ČASOVÉ ŘADY

Mgr. et Mgr. Jiří Kalina, PhD.

prof. Ing. Jiří Holčík, CSc.

UKB, pavilon D29 (Recetox), kancelář 123
kalina@mail.muni.cz

© Institut biostatistiky a analýz

VIII. VZORKOVÁNÍ

© Institut biostatistiky a analýz

DEFINICE

Vzorkování je postup výběru jednotlivých pozorování, na jehož základě získáváme informaci o vlastnostech sledované skutečnosti či jevu. Každé pozorování může obecně zahrnovat více vlastností (věk, diagnóza onemocnění, velikost napětí, …), které mohou být použity k identifikaci daného jevu či jeho části.

Vzorkováním rozumíme postup výběru určité podmnožiny (vzorku) dané množiny (veličiny, populace, dat, materiálu) tak, aby vlastnosti vybraného vzorku (dostatečně) přesně reprezentovaly vlastnosti celé množiny (signálu, populace, dat, materiálu).

Vzorkování je postup selekce jednotlivých pozorování s cílem získat určitou znalost o dané populaci, zejména pro účely statistické inference.

© Institut biostatistiky a analýz

PŘÍKLADY

- ☑ volba frekvence a místa odběru pro hodnocení úrovně znečištění vodních toků;
- ☑ volba parametrů digitalizace obrazu pro jeho přenos či archivaci;
- ☑ volba tématu a množiny respondentů při průzkumu veřejného mínění;
- ☑ výběr výrobků při výstupní kontrole kvality výroby;
- ☑ výběr pacientů pro odhad vývoje daného onemocnění.

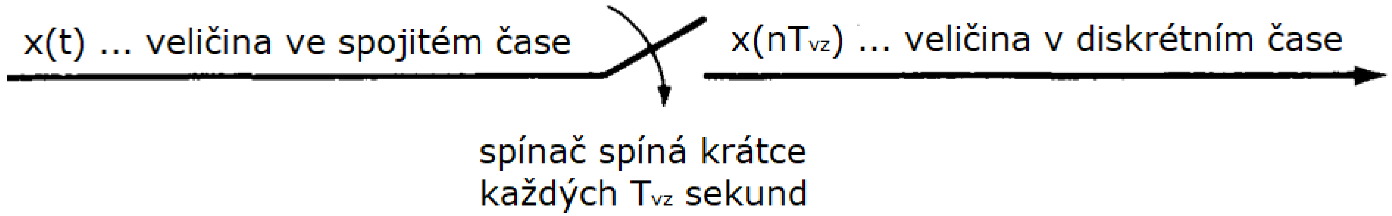
DEFINICE

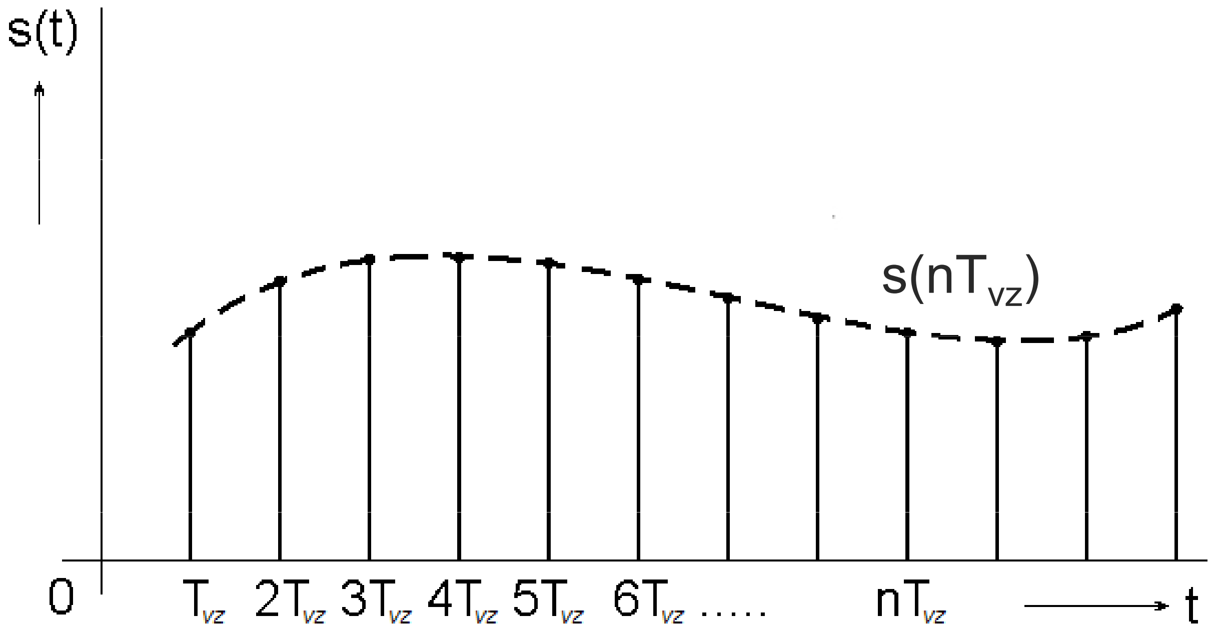
Vzorkováním dané časově proměnné veličiny rozumíme činnost, při které z průběhu určité veličiny, která je definovaná na spojitém definičním oboru, vybíráme hodnoty pouze v určitých časových okamžicích.

Hodnoty časových či prostorových souřadnic mohou být rozmístěny v definičním prostoru obecně nerovnoměrně, z hlediska práce s daty je ale výhodnější, pokud jsou souřadnice vzorků rozmístěny rovnoměrně (a v tom případě lze i teoreticky dovodit pravidlo pro maximální vzdálenost mezi každými dvěma vzorky).

© Institut biostatistiky a analýz

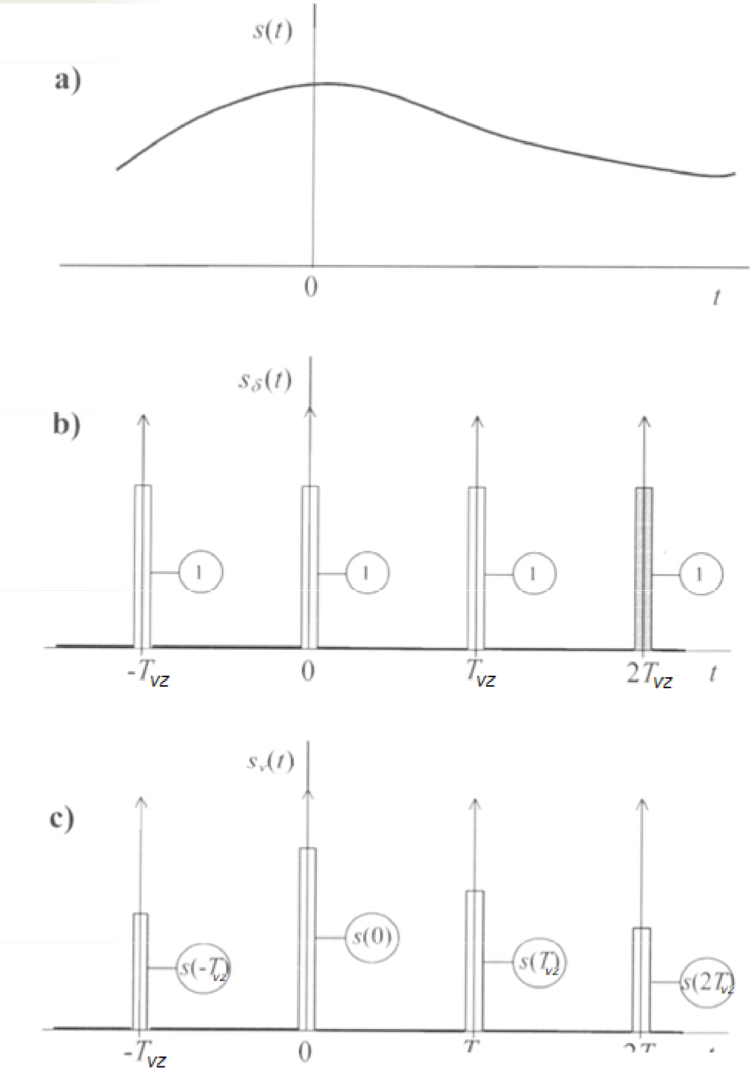
VZORKOVÁNÍ SPOJITÉ VELIČINY





© Institut biostatistiky a analýz

IDEÁLNÍ VZORKOVÁNÍ



Aby bylo možné zjednodušit analýzu vlivu vzorkování na vlastnosti vzorkované veličiny, je navzorkovaná verze původní spojité veličiny s(t) vyjadřována ve tvaru $s(t).s_\delta(t)$, kde $s_\delta(t)$ je periodický sled jednotkových impulsů definovaný jako

$$s_\delta(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_{vz})$$

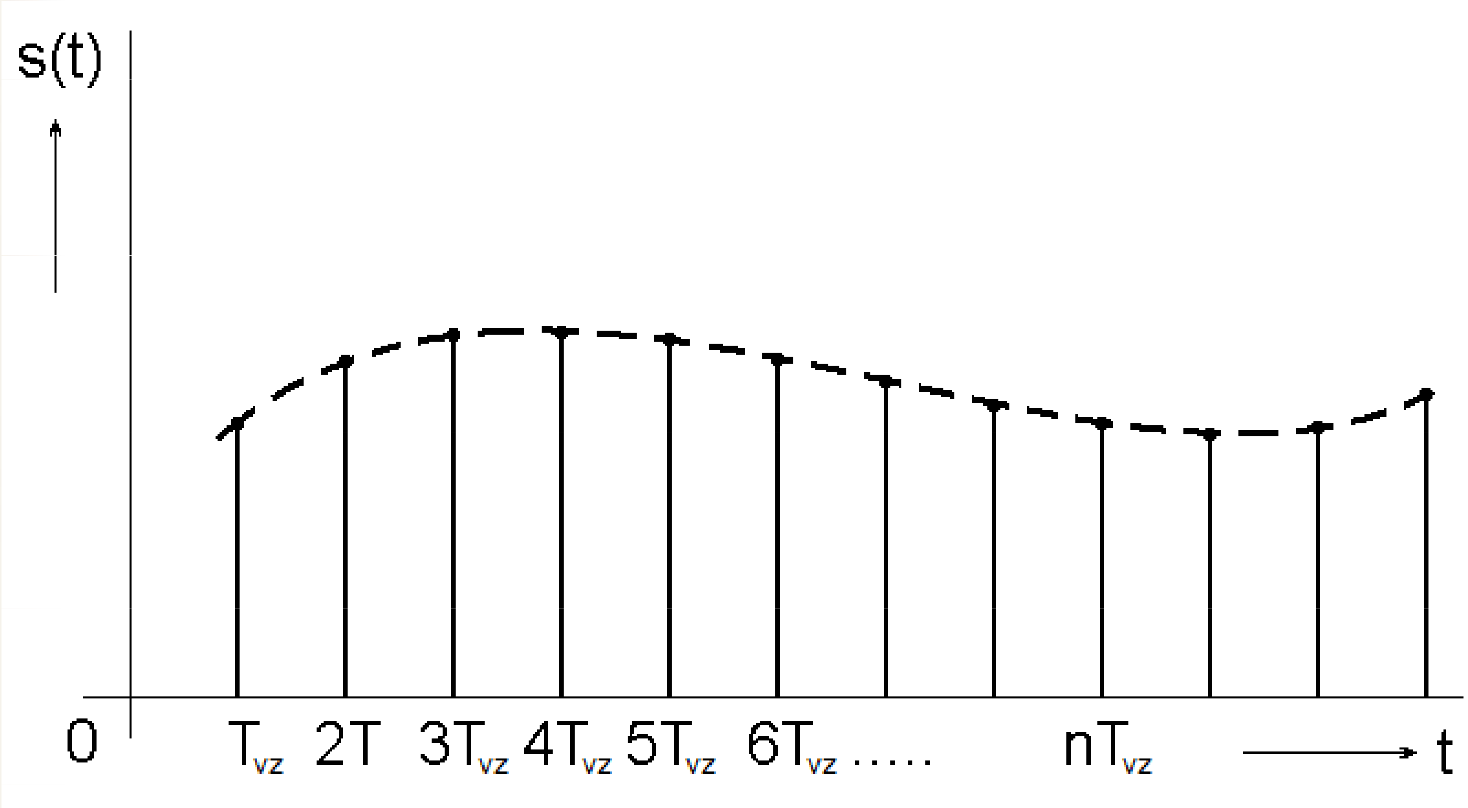
Z toho pro navzorkovanou veličinu platí

$$s_v(t) = s(t).s_\delta(t) = s(t).\sum_{n=-\infty}^{\infty} \delta(t - nT_{vz}) = \sum_{n=-\infty}^{\infty} s(nT_{vz}).\delta(t - nT_{vz})$$

VZORKOVACÍ TEORÉM

$$s(t) \rightarrow s(T_1), s(T_2), s(T_3), \ldots, s(T_n), \ldots$$

$$s(t) \rightarrow s(T_{vz}), s(2T_{vz}), s(3T_{vz}), \ldots, s(nT_{vz}), \ldots$$

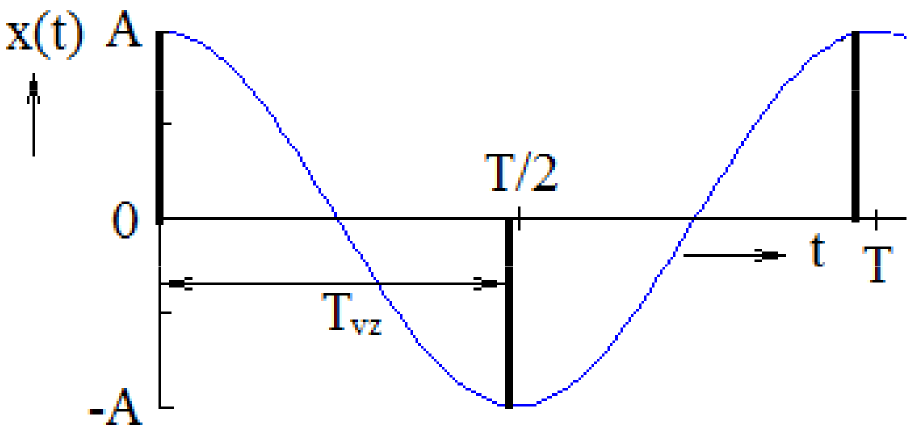


© Institut biostatistiky a analýz

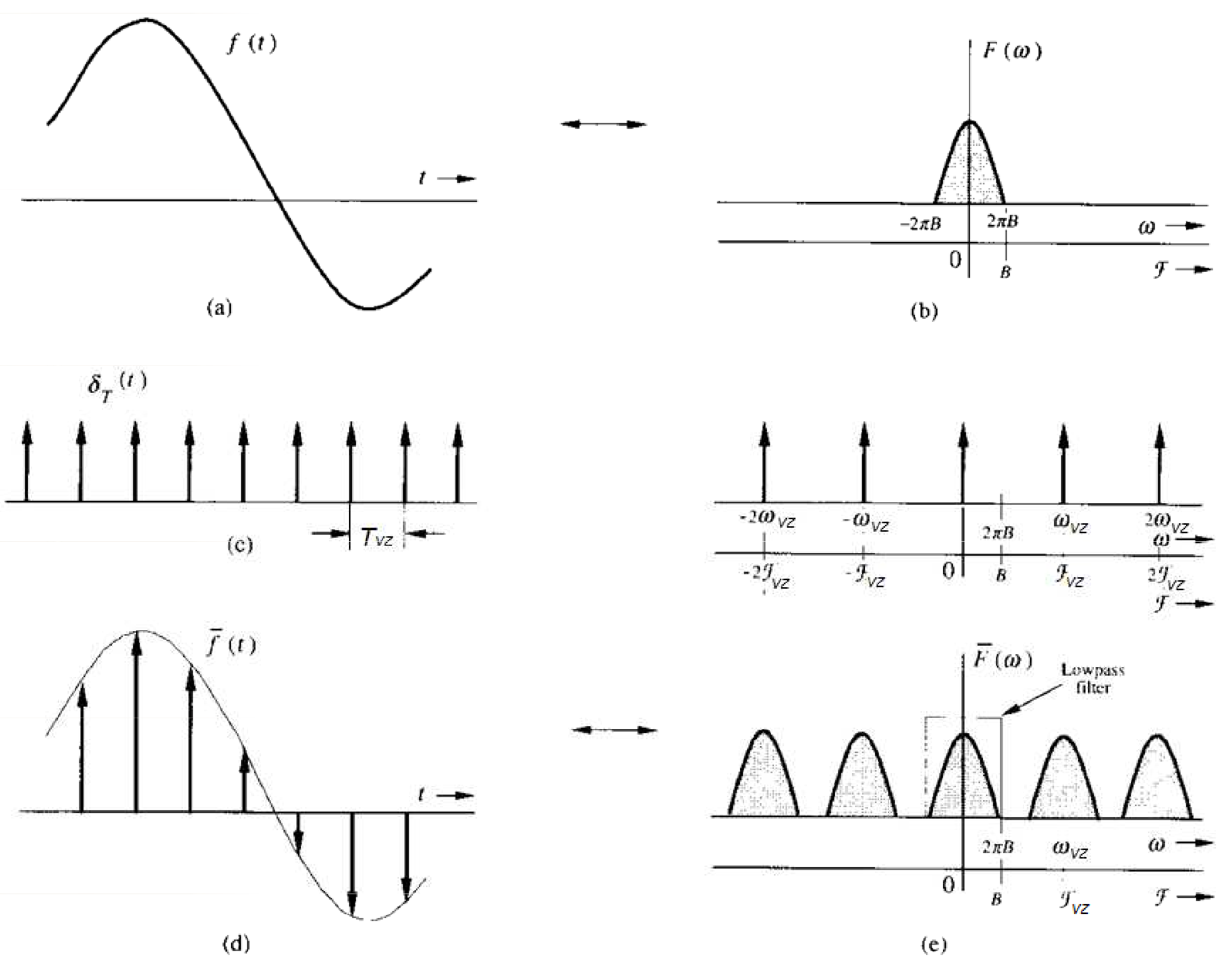
VZORKOVACÍ TEOREM

intuitivní (?) zdůvodnění minimální vzorkovací frekvence

Reálné vzorkování



VZORKOVACÍ TEORÉM



Vzorkovací frekvence:

$$f_{vz} > 2B = f_N,$$

kde B je maximální kmitočet ve vzorkované veličině

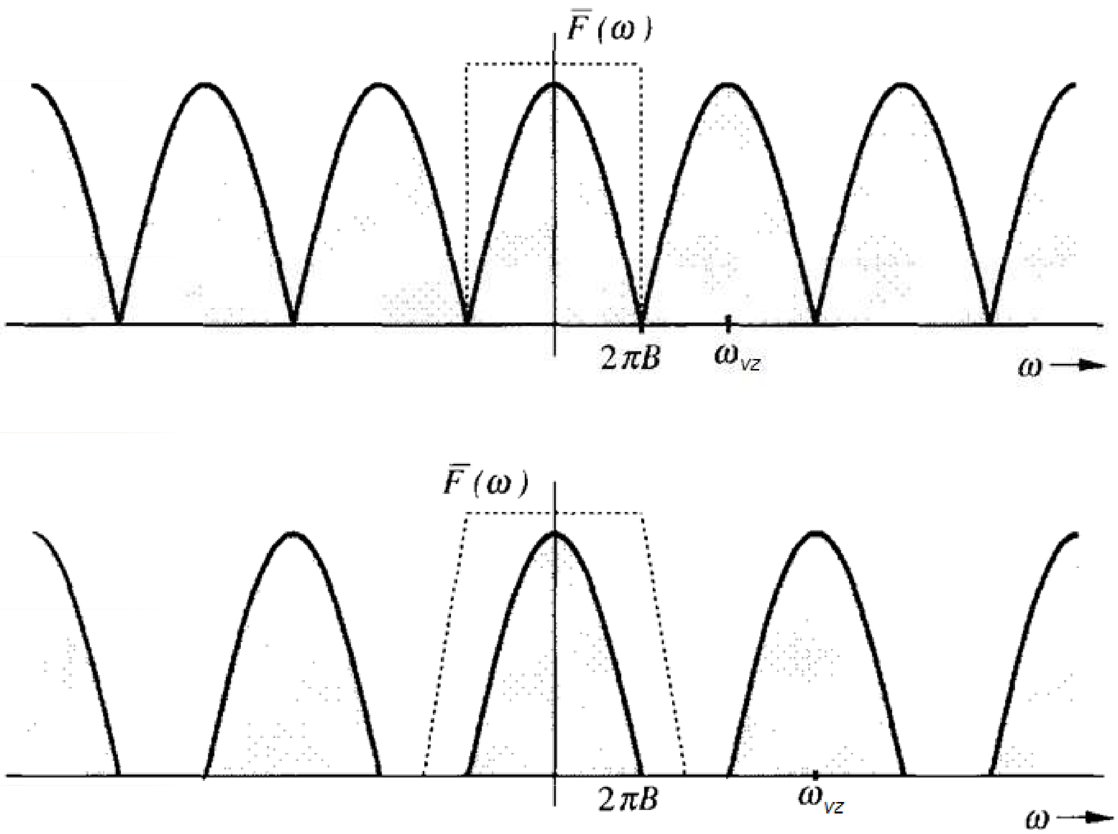
f_N –

Nyquistův, (Shannonův, Kotelnikovův) kmitočet

$$T_N = 1/f_N = 1/2B$$

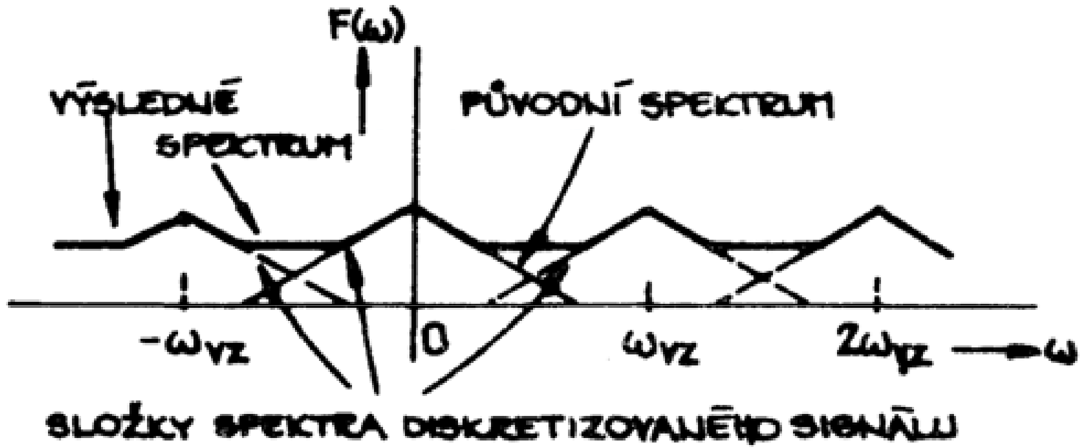
Nyquistův interval (perioda), vzorkovací interval (perioda)

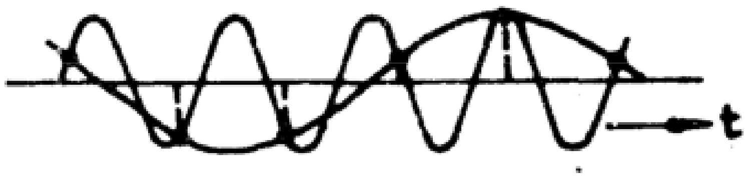
VZORKOVACÍ TEOREM



$f_{vzr} = (4 \div 5).f_N$

VZORKOVACÍ TEORÉM





překrývání spekter - aliasing

HARMONICKÁ POSLOUPNOST

$$x(nT_{vz}) = A.\cos(2\pi f\, nT_{vz} + \varphi_0) = A.\cos(\frac{2\pi\, nT_{vz}}{NT_{vz}} + \varphi_0) = A.\cos(\frac{2\pi\, n}{N} + \varphi_0),$$

když $T = NT_{vz}$ a tedy $f = 1/NT_{vz}$ nebo

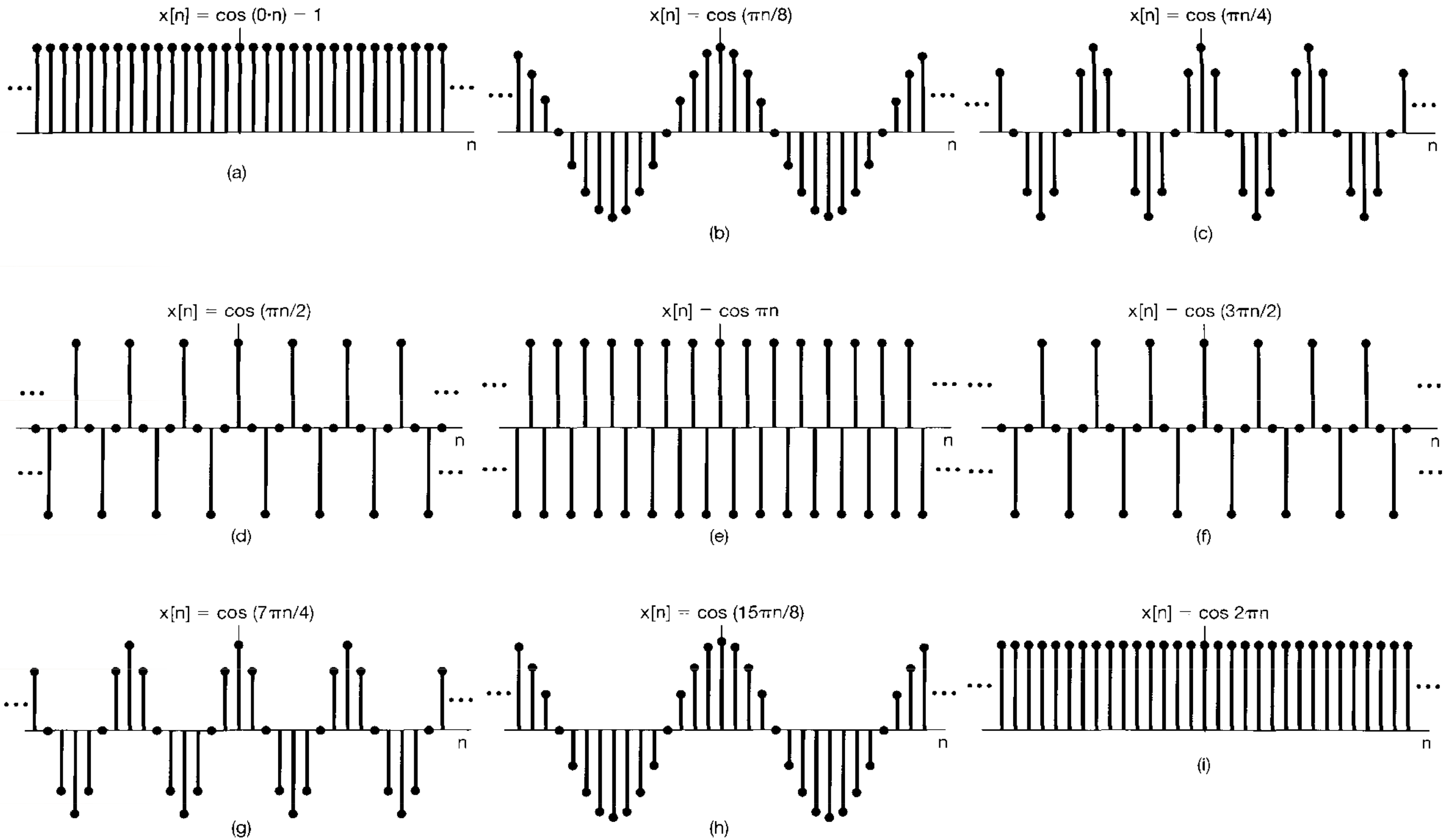
$$x(nT_{vz}) = A.\exp(\frac{2\pi i\, n}{N} + \varphi_0).$$

Komplexní exponenciála samozřejmě rovněž reprezentuje periodickou veličinu, protože platí

$$x\left[(k+N)T_{vz}\right] = \exp\frac{i2\pi(k+N)}{N} = \exp\frac{i2\pi k}{N}.\exp(i2\pi),$$

kdy $\exp(2\pi i) = \cos(2\pi) + i.\sin(2\pi) = 1 + i.0 = 1$

HARMONICKÁ FUNKCE A VZORKOVÁNÍ



REKONSTRUKCE SPOJITÉ FUNKCE Z NAVZORKOVANÉ POSLOUPNOSTI

Předpokládejme, že původní spojitá funkce $x_a(t)$ měla frekvenčně omezené spektrum $X_a(f)$, tj. platí pro ni

$$X_a(f) = \begin{cases} X(f) & |f| \le f_{vz}/2 \\ 0 & |f| > f_{vz}/2 \end{cases}$$

kde je X(f) frekvenční spektrum dané posloupnosti. Protože víme, že spektrum navzorkované posloupnosti je periodické s periodou danou vzorkovací frekvencí, zajímá nás pouze její jedna (první) perioda, pro kterou v rozsahu frekvencí $|f| \le f_{vz}/2$ platí

$$X_a(f) = X(f) = T_{vz} \sum_{n=-\infty}^{\infty} x(nT_{vz}).e^{-i2\pi f n T_{vz}}$$

© Institut biostatistiky a analýz

REKONSTRUKCE SPOJITÉ FUNKCE Z NAVZORKOVANÉ POSLOUPNOSTI

Potom pro původní funkci $x_a(t)$ je

$$x_a(t) = \int_{-f_{vz}/2}^{f_{vz}/2} X_a(f).e^{i2\pi ft}df = \int_{-f_{vz}/2}^{f_{vz}/2} \left[T_{vz} \sum_{n=-\infty}^{\infty} x(nT_{vz}).e^{-i2\pi fnT_{vz}} \right].e^{i2\pi ft}df =$$

$$= T_{vz} \sum_{n=-\infty}^{\infty} x(nT_{vz}). \int_{-f_{vz}/2}^{f_{vz}/2} e^{i2\pi f(t-nT_{vz})}df = \left| \int e^{ax}dx = \frac{1}{a}e^{ax} + C \right| =$$

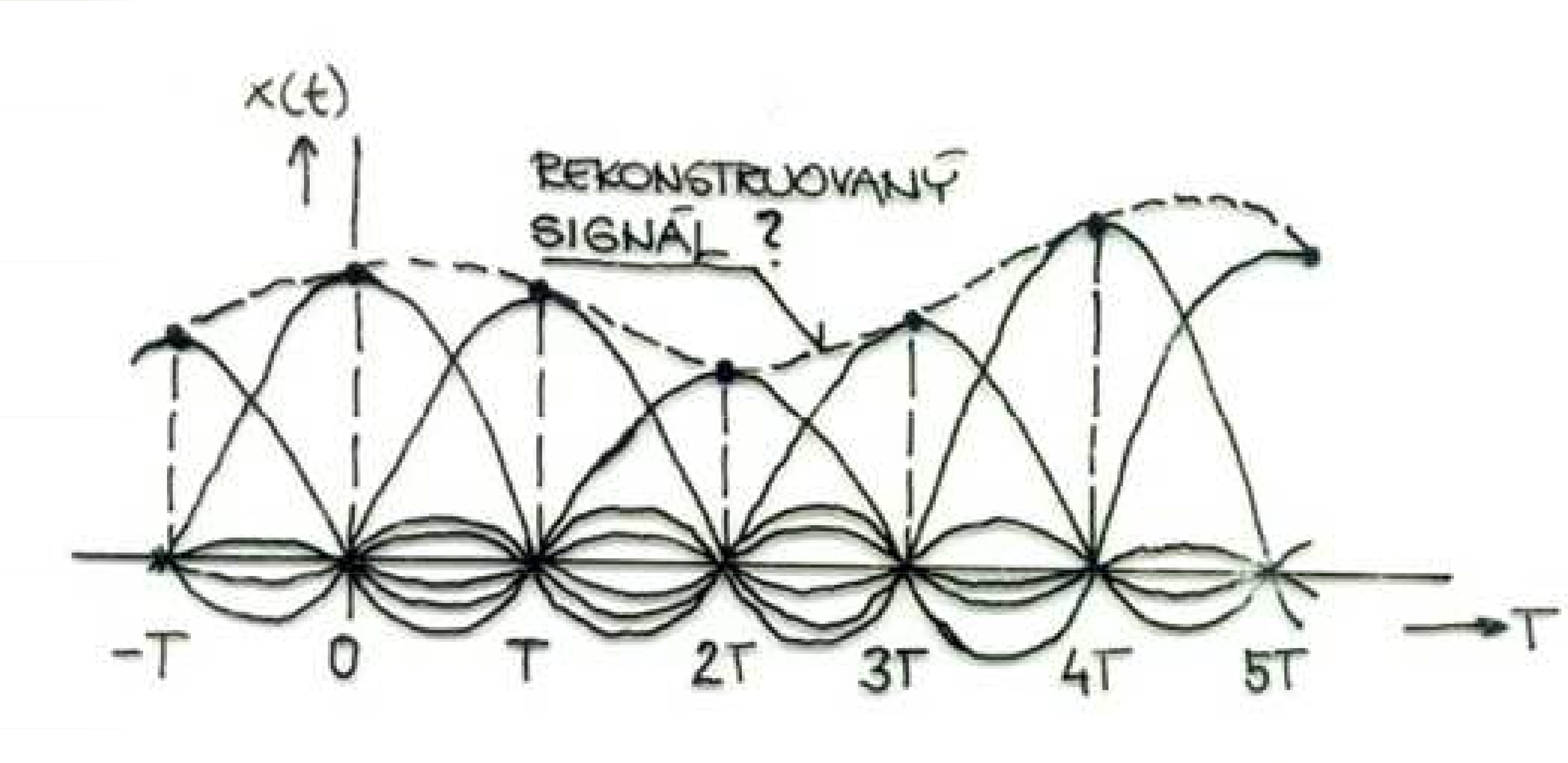
$$= \frac{1}{f_{vz}} \sum_{n=-\infty}^{\infty} x(nT_{vz}). \frac{e^{i2\pi(t-nT_{vz}).f_{vz}/2} - e^{-i2\pi(t-nT_{vz}).f_{vz}/2}}{i2\pi(t-nT_{vz})} =$$

$$= \sum_{n=-\infty}^{\infty} x(nT_{vz}).Si\left(\pi(t-nT_{vz}).\frac{1}{T_{vz}} \right)$$

tj. původní funkce je dána nekonečným součtem vzorkovacích funkcí, které procházejí každou hodnotou $x_a(t)$ z nekonečného počtu vzorků navzorkované posloupnosti.

© Institut biostatistiky a analýz

REKONSTRUKCE SPOJITÉ FUNKCE Z NAVZORKOVANÉ POSLOUPNOSTI



© Institut biostatistiky a analýz

IX. FREKVENČNÍ TRANSFORMACE

ČASOVÉ ŘADY

© Institut biostatistiky a analýz

HARMONICKÁ ANALÝZA DISKRÉTNÍCH POSLOUPNOSTÍ

- ☑ diskrétní Fourierova řada
- ☑ Fourierova transformace s diskrétním časem - DTFT
- ☑ diskrétní Fourierova transformace - DFT
- ☑ rychlá Fourierova transformace - FFT

© Institut biostatistiky a analýz

ROZKLAD PERIODICKÉ ČASOVÉ ŘADY

☑ spojitá veličina – opakování

Fourierova řada (v komplexním tvaru)

$$f(t) = \sum_{n=-\infty}^{\infty} \dot{c}_n e^{in\Omega t} \qquad \Omega = 2\pi / T$$

kde $\dot{c}_n$ jsou komplexní Fourierovy koeficienty

$$\dot{c}_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t).e^{-in\Omega t} dt$$

Ω – úhlový kmitočet základní harmonické složky (základní harmonická);

© Institut biostatistiky a analýz

FOURIEROVA ŘADA PRO DISKRÉTNÍ posloupnosti

- ☑ nechť $x(kT_{vz})$ je periodická posloupnost s periodou NT_{vz}; pak $x(kT_{vz})$ lze rozložit pomocí komplexní exponenciální Fourierovy řady

$$x(kT_{vz}) = \sum_{n=0}^{N-1} c_n . \exp \frac{i2\pi nk}{N}, \quad k = 0, \pm 1, \pm 2, \ldots$$

kde

$$c_n = \frac{1}{N} \sum_{k=0}^{N-1} x(kT_{vz}) . \exp\left(- \frac{i2\pi kn}{N} \right), \quad n = 0, 1, \ldots, N-1$$

© Institut biostatistiky a analýz

FOURIEROVA ŘADA PRO DISKRÉTNÍ posloupnosti

- nechť $x(kT_{vz})$ je periodická posloupnost s periodou NT_{vz}; pak $x(kT_{vz})$ lze rozložit pomocí komplexní exponenciální Fourierovy řady

$$x(kT_{vz}) = \sum_{n=0}^{N-1} c_n . \exp \frac{i2\pi nk}{N}, \quad k = 0, \pm 1, \pm 2, \ldots$$

kde

$$c_n = \frac{1}{N} \sum_{k=0}^{N-1} x(kT_{vz}) . \exp\left(- \frac{i2\pi kn}{N} \right), \quad n = 0, 1, \ldots, N-1$$

ZÁVĚR

je-li časová řada periodická, je frekvenční spektrum diskrétní; tj. nezáleží na spojitosti či diskrétnosti časových dat, **důležitá je jejich periodičnost.**

© Institut biostatistiky a analýz

FOURIEROVA ŘADA
DŮKAZ

- změňme index sumace ve vztahu pro výpočet koeficientu $\dot{c}_n$

$$\dot{c}_n = \frac{1}{N}\sum_{m=0}^{N-1} x(mT_{vz}).\exp(-2i\pi mn/N)$$

$$x(kT_{vz}) = \sum_{n=0}^{N-1} \dot{c}_n.\exp(2i\pi nk/N) = \sum_{n=0}^{N-1}\frac{1}{N}\left(\sum_{m=0}^{N-1} x(mT_{vz}).\exp(-2i\pi mn/N)\right).\exp(2i\pi nk/N) =$$

$$= \frac{1}{N}\sum_{m=0}^{N-1} x(mT_{vz}).\sum_{n=0}^{N-1}\exp[2i\pi n(k-m)/N],$$

© Institut biostatistiky a analýz

FOURIEROVA ŘADA
DŮKAZ

potom

$$\text{pro } k = m \text{ je } \sum_{n=0}^{N-1} \exp[2i\pi n(k-m)/N] = N$$

$$\text{pro } k \neq m \text{ je } \sum_{n=0}^{N-1} \exp[2i\pi n(k-m)/N] = \frac{1-\exp[2i\pi N(k-m)/N]}{1-\exp[2i\pi(k-m)/N]} = 0$$

(součet N členů geometrické posloupnosti $s_n = a_1 \frac{1-q^n}{1-q}$)

$$\boxed{x(kT_{vz})} = \frac{1}{N}\sum_{m=0}^{N-1} x(mT_{vz}).\sum_{n=0}^{N-1} \exp[2i\pi n(k-m)/N] =$$

$$= \frac{1}{N} x(kT_{vz}).N = \boxed{x(kT_{vz}) \qquad \text{c.b.d.}}$$

FOURIEROVA ŘADA
PŘÍKLAD

$x(kT_{vz}) = A.\cos(2\pi k/N)$ je periodická posloupnost s periodou N

$$A.\cos\frac{2\pi k}{N} = \frac{A}{2}.\left[\exp\frac{2i\pi k}{N} + \exp\left(-\frac{2i\pi k}{N}\right)\right]$$

Nyní, protože

$$\exp\frac{2i\pi k(N-1)}{N} = \exp\frac{2i\pi kN}{N}.\exp\left(-\frac{2i\pi k}{N}\right) = \exp\left(-\frac{2i\pi k}{N}\right);$$

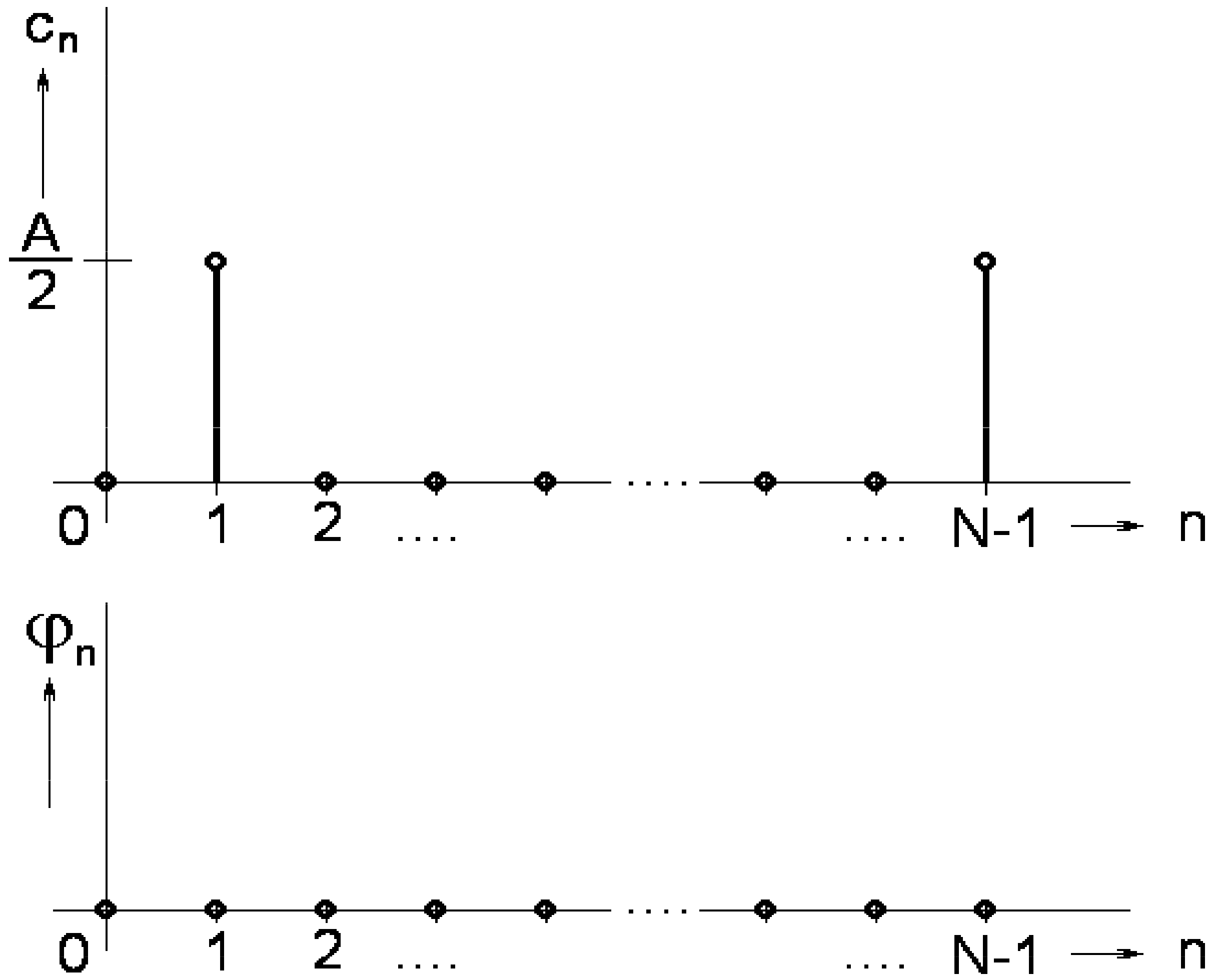
proto

$$A.\cos\frac{2\pi k}{N} = \frac{A}{2}.\left[\exp\frac{2i\pi k}{N} + \exp\left(\frac{2i\pi(N-1)k}{N}\right)\right]$$

$$a_1 = \frac{A}{2}, \quad a_{N-1} = \frac{A}{2}, \quad a_n = 0 \text{ pro všechna jiná n}$$

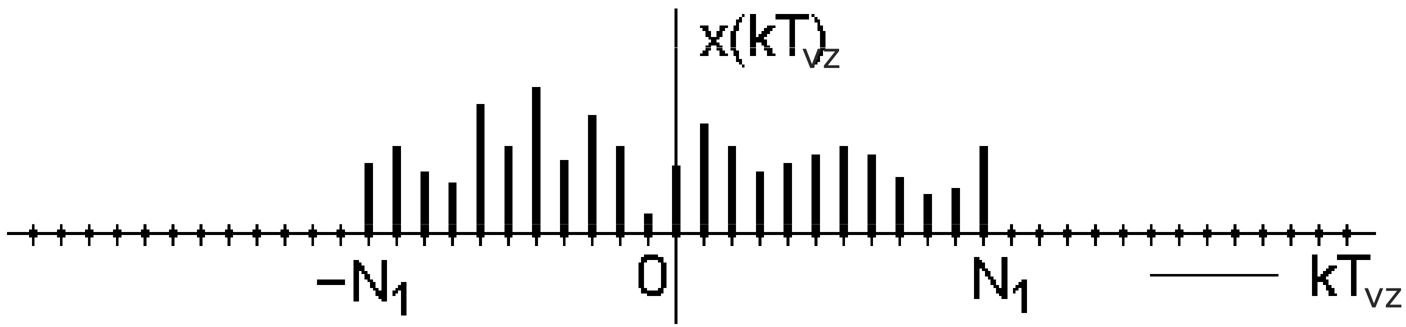
© Institut biostatistiky a analýz

FOURIEROVA ŘADA
PŘÍKLAD



FOURIEROVA TRANSFORMACE S DISKRÉTNÍM ČASEM - DTFT

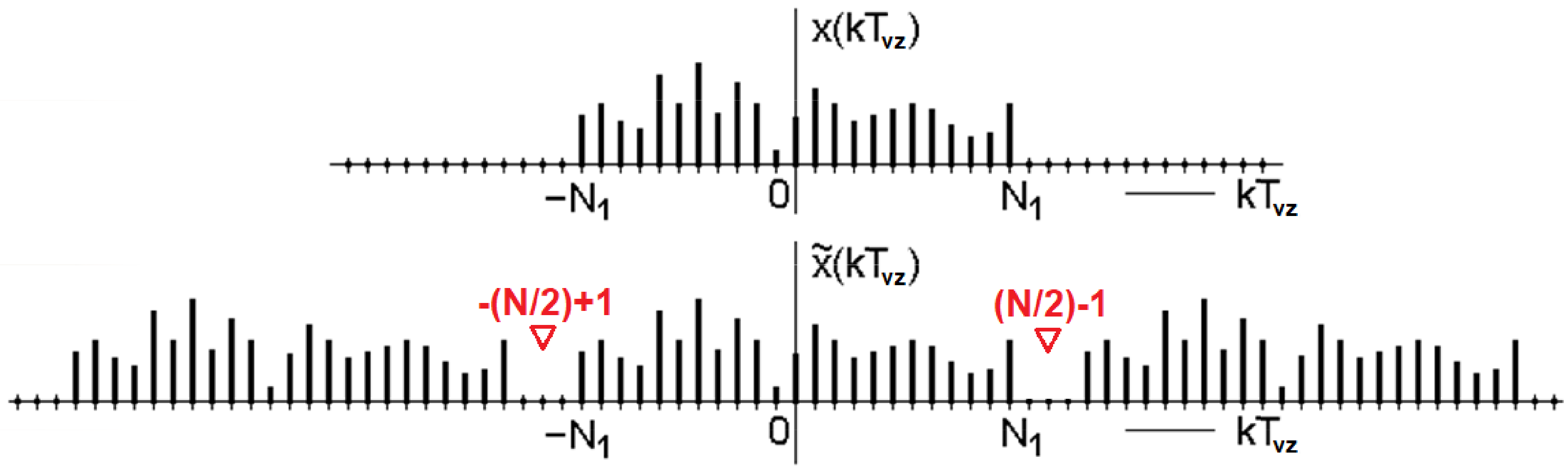
nechť $x(kT_{vz})$ je časově omezená posloupnost s diskrétním časem s $x(kT_{vz})=0$ pro všechna celá $k>N_1$ a $k<-N_1$, kde N_1 je celočíselná konstanta



FOURIEROVA TRANSFORMACE S DISKRÉTNÍM ČASEM - DTFT

dále, nechť pro kladné sudé celé číslo $N>2N_1$ označíme $\tilde{x}_N(kT_{vz})$ periodický signál s periodou NT, který je $x(kT_{vz})$ pro k = -N/2, -(N/2)+1, …, -1, 0, 1, …, (N/2)-1.

z definice $\tilde{x}_N(kT_{vz})$ máme $x(kT_{vz}) = \lim_{N\to\infty} \tilde{x}_N(kT_{vz})$



FOURIEROVA TRANSFORMACE S DISKRÉTNÍM ČASEM - DTFT

- ☑ protože $\tilde{x}_N(kT_{vz})$ je periodická posloupnost s periodou NT_{vz}, má Fourierovu řadu

$$\tilde{x}_N(kT_{vz}) = \sum_{n=0}^{N-1} \dot{c}_n . \exp \frac{2i\pi nk}{N}$$

kde

$$\dot{c}_n = \frac{1}{N} \sum_{k=-N/2}^{(N/2)-1} \tilde{x}_N(kT_{vz}) . \exp\left(-\frac{2i\pi kn}{N}\right), \quad n = 0,1,\ldots,N-1$$

© Institut biostatistiky a analýz

FOURIEROVA TRANSFORMACE S DISKRÉTNÍM ČASEM - DTFT

- Z definice $\tilde{x}_N(kT_{vz})$ vyplývá, že poslední uvedený vztah lze přepsat do tvaru

$$\dot{c}_n = \frac{1}{N}\sum_{k=-\infty}^{\infty} x(kT_{vz}).\exp\left(-\frac{2i\pi knT_{vz}}{NT_{vz}}\right), \quad n = 0,1,\ldots,N-1,$$

a potom

$$X(\omega) = \sum_{k=-\infty}^{\infty} x(kT_{vz}).\exp(-ik\omega T_{vz}), \qquad \omega = 2\pi n/NT_{vz}$$

kde ω je pro N→∞ spojitá (nediskrétní) veličina.

© Institut biostatistiky a analýz

FOURIEROVA TRANSFORMACE S DISKRÉTNÍM ČASEM - DTFT

ZÁVĚR

je-li časová řada neperiodická a nekonečná, je frekvenční spektrum spojité; tj. nezáleží na spojitosti či diskrétnosti časových dat, **důležitá je jejich neperiodičnost a nekonečná doba trvání.**

DISKRÉTNÍ FOURIEROVA TRANSFORMACE - DFT

- aby bylo možné počítat s frekvenčním spektrem na počítači, je třeba spektrální funkci diskretizovat;
- předpokládejme, že diskrétní veličina $x(nT_{vz})=0$ pro $n < 0$ a $n \geq N-1$, pak DFT je definována vztahem

$$X(k\Omega) = \sum_{n=0}^{N-1} x(nT_{vz}).e^{-ik\Omega nT} = \sum_{n=0}^{N-1} x(nT_{vz}).e^{-ik\frac{2\pi}{NT_{vz}}nT_{vz}} = \sum_{n=0}^{N-1} x(nT_{vz}).e^{-i2\pi kn/N}$$

$$X(k) = \sum_{n=0}^{N-1} x(n).e^{-i2\pi kn/N}$$

© Institut biostatistiky a analýz

ZPĚTNÁ DISKRÉTNÍ FT – DFT^{-1}

$$x(nT_{vz}) = \frac{1}{N}\sum_{k=0}^{N-1} X(k\Omega).e^{inT_{vz}k\Omega} = \frac{1}{N}\sum_{k=0}^{N-1} X(k\Omega).e^{i2\pi kn/N}$$

$$x(n) = \frac{1}{N}\sum_{k=0}^{N-1} X(k).e^{i2\pi kn/N}$$

© Institut biostatistiky a analýz

INVERZIBILITA DFT

$$\mathcal{DFT}^{-1}\{\mathcal{DFT}(x)\} = x$$

$$x(mT_{vz}) = \frac{1}{N}\sum_{k=0}^{N-1} X(k\Omega).e^{imT_{vz}k\Omega} = \frac{1}{N}\sum_{k=0}^{N-1}\left(\sum_{n=0}^{N-1} x(nT_{vz}).e^{-ik\Omega nT_{vz}}\right).e^{imT_{vz}k\Omega} =$$

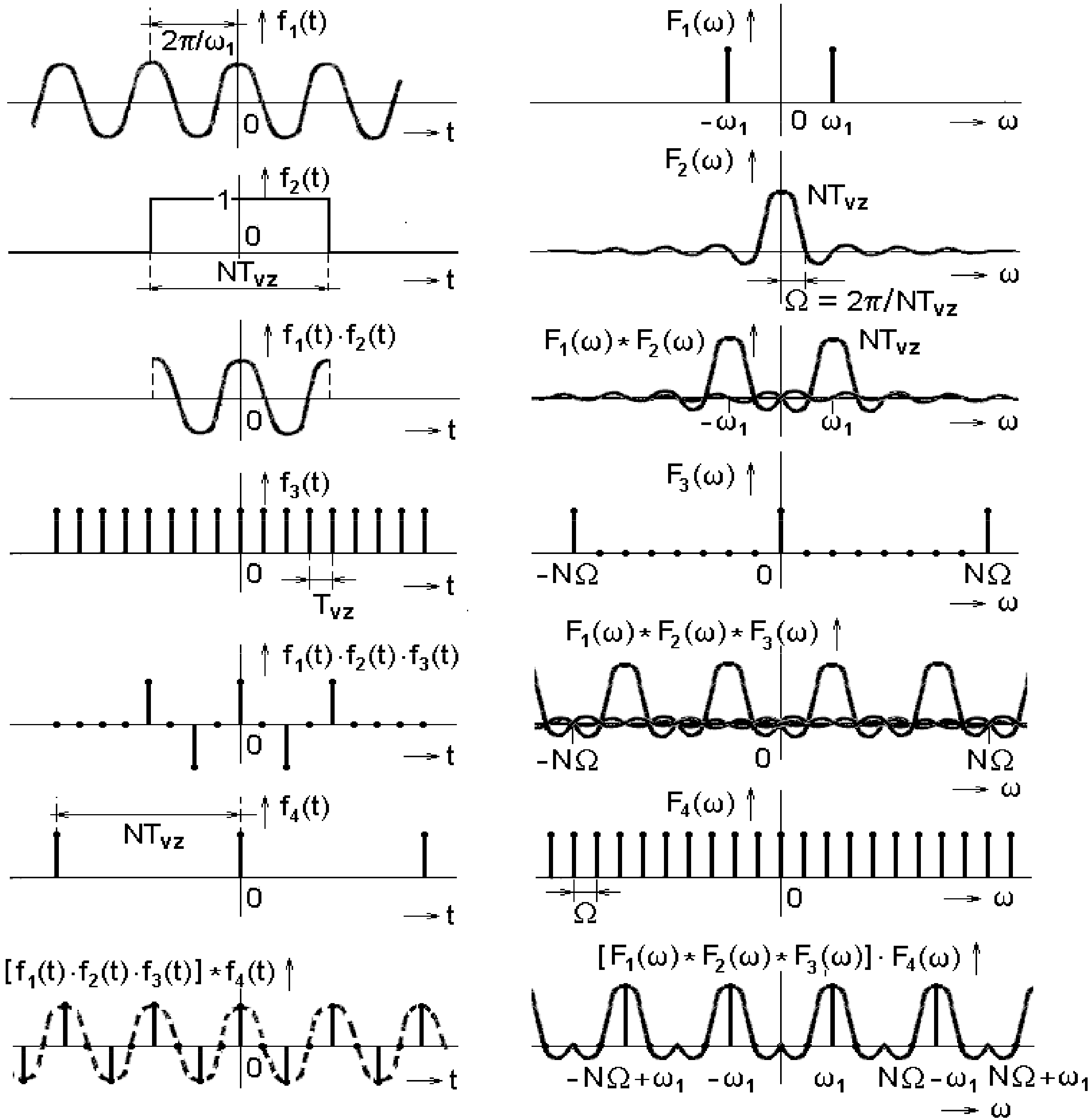
$$= \frac{1}{N}\sum_{n=0}^{N-1} x(nT_{vz}).\sum_{k=0}^{N-1}.e^{i(m-n)k\Omega T_{vz}} = \left| \begin{array}{c} \text{pro } m = n \text{ je } \sum_{k=0}^{N-1}.e^{i(m-n)k\Omega T_{vz}} = N \\ \text{pro } m \neq n \text{ je } \sum_{k=0}^{N-1}.e^{i(m-n)k\Omega T_{vz}} = \sum_{n=0}^{N-1}\frac{1-e^{i(m-n)N\Omega T_{vz}}}{1-e^{i(m-n)\Omega T_{vz}}} = 0 \end{array} \right| =$$

$$= \frac{1}{N} x(mT_{vz}).N = x(mT_{vz}) \qquad \Omega = 2\pi / NT_{vz}$$

© Institut biostatistiky a analýz

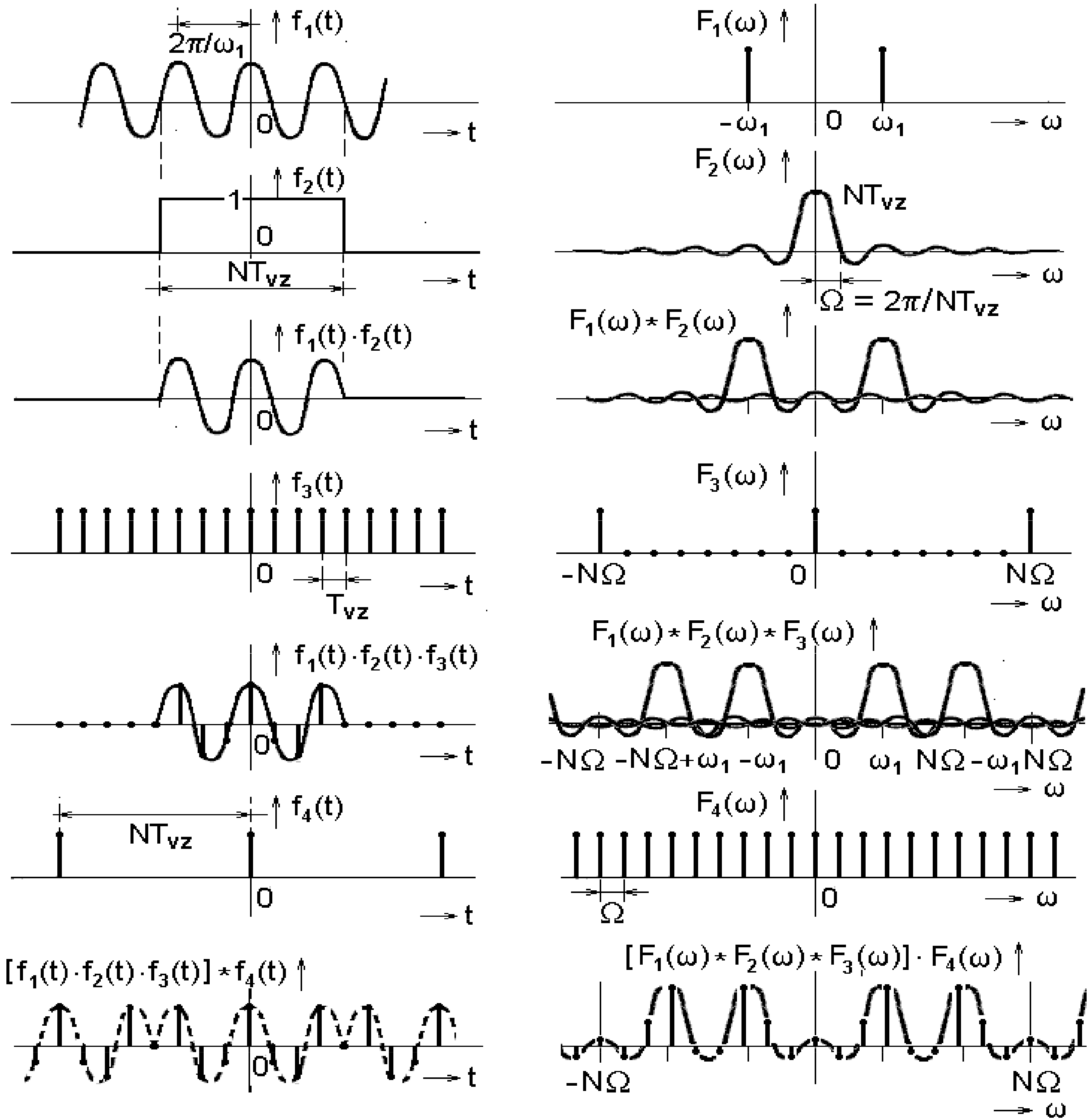
DFT

$\omega_1 = 2\Omega$
$= 4\pi/NT_{vz}$



DFT

$\omega_1 = 2{,}5\Omega = 5\pi/NT_{vz}$



© Institut biostatistiky a analýz

DISKRÉTNÍ FOURIEROVA TRANSFORMACE - DFT

ZÁVĚR

je-li časová řada konečná a frekvenční spektrum diskrétní, pak její frekvenční spektrum je rovno spektru diskrétní Fourierovy řady s periodou rovnou době trvání časové řady

© Institut biostatistiky a analýz