

B2M31SYN – 9. PŘEDNÁŠKA

7. prosince 2016

Granulační syntéza

Konkatenační syntéza

Nelineární funkce

Tvarovací syntéza

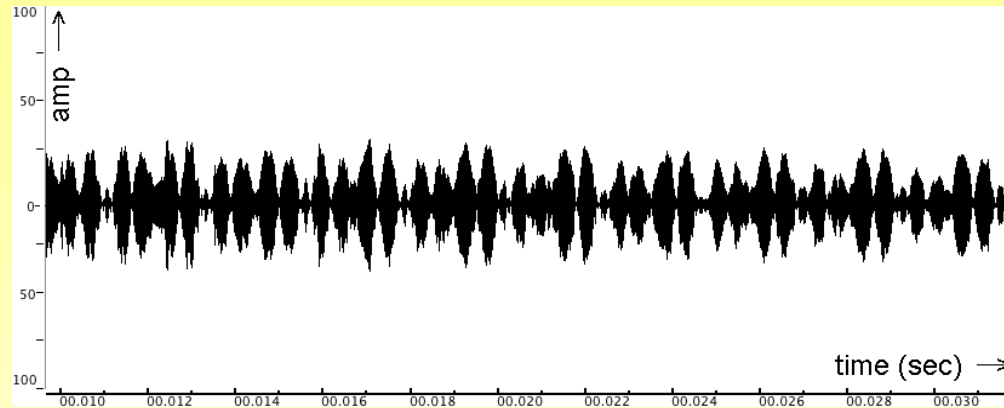
Granulační syntéza

- Jako alternativu k popisu pomocí sinusovek při úvahách o zvuku navrhl Gabor (1947) "akustická kvanta".
- Klasická definice *granulační syntézy* říká, že akustické kvanta jsou malé samostané exploze "harmonických oscilací" (např. sinusových vln) s amplitudovou obálkou založenou na Gaussově distribuci.
- Barry Truax jako první vytvořil program umožňující realizaci granulační syntézy v reálném čase. Za nejlepší jeho skladbu používající tento systém se považuje *Riverrun (1986)*.

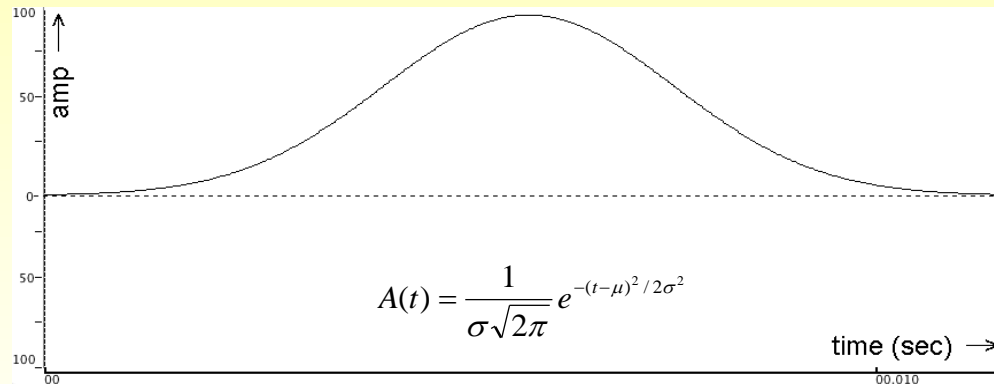


Postup při granulační syntéze

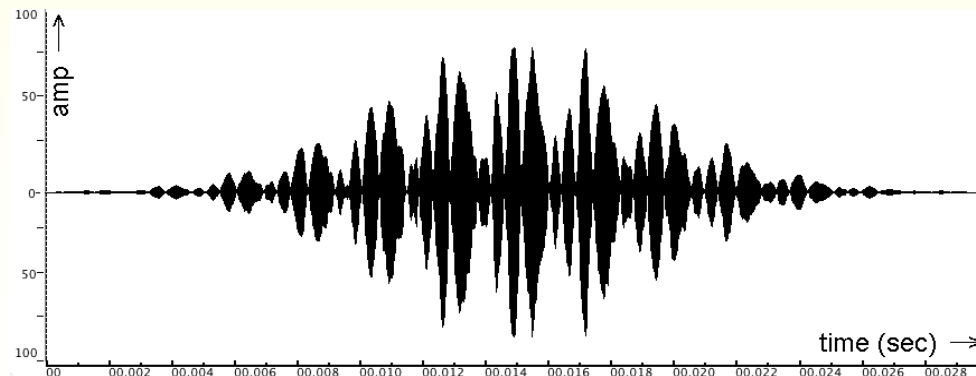
Vzorkovaný
audio
signál



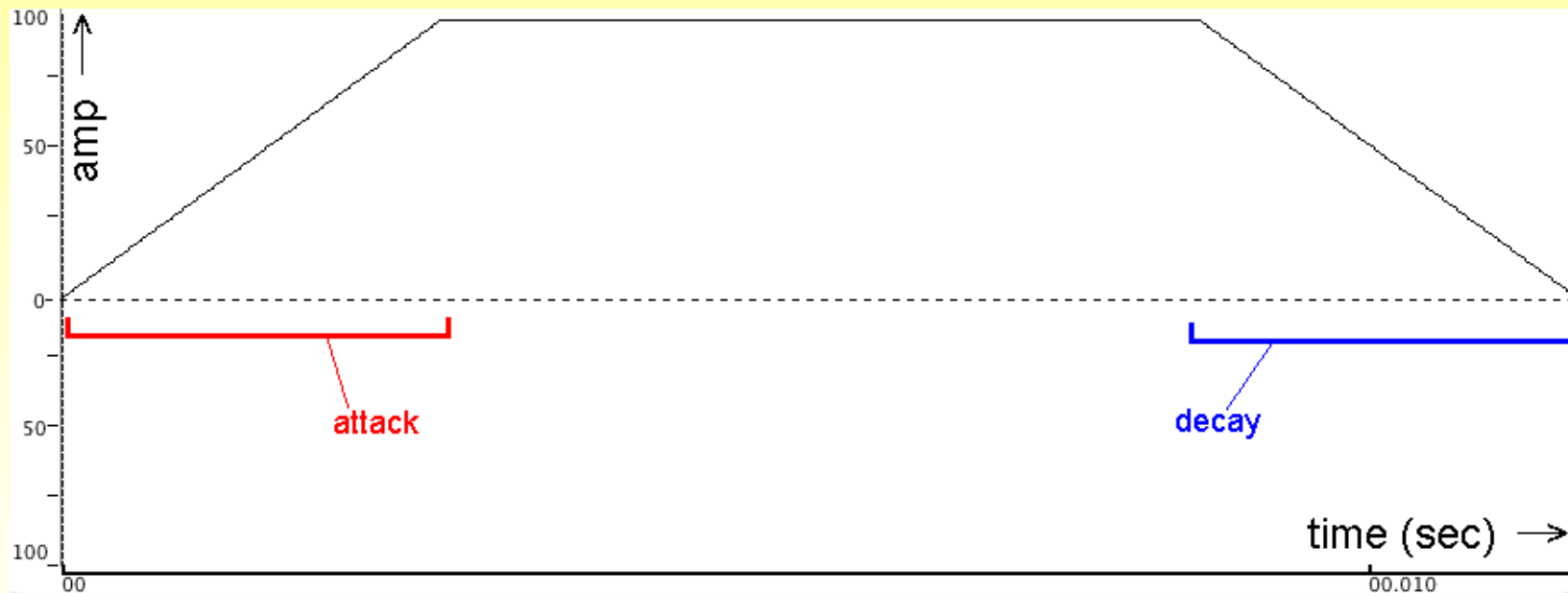
Gaussovská
obálka



zvukové
zrno

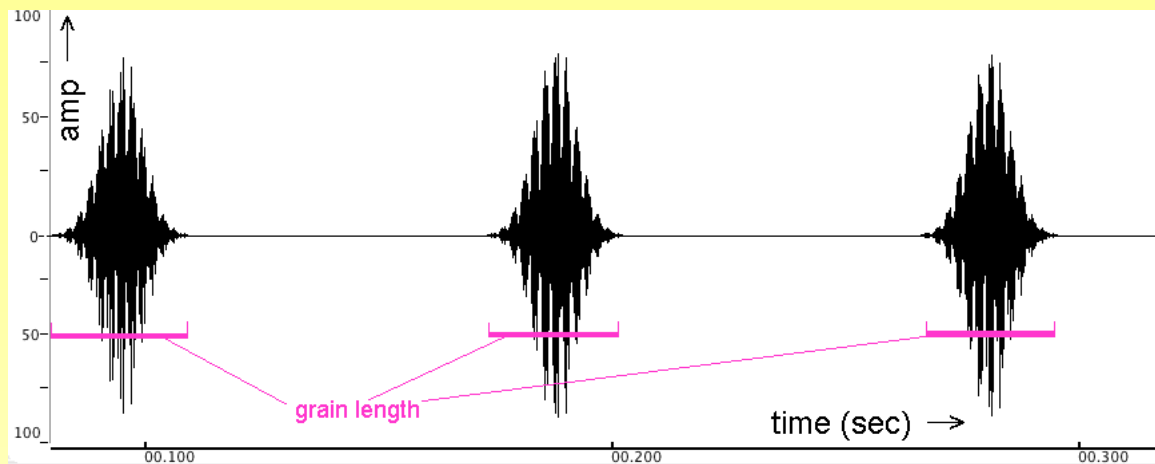


Další obálky

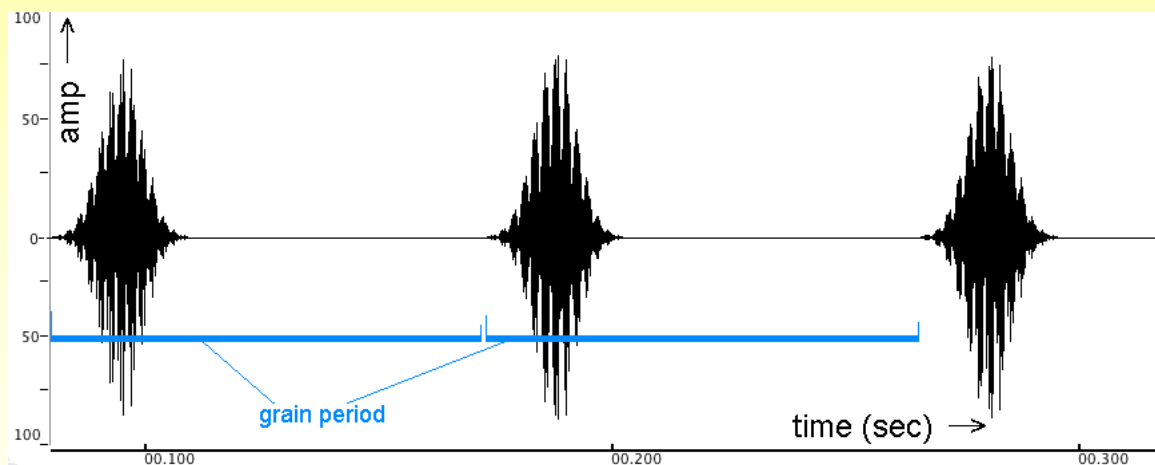


Trapezoidální obálka (Truax)

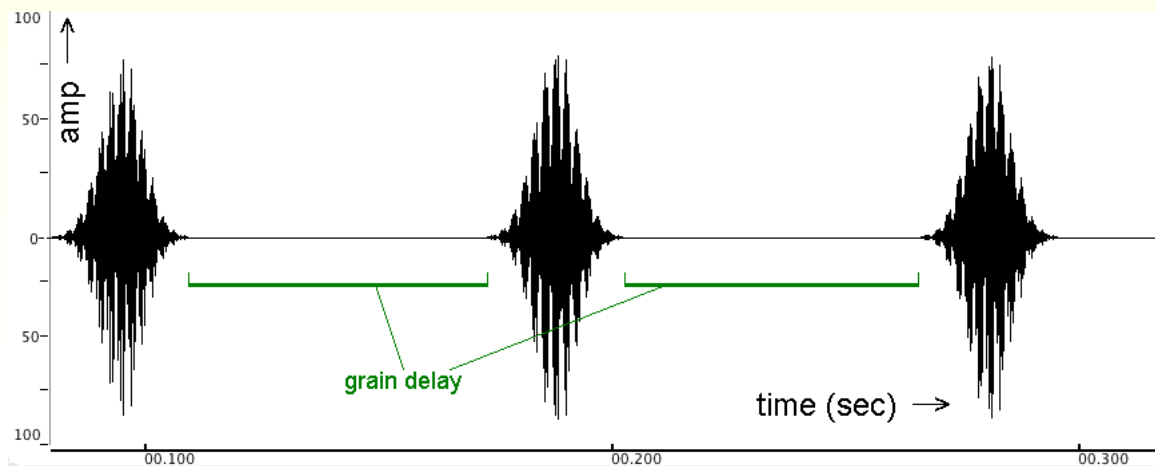
Délka
zrna



Perioda



Zpoždění



Délka zrna

- Typická délka zrn je mezi 1 a 50 ms.
- Abychom slyšeli události zněle, je na vysokých frekvencích jejich minimální délka 13 ms a na nízkých frekvencích 45 ms.
- Délky větší než 50 ms vytváří dojem separovaných zvukových událostí.

Granulační syntéza

Technika není vhodná pro klasické hudební barvy zvuku

- Vyniká v některých přirozených zvucích, které je obtížné generovat jinou metodou (např. praskající oheň, bublající voda, poryvy větru, výbuchy).

http://en.wikipedia.org/wiki/Granular_synthesis

Historie syntézy řeči

- **Mechanické syntetizéry**

- Kratzenstein (1779) (první pokusy - samohlásky)
- Kempelen (1791) (mluvící stroj)

- **Elektronické syntetizéry**

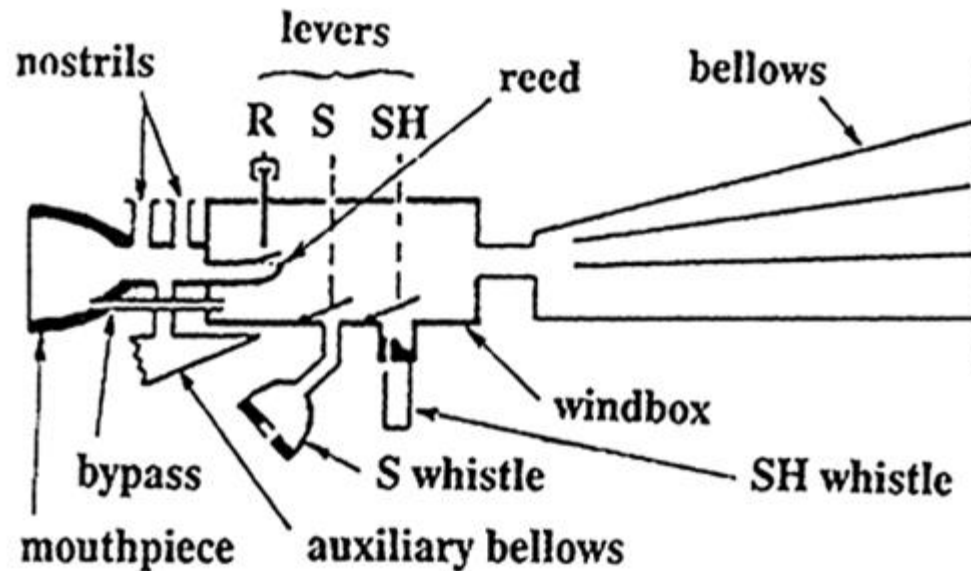
- Stewart (1922) (dva rezonátory)
- Dudley (1939) (Voder - první syntetizér souvislé řeči)
- Fant (1953) (OVE I - kaskádní formantový syntetizér)

- **Digitální syntetizéry**

- Klatt (1979) (MITalk, Klattalk - nejlepší formantová syntéza)

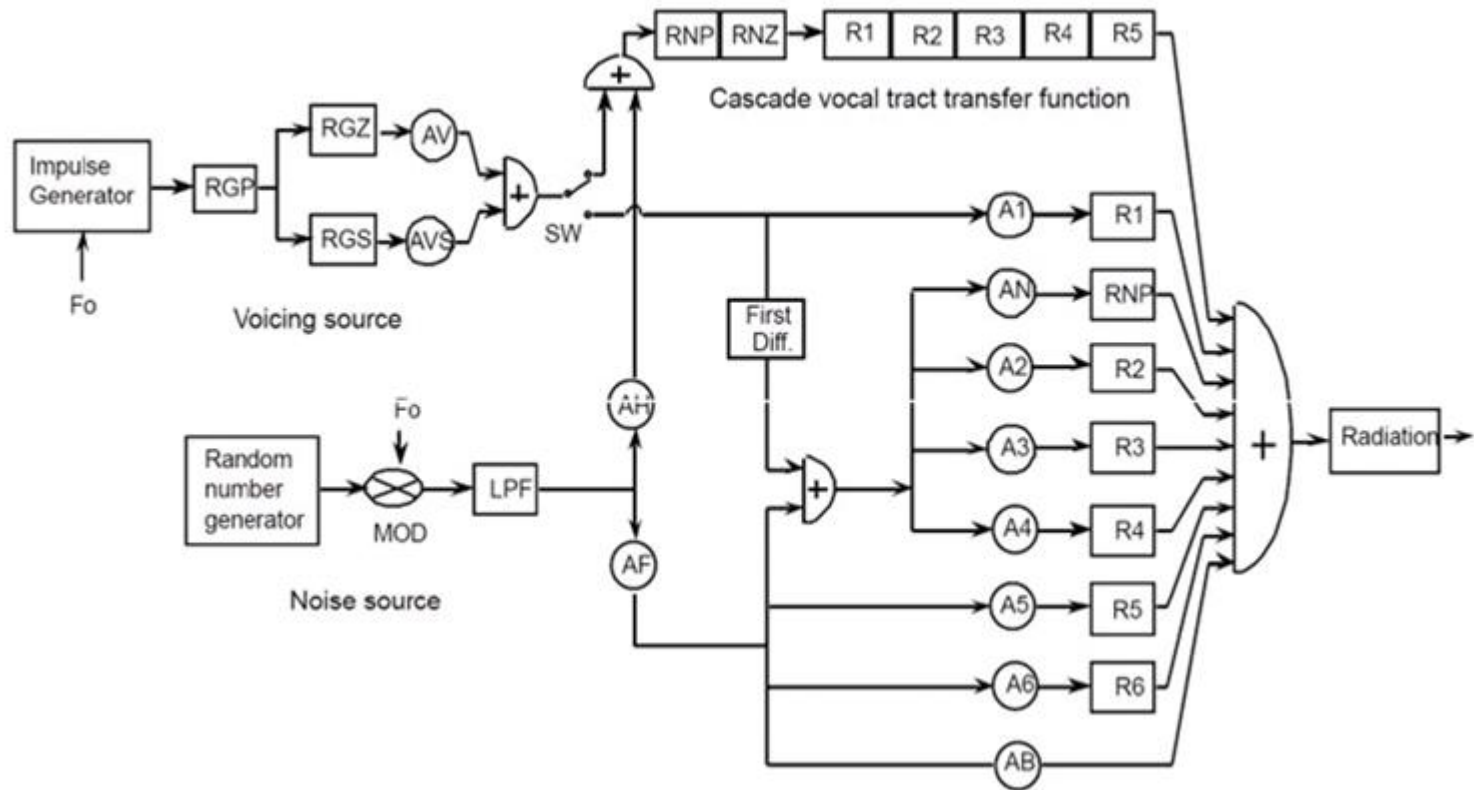
Kempelenův syntezátor

- Jednalo se o mechanické zařízení napodobující artikulační orgány člověka



Klatt

- Blokové schéma Klattova syntezátoru



Historie české syntézy řeči

- **Kaňka**

- první pokusy (20. léta 20. století)

- **Janota**

- první český syntetizér řeči (1964)

- **Ptáček, Borovičková, Maláč**

- OVED 1 - první syntetizér souvislé řeči (1968)

- HO2 - první český TTS systém (1977)

- **ÚŘE AV ČR, Ptáček**

- PCVOX - první český LPC TTS systém (1990)

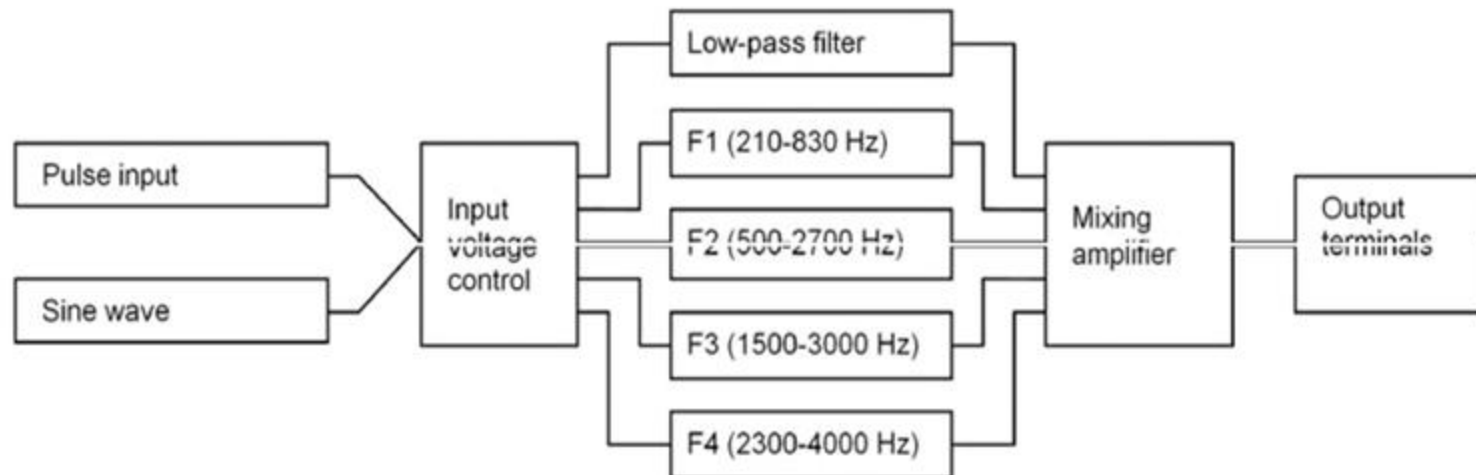
- Epos - první český open source TTS (1996)

- **ZČU Plzeň**

- ARTIC - první český Unit Selection TTS (2000)

Janota (1964)

- Jednoduchý paralelní formantový syntetizér češtiny



Metody syntézy řeči

- **Artikulační syntéza**
 - Modelování kompletního hlasového traktu
 - Výpočetně náročné, špatná kvalita
 - Prakticky se zatím nevyužívá, pouze ve výuce
- **Formantová syntéza**
 - Zjednodušené modelování hlasového traktu pomocí formantů a antiformantů
 - Praktické aplikace - nejpoužívanější v 60.- 80. letech
- **Konkatenační syntéza**
 - Řetězení segmentů řeči z inventáře řečových jednotek
 - Používá většina současných TTS

Formantová syntéza

- **Založena na akustické teorii vytváření řeči**
- **Zjednodušená simulace vytváření řeči člověkem**
 - **Zdroj buzení - generátor impulsů pro znělé zvuky a šum nebo smíšené buzení pro neznělé zvuky**
 - **Hlasový trakt - modelování pomocí filtru, jehož parametry odpovídají formantům a antiformantům hlasového traktu**
- **Syntéza podle pravidel – parametry se nastavují podle manuálně nalezených pravidel**
- **Dříve velmi úspěšná a používaná metoda**

Formantová syntéza

- **Výhody:**

- Jednoduchý model
- Snadné řízení prozodických charakteristik
- Konstantní kvalita
- Změny hlasu a emoce možno řídit podle pravidel
- Schopnost vytvářet plynulou kvalitní řeč

- **Nevýhody:**

- Pracné hledání a nastavování pravidel
- Vzájemná interakce mezi hodnotami parametrů
- Složité vytváření některých zvuků (explozívy)
- Nízká přirozenost řeči

Konkatenáční syntéza

- Používá přímo části přirozeného řečového signálu

Konkatenáční syntéza

- Používá přímo části přirozeného řečového signálu
- Předpokládá, že se řeč skládá z řečových (akustických) jednotek

Konkatenáční syntéza

- Používá přímo části přirozeného řečového signálu
- Předpokládá, že se řeč skládá z řečových (akustických) jednotek
- Řeč je možné rozdělit na segmenty odpovídající těmto jednotkám a uložit je do inventáře řečových jednotek

Konkatenáční syntéza

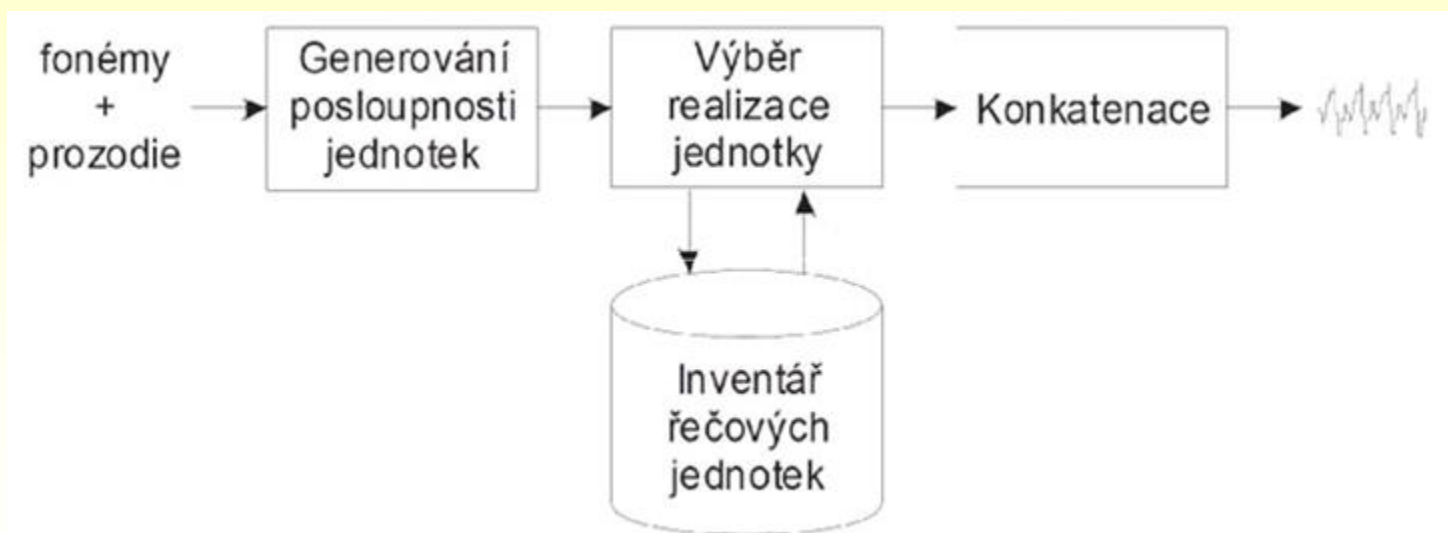
- Používá přímo části přirozeného řečového signálu
- Předpokládá, že se řeč skládá z řečových (akustických) jednotek
- Řeč je možné rozdělit na segmenty odpovídající těmto jednotkám a uložit je do inventáře řečových jednotek
- Řeč se vytváří řetězením (konkatenací) řečových segmentů uložených v inventáři řečových jednotek

Konkatenáční syntéza

- Používá přímo části přirozeného řečového signálu
- Předpokládá, že se řeč skládá z řečových (akustických) jednotek
- Řeč je možné rozdělit na segmenty odpovídající těmto jednotkám a uložit je do inventáře řečových jednotek
- Řeč se vytváří řetězením (konkatenací) řečových segmentů uložených v inventáři řečových jednotek
- Syntetická řeč napodobuje řečníka z inventáře

Konkatenáční syntéza

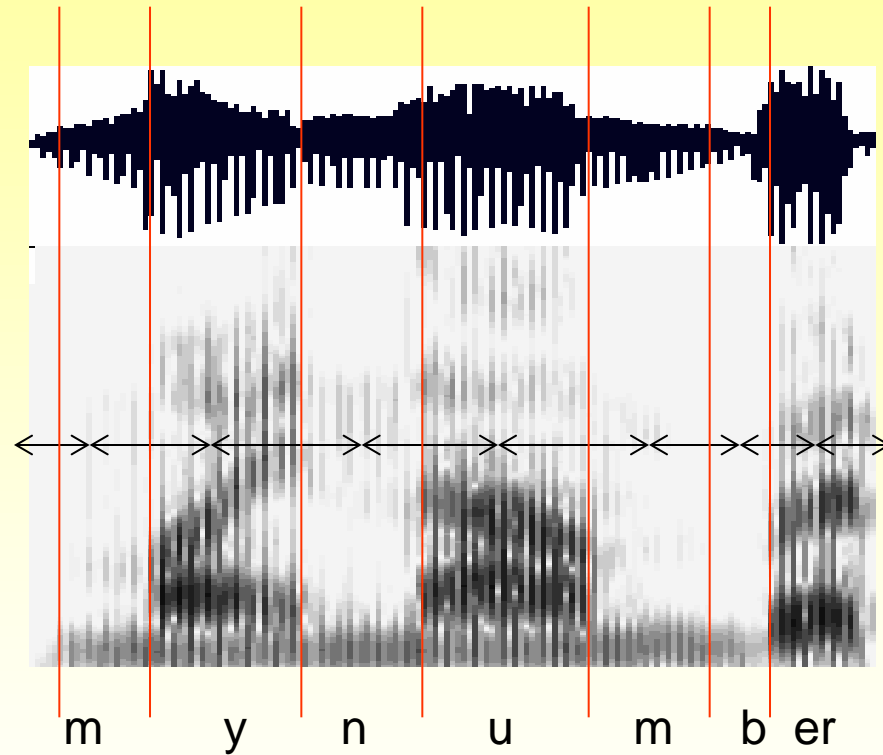
- Generování posloupnosti řečových jednotek
- Výběr vhodné realizace řečové jednotky
- Vlastní řetězení (konkatenace)
- Syntéza řízená daty – parametry syntetizéru se nastavují automaticky z řečových dat



Konkatenáční syntéza

- **Vytváření inventáře řečových jednotek**
 - Ruční vytváření
 - Automatické vytváření
- **Způsob reprezentace řečových jednotek**
 - Neparametrická (přímo vzorky řeči)
 - Parametrická (LPC, keprální, HMM)
- **spektrální/prozodické modifikace jednotek**
 - Bez modifikací (pouze řetězení)
 - S modifikacemi (snaha o minimalizaci nespojitostí na hranicích řetězených jednotek)
- **Možnosti generování řeči**
 - S omezeným slovníkem (věty ze specifické oblasti)
 - S neomezeným slovníkem (libovolné věty)

Konkatenáční syntéza



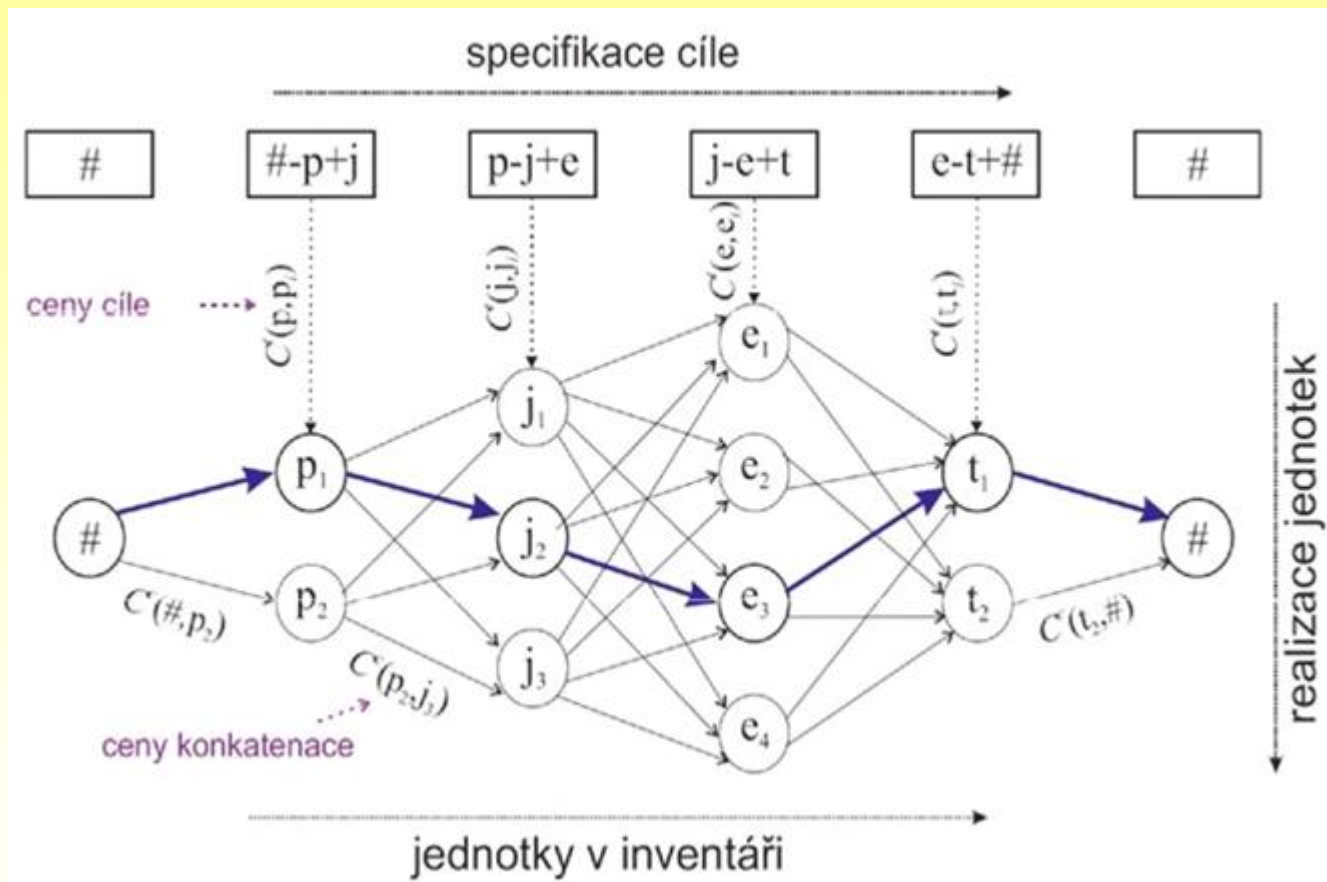
Porovnání technik syntézy řeči

Konkatenační	Formantová	Artikulační
Jednoduché	Střední složitost	Složitě
Slyšitelné přechody	Trochu umělé	Nejpřirozenější
Požadavek ohromných databází pro dosažení plynulosti	Požadavek na minimální množství řídicích dat	Obrovské databáze nejsou podmínkou
Kritické je efektivní vyhledávání	Modelování rezonátorů pomocí nul a pólů	Potřeba přesných a výkonných modelů

Korpusově orientovaná syntéza

- Zvláštní případ konkatenční syntézy
- Využití rozsáhlých a prozodicky pečlivě anotovaných řečových korpusů
- Více realizací každé řečové jednotky (v rozdílných fonetických, spektrálních i prozodických kontextech)
- Plně automatická konkatenční syntéza
- Všechny parametry se určují automaticky na základě dat z řečového korpusu
- Často tzv. neuniformní řečové jednotky (jednotky různého typu) –během syntézy se vybere typ a realizace jednotky
- => syntéza výběrem jednotek – Unit Selection

Korpusově orientovaná syntéza



České řečové syntezátory

- **Komerční**

- **ARTIC - ZČU Plzeň** <http://www.speechtech.cz/>

- <http://www.speechtech.cz/cs/produkty/demo.html#lva210>

- **Acapela Group - Eliška**

- <http://www.acapela-group.com/text-to-speech-interactive-demo.html>

- **RealSpeak - Zuzana**

- <http://www.freedomscientific.com/downloads/RealSpeak-Solo-Direct-Voi>

- **OpenSource**

- **Epos** <http://epos.ufe.cz/> (zde jsou i zajímavé odkazy)

- <http://epos.ufe.cz/cgi-bin/saye.cgi?lang=czech>

- **Festival (s podporou češtiny od Brailcomu)**

- <http://live.freebsoft.org/desktop#speech-synthesis>

- **Mbrola**

- <http://tcts.fpms.ac.be/synthesis/mbrola.html>

Aplikace syntézy řeči

- Pomůcky pro nevidomé a slabozraké
- Telekomunikační služby
- Automatické čtení (SMS, e-maily ...)
- Výuka jazyků
- Multimédia, komunikace člověk-počítač
- Mluvící hračky pro děti
- Výzkum (fonetika, lingvistika, akustika)

Aplikace syntézy řeči

- Pomůcky pro nevidomé a slabozraké
- Telekomunikační služby
- Automatické čtení (SMS, e-maily ...)
- Výuka jazyků
- Multimédia, komunikace člověk-počítač
- Mluvící hračky pro děti
- Výzkum (fonetika, lingvistika, akustika)

Literatura

- **J. Psutka, L. Müller, J. Matoušek, V. Radová:
Mluvíme s počítačem česky. Academia 2006**

Nelineární a tvarovací syntézy

- Metody vyvinuté J. C. Rissetem

Nelineární a tvarovací syntézy

- Metody vyvinuté J. C. Rissetem
- Rovněž nazývané “nelineární zkreslení”

Nelineární a tvarovací syntézy

- Metody vyvinuté J. C. Rissetem
- Rovněž nazývané “nelineární zkreslení”
- Modifikují (zkreslují) audio signál prostřednictvím “přenosové funkce”

Nelineární a tvarovací syntézy

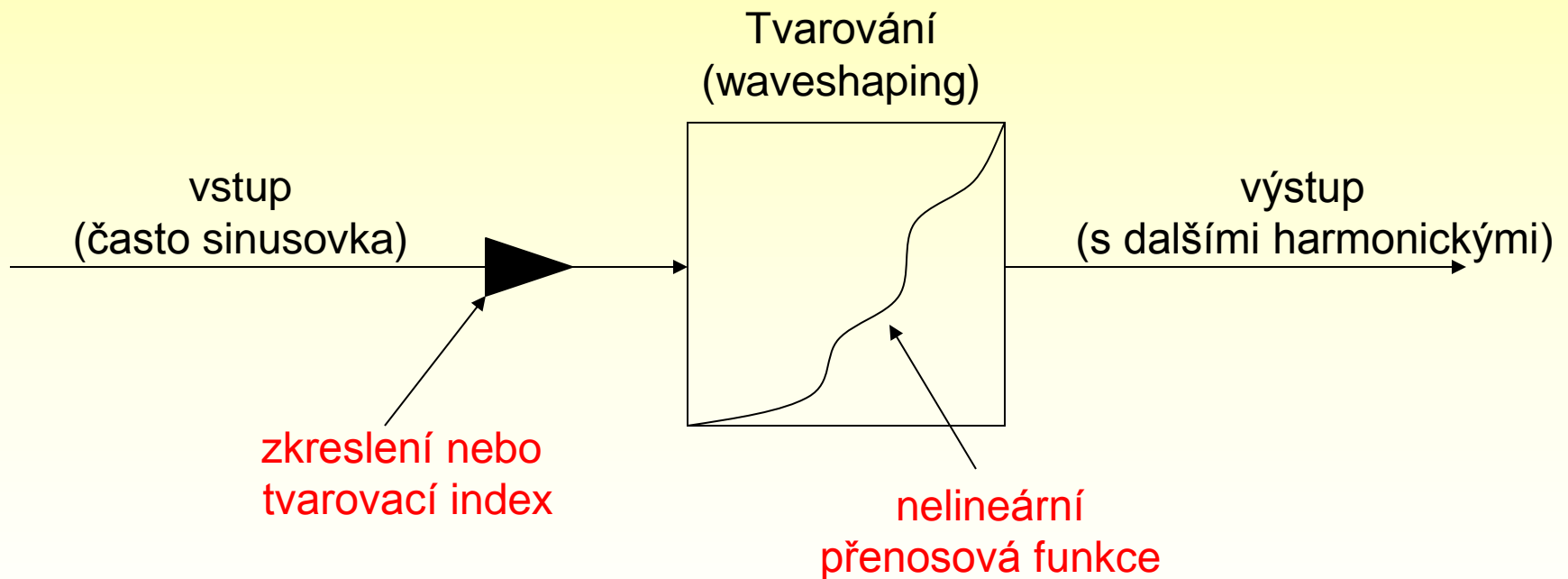
- Metody vyvinuté J. C. Rissetem
- Rovněž nazývané “nelineární zkreslení”
- Modifikují (zkreslují) audio signál prostřednictvím “přenosové funkce”
- Účinnost metody a možnosti tvorby dynamických změn barev je srovnatelná s FM

Nelineární a tvarovací syntézy

- Metody vyvinuté J. C. Rissetem
- Rovněž nazývané “nelineární zkreslení”
- Modifikují (zkreslují) audio signál prostřednictvím “přenosové funkce”
- Účinnost metody a možnosti tvorby dynamických změn barev je srovnatelná s FM
- Předností je přesný návrh pásmově omezeného spektra (bez “FM cvrlikání”)

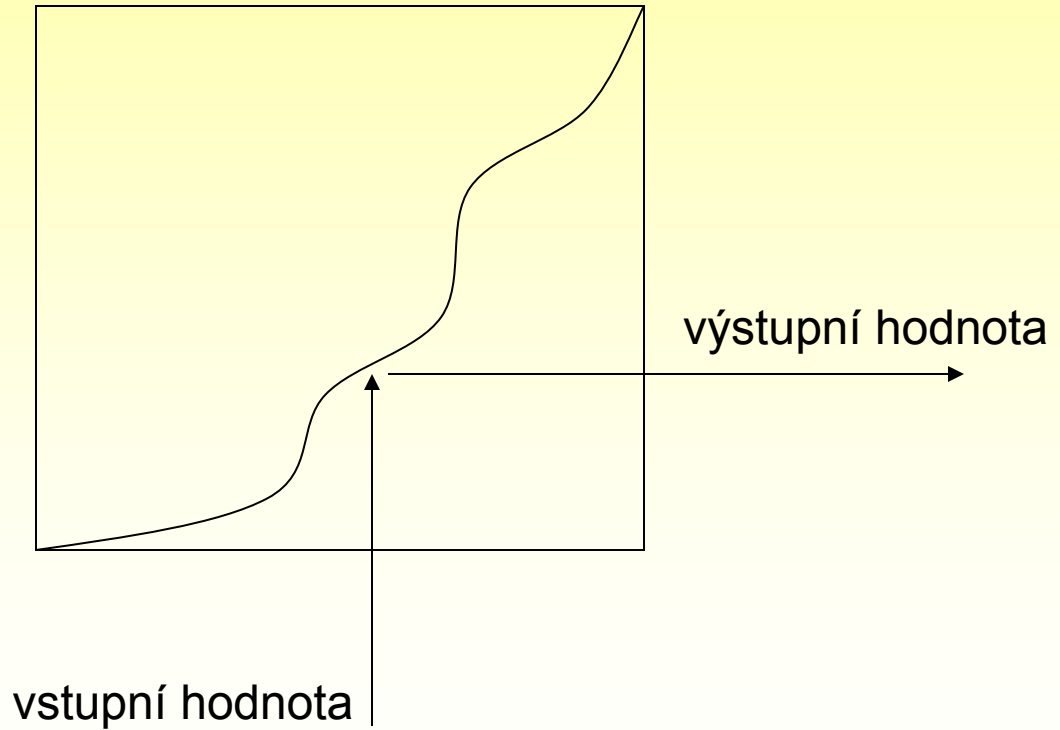
Tvarování

Technika, kterou lze vytvořit dynamická spektra

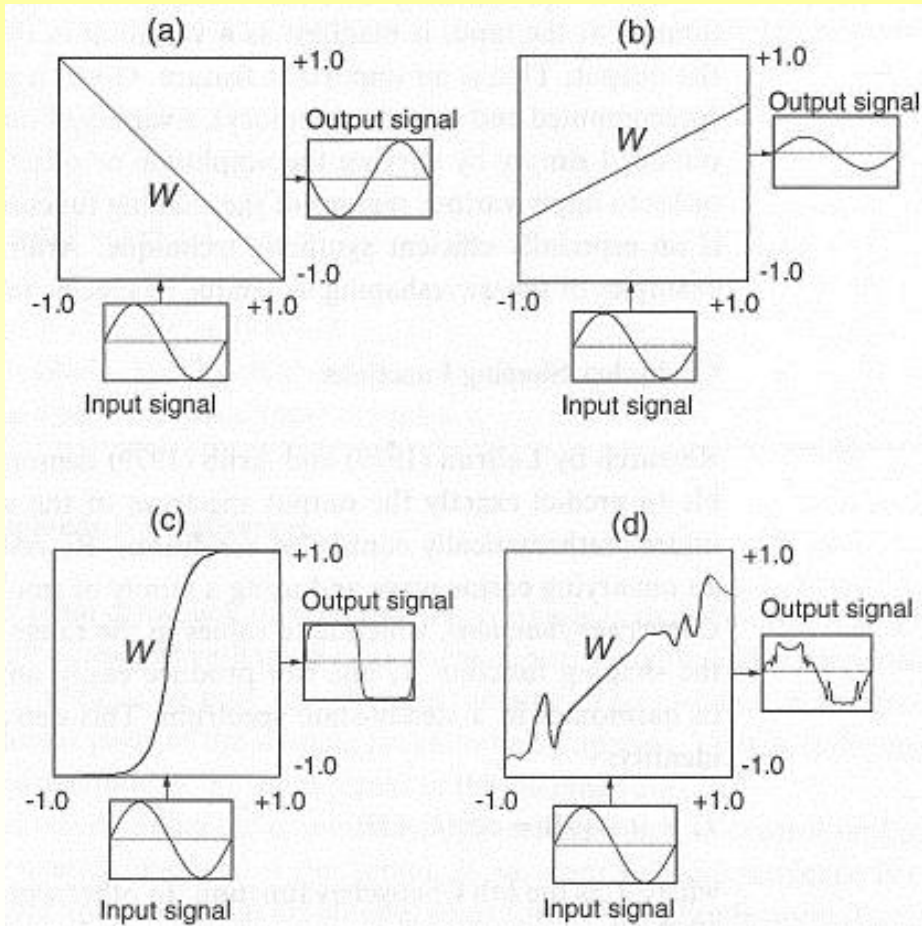


Tvarování

Přenosová funkce



Přenosové funkce



Tvarování může modifikovat vstupní signál v závislosti na přenosové funkci různými způsoby.

Harmonický průběh je:
(a) invertován,
(b) potlačen,
(c) oříznut,
(d) výrazně pozměněn.

Příklady přenosových funkcí

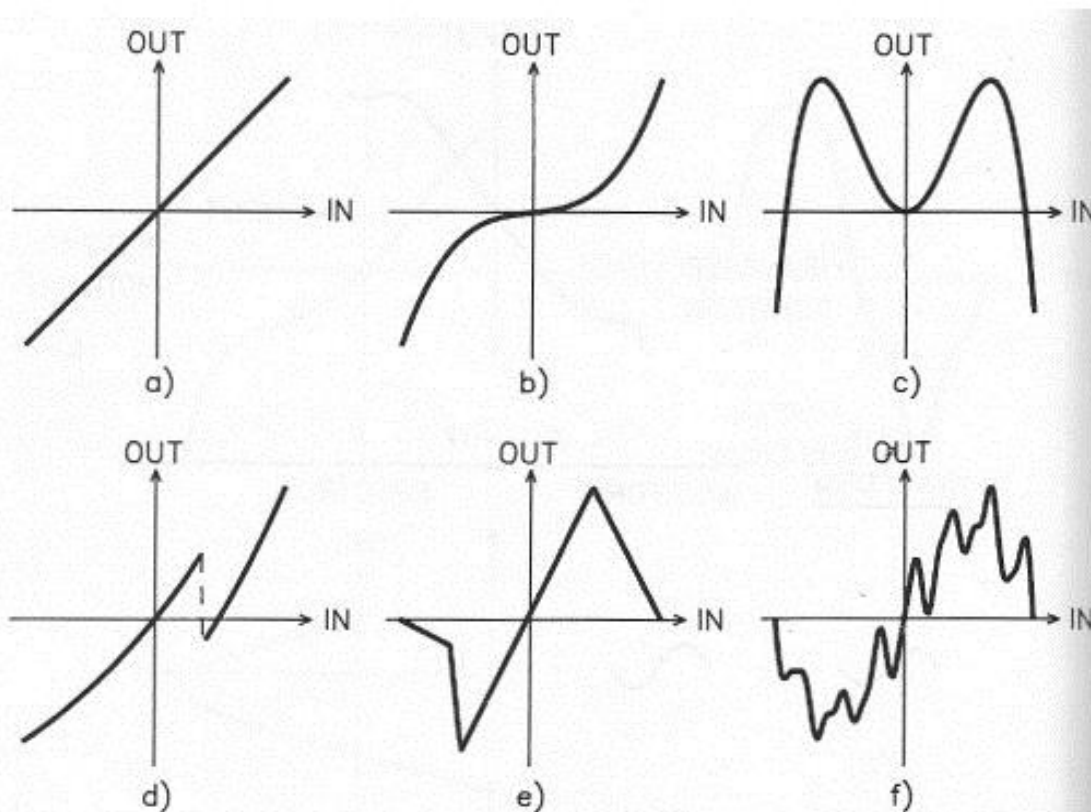


FIGURE 4.22 Six transfer functions: (a) linear, (b) odd, (c) even, (d) with an abrupt jump, (e) with sharp points, and (f) with ripple.

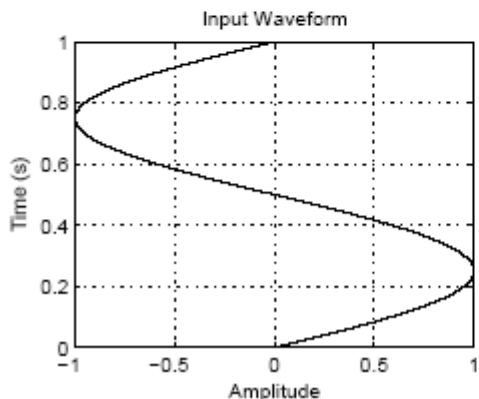
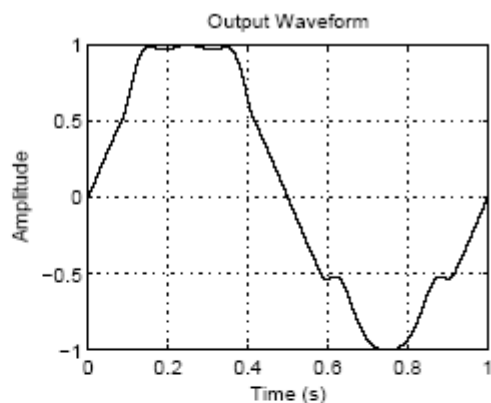
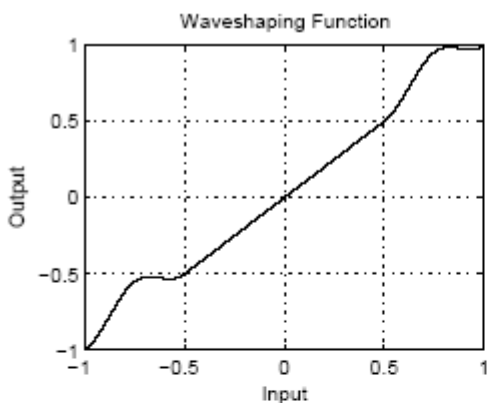
(a) lineární funkce nemění výsledné spektrum (na rozdíl od nelineární)

(b) lichá symetrická funkce generuje pouze liché harmonické

(c) sudá symetrická funkce generuje sudé harmonické

(d-f) funkce s prudkými změnami mohou zapříčinit aliasing

Příklady přenosových funkcí



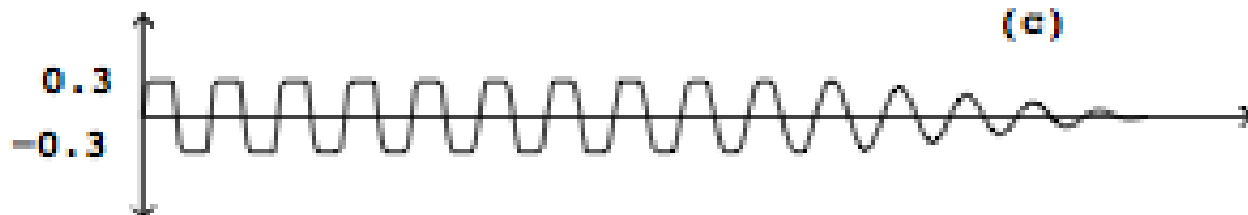
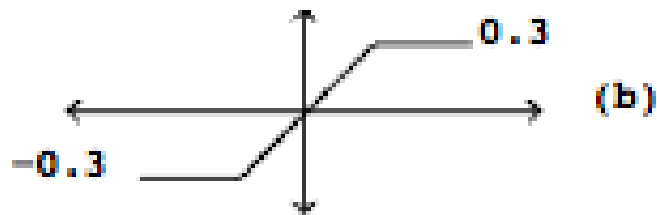
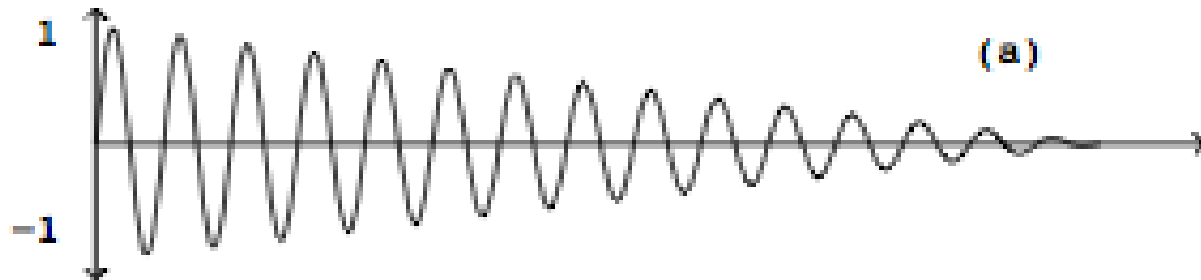
(a) lineární funkce nemění výsledné spektrum (na rozdíl od nelineární)

(b) lichá symetrická funkce generuje pouze liché harmonické

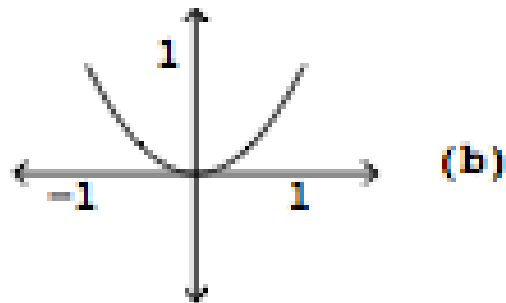
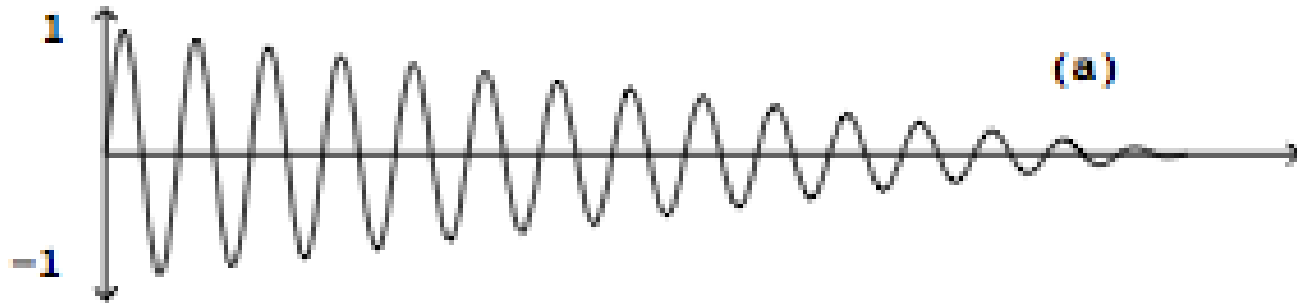
(c) sudá symetrická funkce generuje sudé harmonické

(d-f) funkce s prudkými změnami mohou zapříčinit aliasing

Dynamické nelineární zkreslení



Dynamické nelineární zkreslení



quadratic transfer function
 $y = x^2$



Intermodulace

- Pokud do přenosové funkce vstupuje více harmonických, vznikají další harmonické (intermodulační produkty)
- Čím více vstupních složek, tím více intermodulačních produktů

Návrh přenosových funkcí

- Přenosové funkce lze navrhovat různými způsoby včetně grafických

Návrh přenosových funkcí

- Přenosové funkce lze navrhovat různými způsoby včetně grafických
- Nicméně pro návrh přenosové funkce, která přesně definuje spektrum výstupního signálu, je nejlepší použít polynomů.

Návrh přenosových funkcí

- Přenosové funkce lze navrhovat různými způsoby včetně grafických
- Nicméně pro návrh přenosové funkce, která přesně definuje spektrum výstupního signálu, je nejlepší použít polynomů.
- Polynomy neprodukují harmonické výše než N:

$$F(x) = d_0 + d_1x + d_2x^2 + \dots + d_Nx^N$$

Čebyševovy polynomy

- Jsou užitečné při návrhu přenosové funkce

Čebyševovy polynomy

- Jsou užitečné při návrhu přenosové funkce
- Generují harmonické parciály

Čebyševovy polynomy

- Jsou užitečné při návrhu přenosové funkce
- Generují harmonické parciály
- Specifikace relativních amplitud

Čebyševovy polynomy

- Jsou užitečné při návrhu přenosové funkce
- Generují harmonické parciály
- Specifikace relativních amplitud
- Na vstupu je harmonický signál

Čebyševovy polynomy

- Jsou užitečné při návrhu přenosové funkce
- Generují harmonické parciály
- Specifikace relativních amplitud
- Na vstupu je harmonický signál
- Čebyševův polynom $k^{\text{tého}}$ řádu generuje právě k^{tou} harmonickou

Čebyševovy polynomy

- Jsou užitečné při návrhu přenosové funkce
- Generují harmonické parciály
- Specifikace relativních amplitud
- Na vstupu je harmonický signál
- Čebyševův polynom $k^{\text{tého}}$ řádu generuje právě k^{tou} harmonickou
- Čebyševovy polynomy různých řádů lze sčítat a vytvořit tak přenosovou funkci, která bude definovat amplitudové spektrum