. Je tedy $n < p$. Protože $p \mid (n! - 1)$, platí $p \leq n! - 1$, tedy $p < n!$. Prvočíslo p splňuje podmínky úlohy. □

Z této věty rovněž vyplývá nekonečnost prvočísel, její tvrzení je ale velice slabé. Následující tvrzení, uvedené bez důkazu, je podstatně silnější.

Věta 15 (Čebyševova, Bertrandův postulát)**.** *Pro libovolné číslo $n > 1$ existuje alespoň jedno prvočíslo p splňující $n < p < 2n$.*

Prvočísel je vcelku málo

Příklad. Dokažte, že pro libovolné přirozené číslo n existuje n po sobě jdoucích přirozených čísel, z nichž žádné není prvočíslo.

Řešení. Zkoumejme čísla $(n+1)! + 2, (n+1)! + 3, \ldots, (n+1)! + (n+1)$. Mezi těmito n po sobě jdoucími čísly není žádné prvočíslo, protože pro libovolné $k \in \{2, 3, \ldots, n+1\}$ platí $k \mid (n+1)!$, a tedy $k \mid (n+1)! + k$, a proto $(n+1)! + k$ nemůže být prvočíslo. □

Prvočísla jsou relativně rovnoměrně rozložena v tom smyslu, že v libovolné „rozumné“ aritmetické posloupnosti je jich nekonečně mnoho. Například zbytek 1 po dělení čtyřmi, stejně jako zbytek 3 po dělení čtyřmi dá vždy nekonečně mnoho prvočísel (zbytek 0 nedá samozřejmě žádné a zbytek 2 pouze jediné). Obdobná situace je pak při uvažování zbytků po dělení libovolným jiným přirozeným číslem, jak uvádí následující věta, jejíž důkaz je ovšem velmi obtížný.

Věta 16 (Dirichletova o prvočíslech v aritmetické posloupnosti)**.** *Jsou-li a, m nesoudělná přirozená čísla, existuje nekonečně mnoho přirozených čísel k tak, že $mk + a$ je prvočíslo. Jinými slovy, mezi čísly $1 \cdot m + a, 2 \cdot m + a, 3 \cdot m + a, \ldots$ existuje nekonečně mnoho prvočísel.*

Uveďme proto alespoň důkaz ve speciálním případě.

Prvočísel tvaru $3k + 2$ je nekonečně mnoho

Příklad. Dokažte, že existuje nekonečně mnoho prvočísel tvaru $3k + 2$, kde $k \in \mathbb{N}_0$.

Řešení. Předpokládejme naopak, že existuje pouze konečně mnoho prvočísel tohoto tvaru a označme je $p_1 = 2$, $p_2 = 5$, $p_3 = 11$, …, p_n. Položme $N = 3p_2 \cdot p_3 \cdot \cdots \cdot p_n + 2$. Rozložíme-li N na součin prvočísel, musí v tomto rozkladu vystupovat aspoň jedno prvočíslo p tvaru $3k+2$, neboť v opačném případě by bylo N součinem prvočísel tvaru $3k + 1$ (uvažte, že N není dělitelné třemi), a tedy podle dřívějšího příkladu by bylo i N tvaru $3k + 1$, což není pravda. Prvočíslo p ovšem nemůže být žádné z prvočísel $p_1, p_2, \ldots, p_n$, jak plyne z tvaru čísla N, a to je spor.

Asymptotické chování prvočísel

Z tvrzení uvedených v této kapitole je možné si udělat hrubou představu o tom, jak „hustě“ se mezi přirozenými čísla prvočísla vyskytují. Přesněji (i když „pouze“ asymptoticky) to popisuje velmi důležitá tzv. „Prime Number Theorem“:

Věta 17 (Prime Number Theorem, věta o hustotě prvočísel)**.** *Nechť $\pi(x)$ udává počet prvočísel menších nebo rovných číslu $x \in \mathbb{R}$. Pak*

$$\pi(x) \sim \frac{x}{\ln x},$$

tj. podíl funkcí $\pi(x)$ a $x/\ln x$ se pro $x \to \infty$ limitně blíží k 1.

Příklad. O tom, jak odpovídá asymptotický odhad $\pi(x) \sim x/\ln(x)$, v některých konkrétních příkladech vypovídá následující tabulka:

x	$\pi(x)$	$x/\ln(x)$	relativní chyba
100	25	21.71	0.13
1000	168	144.76	0.13
10000	1229	1085.73	0.11
100000	9592	8685.88	0.09
1000000	78498	72382.41	0.08

8 Aritmetické funkce

Aritmetické funkce

Aritmetickou funkcí zde rozumíme funkci, jejímž definičním oborem je množina kladných celých čísel.

Definice. Multiplikativní funkcí přirozených čísel rozumíme takovou aritmetickou funkci, která splňuje, že pro všechny dvojice **nesoudělných** čísel $a, b \in \mathbb{N}$ platí

$$f(a \cdot b) = f(a) \cdot f(b).$$

Příklad. Multiplikativními funkcemi jsou např. funkce $f(n) = \sigma(n)$, $f(n) = \tau(n)$ nebo, jak brzy dokážeme i tzv. Eulerova funkce $f(n) = \phi(n)$.

V podstatě přímo z definice plyne, že multiplikativní funkce je jednoznačně určena svými hodnotami na mocninách prvočísel:

$$f(n) = f(p_1^{\alpha_1} \cdots p_k^{\alpha_k}) = f(p_1^{\alpha_1}) \cdots f(p_k^{\alpha_k}).$$

Eulerova funkce

Definice. Nechť $n \in \mathbb{N}$. Definujme Eulerovu funkci ϕ předpisem

$$\phi(n) = |\{a \in \mathbb{N} \,|\, 0 < a \leq n,\ (a,n) = 1\}|$$

(lépe počet zbytkových tříd nesoudělných s n nebo také těch, které mají modulo n inverzi).

Příklad. $\phi(1) = 1, \phi(5) = 4, \phi(6) = 2$, je-li p prvočíslo, je zřejmě $\phi(p) = p - 1$.

Nyní dokážeme několik důležitých tvrzení o funkci ϕ:

Lemma. *Platí* $\phi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^{\alpha-1}(p-1) = p^\alpha(1 - \frac{1}{p})$.

Důkaz. Mezi čísly $\{1, \ldots, p^\alpha\}$ jsou soudělná s p^α právě násobky p, tedy

$$1 \cdot p, 2 \cdot p, \ldots, p^{\alpha-1} \cdot p$$

a těch je $p^{\alpha-1}$. Nesoudělných je proto $p^\alpha - p^{\alpha-1}$. □

Věta 18. *Eulerova funkce ϕ je multiplikativní.*
Nechť $n \in \mathbb{N}$, jehož rozklad je tvaru $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$. Pak

$$\phi(n) = (p_1^{\alpha_1} - p_1^{\alpha_1 - 1}) \cdots (p_k^{\alpha_k} - p_k^{\alpha_k - 1}) = n \cdot \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_k}\right).$$

Důkaz. Nechť a, b jsou nesoudělná. Připomeňme bijekci

$$x \pmod{a \cdot b} \longmapsto (x \pmod{a}, x \pmod{b})$$

Stačí proto ukázat, že $x \pmod{a \cdot b}$ má inverzi, právě když obě složky obrazu mají inverzi – takových dvojic je totiž přesně $\phi(a) \cdot \phi(b)$. To plyne z Čínské zbytkové věty – inverze modulo $a \cdot b$ je inverzí modulo a i modulo b, naopak k inverzím modulo a a modulo b lze najít jedinou reprezentující zbytkovou třídu modulo $a \cdot b$; ta je pak inverzí modulo a i modulo b, tedy modulo $a \cdot b$. □

Příklad. Vypočtěte $\phi(72)$.

Řešení. $72 = 2^3 \cdot 3^2 \implies \phi(72) = 72 \cdot (1 - \frac{1}{2}) \cdot (1 - \frac{1}{3}) = 24$, alternativně $\phi(72) = \phi(8) \cdot \phi(9) = 4 \cdot 6 = 24$. □

Příklad. Dokažte, že $\forall n \in \mathbb{N} :\ \phi(4n+2) = \phi(2n+1)$.

Řešení. $\phi(4n+2) = \phi(2 \cdot (2n+1)) = \phi(2) \cdot \phi(2n+1) = \phi(2n+1)$. □

9 Malá Fermatova věta, Eulerova věta

Úplná a redukovaná soustava zbytků

Definice. *Úplná soustava zbytků* modulo m je libovolná m-tice čísel po dvou nekongruentních modulo m (nejčastěji $0, 1, \ldots, m-1$).

Redukovaná soustava zbytků modulo m je libovolná $\phi(m)$-tice čísel nesoudělných s m a po dvou nekongruentních modulo m.

Lemma. *Nechť $x_1, x_2, \ldots, x_{\phi(m)}$ tvoří redukovanou soustavu zbytků modulo m. Je-li $a \in \mathbb{Z}$, $(a,m) = 1$ pak i čísla $a \cdot x_1, \ldots, a \cdot x_{\phi(m)}$ tvoří redukovanou soustavu zbytků modulo m.*

Důkaz. Protože $(a,m) = 1$ a $(x_i, m) = 1$, platí $(a \cdot x_i, m) = 1$ (buď přes rozklad na prvočísla nebo přes invertibilnost modulo m – platí $(a \cdot x_i)^{-1} = x_i^{-1} \cdot a^{-1}$). Kdyby pro nějaká i, j platilo $a \cdot x_i \equiv a \cdot x_j \pmod{m}$, po vydělení obou stran kongruence číslem a nesoudělným s m dostaneme $x_i \equiv x_j \pmod{m}$. □

Eulerova věta

Věta 19 (Eulerova)**.** *Nechť $a \in \mathbb{Z}$, $m \in \mathbb{N}$, $(a, m) = 1$. Pak*

$$a^{\phi(m)} \equiv 1 \pmod{m}.$$

Důkaz. Buď $x_1, x_2, \ldots, x_{\phi(m)}$ libovolná redukovaná soustava zbytků modulo m. Podle předchozího lemmatu je i $a \cdot x_1, \ldots, a \cdot x_{\phi(m)}$ redukovaná soustava zbytků modulo m. Platí tedy, že pro každé i existuje j ($i, j \in \{1, 2, \ldots, \phi(m)\}$) tak, že $a \cdot x_i \equiv x_j \pmod{m}$. Vynásobením dostáváme $(a \cdot x_1) \cdot (a \cdot x_2) \cdots (a \cdot x_{\phi(m)}) \equiv x_1 \cdot x_2 \cdots x_{\phi(m)} \pmod{m}$. Po úpravě

$$a^{\phi(m)} \cdot x_1 \cdot x_2 \cdots x_{\phi(m)} \equiv x_1 \cdot x_2 \cdots x_{\phi(m)} \pmod{m}$$

vydělení číslem $x_1 \cdot x_2 \cdots x_{\phi(m)}$ dostaneme požadované. □

Speciálním případem pro prvočíselný modul je pak tzv. Fermatova věta.

Věta 20 (Fermatova, Malá Fermatova)**.** *Nechť $a \in \mathbb{Z}$, p prvočíslo, $p \nmid a$. Pak*

$$a^{p-1} \equiv 1 \pmod{p}.$$

V tomto případě se dá věta přeformulovat bez předpokladu $p \nmid a$:

Důsledek. *Nechť $a \in \mathbb{Z}$, p prvočíslo. Pak*

$$a^p \equiv a \pmod{p}.$$

Řád čísla

S Eulerovou funkcí a Eulerovou větou úzce souvisí důležitý pojem *řád čísla modulo m*:

Definice. Nechť $a \in \mathbb{Z}$, $m \in \mathbb{N}$ $(a, m) = 1$. *Řádem čísla a modulo m* rozumíme nejmenší kladné celé číslo n splňující

$$a^n \equiv 1 \pmod{m}.$$

Poznámka. To, že je řád definován, plyne z Eulerovy věty – pro každé číslo nesoudělné s modulem je totiž jistě jeho řád nejvýše roven $\phi(m)$. Jak později uvidíme, velmi důležitá jsou právě ta čísla, jejichž **řád je roven právě** $\phi(m)$ – tato čísla nazýváme primitivními kořeny modulo m.

Příklad. Pro libovolné $m \in \mathbb{N}$ má číslo 1 modulo m řád 1. Číslo -1 má řád

- 1 pro $m = 1$ nebo $m = 2$
- 2 pro $m > 2$

Příklad. Určete řád čísla 2 modulo 7.

Řešení.

$$2^1 = 2 \not\equiv 1 \pmod{7}$$
$$2^2 = 4 \not\equiv 1 \pmod{7}$$
$$2^3 = 8 \equiv 1 \pmod{7}$$

Řád čísla 2 modulo 7 je tedy roven 3. □

Věta 21. *Nechť $m \in \mathbb{N}$, $a \in \mathbb{Z}$, $(a, m) = 1$. Označme r řád čísla a modulo m. Pak pro libovolná $t, s \in \mathbb{N}$ platí*

$$a^t \equiv a^s \pmod{m} \iff t \equiv s \pmod{r}.$$

Důkaz. Díky tomu, že $a^s \equiv a^s \cdot a^r \equiv a^{s+r} \pmod m$, opakují se hodnoty mocnin a^s s periodou r, což dává implikaci "$\Leftarrow$". Zbývá ukázat, že zbytkové třídy $a^0 \pmod m, \ldots, a^{r-1} \pmod m$ jsou všechny různé. Přitom pokud pro $0 \leq t \leq s \leq r-1$ platí $a^s \equiv a^t \pmod m$, pak vydělením a^t dostaneme $a^{s-t} \equiv 1 \pmod m$ a vzhledem k tomu, že $s - t < r$, musí být $s - t = 0$, tj. $s = t$. □

Zřejmým důsledkem předchozí věty a Eulerovy věty je následující tvrzení

Důsledek. *Nechť $m \in \mathbb{N}$, $a \in \mathbb{Z}$, $(a, m) = 1$. Označme r řád čísla a modulo m. Pak platí $r \mid \phi(m)$.*

Naším dalším cílem bude důkaz existence primitivního kořene modulo prvočíslo p. Budeme tedy chtít z prvků různých řádů vyrábět prvky nových řádů, k čemuž nám poslouží několik následujících tvrzení, pak se vrhneme na samotný důkaz existence primitivního kořene.

Věta 22. *Nechť $m \in \mathbb{N}$, $a \in \mathbb{Z}$, $(a, m) = 1$. Je-li řád čísla a modulo m roven $r \in \mathbb{N}$ a $k \mid r$, je řád čísla a^k modulo m roven r/k.*

Důkaz. Podle předchozí věty je $(a^k)^n \equiv 1 \pmod m$, právě když $kn \equiv 0 \pmod r$. Vydělením obou stran i modulu číslem k dostaneme ekvivalentní podmínku $n \equiv 0 \pmod{r/k}$ a nejmenší takové kladné n pak je právě r/k. □

Poznámka. Předchozí lemma a jeho důkaz lze snadno rozšířit i na případ obecného k, které nemusí nutně dělit r. Označme $d = (k, r)$ jejich největší společný dělitel. Opět dostaneme $kn \equiv 0 \pmod r$, nyní však můžeme dělit celou kongruenci i s modulem pouze d a dostaneme tak ekvivalentní kongruenci $(k/d)n \equiv 0 \pmod{r/d}$; protože už jsou k/d a r/d nesoudělné, můžeme vydělit obě strany k/d a dostaneme ekvivalentní kongruenci $n \equiv 0 \pmod{r/d}$. Ve výsledku je tedy řád roven r/d.

Poslední z této řady tvrzení dává do souvislosti řády dvou čísel a řád jejich součinu:

Lemma. *Nechť $m \in \mathbb{N}$, $a, b \in \mathbb{Z}$, $(a, m) = (b, m) = 1$. Jestliže a je řádu s a b je řádu t modulo m, kde $(s, t) = 1$, pak číslo $a \cdot b$ je řádu $s \cdot t$ modulo m.*

Důkaz. Označme r řád čísla $a \cdot b$. Zřejmě platí $(ab)^{st} \equiv 1 \pmod m$, proto $r \mid st$. Naopak podle definice řádu $(ab)^r \equiv 1 \pmod m$ a umocněním obou stran kongruence dostaneme $a^{rs}b^{rs} \equiv 1 \pmod m$. Protože je s řádem čísla a, je $a^s \equiv 1 \pmod m$, tj. $b^{rs} \equiv 1 \pmod m$, a proto $t \mid rs$. Z nesoudělnosti t a s plyne $t \mid r$. Analogicky dostaneme i $s \mid r$, a tedy (opět s využitím nesoudělnosti s, t) $s \cdot t \mid r$. Celkem tedy $r = st$. □

Důsledek. *Nechť $m \in \mathbb{N}$ a r je nejmenší společný násobek všech řádů modulo m. Pak existuje číslo řádu r modulo m.*

Důkaz. Stačí pro a řádu s, b řádu t najít prvek řádu $[s, t]$. Nechť $d = (s, t)$. Definujme k jako součin těch prvočíselných mocnin p^α z rozkladu a, které se v b vyskytují ve stejné nebo vyšší mocnině (tj. $p^\alpha \mid b$) a položme $d = k \cdot l$. Potom a^k je řádu s/k a b^l je řádu t/l a snadno se vidí, že $(s/k, t/l) = 1$ (každé prvočíslo z rozkladu d se vyskytuje pouze v s/k nebo t/l ale ne v obou) a také $s/k \cdot t/l = [s, t]$, takže podle předchozího lemmatu $a^k \cdot b^l$ má řád $[s, t]$, jak jsme chtěli. □

10 Primitivní kořeny

Podle předchozího důsledku existuje zbytková třída maximálního možného řádu r a řády všech ostatních zbytkových tříd jsou děliteli r. Pak všechna $(a, m) = 1$ splňují $a^r \equiv 1 \pmod m$, tj. jsou to řešení kongruence

$$x^r \equiv 1 \pmod m.$$

Jedná se tedy o polynomiální kongruenci stupně r mající $\phi(m)$ kořenů. V dalším nás tedy bude zajímat vztah mezi stupněm polynomu a počtu jeho kořenů modulo m. Z oboru reálných čísel jsme zvyklí, že polynomiální rovnice stupně r má maximálně r kořenů. Pro zbytkové třídy modulo m tomu tak obecně být nemusí, v případě prvočíselného modulu to ale bude opět platit.

Primitivní kořeny modulo součin

Příklad. Nechť $m = 35$ a nechť $(a, m) = 1$. Pak podle Eulerovy věty

$$\begin{aligned} a^4 &\equiv 1 \pmod 5,\ a^6 \equiv 1 \pmod 7 \\ a^{12} &\equiv 1 \pmod 5,\ a^{12} \equiv 1 \pmod 7 \quad \Rightarrow \quad a^{12} \equiv 1 \pmod{35}. \end{aligned}$$

Je tedy každé číslo řádu 12 (případně menšího).

Tedy kongruence $x^{12} \equiv 1 \pmod{35}$ stupně 12 má $\phi(35) = 4 \cdot 6 = 24$ řešení.

Věta 23. *Pokud je m dělitelné aspoň dvěma lichými prvočísly, primitivní kořen modulo m neexistuje.*

Důkaz. Hlavní myšlenka je stejná jako výše. Rozložme $m = k \cdot l$ na součin dvou nesoudělných čísel větších než 2, např. k může být nejvyšší možná mocnina jednoho z lichých prvočísel z předpokladu. Označíme-li $n = [\phi(k), \phi(l)]$ nejmenší společný násobek, platí

$$a^{\phi(k)} \equiv 1 \pmod k \quad \Rightarrow \quad a^n \equiv 1 \pmod k$$

a podobně $a^n \equiv 1 \pmod l$; protože $m = k \cdot l$ a činitelé jsou nesoudělní, dostaneme dohromady $a^n \equiv 1 \pmod m$ a každý prvek má řád maximálně $n = [\phi(k), \phi(l)] < \phi(k)\phi(l) = \phi(m)$, protože hodnoty $\phi(k)$, $\phi(l)$ jsou obě sudé a tudíž soudělné. □

Podobně, pokud je m dělitelné čtyřmi a aspoň jedním lichým prvočíslem, primitivní kořen modulo m neexistuje.

Polynomiální kongruence modulo prvočíslo

Vydělme polynom $f(x)$ se zbytkem kořenovým činitelem $(x - a)$:

$$f(x) = (x - a) \cdot g(x) + r \qquad \Rightarrow \qquad r = f(a)$$

Pokud je a kořenem kongruence $f(x) \equiv 0 \pmod p$, dostaneme

$$f(x) \equiv (x - a) \cdot g(x) \pmod p.$$

Protože je p prvočíslo, jsou kořeny $f(x)$ právě kořen a lineárního činitele $x - a$ a kořeny $g(x)$, který je stupně o jedna menšího. Protože konstantní polynomy nemají kořeny, má $f(x)$ maximálně tolik kořenů, kolik je jeho stupeň (bacha na $f(x) \equiv 0$).

Věta 24. $x^2 \equiv 1 \pmod p$ *má kořeny právě* ± 1.

Primitivní kořen modulo prvočíslo

Věta 25. *Nechť $p \in \mathbb{N}$ je prvočíslo. Pak existuje primitivní kořen modulo p.*

Důkaz. Nechť r je maximální řád, podle Eulerovy věty $r \mid p - 1$. Pak všech $p - 1$ nenulových zbytkových tříd jsou kořeny

$$x^r \equiv 1 \pmod p$$

a podle předchozího $p - 1 \leq r$. □

Věta 26. *Nechť $p \in \mathbb{N}$ je prvočíslo. Pak existuje primitivní kořen modulo p^k.*

Důkaz. Platí $\phi(p^k) = p^{k-1} \cdot (p - 1)$ a tito činitelé jsou nesoudělní.

Nechť $g \pmod p$ je primitivní kořen, tj. prvek řádu $p - 1$. Pak řád $g \pmod{p^k}$ bude násobkem $p - 1$.

Stačí najít prvek řádu p^{k-1}. Ukážeme, že je jím $1 + p$; konkrétně níže ukážeme:

$$(1 + p)^{p^{k-2}} \equiv 1 + p^{k-1} \not\equiv 1 \pmod{p^k}$$

instance pro $k+1$ dá $(1+p)^{p^{k-1}} \equiv 1+p^k \pmod{p^{k+1}}$ a po redukci modulo p^k pak vyjde $(1+p)^{p^{k-1}} \equiv 1 + p^k \equiv 1 \pmod{p^k}$. □

Lemma. $a \equiv b \pmod{p^k} \qquad \Rightarrow \qquad a^p \equiv b^p \pmod{p^{k+1}}$.

Důkaz. Platí $a^p - b^p = (a-b)(a^{p-1} + \cdots + a^k b^l + \cdots + b^{p-1})$, kde první člen je dělitelný p^k podle předpokladu, zatímco druhý člen je modulo p kongruentní

$$a^{p-1} + \cdots + a^k b^l + \cdots + b^{p-1} \equiv \underbrace{a^{p-1} + \cdots + a^{p-1}}_{p \text{ členů}} \equiv p \cdot a^{p-1} \equiv 0 \pmod{p},$$

tedy je dělitelný p a součin je dělitelný p^{k+1}. □

Důsledek. $(1+p)^{p^{k-2}} \equiv 1 + p^{k-1} \pmod{p^k}$.

Důkaz. Indukční krok (v druhé kongruenci aplikujeme předchozí lemma na indukční předpoklad):

$$\begin{aligned}(1+p)^{p^{k-1}} &\equiv ((1+p)^{p^{k-2}})^p \equiv (1+p^{k-1})^p \equiv 1 + \binom{p}{1} p^{k-1} + \binom{p}{2} p^{2(k-1)} + \cdots \\ &\equiv 1 + p^k \pmod{p^{k+1}}.\end{aligned}$$

□

Hledání primitivního kořene

Sofistikovaná metoda se zatím nezná. Zkoušením po nějaké době uspějeme: Kolik je primitivních kořenů modulo prvočíslo? Jsou to právě $g^a \pmod{p}$, kde $(a, \phi(p)) = 1$, tedy jich je $\phi(\phi(p))$. Přitom platí

$$p/\phi(\phi(p)) \in O(\log\log p),$$

takže pravděpodobnost, že náhodné číslo bude primitivním kořenem je zhruba

$$1/\log\log p$$

a počet pokusů potřebných k nalezení primitivního kořene s předem danou pravděpodobností je úměrný $\log\log p$, tedy logaritmický vzhledem k délce vstupu (ověření toho, zda se vskutku jedná o primitivní kořen trvá déle, viz příště).

Diskrétní logaritmus

Definice. Nechť $m > 0$. Celé číslo $g \in \mathbb{Z}$, $(g, m) = 1$ nazveme *primitivním kořenem* modulo m, pokud je jeho řád modulo m roven $\phi(m)$.

Lemma. *Je-li g primitivní kořen modulo m, pak pro každé číslo $a \in \mathbb{Z}$, $(a, m) = 1$ existuje jediné $\alpha \in \mathbb{Z}$, $0 \le \alpha < \phi(m)$ s vlastností $g^\alpha \equiv a \pmod{m}$. Označujeme $\alpha = \log_g a$ a nazýváme* ***diskrétním logaritmem****, příp. indexem čísla a (vzhledem k danému m a zafixovanému primitivnímu kořeni g) a je bijekcí mezi množinami*
$\{a \in \mathbb{Z};\ (a, m) = 1, 0 < a < m\}$ *a* $\{x \in \mathbb{Z};\ 0 \le x < \phi(m)\}$.

Důkaz. Zobrazení $\alpha \pmod{\phi(m)} \mapsto g^\alpha \pmod{m}$ je dobře definované a injektivní podle valstností řádu. Protože mají množiny stejný počet prvků (vpravo bereme pouze invertibilní zbytkové třídy), jedná se o bijekci. □

11 Kvadratické zbytky a nezbytky

Kvadratické kongruence modulo prvočíslo

Věta 27. *Nechť p je liché prvočíslo a $(a, p) = 1$. Kongruence $x^2 \equiv a \pmod{p}$ má řešení, právě když $a^{\frac{p-1}{2}} \equiv 1 \pmod{p}$.*

Důkaz. Použijeme primitivní kořen g a vyjádříme $x^2 \equiv a \pmod p$ pomocí něj: nechť $x \equiv g^\xi$, $a \equiv g^\alpha$, pak kongruence je ekvivalentní

$$(g^\xi)^2 \equiv g^\alpha \pmod p \quad \Leftrightarrow \quad 2\xi \equiv \alpha \pmod{p-1}.$$

Protože je $p-1$ sudé, řešení existuje, právě když α je sudé:

$$\begin{aligned} \alpha \equiv 0 \pmod 2 \quad &\Leftrightarrow \quad \tfrac{p-1}{2} \cdot \alpha \equiv 0 \pmod{p-1}. \\ &\Leftrightarrow \quad a^{\frac{p-1}{2}} \equiv g^{\frac{p-1}{2}\cdot\alpha} \equiv g^0 \equiv 1 \pmod p. \end{aligned}$$

□

Legendreův symbol

Definice. Definujeme $(\frac{a}{p}) = \begin{cases} 1 & a \text{ kvadratický zbytek modulo } p \\ -1 & a \text{ kvadratický nezbytek modulo } p \\ 0 & a \text{ soudělné s } p \end{cases}$.

Jednoduchým důsledkem věty dostáváme $(\frac{a}{p}) \equiv a^{\frac{p-1}{2}} \pmod p$: protože $(a^{\frac{p-1}{2}})^2 \equiv a^{p-1} \equiv 1$, je $a^{\frac{p-1}{2}}$ rovno ± 1.

Důsledek. $(\frac{-1}{p}) = +1$*, resp.* -1*, pokud* $p \equiv 1 \pmod 4$*, resp.* $p \equiv 3 \pmod 4$*.*
Tedy kongruence $x^2 \equiv -1 \pmod p$ *má řešení, právě když* p *dává po dělení čtyřmi zbytek 1.*

Počítání Legendreova symbolu je jednoduché s následujícími pravidly:

- $(\frac{ab}{p}) = (\frac{a}{p})(\frac{b}{p})$,
- $a \equiv b \pmod p \Rightarrow (\frac{a}{p}) = (\frac{b}{p})$,
- $(\frac{2}{p}) = (-1)^{\frac{p^2-1}{8}}$,
- $(\frac{q}{p}) = (\frac{p}{q}) \cdot (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$.

Důkaz. První dvě položky plynou velice snadno z vyjádření $(\frac{a}{p}) \equiv a^{\frac{p-1}{2}} \pmod p$. Druhé dva poznatky jsou značně hlubší, ukážeme hlavní ideu jejich důkazu, která spočívá v „geometrické“ interpretaci Legendreova symbolu $(\frac{a}{p}) \equiv a^{\frac{p-1}{2}} \pmod p$. Při interpretaci Legendreova symbolu vyjdeme z důkazu Eulerovy věty, ilustrujeme na příkladu $(\frac{7}{11}) \equiv 7^5 \pmod{11}$. Pišme tentokrát zbytky jako $\pm 1, \ldots, \pm 5$ a uvažme opět násobení sedmi:

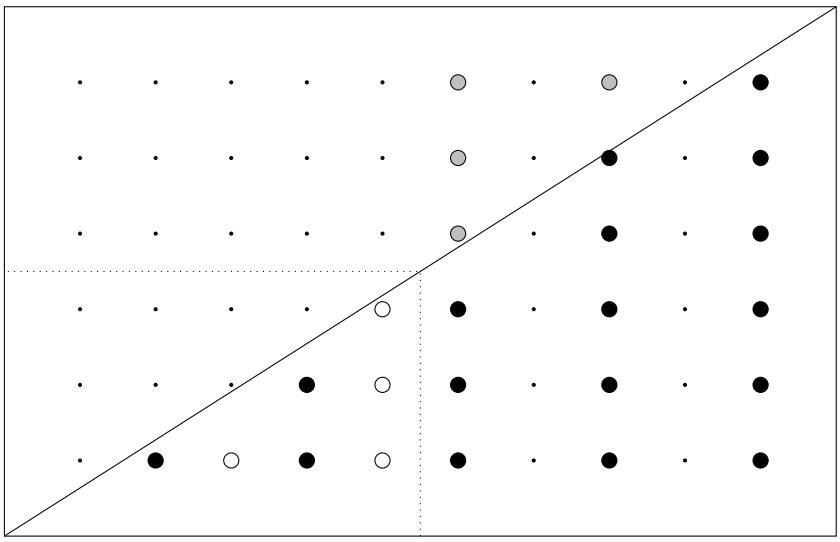
$$\begin{aligned}
7 \cdot 1 &\equiv -4 \pmod{11} & 1 \cdot \frac{7}{11} &= 1 - \frac{4}{11} & 2 \cdot \frac{7}{11} &= 1 + \frac{3}{11} \\
7 \cdot 2 &\equiv +3 \pmod{11} & 2 \cdot \frac{7}{11} &= 1 + \frac{3}{11} & 4 \cdot \frac{7}{11} &= 2 + \frac{6}{11} \\
7 \cdot 3 &\equiv -1 \pmod{11} & 3 \cdot \frac{7}{11} &= 2 - \frac{1}{11} & 6 \cdot \frac{7}{11} &= 3 + \frac{9}{11} \\
7 \cdot 4 &\equiv -5 \pmod{11} & 4 \cdot \frac{7}{11} &= 3 - \frac{5}{11} & 8 \cdot \frac{7}{11} &= 5 + \frac{1}{11} \\
7 \cdot 5 &\equiv +2 \pmod{11} & 5 \cdot \frac{7}{11} &= 3 + \frac{2}{11} & 10 \cdot \frac{7}{11} &= 6 + \frac{4}{11}
\end{aligned}$$

a vynásobením dostaneme $7^5 \cdot 5! \equiv \pm 5! \pmod{11}$, takže výsledné znaménko je přesně Legendreův symbol $7^5 \pmod{11}$. Uvážíme-li namísto zbytků po dělení jedenácti příslušné zlomky se jmenovatelem 11 jako v prostředním sloupci, budou kladné zbytky odpovídat „oznaménkové desetinné části“ menší než $1/2$, zatímco záporné zbytky oznaménkové desetinné části větší než $1/2$. Ještě jinak, po vynásobení zlomků dvěma jako v pravém sloupci budou kladné zbytky odpovídat sudé

celé části (standardní – desetinná část bez znamének), zatímco záporné zbytky liché celé části. Dostáváme tak

$$\left(\frac{7}{11}\right) = (-1)^{[2\cdot\frac{7}{11}]+[4\cdot\frac{7}{11}]+[6\cdot\frac{7}{11}]+[8\cdot\frac{7}{11}]+[10\cdot\frac{7}{11}]}.$$

Podstatná je tedy parita exponentu, přičemž exponent je „zjevně“ počet černě zvýrazněných mřížových bodů v následujícím obrázku (sudé sloupce, pod diagonálou):



Protože je v každém sloupečku dohromady sudý počet (šest) mřížových bodů, je parita černě zvýrazněných bodů v pravé polovině stejná jako počet šedě zvýrazněných bodů (komplement v každém sloupečku) a díky středové symetrii je pak tento počet stejný jako počet bíle zvýrazněných bodů. Dohromady dostáváme, že $(\frac{7}{11})$ je rovno -1 na počet mřížových bodů pod diagonálou v levé části. Symetricky dostaneme, že $(\frac{11}{7})$ je rovno -1 na počet mřížových bodů vlevo od diagonály ve spodní části. Vynásobením pak $(\frac{7}{11})\cdot(\frac{11}{7})$ je rovno -1 na počet mřížových bodů v levé spodní části; těch je zjevně $5\cdot 3$. Obecně pak

$$\left(\frac{q}{p}\right)\cdot\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}.$$

V důkazu jsme nepoužili, že 7 je prvočíslo, ale jen, že je liché (protože jsme potřebovali v každém sloupečku sudý počet mřížových bodů). Proto můžeme počítat (směrnice této diagonály je přibližně 1 a počty mřížových bodů v jednotlivých sloupečcích budou narůstat o jedna přesně do poloviny):

$$\left(\frac{2}{p}\right) = \left(\frac{p+2}{p}\right) = (-1)^{1+2+\cdots+\frac{p-1}{2}} = (-1)^{\frac{p-1}{2}\cdot\frac{p+1}{4}} = (-1)^{\frac{p^2-1}{8}}. \qquad \square$$

Ukážeme nyní, jak spočítat druhé odmocniny z kvadratického zbytku $a \pmod p$ modulo prvočíslo splňující $p \equiv 3 \pmod 4$. V takovém případě

$$a \equiv \left(\frac{a}{p}\right)\cdot a \equiv a^{\frac{p-1}{2}}\cdot a \equiv a^{\frac{p+1}{2}}$$

a odmocninu lze snadno spočítat jako $\pm a^{\frac{p+1}{4}}$. V případě, že a není kvadratický zbytek, tj. $(\frac{a}{p}) = -1$, dostaneme takto druhé odmocniny z $-a$, jak lze snadno z postupu vidět.

Jacobiho symbol

Definice. a libovolné, $n = p_1\cdots p_k$ liché číslo vyjádřené jako součin ne nutně různých prvočísel; definujeme

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)\cdots\left(\frac{a}{p_k}\right).$$

Počítání Jacobiho symbolu je jednoduché s následujícími pravidly:

- $(\frac{ab}{n}) = (\frac{a}{n})(\frac{b}{n})$,
- $a \equiv b \pmod{n} \Rightarrow (\frac{a}{n}) = (\frac{b}{n})$,
- $(\frac{2}{n}) = (-1)^{\frac{n^2-1}{8}}$,
- $(\frac{m}{n}) = (\frac{n}{m}) \cdot (-1)^{\frac{m-1}{2}\frac{n-1}{2}}$ pro m, n lichá.

Důkaz. Toto plyne z vlastností Legendreova symbolu poměrně přímočarým překladem (využívající identity typu $(m^2-1)+(n^2-1) \equiv (mn)^2 - 1 \pmod{16}$ pro m, n lichá). □

Příklad. Spočtěte $(\frac{219}{383})$.

Příklad. Spočtěte $(\frac{3}{p})$ pro prvočíslo p, tj. rozhodněte, kdy je 3 kvadratický zbytek modulo prvočíslo p.

Testování prvočíselnosti

Test (test)**.** *Je-li p prvočíslo, pak $a^{p-1} \equiv 1 \pmod{p}$ (pro a náhodné nesoudělné s p).*

Test (složitější test)**.** *Je-li p prvočíslo, pak $a^{\frac{p-1}{2}} \equiv \pm 1 \pmod{p}$ (pro a náhodné nesoudělné s p).*

Test (ještě složitější test)**.** *Je-li p prvočíslo, pak $a^{\frac{p-1}{2}} \equiv (\frac{a}{p}) \pmod{p}$ (pro a náhodné nesoudělné s p).*

12 Výpočetní aspekty teorie čísel

Základní úlohy výpočetní teorie čísel

V mnoha praktických úlohách využívajících výsledky teorie čísel je zapotřebí umět rychle provést jeden či více z následujících výpočtů:

1. běžné aritmetické operace (součet, součin, dělení se zbytkem) na celých číslech,
2. zbytek mocniny celého čísla a na přirozené číslo n po dělení daným m.
3. inverzi celého čísla a modulo $m \in \mathbb{N}$,
4. největší společný dělitel dvou celých čísel (a případně koeficienty do Bezoutovy rovnosti),
5. rozhodnout o daném čísle, je-li prvočíslo nebo složené,
6. v případě složenosti rozložit dané číslo na součin prvočísel.

Základní aritmetické operace

Základní aritmetické operace se i na velkých číslech obvykle provádějí obdobně jako jsme se to učili na základní a střední škole, kdy umíme sčítat v *lineárním*, násobit a dělit se zbytkem v *kvadratickém* čase. Pro **násobení**, které je základem mnoha dalších operací, existují asymptoticky rychlejší algoritmy (typu *rozděl a panuj*) - např. první takový Karatsubův (1960) časové náročnosti $\Theta(n^{\log_2 3})$ nebo algoritmus Schönhage-Strassenův (1971) časové náročnosti $\Theta(n \log n \log\log n)$, který využívá tzv. Fast Fourier Transform. Ten je ale přes svou asymptotickou převahu výhodný až pro násobení čísel majících alespoň desítky tisíc cifer (a používá se tak např. v GIMPS).

Pěkný přehled je např. na `http://en.wikipedia.org/wiki/Computational_complexity_of_mathematical_operations`

GCD a modulární inverze

Jak už jsme ukazovali dříve, výpočet řešení kongruence $a \cdot x \equiv 1 \pmod{m}$ s neznámou x lze snadno (díky Bezoutově větě) převést na výpočet největšího společného dělitele čísel a a m a na hledání koeficientů k, l do Bezoutovy rovnosti $k \cdot a + l \cdot m = 1$ (nalezené k je pak onou hledanou inverzí a modulo m).

```
function extended_gcd(a, m)
    if m == 0
        return (1, 0)
    else
        (q, r) := divide (a, m)
        (k, l) := extended_gcd(m, r)
        return (l, k - q * l)
```

Podrobná analýza (viz např. `[Knuth] nebo [Wiki]`) ukazuje, že tento algoritmus je **kvadratické** časové složitosti.

Modulární umocňování

Modulární umocňování je, jak jsme již viděli dříve, velmi využívaná operace mj. při ověřování, zda je dané číslo prvočíslo nebo číslo složené. Jedním z efektivních algoritmů je tzv. **modulární umocňování zprava doleva**:

```
function modular_pow(base, exponent, modulus)
    result := 1
    while exponent > 0
      if (exponent mod 2 == 1):
        result := (result * base) mod modulus
      exponent := exponent >> 1
      base = (base * base) mod modulus
    return result
```

Algoritmus modulárního umocňování je založen na myšlence, že např. při počítání $2^{64} \pmod{1000}$

- není třeba nejprve počítat 2^{64} a poté jej vydělit se zbytkem číslem 1000, ale lépe je postupně násobit „dvojky“ a kdykoliv je výsledek větší než 1000, provést redukci modulo 1000,
- ale zejména, že není třeba provádět takové množství násobení (v tomto případě 63 naivních násobení je možné nahradit pouze šesti umocněními na druhou, neboť

$$2^{64} = (((((2^2)^2)^2)^2)^2)^2.$$

Příklad (Ukázka průběhu algoritmu). Vypočtěme $2^{560} \pmod{561}$. Protože $560 = (1000110000)_2$, dostaneme uvedeným algoritmem

exponent	base	result	exp's last digit
560	2	1	0
280	4	1	0
140	16	1	0
70	256	1	0
35	460	1	1
17	103	460	1
8	511	256	0
4	256	256	0
2	460	256	0
1	103	256	1
0	511	1	0

A tedy $2^{560} \equiv 1 \pmod{561}$.

Efektivita modulárního umocňování

V průběhu algoritmu se pro každou binární číslici exponentu provede umocnění základu na druhou modulo n (což je operace proveditelná v nejhůře kvadratickém čase), a pro každou „jedničku“ v binárním zápisu navíc provede jedno násobení. Celkově jsme tedy schopni provést modulární umocňování nejhůře v **kubickém** čase.

13 Diofantické rovnice

Příklad. Vyřešte diofantickou rovnici

$$72x + 100y = 16.$$

Příklad. Vyřešte diofantickou rovnici

$$72x + 100y + 45z = 1.$$

14 Kryptografie s veřejným klíčem

Kryptografie s veřejným klíčem (PKC)

Dva hlavní úkoly pro PKC jsou zajistit

- šifrování, kdy zprávu **zašifrovanou** veřejným klíčem příjemce není schopen rozšifrovat nikdo kromě něj (resp. držitele jeho soukromého klíče)
- podepisování, kdy integrita zprávy **podepsané** soukromým klíčem odesílatele může být ověřena kýmkoliv s přístupem k veřejnému klíči odesílatele

Nejčastěji používané systémy PKC:

- RSA (šifrování) a odvozený systém pro podepisování zpráv
- Digital signature algorithm (DSA) a varianta založená na eliptických křivkách (ECDSA)
- Rabinův kryptosystém (a podepisování)
- ElGamal kryptosystém (a podepisování)
- Kryptografie eliptických křivek (ECC)
- Diffie-Hellmanův protokol na výměnu klíčů (DH)

RSA

Ron Rivest, Adi Shamir, Leonard Adleman (1977; C. Cocks, GCHQ – 1973)

- každý účastník A potřebuje dvojici klíčů – veřejný V_A a soukromý S_A
- generování klíčů: zvolí se dvě velká prvočísla p, q, vypočte se $n = pq$, $\phi(n) = (p-1)(q-1)$, dále se zvolí e a ověří, že $(e, \phi(n)) = 1$, např. pomocí Euklidova algoritmu se spočítá d tak, aby $e \cdot d \equiv 1 \pmod{\phi(n)}$
- $V_A = (n, e)$, $S_A = d$
- zašifrování numerického *kódu* zprávy M: $C \equiv V_A(M) \equiv M^e \pmod{n}$
- dešifrování šifry C: $M \equiv S_A(C) \equiv C^d \pmod{n}$

Rabinův kryptosystém

Prvním veřejným kryptosystémem, k jehož prolomení je prokazatelně potřeba faktorizovat modul, je *Rabinův kryptosystém*, který si uvedeme ve zjednodušené verzi:

- každý účastník A potřebuje dvojici klíčů – veřejný V_A a soukromý S_A
- generování klíčů: A zvolí dvě podobně velká prvočísla $p, q \equiv 3 \pmod 4$, vypočte $n = pq$.
- $V_A = n$, $S_A = (p, q)$
- zašifrování numerického *kódu* zprávy M: $C = V_A(M) \equiv M^2 \pmod n$
- dešifrování šifry C: vypočtou se (čtyři) odmocniny z C modulo n a snadno se otestuje, která z nich byla původní zprávou.

Poznámka. Výpočet druhé odmocniny z C modulo $n = pq$, kde $p \equiv q \equiv 3 \pmod 4$

- vypočti odmocniny modulo p a modulo q, konkrétně $r \equiv \pm C^{(p+1)/4} \pmod p$ a $s \equiv \pm C^{(q+1)/4} \pmod q$
- pomocí Čínské zbytkové věty spočti pro každou kombinaci odmocnin modulo p a modulo q odpovídající odmocninu modulo $n = pq$

Příklad. V Rabinově kryptosystému Alice zvolila za svůj soukromý klíč $p = 23, q = 31$, veřejným klíčem je pak $n = pq = 713$. Zašifrujte zprávu $M = 327$ pro Alici a ukažte, jak bude Alice tuto zprávu dešifrovat.

Řešení. $C = 692$, kandidáti původní zprávy jsou $\pm 138 \pm 248 \pmod{713}$.

Princip digitálního podpisu

Poznámka. Podepisování

1. Vygeneruje se otisk (hash) H_M zprávy pevně stanovené délky (např. 160 nebo 256 bitů).
2. Podpis zprávy $S_A(H_M)$ je vytvořen (pomocí dešifrování) z tohoto hashe s nutností znalosti soukromého klíče podepisujícího.
3. Zpráva M (případně zašifrovaná veřejným klíčem příjemce) je spolu s podpisem odeslána.

Poznámka. Ověření podpisu

1. K přijaté zprávě M se (po jejím případném dešifrování) vygeneruje otisk H'_M
2. S pomocí veřejného klíče (deklarovaného) odesílatele zprávy se rekonstruuje původní otisk zprávy $V_A(S_A(H_M)) = H_M$.
3. Oba otisky se porovnají $H_M = H'_M$?.

Diffie-Hellman key exchange

Whitfield Diffie, Martin Hellman (1976; M. Williamson, GCHQ - 1974)

Výměna klíčů pro symetrickou kryptografii bez předchozího kontaktu (tj. náhrada jednorázových klíčů, kurýrů s kufříky, ...).

- Dohoda stran na **prvočísle** p a primitivním kořenu g modulo p (veřejné)
- Alice vybere náhodné a a pošle $g^a \pmod p$
- Bob vybere náhodné b a pošle $g^b \pmod p$
- Společným klíčem pro komunikaci je $g^{ab} \pmod p$.

Poznámka.
- Problém diskrétního logaritmu (DLP)
- Nezbytná autentizace (*man in the middle attack*)

Kryptosystém ElGamal

Z protokolu DH na výměnu klíčů odvozen šifrovací algoritmus ElGamal:

- Alice zvolí prvočíslo p spolu s primitivním kořenem g
- Alice zvolí a a spočítá $h \equiv g^a \pmod{p}$
- $V_A = (p, g, h)$, $S_A = a$
- šifrování zprávy M: Bob zvolí náhodné b a vypočte $C_1 \equiv g^b \pmod{p}$ a $C_2 \equiv M \cdot h^b \pmod{p}$ a pošle $C = (C_1, C_2)$
- dešifrování zprávy: $M \equiv C_2/C_1^a \pmod{p}$

Poznámka. Analogicky jako v případě RSA lze odvodit podepisování.

Eliptické křivky

Eliptické křivky jsou rovinné křivky o rovnici tvaru $y^2 = x^3 + ax + b$ a zajímavé jsou tím, že na jejich bodech lze definovat operace tak, že výslednou strukturou bude komutativní grupa.

Přitom uvedené operace lze efektivně provádět a navíc se ukazuje, že mají (nejen) pro kryprografii zajímavé vlastnosti – srovnatelné bezpečnosti jako RSA lze dosáhnout již s podstatně kratšími klíči. Výhodou je rovněž velké množství použitelných eliptických křivek (a tedy grup různé struktury) podle volby parametru a, b .

Protokoly:

- ECDH - přímá varianta DH na eliptické křívce (jen místo generátoru se vybere *vhodný* bod na křivce)
- ECDSA - digitální podpis pomocí eliptických křivek.

Poznámka. Problém diskrétního logaritmu (ECDLP).

Navíc se ukazuje, že eliptické křivky jsou velmi dobře použitelné při faktorizaci prvočísel.

15 (n, k)–kódy

Při přenosu informace zpravidla dochází k její deformaci. Budeme pro jednoduchost pracovat s modelem, kdy jednotlivé částečky informace jsou buď nuly nebo jedničky a budeme je interpretovat jako prvky $\mathbb{Z}_2$, tj. jako zbytkové třídy modulo 2, přičemž podstatné pro nás bude, že s nimi můžeme provádět algebraické manipulace: sčítat a násobit. Prvky $\mathbb{Z}_2$ budeme pro jednoduchost značit pomocí reprezentatů, tj. 0 a 1 namísto přesnějšího [0] a [1]. Protože operace splňují axiomy „tělesa“, můžeme uvažovat vektorové prostory se skaláry v $\mathbb{Z}_2$, budeme však využívat pouze dva příklady: standardní (sloupcové) vektory jsou prvky kartézské mocniny $(\mathbb{Z}_2)^k$ a polynomy jsou prvky $\mathbb{Z}_2[x]$. Sloupcové vektory sčítáme po složkách, přičemž je dobré mít na paměti, že v $\mathbb{Z}_2$ platí $1 + 1 = 0$. To stejné platí pro polynomy – při jejich sčítání sčítáme odpovídající si koeficienty a opět využíváme $1 + 1 = 0$. Ještě jedna ekvivalentní interpretace tohoto pravidla se nám bude hodit: $-v = v$. (Násobení skaláry není moc zajímavé, protože násobení 0 i násobení 1 je vynucené axiomy vektorového prostoru.)

Přenášíme slova o k bitech. Protože přenosové chyby chceme rozpoznávat a ideálně i opravovat, přidáváme za tím účelem dodatečných $n - k$ bitů informace pro pevně zvolené $n > k$. Mluvíme pak o (n, k)-*kódu*.

Všech slov o k bitech je 2^k a každé z nich má jednoznačně určovat jedno *kódové slovo* z 2^n možných. Máme tedy ještě

$$2^n - 2^k = 2^k(2^{n-k} - 1)$$

slov, které jsou chybové. Lze tedy tušit, že pro veliké k nám i malý počet přidaných bitů dává hodně redundantní informace.

Úplně jednoduchým příkladem je *kód kontrolující paritu*. Kódové slovo o $k+1$ bitech je určené tak, aby přidáním prvního bitu byl zaručen sudý počet jedniček ve slově.

Pokud při přenosu dojde k lichému počtu chyb, přijdeme na to. Dvě různá kódová slova se při tomto kódu vždy liší alespoň ve dvou pozicích, chybové slovo se ale od dvou různých kódových slov liší pouze v pozici jedné. Nemůžeme proto umět chyby opravovat ani kdybychom věděli, že došlo k právě jedné.

Navíc neumíme detekovat tak obvyklé chyby, jako je záměna dvou sousedních hodnot ve slově.

Definice. *Hammingova vzdálenost* dvou slov je rovna počtu bitů, ve kterých se liší.

Věta 28. *1. Kód odhaluje r a méně chyb právě, když je minimální Hammingova vzdálenost kódových slov právě $r+1$.*

2. Kód opravuje r a méně chyb právě, když je minimální Hammingova vzdálenost kódových slov právě $2r+1$.

16 Lineární kódy

Definice. *Lineární kód* je injektivní lineární zobrazení $g : (\mathbb{Z}_2)^k \to (\mathbb{Z}_2)^n$. Matice G typu n/k reprezentující toto zobrazení v standardních bazích se nazývá generující *matice kódu*.

Pro každé slovo u je

$$v = G \cdot u$$

příslušné kódové slovo. Pro jednoduchost budeme předpokládat, že kód přidává dodatečnou informaci a pro konkrétnost tedy, že v má blokový tvar $v = \binom{Pu}{u}$, kde Pu je ona dodatečná informace a u je původní vektor. Proto matice G bude mít blokový tvar

$$G = \begin{pmatrix} P \\ \mathbb{I}_k \end{pmatrix}.$$

Věta 29. *Je-li $g : (\mathbb{Z}_2)^k \to (\mathbb{Z}_2)^n$ lineární kód s maticí jako výše, potom zobrazení $h : (\mathbb{Z}_2)^n \to (\mathbb{Z}_2)^{n-k}$ s maticí*

$$H = \begin{pmatrix} \mathbb{I}_{n-k} & P \end{pmatrix}$$

má následující vlastnost: slovo v je kódové, právě když $H \cdot v = 0$.

Důkaz. Slovo $v = \binom{x}{y}$ je kódové, právě když $v = Gy = \binom{Py}{y}$, protože v takovém případě jsme schopni původní slovo y dekódovat z informačních bitů. Slovo v je tedy kódové, právě když $x = Py$, což je přesně podmínka $H\binom{x}{y} = 0$. □

Matici H z věty se říká *matice kontroly parity* přílušného (n,k)–kódu, součinu Hv říkáme *syndrom* slova v. Nyní vysvětlíme význam syndromu i v případě, že je nenulový. Pokud neplatí $Hv = 0$, chceme přičtením co nejmenšího chybového vektoru k v dosáhnout toho, aby podmínka $Hv = 0$ platila. Přitom přičtení jedničky na i-tém místě způsobí změnu hodnoty přesně o i-tý sloupec kontrolní matice H. Chceme tedy hodnotu Hv vynulovat co nejmenším počtem sloupců matice H, přičemž v levé submatici máme právě sloupce s jednou jedničkou a v pravé právě kontrolní bity sloupců matice kódu. Kdybychom využívali pouze levou submatici (opravovali pouze kontrolní bity), potřebujeme opravit přesně tolik bitů, kolik obsahuje Hv jedniček (navíc přesně na stejných pozicích), syndrom tedy udává přesně tu jedinou možnou chybu, ke které mohlo dojít čistě na kontrolních bitech. Při použití informačních bitů můžeme počet chyb snížit, je tedy sice Hv jistá míra chybovosti, ale ne minimální.

Naopak, nechť h je lineární zobrazení, jehož matice H je tvaru jako ve větě. Pak můžeme sestrojit lineární kód s maticí G jako před zněním věty a podle věty pak H bude matice kontroly parity. Jsme tedy schopni lineární kód zadat jeho (lineární) kontrolou parity. Jednoduše to lze ilustrovat na klasické kontrole parity, kde $h(x_0, \ldots, x_k) = x_0 + \cdots + x_k$ je zjevně lineární s maticí

$$H = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix},$$

kýženého tvaru. Nemusíme tedy specifikovat matici kódu, tu lze z matice kontroly parity odvodit. V další části uvedeme konstrukci polynomiálních kódů skrze jejich matici kontroly parity.

Poznámka. Jak jsme viděli, přenos zprávy Gu s chybou e dává výsledek

$$v = Gu + e,$$

kde ale neznáme u, e a hledáme takový "rozklad", kde e obsahuje co nejméně jedniček (oprava chyby za předpokladu co nejmenšího počtu chyb).

Je-li $v = \binom{x}{y}$, pak jednou z možností na odeslanou zprávu je $Gy = \binom{Py}{y}$ (ne nutně optimální), tedy

$$\binom{x}{y} = \binom{Py}{y} + \binom{x+Py}{0}$$

kde $s = x + Py = (\mathbb{I}_{n-k}\ P)\binom{x}{y} = Hv$ je právě syndrom slova v a jedná se tedy o chybu za předpokladu, že k ní došlo pouze na kontrolních bitech (z informačních bitů lze proto přečíst původní slovo).

Protože kódová slova jsou právě součty sloupců matice G, lze všechny další možnosti obdržet přičítáním sloupců g_i ke kódovému slovu i chybě, buď přičítáním jednoho sloupce:

$$\binom{x}{y} = \left(\binom{Py}{y} + g_i\right) + \left(\binom{x+Py}{0} + g_i\right)$$

případně dvou sloupců, atd.

Přitom přičtení každého sloupce vyrobí jednu jedničku v informačních bitech chyby, snažíme se jejich počet kompenzovat snížením počtu jedniček v kontrolních bitech. Toto budeme zkoušet pouze pro malý počet chyb, viz cvičení.

17 Polynomiální kódy

Jak konstruovat kódová slova, abychom je snadno rozpoznali? Kontrolu parity jsme už viděli, další triviální možnost je prosté opakování bitů – např. $(3,1)$–kód bere jednotlivé bity a posílá je třikrát po sobě.

Docela systematickou cestou je využití dělitelnosti polynomů. Zpráva $b_0 b_1 \dots b_{k-1}$ je reprezentována jako polynom $m(x) = b_0 + b_1 x + \cdots + b_{k-1}x^{k-1} \in \mathbb{Z}_2[x]$.

Definice. Nechť $p(x) = a_0 + \cdots + a_{n-k}x^{n-k} \in \mathbb{Z}_2[x]$ je polynom s $a_0 = 1$, $a_{n-k} = 1$. *Polynomiální kód generovaný polynomem* $p(x)$ je lineární (n,k)–kód, jehož slova jsou polynomy stupně menšího než n dělitelné $p(x)$ a jehož kontrola parity je lineární zobrazení h posílající polynom na jeho zbytek po dělení $p(x)$.

Zobrazení h je opravdu lineární (zbytek součtu je součet zbytků). Polynomy stupně menšího než $n-k$ jsou automaticky zbytkem po dělení $p(x)$, takže je na nich h identita a matice H je tedy tvaru

$$H = \begin{pmatrix} \mathbb{I}_{n-k} & P \end{pmatrix},$$

jak je potřeba.

Nastíníme nyní, jak matici P sestavit. Její první sloupec je zbytek x^{n-k} a sestává se tedy z koeficientů $p(x)$ stupně menšího než $n-k$. Další sloupec je zbytkem $x^{n-k+1} = x \cdot x^{n-k}$ a získáme jej tedy z předchozího sloupce vynásobením x, přičemž případný výskyt x^{n-k} nahradíme jeho zbytkem, tedy prvním sloupcem P. Konkrétně tedy sloupec posuneme dolů a případná jednička se při přetečení nahradí přičtením prvního sloupce. Takto postupujeme i pro další sloupce – posouváme předchozí, při přetečení přičítáme první!

Příklad.
1. Polynom $p(x) = 1 + x$ generuje $(n, n-1)$–kód kontroly parity pro všechna $n \geq 3$.
2. Polynom $p(x) = 1 + x + x^2$ generuje $(3,1)$–kód opakování bitů.

První tvrzení plyne z toho, že $1 + x$ dělí polynom $v(x)$ tehdy a jen tehdy, když $v(1) = 0$ a to nastane tehdy, když je ve $v(x)$ sudý počet nenulových koeficientů. Druhé je zřejmé.

Matice příslušná k polynomu $p(x) = 1 + x + x^3$ a jím určenému $(7, 4)$–kódu je

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Přenos slova $v \in \mathbb{Z}_2[x]$ dopadne příjmem polynomu

$$u(x) = v(x) + e(x)$$

kde $e(x)$ je tzv. *chybový polynom* reprezentující vektor chyby přenosu.

Chyba je rozpoznatelná pouze, když generátor kódu $p(x)$ nedělí $e(x)$. Máme proto zájem o polynomy, které nevystupují jako dělitelé zbytečně často. Připomeňme, že matice kontroly parity obsahuje ve sloupcích zbytky po dělení x^i polynomem $p(x)$. Pokud chceme, aby kód rozpoznal jednoduché chyby, nesmí matice kontroly parity obsahovat žádný nulový sloupec – to totiž odpovídá $p(x) \mid x^i$. Pokud chceme, aby kód rozpoznal dvojité chyby, nesmí matice kontroly parity obsahovat žádný sloupec dvakrát – to totiž odpovídá $p(x) \mid x^i + x^j$. Při počtu řádků $m = n - k$, tak může P obsahovat maximálně $2^m - 1$ sloupců.

Definice. Ireducibilní polynom $p(x) \in \mathbb{Z}_2[x]$ stupně m se nazývá *primitivní*, jestliže $p(x) \nmid (1+x^\ell)$ pro $\ell < 2^m - 1$, a teprve $p(x) \mid (1 + x^\ell)$ pro $\ell = 2^m - 1$.

Věta 30. *Je-li $p(x)$ primitivní polynom stupně m, pak pro všechna $n \leq 2^m - 1$ rozpoznává příslušný (n, k)–kód všechny jednoduché a dvojité chyby.* □

Důsledek. *Je-li $q(x)$ primitivní polynom stupně m, pak pro všechna $n \leq 2^m - 1$ rozpoznává (n, k)–kód generovaný polynomem $p(x) = q(x)(1 + x)$ všechny dvojité chyby a všechna slova s lichým počtem chyb.* □

Tabulka dává informace o výsledcích předchozích dvou vět pro několik polynomů:

primitivní polynom	kontrolní bity	délka slova
$1 + x + x^2$	2	3
$1 + x + x^3$	3	7
$1 + x + x^4$	4	15
$1 + x^2 + x^5$	5	31
$1 + x + x^6$	6	63
$1 + x^3 + x^7$	7	127
$1 + x^2 + x^3 + x^4 + x^8$	8	255
$1 + x^4 + x^9$	9	511
$1 + x^3 + x^{10}$	10	1023

Nástroje pro konstrukci primitivních polynomů dává teorie konečných polí. Souvisí s tzv. primitivními prvky v Galoisových polích $G(2^m)$.

Ze stejné teorie lze také dovodit příjemnou realizaci dělení se zbytkem, tj. ověřování, zda je přijaté slovo kódové, pomocí zpožďovacích registrů. Jde o jednoduchý obvod s tolika prvky, kolik je stupeň polynomu.

18 Kombinatorika – motivace

Kombinatorika = umění počítat

Často potřebujeme umět spočítat, kolika možnými způsoby se něco může stát!

Nebývá to jednoduché a naším cílem nyní bude vybudovat základní prostředky pro řešení úloh obdobných těmto:

Analýza algoritmů: Např. chceme zjistit očekávaný počet porovnání během algoritmu `Quicksort`.

Odvození Cayleyho formule: Chceme znát počet různých stromů na daných n vrcholech.

Quicksort – analýza průměrného případu

Ukázka implementace (*divide and conquer*):

```
if L == []: return []
return qsort([x for x in L[1:] if x < L[0]])
     + L[0:1]
     + qsort([x for x in L[1:] if x >= L[0]])
```

1. Počet porovnání při rozdělení (*divide*): $n-1$.
2. (Předpoklad náhodnosti): Pravděpodobnost toho, že prvek $L[0]$ je k-tý největší, je $\frac{1}{n}$.
3. Velikost tříděných polí ve fázi *conquer*: $k-1$ a $n-k$.

Pro střední hodnotu počtu porovnání tak dostáváme rekurentní vztah:

$$C_n = n - 1 + \sum_{k=1}^{n} \frac{1}{n} \left(C_{k-1} + C_{n-k}\right).$$

Zjednodušení rekurence

$$C_n = n - 1 + \sum_{k=1}^{n} \frac{1}{n} \left(C_{k-1} + C_{n-k}\right),\ C_0 = 0.$$

$$C_n = n - 1 + \frac{2}{n} \sum_{k=1}^{n} C_{k-1} \qquad \textit{symetrie obou sum}$$

$$nC_n = n(n-1) + 2\sum_{k=1}^{n} C_{k-1} \qquad \textit{vynásob } n$$

$$(n-1)C_{n-1} = (n-1)(n-2) + 2\sum_{k=1}^{n-1} C_{k-1} \qquad \textit{tentýž výraz pro } C_{n-1}$$

$$nC_n = (n+1)C_{n-1} + 2(n-1) \qquad \textit{odečteno a upraveno}$$

Vyřešení rekurence

$$nC_n = (n+1)C_{n-1} + 2(n-1)$$

Přestože jsme již rekurenci výrazně zjednodušili, takže je možné jednoduše iterativně hodnoty C_n dopočítat, je často žádoucí tyto hodnoty konkrétně (nebo alespoň přibližně) vyjádřit explicitně jako funkci n.

Nejprve si pomůžeme drobným trikem, kdy vydělíme obě strany výrazem $n(n+1)$:

$$\frac{C_n}{n+1} = \frac{C_{n-1}}{n} + \frac{2(n-1)}{n(n+1)}$$

Nyní tento vztah „rozbalíme“ (*telescope*, příp. si pomůžeme substitucí $B_n = C_n/n+1$):

$$\frac{C_n}{n+1} = \frac{2(n-1)}{n(n+1)} + \frac{2(n-2)}{(n-1)n} + \cdots + \frac{2\cdot 1}{2\cdot 3} + \frac{C_1}{2}$$

Odkud

$$\frac{C_n}{n+1} = 2\sum_{k=1}^{n-1} \frac{k}{(k+1)(k+2)}.$$

Výraz sečteme např. pomocí rozkladu na parciální zlomky $\frac{k}{(k+1)(k+2)} = \frac{2}{k+2} - \frac{1}{k+1}$ a dostaneme

$$\frac{C_n}{n+1} = 2\left(H_{n+1} - 2 + \frac{1}{n+1}\right),$$

odkud

$$C_n = 2(n+1)H_{n+1} - 4(n+1) + 2$$

($H_n = \sum_{k=1}^{n} \frac{1}{k}$ je součet prvních n členů harmonické řady). Přitom je možné odhadnout $H_n \sim \int_1^n \frac{\mathrm{d}x}{x} + \gamma = \ln n + \gamma$, odkud

$$C_n \sim 2(n+1)(\ln(n+1) + \gamma - 2) + 2.$$

19 Elementární kombinatorické metody

Pravidlo součtu a součinu

Vylučující se možnosti sčítáme, vzájemně nezávislé a současně se vyskytující případy se násobí.

Na dané konečné množině S s n prvky je právě $n!$ různých pořadí.

Počet *kombinací k-tého stupně* z n prvků je ($k \leq n$)

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

Pro počet variací platí

$$v(n,k) = n(n-1)\cdots(n-k+1)$$

pro všechny $0 \leq k \leq n$ (a nula jinak).

Příklad. Určete, kolika způsoby lze z 15 poslanců vybrat čtyřčlennou komisi, není-li možné, aby jistí 2 poslanci pracovali spolu.

Řešení. Výsledek je $\binom{15}{4} - \binom{13}{2} = 1287$.

Kombinace a variace s opakováním

Nechť je mezi n danými prvky p_1 prvků prvního druhu, p_2 prvků druhého druhu, …, p_k prvků k-tého druhu, $p_1+p_2+\cdots+p_k = n$, potom pro počet pořadí těchto prvků s opakováním, $P(p_1,\dots,p_k)$, platí

$$P(p_1,\dots,p_k) = \frac{n!}{p_1!\cdots p_k!}.$$

Pro *variace k-tého stupně s opakováním* z n prvků platí

$$V(n,k) = n^k.$$

Pro *kombinace s opakováním*, $C(n,k)$, platí

Věta 31. *Počet kombinací s opakováním k-té třídy z n prvků je pro všechna* $0 \leq k$ *a* $0 < n$

$$C(n,k) = \binom{n+k-1}{k}.$$

Princip inkluze a exkluze

Poznámka. Uvažujme obecnou konečnou množinu M a její podmnožiny $A_1, \ldots, A_k$. Budeme psát $|M|$ pro počet prvků množiny M, tj. pro *mohutnost* množiny M.

$$|M \setminus (\cup_{i=1}^{k} A_i)| = |M| + \sum_{j=1}^{k} \left((-1)^i \sum_{1 \le i_1 < \cdots < i_j \le k} |A_{i_1} \cap \cdots \cap A_{i_j})| \right).$$

Příklady kombinatorických rovností

Dokážeme (pokud možno kombinatorickou úvahou):

Aritmetická řada $$\sum_{k=0}^{n} k = \frac{n(n+1)}{2} = \binom{n+1}{2}$$

Geometrická řada $$\sum_{k=0}^{n} x^k = \frac{x^{n+1}-1}{x-1}$$

Binomická věta $$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Horní binomická řada $$\sum_{k=0}^{n} \binom{k}{m} = \binom{n+1}{m+1}$$

Vandermondova konvoluce $$\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k} \binom{n}{r-k}$$

Odkazovací strategie

Ve vězení je 100 vězňů s čísly jedna až sto. V uzavřené místnosti je 100 krabiček s čísly jedna až sto a v každé z nich náhodně rozdělené papírky s čísly také 1 až sto. Do místnosti budou po jednom postupně vcházet vězni a každý smí otevřít 50 krabic, vězni přitom spolu nekomunikují. Jestliže každý z nich najde svoje číslo, všechny pustí, v opačném případě všechny popraví.

Doporučte nějakou rozumnou strategii pro vězně ...

20 Vytvořující funkce

Motivace

Motto: *spojité a diskrétní modely se vzájemně potřebují a doplňují.*

Příklad. Máme v peněžence 4 korunové mince, 5 dvoukorunových a 3 pětikorunové. Z automatu, který nevrací, chceme minerálku za 22 Kč. Kolika způsoby to umíme, aniž bychom ztratili přeplatek?

Hledáme zjevně čísla i, j a k taková, že $i + j + k = 22$ a zároveň

$$i \in \{0,1,2,3,4\},\ j \in \{0,2,4,6,8,10\},\ k \in \{0,5,10,15\}.$$

Uvažme součin polynomů (třeba nad reálnými čísly)

$$(x^0 + x^1 + x^2 + x^3 + x^4)(x^0 + x^2 + x^4 + x^6 + x^8 + x^{10})(x^0 + x^5 + x^{10} + x^{15}).$$

Mělo by být zřejmé, že hledaný počet řešení je díky (Cauchyovskému) způsobu násobení polynomů právě koeficient u x^{22} ve výsledném polynomu.

Skutečně tak dostáváme **čtyři možnosti** $3\times5+3\times2+1\times1$, $3\times5+2\times2+3\times1$, $2\times5+5\times2+2\times1$ a $2 \times 5 + 4 \times 2 + 4 \times 1$.

Předchozí příklad asi vypadal spíš jako složitý zápis jednoduchých „backtrackingových úvah“. Následující příklad ukazuje, že tento postup lze ale s výhodou zobecnit.

Nechť I, J jsou konečné množiny nezáporných celých čísel. Potom je pro dané $r \in \mathbb{N}$ počet řešení (i, j) rovnice $i + j = r$ splňujících $i \in I, j \in J$ roven koeficientu u x^r v polynomu $(\sum_{i\in I} x^i)(\sum_{j\in J} x^j)$.

Příklad. Kolika způsoby můžeme pomocí mincí (1, 2, 5, 10, 20 a 50 Kč) zaplatit platbu 100 Kč?

Hledáme přirozená čísla $a_1, a_2, a_5, a_{10}, a_{20}$ a a_{50} taková, že a_i je násobkem i pro všechna $i \in \{1, 2, 5, 10, 20, 50\}$ a zároveň $a_1 + a_2 + a_5 + a_{10} + a_{20} + a_{50} = 100$.

Podobně jako výše je vidět, že požadovaný počet lze získat jako koeficient u x^{100} v

$$(1 + x + x^2 + \dots)(1 + x^2 + x^4 + \dots)(1 + x^5 + x^{10} + \dots)$$
$$(1 + x^{10} + x^{20} + \dots)(1 + x^{20} + x^{40} + \dots)(1 + x^{50} + x^{100} + \dots)$$

Drobný rozdíl je v tom, že se nyní nejedná o polynom, ale nekonečnou řadu; přitom ale není problém všechny výpočty končit u x^{100}, čímž se opět dostaneme k polynomům (nicméně nekonečná geometrická řada má jednodušší vzorec pro součet!).

Příklad. V krabici je 5 červených, 10 modrých a 15 bílých míčků, míčky stejné barvy přitom nelze rozeznat. Kolika způsoby je možné vybrat soubor 7 míčků k vyzkoušení? A o kolik míň to bude, když chceme aspoň 1 červený, aspoň 2 modré a aspoň 3 bílé?

Řešení. Hledaný počet je roven koeficientu u x^7 v součinu

$$(1 + x + x^2 + \cdots + x^5)(1 + x + x^2 + \cdots + x^{10})(1 + x + x^2 + \cdots + x^{15}).$$

Když máme předepsaný nějaký počet jako nejmenší možný, prostě začneme až od příslušných mocnin.

Kombinatorické vztahy

Využitím operací s polynomy lze velmi snadno odvodit také některé kombinatorické vztahy, které známe již z dřívějška. Využijeme přitom **binomickou větu.**

Věta 32 (binomická)**.** *Pro $n \in \mathbb{N}$ a $r \in \mathbb{R}$ platí*

$$\sum_{k=0}^{n} \binom{n}{k} x^k = (1 + x)^n.$$

Na pravou stranu se můžeme dívat jako na součin n polynomů, levá je zápisem polynomu vzniklého jejich roznásobením. Dosazením čísel $x = 1$, resp. $x = -1$ dostáváme známé vzorce:

Důsledek.
- $\sum_{k=0}^{n} \binom{n}{k} = 2^n$,
- $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$.

Podíváme se teď na obě strany v binomické větě „spojitýma očima“ a s využitím vlastností derivací odvodíme další vztah mezi kombinačními čísly.

Důsledek. *Platí*

$$\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}.$$

Důkaz. Na obě strany binomické věty se podíváme jako na polynomiální funkce. Derivací pravé strany dostaneme $n(1+x)^{n-1}$, derivací levé strany (člen po členu) pak $\sum_{k=1}^{n} k\binom{n}{k}x^{k-1}$. Dosazením $x = 1$ dostaneme tvrzení. □

21 (Formální) mocninné řady

(Formální) mocninné řady

Definice. Buď dána nekonečná posloupnost $a = (a_0, a_1, a_2, \dots)$. Její *vytvořující funkcí* rozumíme (formální) mocninnou řadu tvaru

$$\sum_{k=0}^{\infty} a_k x^k = a_0 + a_1 x + a_2 x^2 + \cdots .$$

Poznámka. O **formální** mocninné řadě hovoříme proto, že se zatím na tuto řadu díváme čistě formálně jako na jiný zápis dané posloupnosti a nezajímáme se o konvergenci. Na druhou stranu to ale znamená, že formální mocninná řada není funkce a nemůžeme do ní dosazovat. To ovšem vzápětí napravíme, když s využitím znalostí z analýzy nekonečných řad přejdeme od formálních mocninnných řad k příslušným funkcím.

Příklad. Posloupnosti samých jedniček odpovídá formální mocninná řada $1 + x + x^2 + x^3 + \cdots$. Z analýzy víme, že stejně zapsaná mocninná řada konverguje pro $x \in (-1, 1)$ a její součet je roven funkci $1/(1 - x)$. Stejně tak obráceně, rozvineme-li tuto funkci do Taylorovy řady v bodě 0, dostaneme zřejmě původní řadu. Takovéto „zakódování“ posloupnosti čísel do funkce a zpět je klíčovým obratem v teorii vytvořujících funkcí.

Jak jsme již zmínili, tento obrat lze ale použít pouze tehdy, pokud víme, že řada alespoň v nějakém okolí 0 konverguje. Často ale „diskrétní“ matematici používají následující „podvod“:

- pomocí formálních mocninných řad odvodí nějaký vztah (formuli, rekurenci,...) bez toho, aby se zajímali o konvergenci
- jinými prostředky (často matematickou indukcí) tento vztah dokážou

Vytvořující funkce v praxi využíváme:

- k nalezení **explicitní formule** pro k-tý člen posloupnosti;
- často vytvořující funkce vycházejí z rekurentních vztahů, občas ale díky nim odvodíme rekurentní vztahy nové;
- výpočet průměrů či jiných statistických závislostí (např. průměrná složitost algoritmu);
- důkaz různých identit;
- často je nalezení přesného vztahu příliš obtížné, ale mnohdy stačí vztah přibližný nebo alespoň asymptotické chování.

Součet formální mocninné řady

Následující větu znáte z matematické analýzy z loňského semestru:

Věta 33. *Buď $(a_0, a_1, a_2, \dots)$ posloupnost reálných čísel. Platí-li pro nějaké $R \in \mathbb{R}$, že pro všechna $k \gg 0$ je $|a_k| \leq R^k$, pak řada*

$$a(x) = \sum_{k \geq 0} a_k x^k$$

konverguje pro každé $x \in (-\frac{1}{R}, \frac{1}{R})$. Součet této řady tedy definuje funkci na uvedeném intervalu, tuto funkci označujeme rovněž $a(x)$.

Hodnotami funkce $a(x)$ na libovolném okolí 0 je jednoznačně určena původní posloupnost, neboť má $a(x)$ v 0 derivace všech řádů a platí

$$a_k = \frac{a^{(k)}(0)}{k!}.$$

Přehled mocninných řad

$$\frac{1}{1-x} = \sum_{k\geq 0} x^k,$$
$$\ln \frac{1}{1-x} = \sum_{k\geq 1} \frac{x^k}{k},$$
$$e^x = \sum_{k\geq 0} \frac{x^k}{k!},$$
$$(1+x)^r = \sum_{k\geq 0} \binom{r}{k} x^k.$$

Poznámka.

- Poslední vzorec
$$(1+x)^r = \sum_{k\geq 0} \binom{r}{k} x^k$$
je tzv. **zobecněná binomická věta**, kde pro $r \in \mathbb{R}$ je binomický koeficient definován vztahem
$$\binom{r}{k} = \frac{r(r-1)(r-2)\cdots(r-k+1)}{k!}.$$
Speciálně klademe $\binom{r}{0} = 1$.

- Pro $n \in \mathbb{N}$ z uvedeného vztahu snadno dostaneme
$$\frac{1}{(1-x)^n} = \sum_{k\geq 0} \binom{k+n-1}{n-1} x^k.$$
To plyne ze vztahu
$$\binom{-n}{k} = (-1)^k \cdot \binom{k+n-1}{n-1},$$
který odvodíte na cvičení.

- Ještě o něco obecněji můžeme substitucí αx za x obdržet obecnější vzorec, vhodný pro počítání konkrétních příkladů
$$\frac{1}{(1-\alpha x)^n} = \sum_{k\geq 0} \binom{k+n-1}{n-1} \alpha^k \cdot x^k.$$

22 Operace s vytvořujícími funkcemi

Některým jednoduchým operacím s posloupnostmi odpovídají jednoduché operace nad mocninnými řadami:

- Sčítání $(a_i + b_i)$ posloupností člen po členu odpovídá součet $a(x) + b(x)$ příslušných vytvořujících funkcí.
- Vynásobení $(\alpha \cdot a_i)$ všech členů posloupnosti stejným skalárem α odpovídá vynásobení $\alpha \cdot a(x)$ příslušné vytvořující funkce.
- Vynásobení vytvořující funkce $a(x)$ monomem x^k odpovídá posunutí posloupnosti doprava o k míst a její doplnění nulami.

- Pro posunutí posloupnosti doleva o k míst (tj. vynechání prvních k míst posloupnosti) nejprve od $a(x)$ odečteme polynom $b_k(x)$ odpovídají posloupnosti $(a_0, \dots, a_{k-1}, 0, \dots)$ a poté podělíme vytvořující funkci x^k.

Dalšími důležitými operacemi, které se při práci s vytvořujícími funkcemi často objevují, jsou:

- Derivování podle x: funkce $a'(x)$ je vytvořující fukcí posloupnosti $(a_1, 2a_2, 3a_3, \dots)$, člen s indexem k je $(k+1)a_{k+1}$ (tj. mocninnou řadu derivujeme člen po členu).
- Integrování: funkce $\int_0^x a(t)\,\mathrm{d}t$ je vytvořující fukcí posloupnosti $(0, a_0, \frac{1}{2}a_1, \frac{1}{3}a_2, \frac{1}{4}a_3, \dots)$, pro $k \geq 1$ je člen s indexem k roven $\frac{1}{k}a_{k-1}$ (zřejmě je derivací příslušné mocninné řady člen po členu původní funkce $a(x)$).
- Násobení řad: součin $a(x)b(x)$ je vytvořující funkcí posloupnosti $(c_0, c_1, c_2, \dots)$, kde

$$c_k = a_0b_k + a_1b_{k-1} + \cdots + a_kb_0 = \sum_{i+j=k} a_ib_j,$$

  tj. členy v součinu až po c_k jsou stejné jako v součinu $(a_0+a_1x+a_2x^2+\cdots+a_kx^k)(b_0+b_1x+b_2x^2+\cdots b_kx^k)$. Posloupnost (c_k) bývá také nazývána *konvolucí* posloupností $(a_k), (b_k)$.

V dalším bude výhodné položit $a_{-1} = 0$, $a_{-2} = 0$, atd. (pak můžeme sčítat přes všechna k), jenom bacha na konkrétní součty typu $\sum_{k\geq 0} x^k = \frac{1}{1-x}$, pro ty je potřeba sčítat pouze přes nezáporná k. Operace s vytvořujícími funkcemi přepíšeme:

- $\sum a_kx^k + \sum b_kx^k = \sum(a_k+b_k)x^k$.
- $\alpha \cdot \sum a_kx^k = \sum(\alpha a_k)x^k$.
- $x^n \cdot \sum a_kx^k = \sum a_{k-n}x^k$.
- $(\sum a_kx^k) \cdot (\sum b_kx^k) = \sum c_kx^k$, kde

$$c_k = \sum_{i+j=k} a_ib_j.$$

Ukažme si důležitý příklad využívající konvoluci posloupností:

Příklad. $\frac{1}{1-x}a(x)$ je v.f.p. $(a_0, a_0+a_1, a_0+a_1+a_2, \dots)$.

Odtud např. dostáváme, že

$$\frac{1}{1-x}\ln\frac{1}{1-x} \quad \text{je v.f.p. harmonických čísel } H_k.$$

Příklad. Protože $\frac{1}{1-x} = \sum_{k\geq 0} x^k$, dostáváme konvolucí posloupnosti $(1, 1, \dots)$ se sebou vztahy

$$\frac{1}{(1-x)^2} = \sum_{k\geq 0}(k+1)x^k, \quad \frac{1}{(1-x)^3} = \sum_{k\geq 0}\binom{k+2}{2}x^k,$$

což máme dokázáno z dřívějška jako důsledek zobecněné binomické věty.

Příklad. V krabici je 30 červených, 40 modrých a 50 bílých míčků, míčky stejné barvy přitom nelze rozeznat. Kolika způsoby je možné vybrat soubor 70 míčků?

Řešení. Hledaný počet je roven koeficientu u x^{70} v součinu

$$(1+x+x^2+\cdots+x^{30})(1+x+x^2+\cdots+x^{40})(1+x+x^2+\cdots+x^{50}).$$

Tento součin upravíme na tvar $(1-x)^{-3}(1-x^{31})(1-x^{41})(1-x^{51})$, odkud pomocí zobecněné binomické věty dostaneme

$$\left(\binom{2}{2} + \binom{3}{2}x + \binom{4}{2}x^2 + \cdots\right)(1 - x^{31} - x^{41} - x^{51} + x^{72} + \dots)$$

a tedy koeficientem u x^{70} je zřejmě $\binom{70+2}{2} - \binom{70+2-31}{2} - \binom{70+2-41}{2} - \binom{70+2-51}{2} = 1061$.

Příklad. Zkusme pomocí vytvořujících funkcí najít explicitní vzoreček pro $1+2+\cdots+2^k$. Protože je $\frac{1}{1-2x}$ vytvořující funkce posloupnosti (2^k), je vytvořující funkcí pro posloupnost $(1+2+\cdots+2^k)$ funkce

$$\tfrac{1}{1-x}\cdot\tfrac{1}{1-2x} = 2\cdot\tfrac{1}{1-2x} - \tfrac{1}{1-x}.$$

Proto je zpětně tato posloupnost rovna $(2\cdot 2^k - 1)$.

Rozklad na parciální zlomky!

Rozklad na parciální zlomky – připomenutí

Rozklad na parciální zlomky jsme již viděli dříve při integraci racionálních lomených funkcí, přesto připomeneme:

- Předpokládáme, že $P(x)/Q(x)$ je podíl polynomů, kde $\deg P < \deg Q$ (jinak vydělíme se zbytkem) a $P(x), Q(x)$ nemají společné kořeny.
- Polynom $Q(x)$ rozložíme na kořenové činitele.
- Jsou-li všechny kořeny $\alpha_1, \ldots, \alpha_\ell$ jednoduché, pak

$$\frac{P(x)}{Q(x)} = \frac{A_1}{x-\alpha_1} + \cdots + \frac{A_\ell}{x-\alpha_\ell}.$$

- Má-li kořen α_i násobnost k_i, pak se příslušný parciální zlomek nahradí součtem parciálních zlomků tvaru

$$\frac{A_{i1}}{(x-\alpha_i)} + \frac{A_{i2}}{(x-\alpha_i)^2} + \cdots + \frac{A_{ik_i}}{(x-\alpha_i)^{k_i}}.$$

- V případě dvojice komplexně sdružených kořenů nahrazujeme sčitanec $A/(x-\alpha)$ sčítancem $(Ax+B)/(x^2+px+q)$ včetně příslušných mocnin jmenovatele.
- Neznámé dopočítáme roznásobením a buď porovnáním koeficientů u jednotlivých mocnin x nebo dosazením jednotlivých kořenů.
- Výrazy $A/(x-\alpha)^k$ převedeme na výrazy tvaru $B/(1-\beta x)^k$ vydělením čitatele i jmenovatele výrazem $(-\alpha)^k$. Tento výraz již umíme rozvinout do mocninné řady.

Rozklad na parciální zlomky – vychytávka

Protože preferujeme $1-\beta x$, bude lepší jmenovatel rozložit rovnou na součin takovýchto činitelů, např.

$$1 - 5x + 6x^2 = (1-2x)(1-3x),$$

který obecně získáme "otočením" polynomu – provedeme substituci $x = \frac{1}{t}$ a vynásobme t^2:

$$1 - 5\tfrac{1}{t} + 6\tfrac{1}{t^2} = (1-2\tfrac{1}{t})(1-3\tfrac{1}{t})$$
$$t^2 - 5t + 6 = (t-2)(t-3)$$

Přitom poslední tvar je již klasický rozklad na kořenové činitele, ve kterém můžeme použít např. známé vzorečky pro kořeny kvadratického polynomu.

23 Řešení rekurencí

Fibonacciho čísla a zlatý řez

Připomeňme, že Fibonacciho čísla jsou dána rekurentním předpisem

$$F_0 = 0, F_1 = 1, F_k = F_{k-1} + F_{k-2} \text{ pro } k \geq 2.$$

Již dříve jste si uváděli všemožné výskyty této posloupnosti v přírodě, v matematice nebo v teoretické informatice (podrobně viz `http://is.muni.cz/th/41281/prif_d/disertace.pdf`). Naším cílem bude (opět) najít formuli pro výpočet n-tého členu posloupnosti.

Poznámka. (Nejen) pro manipulace se sumami používají autoři *Concrete mathematics* velmi vhodné označení `[logický predikát]` – výraz je roven 1 v případě splnění predikátu, jinak 0., např. $[k = 1], [2|k]$ apod. Pro vyjádření koeficientu u x^k ve vytvořující funkci $F(x)$ se pak často používá zápis $[x^k]F(x)$.

Uvažme vytvořující funkci $F(x)$ Fibonacciho posloupnosti. Při podmínkách $F_0 = 0$, $F_1 = 1$ je to $F(x) - xF(x) - x^2F(x) = x$, a tedy

$$F(x) = \frac{x}{1 - x - x^2}.$$

Našim cílem je tedy odvodit vztah pro k-tý člen posloupnosti.

Využijeme k tomu rozklad na parciální zlomky a dostaneme

$$\frac{x}{1 - x - x^2} = \frac{A}{1 - \lambda x} + \frac{B}{1 - \mu x},$$

kde λ, μ jsou kořeny $t^2 - t - 1$ a A, B vhodné konstanty odvozené z počátečních podmínek. Odtud už vcelku snadno vyjde $F_k = A \cdot \lambda^k + B \cdot \mu^k$, jak to známe z dřívějška.

S využitím počátečních podmínek dostáváme

$$F_k = \frac{1}{\sqrt{5}} \left[\left(\frac{1 + \sqrt{5}}{2} \right)^k - \left(\frac{1 - \sqrt{5}}{2} \right)^k \right].$$

Jistě je zajímavé, že tento výraz plný iracionálních čísel je vždy celočíselný.

Uvážíme-li navíc, že $(1 - \sqrt{5})/2 \approx -0.618$, vidíme, že pro všechna přirozená čísla lze F_k snadno spočítat zaokrouhlením čísla $\frac{1}{\sqrt{5}}(\frac{1+\sqrt{5}}{2})^k$.

Navíc je vidět, že $\lim_{k\to\infty} F_{k+1}/F_k = \lambda \approx 1.618$, což je poměr známý jako zlatý řez – objevuje se již od antiky v architektuře, výtvarném umění i hudbě.

Řešení rekurencí

Mocninné řady jsou velmi silným nástrojem pro řešení rekurencí (a to nejen lineárních!). Tím je míněno vyjádření členu a_k jako funkci k. Často se s pomocí řad podaří vyřešit na první pohled velmi složité rekurence.

Obvyklý (takřka mechanický) postup pro řešení rekurencí se skládá ze 4 kroků:

1. Zapíšeme jedinou rovnicí závislost a_k na ostatních členech posloupnosti. Tento vztah musí platit pro všechna $k \in \mathbb{N}_0$ (předpokládajíce $a_{-1} = a_{-2} = \cdots = 0$).

2. Obě strany rovnice vynásobíme x^k a sečteme přes všechna $k \in \mathbb{N}_0$. Na jedné straně tak dostaneme $\sum_{k\geq 0} a_k x^k$, což je vytvořující funkce $A(x)$. Pravou stranu vztahu je pak třeba upravit na výraz rovněž obsahující $A(x)$.

3. Zjištěná rovnice se vyřeší vzhledem k $A(x)$.

4. Výsledné $A(x)$ se rozvine do mocninné řady, přičemž koeficient u x^k udává a_k, tj. $a_k = [x^k]A(x)$.

Příklad. Řešte rekurenci

$$a_0 = 0, a_1 = 1$$
$$a_k = 5a_{k-1} - 6a_{k-2}$$

Řešení.

- Krok 1: $a_k = 5a_{k-1} - 6a_{k-2} + [k = 1]$.
- Krok 2: $A(x) = 5xA(x) - 6x^2A(x) + x$.
- Krok 3:
$$A(x) = \frac{x}{1 - 5x + 6x^2} = \frac{1}{1 - 3x} - \frac{1}{1 - 2x}.$$
- Krok 4: $a_k = 3^k - 2^k$.

Quicksort – analýza průměrného případu

Ukázka implementace (*divide and conquer*, rozmyslete, proč není optimální):

```
if L == []: return []
return qsort([x for x in L[1:] if x<L[0]])
     + L[0:1]
     + qsort([x for x in L[1:] if x>=L[0]])
```

1. Počet porovnání při rozdělení (*divide*): $k - 1$.
2. (Předpoklad náhodnosti): Pravděpodobnost toho, že prvek $L[0]$ je i-tý největší, je $\frac{1}{k}$.
3. Velikost tříděných polí ve fázi *conquer*: $i - 1$ a $k - i$.

Pro střední hodnotu počtu porovnání tak dostáváme rekurentní vztah:

$$C_k = k - 1 + \sum_{i=1}^{k} \frac{1}{k} \left(C_{i-1} + C_{k-i}\right).$$

Analýza Quicksortu pomocí vytvořujících funkcí

Vyřešme nyní rekurenci

$$kC_k = k(k-1) + 2\sum_{i=1}^{k} C_{i-1}, C_0 = C_1 = 0$$

pomocí uvedeného postupu.

- $\sum_{k\geq 0} kC_k x^k = \sum_{k\geq 0} k(k-1)x^k + 2\sum_{k\geq 0}\sum_{i=1}^{k} C_{i-1}x^k$
- $xC'(x) = \frac{2x^2}{(1-x)^3} + 2\frac{xC(x)}{1-x}$
- Vyřešíme tuto lineární diferenciální rovnici prvního řádu $\left((1-x)^2C(x)\right)' = \frac{2x}{1-x}$, a tedy
$$C(x) = \frac{2}{(1-x)^2}\left(\ln\frac{1}{1-x} - x\right),$$
odkud konečně $C_k = 2(k+1)(H_{k+1} - 1) - 2k$.

24 Binární stromy a Catalanova čísla

S využitím standardních vytvořujících funkcí určíme formuli pro počet b_n tzv. pěstovaných binárních stromů na n vrcholech, které je pro naše účely možné definovat jako kořen s uspořádanou dvojicí [*levý binární podstrom, pravý binární podstrom*].

Prozkoumáním případů pro malá n vidíme, že

$$b_0 = 1, b_1 = 1, b_2 = 2, b_3 = 5.$$

Dělením problému na levý a pravý strom dostaneme pro $n \geq 1$

$$b_n = b_0 b_{n-1} + b_1 b_{n-2} + \cdots + b_{n-1} b_0.$$

Vidíme, že jde vlastně o konvoluci posloupností. Vztah upravíme, aby platil pro všechna $n \in N_0$:

$$b_n = \sum_{0 \leq k < n} b_k b_{n-k-1} + [n = 0].$$

Tím máme hotov krok 1 (obecného postupu z minula).

V kroku 2 vynásobíme obě strany x^n a sečteme. Je-li $B(x)$ odpovídající vytvořující funkce, pak:

$$\begin{aligned} B(x) &= \sum_n b_n x^n = \sum_{n,k} b_k b_{n-k-1} x^n + \sum_{n,k} [n = 0] x^n = \\ &= \sum_k b_k x^k \left(\sum_n b_{n-k-1} x^{n-k} \right) + 1 = \\ &= \sum_k b_k x^k (xB(x)) + 1 = B(x) \cdot xB(x) + 1. \end{aligned}$$

Pravá strana rekurence na prvním řádku je koeficientem u x^{n-1} v součinu $B(x) \cdot B(x)$, tj. členem u x^n v $xB(x)^2$.

Je tedy $xB(x)^2$ vytvořující po tutéž posloupnost jako $B(x)$ s výjimkou prvního členu u x^0.

V kroku 3 řešíme kvadratickou rovnici $B(x) = xB(x)^2 + 1$ pro $B(x)$:

$$B(x) = \frac{1 \pm \sqrt{1 - 4x}}{2x}.$$

Znaménko + ale nepřichází v úvahu, protože pak by pro $x \to 0_+$ $B(x)$ měla limitu ∞, zatímco vytvořující funkce pro naši posloupnost musí mít v 0 hodnotu $b_0 = 1$. Naopak pro znaménko − to tak dostaneme.

Pro vytvořující funkci $B(x)$ tedy platí

$$B(x) = \frac{1 - \sqrt{1 - 4x}}{2x}.$$

Zbývá už pouze krok 4, tedy rozvinout $B(x)$ do mocninné řady.

Rozvoj získáme pomocí zobecněné binomické věty

$$(1 - 4x)^{1/2} = \sum_{k \geq 0} \binom{1/2}{k} (-4x)^k = 1 + \sum_{k \geq 1} \frac{1}{2k} \binom{-1/2}{k-1} (-4x)^k$$

a po vydělení $1 - \sqrt{1 - 4x}$ výrazem $2x$ dostaneme

$$\begin{aligned} B(x) &= \sum_{k \geq 1} \frac{1}{k} \binom{-1/2}{k-1} (-4x)^{k-1} = \\ &= \sum_{n \geq 0} \binom{-1/2}{n} \frac{(-4x)^n}{n+1} = \sum_{n \geq 0} \binom{2n}{n} \frac{x^n}{n+1}. \end{aligned}$$

Catalanova čísla

Dokázali jsme, že počet binárních pěstovaných stromů na n vrcholech je roven $b_n = \frac{1}{n+1}\binom{2n}{n}$.

Tato významná posloupnost se nazývá posloupnost *Catalanových čísel*.

Kromě toho, že Catalanova čísla vyjadřují počet binárních pěstovaných stromů, vystupují rovněž jako:

- počet *monotónních* cest z $[0,0]$ do $[n,n]$ podél stran jednotkových čtverců, které nepřekročí diagonálu
- počet slov délky $2n$ obsahujících n znaků X a Y takových, že žádný prefix slova neobsahuje více Y než X
- podobně takové fronty u pokladny (n lidí má 5korunu a m 10korunu, lístek stojí 5 Kč.), že nezásobená pokladna může vždy vrátit
- počet korektně ozávorkovaných výrazů složených z levých a pravých závorek
- počet různých triangulací konvexního $(n+2)$-úhelníku.

25 Caleyho vztah pro počet stromů

Cayleyho formule

Cayleyho formule je vztah z kombinatorické teorie grafů, který udává, že počet stromů (tj. grafů, v nichž jsou libovolné dva vrcholy spojené právě jednou cestou) na n vrcholech je $\kappa(K_n) = n^{n-2}$. Dokážeme tento výsledek pomocí exponenciálních vytvořujících funkcí.

Označme pro jednoduchost $t_n = \kappa(K_n)$. Lze snadno spočítat, že $t_1 = t_2 = 1, t_3 = 3, t_4 = 16$. (Např. víme, že v případě stromů na 4 vrcholech musíme z $\binom{6}{3} = 20$ potenciálních grafů s právě 3 hranami odebrat ty možnosti, kde tyto hrany tvoří trojúhelník. Těch je ale právě $\binom{4}{3} = 4$).

Rekurentní vztah získáme tak, že zafixujeme jeden vrchol v_0 a možné případy rozdělíme podle počtu komponent v grafu, který dostaneme z koster K_n tak, že odstraníme vrchol v a hrany s ním incidentní.

Kvůli jednoduchosti argumentu budeme rekurzivně vyjadřovat počet tzv. kořenových stromů, tj. stromů s vybraným vrcholem, kterému říkáme kořen. Jejich počet je zjevně $u_n = nt_n$. Uvažme kořenový strom na n vrcholech s kořenem v_0. Odstraněním tohoto vrcholu dostaneme disjunktní sjednocení několika stromů (tzv. les), přičemž každý strom, tj. každá z komponent, má vybraný kořen v_i, konkrétně soused v_0 v původním stromu. Naopak, pokud zvolíme vrchol v_0, rozložíme množinu zbylých vrcholů na několik neprázdných disjunktních částí – označme jejich počet – a na každé vybereme strukturu kořenového stromu s kořenem v_i, dostaneme přidáním hran $v_0v_1, \ldots, v_0v_m$ kořenový strom. Proto pro $n > 1$ platí

$$u_n = \sum_{m\geq 0} \frac{1}{m!} \sum_{1+k_1+\cdots+k_m=n} \frac{n!}{1!k_1!\cdots k_m!} u_{k_1}\cdots u_{k_m}$$

(množinu všech vrcholů obarvíme barvami následovně: jeden vrchol v_0 barvou 0, dále k_i vrcholů barvou i, pro každé i; faktor $\frac{1}{m!}$ zaručí, že nezáleží na pořadí barev $1,\ldots,m$, takže se vskutku jedná o rozklad; v komponentě každé barvy zvolíme kořenový strom).

Vydělením $n!$ pak rekurenci výrazně zjednodušíme, zejména při substituci $\widehat{u}_n = \frac{u_n}{n!}$:

$$\frac{u_n}{n!} = \sum_{m\geq 0} \frac{1}{m!} \sum_{k_1+\cdots+k_m=n-1} \frac{u_{k_1}}{k_1!}\cdots\frac{u_{k_m}}{k_m!}$$

a je vidět, že vnitřní sumu dostaneme jako koeficient u x^{n-1} v m-té mocnině řady $\widehat{U}(x) = \sum u_n \frac{x^n}{n!}$. Proto je

$$\frac{u_n}{n!} = [x^{n-1}] \sum_{m\geq 0} \frac{1}{m!} \widehat{U}(x)^m,$$

a tedy

$$\widehat{U}(x) = xe^{\widehat{U}(x)}.$$

Věta 34. *Pokud vytvořující funkce* $g(x) = \sum_{n\geq 1} g_n x^n$ *splňuje vztah*

$$x = f(g(x)),$$

kde $f(0) = 0, f'(0) \neq 0$*, pak*

$$g_n = \frac{1}{n}[u^{n-1}]\left(\frac{u}{f(u)}\right)^n.$$

Důkaz. Nebudeme se pokoušet o rigorózní důkaz, uvedeme jen, proč by vůbec nějaký vztah mezi koeficienty f a g měl existovat a jak lze pro malá n odvodit: Označíme-li $f(u) = \sum_{k\geq 1} a_k u^k$ a $g(x) = \sum_{l\geq 1} b_l x^l$, pak dosazením do sebe dostaneme vztah

$$\begin{aligned} x = f(g(x)) &= a_1(b_1x + b_2x^2 + b_3x^3 + \dots) \\ &\quad + a_2(b_1x + b_2x^2 + \dots)^2 + a_3(b_1x + \dots)^3 + \dots \\ &= a_1b_1x + (a_1b_2 + a_2b_1^2)x^2 + (a_1b_3 + a_2(b_1b_2 + b_2b_1) + a_3b_1^3)x^3 + \dots \end{aligned}$$

ze kterého lze induktivně počítat (první koeficient $a_1b_1 = 1$, další jsou nulové):

$$b_1 = \frac{1}{a_1},\ b_2 = \frac{-a_2}{a_1^3},\ b_3 = \frac{2a_2^2 - a_1a_3}{a_1^5}, \dots$$ □

Řešíme $\widehat{U}(x) = x\,\mathrm{e}^{\widehat{U}(x)}$, tj. $\widehat{U}(x)$ splňuje vztah $x = f(\widehat{U}(x))$, kde $f(u) = \frac{u}{e^u}$. Odtud z Lagrangeovy formule

$$\begin{aligned} [x^n]\widehat{U}(x) &= \frac{1}{n}[u^{n-1}]\left(\frac{u}{u/e^u}\right)^n \\ &= \frac{1}{n}[u^{n-1}]\,\mathrm{e}^{un} = \frac{1}{n}\frac{n^{n-1}}{(n-1)!} = \frac{n^{n-1}}{n!} \end{aligned}$$

Protože $\frac{u_n}{n!} = [x^n]\widehat{U}(x)$, dostáváme odtud

$$t_n = \tfrac{u_n}{n} = n^{n-2}.$$

Další aplikace Lagrangeovy inverzní formule

Vraťme se ještě krátce ke Catalanovým číslům. Chtěli jsme vyřešit rovnici

$$B(x) = xB(x)^2 + 1$$

a výslednou funkci $B(x)$ pak rozvést do mocninné řady. Substitucí $B(x) = C(x) + 1$ rovnici převedeme na $C(x) = x(C(x) + 1)^2$ neboli

$$\frac{C(x)}{(C(x)+1)^2} = x.$$

Označíme-li nyní $f(u) = \frac{u}{u+1}^2$, je hledaná funkce $C(x)$ k této funkci inverzní a podle Lagrangeovy formule její koeficienty jsou

$$\begin{aligned} c_n &= \frac{1}{n}[u^{n-1}](\frac{u}{u/(u+1)^2})^n = \frac{1}{n}[u^{n-1}]((u+1)^2)^n \\ &= \frac{1}{n}[u^{n-1}](1+u)^{2n} = \frac{1}{n}\binom{2n}{n-1} \end{aligned}$$

což lze vskutku ekvivalentně přepsat jako $\frac{1}{n+1}\binom{2n}{n}$.

Ukážeme ještě jeden způsob důkazu využívající exponenciální vytvořující funkce.

Exponenciální vytvořující funkce

Kromě výše zmíněných vytvořujících funkcí se v praxi rovněž často objevují jejich tzv. *exponenciální* varianty[4].

$$g(x) = \sum_{n\geq 0} g_n \frac{x^n}{n!}.$$

Poznámka. Jméno vychází z toho, že exponenciální funkce e^x je (exponenciální) vytvořující funkcí pro základní posloupnost $(1, 1, 1, 1, \dots)$.

V zápětí v důkazu Cayleyho věty uvidíme, že je použití exponenciálních vytvořujících funkcí výhodné.

Opět standardním operacím s posloupnostmi odpovídají jednoduché operace nad mocninnými řadami (coby exponenciálními vytvořujícími funkcemi):

- Sčítání $(a_i + b_i)$ posloupností člen po členu odpovídá součet $a(x) + b(x)$ příslušných vytvořujících funkcí.
- Vynásobení $(\alpha \cdot a_i)$ všech členů posloupnosti stejným skalárem α odpovídá vynásobení $\alpha \cdot a(x)$ příslušné vytvořující funkce.
- Derivování podle x: funkce $a'(x)$ vytvořuje posloupnost $(a_1, a_2, a_3, \dots)$, tj. derivování odpovídá posuvu doleva o jedno místo .
- Integrování $\int_0^x a(t)\mathrm{d}t$ vytvořuje posloupnost $(0, a_0, a_1, a_2, a_3, \dots)$, tj. odpovídá posuvu doprava o jedno místo.
- Součin vytvořujících funkcí vytvořuje posloupnost se členy

$$c_n = \sum_{i+j=n} \binom{n}{i} a_i b_j$$

Alternativní závěr výpočtu

Pro dokončení výpočtu budeme potřebovat tvrzení, které uvedeme bez důkazu.

Definice. Zobecněnou exponenciální mocninnou řadou $\mathcal{E}_t(x)$ nazýváme řadu

$$\mathcal{E}_t(x) = \sum_{k\geq 0} (tk+1)^{k-1} \frac{x^k}{k!}.$$

Snadno je vidět, že $\mathcal{E}_0 = \mathrm{e}^x$, dále označujeme $\mathcal{E}(x) = \mathcal{E}_1(x)$.
Fakt: $\ln \mathcal{E}_t(x) = x \cdot \mathcal{E}_t(x)$, tj. spec. $\mathcal{E}(x) = \mathrm{e}^{x\mathcal{E}(x)}$.
Srovnáním tohoto vztahu s výše uvedeným $\widehat{U}(x) = x\,\mathrm{e}^{\widehat{U}(x)}$ vidíme, že $\widehat{U}(x) = x\mathcal{E}(x)$.
Proto

$$t_n = \frac{u_n}{n} = \frac{n!}{n}[x^n]\widehat{U}(x) = (n-1)![x^{n-1}]\mathcal{E}(x) = n^{n-2}.$$

26 Rekurzivně propojené posloupnosti

Někdy dokážeme snadno vyjádřit hledaný počet jen pomocí více vzájemně provázaných posloupností.

Příklad. Kolika způsoby můžeme pokrýt (nerozlišenými) kostkami domina obdélník $3 \times n$?

[4] Používají se i další typy vytvořujících funkcí (např. v teorii čísel se používají Dirichletovy vytvořující funkce, kde roli faktoru x^n hraje n^{-x}), ale těmi se zde zabývat nebudeme.

Řešení. Snadno zjistíme, že $c_1 = 0, c_2 = 3, c_3 = 0$, dále klademe $c_0 = 1$ (nejde jen o konvenci, má to svou logiku).

Najdeme rekurzívní vztah – diskusí chování „na kraji“ zjistíme, že $c_n = 2r_{n-1} + c_{n-2}$, $r_n = c_{n-1} + r_{n-2}$, $r_0 = 0, r_1 = 1$, kde r_n je počet pokrytí obdélníku $3 \times n$, ze kterého jsme odstranili levý horní roh.

Hodnoty c_n a r_n pro několik malých n jsou:

n	0	1	2	3	4	5	6	7
c_n	1	0	3	0	11	0	41	0
r_n	0	1	0	4	0	15	0	56

- Krok 1: $c_n = 2r_{n-1} + c_{n-2} + [n = 0], \quad r_n = c_{n-1} + r_{n-2}$.
- Krok 2: $C(x) = 2xR(x) + x^2C(x) + 1, \quad R(x) = xC(x) + x^2R(x)$.
- Krok 3:
$$C(x) = \frac{1 - x^2}{1 - 4x^2 + x^4}, \qquad R(x) = \frac{x}{1 - 4x^2 + x^4}.$$
- Krok 4: Vidíme, že funkce $C(x)$ je funkce x^2, ušetříme si práci tím, že uvážíme funkci $D(z) = 1/(1-4z+z^2)$, pak totiž $C(x) = (1-x^2)D(x^2)$, tj. $[x^{2n}]C(x) = [x^{2n}](1-x^2)D(x^2) = [x^n](1-x)D(x)$, a tedy $c_{2n} = d_n - d_{n-1}$.

Kořeny $1 - 4x + x^2$ jsou $2 + \sqrt{3}$ a $2 - \sqrt{3}$ a již standardním způsobem obdržíme

$$c_{2n} = \frac{(2+\sqrt{3})^n}{3-\sqrt{3}} + \frac{(2-\sqrt{3})^n}{3+\sqrt{3}}.$$

Podobně jako u Fibonacciho posloupnosti je druhý sčítanec pro velká n zanedbatelný a pro všechna n leží mezi 0 a 1, proto

$$c_{2n} = \left\lfloor \frac{(2+\sqrt{3})^n}{3-\sqrt{3}} \right\rfloor.$$

Např. $c_{20} = 413403$.

Ještě jeden příklad

Příklad. Vyřešte rekurenci

$$a_0 = a_1 = 1$$
$$a_n = a_{n-1} + 2a_{n-2} + (-1)^n$$

Řešení. Tato rekurence je opět jiného typu než dosud studované. Jako vždy neuškodí vypsání prvních několika členů posloupnosti (teď ale ani moc nepomůže, snad jen pro kontrolu správnosti výsledku).[5]

- Krok 1: $a_n = a_{n-1} + 2a_{n-2} + (-1)^n[n \geq 0] + [n = 1]$.
- Krok 2: $A(x) = xA(x) + 2x^2A(x) + \frac{1}{1+x} + x$.
- Krok 3:
$$A(x) = \frac{1 + x + x^2}{(1 - 2x)(1 + x)^2}.$$
- Krok 4: $a_n = \frac{7}{9}2^n + \left(\frac{1}{3}n + \frac{2}{9}\right)(-1)^n$.

[5] Narozdíl od tvrzení v *Concrete mathematics* je již možné tuto posloupnost nalézt v *The On-Line Encyclopedia of Integer Sequences.*