Einführung in die Signalverarbeitung

Phonetik und Sprachverarbeitung, 2. Fachsemester, Block Sprachtechnologie I

Florian Schiel

Institut für Phonetik und Sprachverarbeitung, LMU München

Signalverarbeitung - Teil 3

Allgemeines

- Unterrichtssprache ist Deutsch (englische Fachbegriffe in Klammern)
- Fragen am besten sofort; besser einmal zuviel gefragt
- Literatur/Videos:
  - William Cox: Video-Tutorial Fourier-Transformation https://www.youtube.com/watch?v=FjmwwDHT98c
  - Barry van Veen: Video-Tutorial zum Küpfmüller-Tiefpass https://www.youtube.com/watch?v=Cb7NmiupCHw bei ca. 6min
  - Pfister B, Kaufmann T (2008): Sprachverarbeitung - Grundlagen und Methoden der Sprachsynthese und Spracherkennung. Springer-Verlag Berlin Heidelberg. Prentice-Hall, New Jersey, USA.

Periodische und nicht-periodische Signale

Wir unterscheiden in Zukunft folgende wichtige Signale:

- Sinusschwingung, definiert durch Frequenz f, Amplitude A und Phase ϕ

$$s(t) = A\sin(2\pi ft + \phi)$$

- Periodisches Signal = Signal wiederholt sich nach der *Periodendauer* (period time) T_0 exakt.
  Die Form der Schwingung spielt keine Rolle.
- Stochastisches Signal (Rauschen, noise) = Abtastwerte sind nicht durch eine Funktion determiniert, sondern nach statistischen Gesetzen.
  Spezialfall: *weißes Rauschen* : alle Frequenzen gleich häufig.
- Impulssignal = ein kurzer Impuls, der sich nicht wiederholt.
  Spezialfall: *Dirac-Impuls* = ein Impuls der Energie 1 bei $t = 0$.

Fouriersynthese und Fourierreihe

Jean Baptiste Joseph Fourier (1768–1830), Mathematiker

Fouriersynthese

Jedes beliebige periodische Signal mit Periodendauer T_0 lässt sich aus einer (unendliche) Summe von Cosinus- und Sinusschwingungen bilden, deren Frequenzen ganzzahlige Vielfache von $f_0 = 1/T_0$ sind.

$$s(t) = a_0 + \sum_{k=1}^{\infty} a_k cos(2\pi k f_0 t) + \sum_{k=1}^{\infty} b_k sin(2\pi k f_0 t)$$

f_0 wird oft *Grundfrequenz* genannt;
a_0 ist der *Gleichanteil* (*off-set, bias*),
d.h. eine Verschiebung des Signals nach oben oder unten.

Fouriersynthese und Fourierreihe

Tutoriat: harsyn

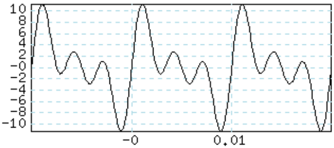
Fourierreihe und Linienspektrum

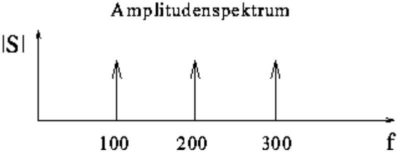
Fourierreihenentwicklung

Jedes periodische Signal lässt sich in (unendlich viele) Cosinus und Sinusschwingungen der Periode T_0 *zerlegen*

Linienspektrum

Die Aneinanderreihung der gefundenen Amplituden der einzelnen Sinusschwingungen nennt man Linienspektrum. Die erste Linie (bei $f = 0$) beschreibt den Gleichanteil.





Fouriertransformation

Gedankenexperiment:
Periodendauer T_0 eines periodischen Signals wird größer
$\rightarrow$ Grundfrequenz $f_0 = 1/T_0$ wird kleiner
$\rightarrow$ Abstände der Linien im Linienspektrum schrumpfen

Für $T_0 \rightarrow \infty$ besteht Signal nur noch aus einer einzelnen (theoretisch unendlich langen) Periode und die Linien im Spektrum wachsen zu einem *Kontinuum* zusammen.

So ein finites Signal ist aber ein beliebiges Signal (Teil eines periodischen, stochastisch, Impuls ...)
Fourierreihenentwicklung angewandt auf finites Signal
$\rightarrow$ *Fouriertransformation*
$\rightarrow$ kontinuierlicher spektraler Verlauf, genannt *Fourierspektrum*

Fouriertransformation

Fouriertransformation

Die Fouriertransformation berechnet aus einem beliebigen Signal $s(t)$ das kontinuierliche (komplexe) Fourierspektrum $S(f)$

$$S(f) = \mathcal{F}\{s(t)\} = \int_{-\infty}^{+\infty} s(t) e^{-j2\pi ft} dt$$

mit *Eulerformel*

$$e^{-j2\pi ft} = cos(2\pi ft) + jsin(2\pi ft)$$

und j ist die komplexe Einheit $j = \sqrt{-1}$

$S(f)$ ist daher auch komplex und besteht aus einen Real- und Imaginärteil (oder: aus Amplitude und Phase).

Fouriertransformation

Inverse Fouriertransformation

Die inverse Fouriertransformation berechnet aus einem beliebigen Spektrum $S(f)$ wieder das reelle Signal $s(t)$

$$s(t) = \mathcal{F}^{-1}\{S(f)\} = \int_{-\infty}^{+\infty} S(f)e^{j2\pi ft}dt$$

d.h. die inverse Fouriertransformation erzeugt exakt wieder das Ausgangssignal $s(t)$.

$$\mathcal{F}^{-1}\{\mathcal{F}\{s(t)\}\} = s(t)$$

Fouriertransformation

Gute Nachricht

Merken Sie sich diese Formeln nicht!
Wichtig ist nur, dass sich ein beliebiges Signal genau wie ein periodisches Signal in Sinusteilschwingungen zerlegen lässt (nur eben für jede mögliche Frequenz!). Den dabei errechneten Energieverlauf über der Frequenz nennen wir *Spektrum*.

Spektrum : in der Phonetik meistens das reelle *Amplitudenspektrum* (Phase wird i.d.R. nicht wahrgenommen):

$$|S(f)| = \sqrt{\mathcal{R}e(S(f))^2 + \mathcal{I}m(S(f))^2}$$

Fourier und Faltung

Spektrum einer Faltung

Besteht ein Signal $s(t_n)$ aus der Faltung zweiter Signale $a(t_n)$ und $b(t_n)$, so ist das Spektrum $S(f)$ das Produkt der beiden Spektren $A(f)$ und $B(f)$

$$a(t_n) \star b(t_n) \qquad \circ\!-\!\bullet \qquad A(f) \cdot B(f)$$

Umgekehrt gilt auch:

$$a(t_n) \cdot b(t_n) \qquad \circ\!-\!\bullet \qquad A(f) \star B(f)$$

Fourier und Faltung

Weitere Rechenregeln für lineare Systeme:

additivity: $a(t_n) + b(t_n) \quad \circ - \bullet \quad A(f) + B(f)$

und

homogeneity: $k\, a(t_n) \quad \circ - \bullet \quad k\, A(f)$

k ist konstante Zahl.

Digitale Filter

Digitales Filter = Faltung des Signals $b(t_n)$ mit der Impulsantwort $a(t_n)$
$\rightarrow$ Digitales Filter = Multiplikation des Spektrum mit dem Spektrum der Impulsantwort

$$g(t_n) = a(t_n) \star b(t_n) \qquad \circ - \bullet \qquad A(f) \cdot B(f) = G(f)$$

Spektrum der Impulsantwort $A(F)$ wird *Übertragungsfunktion* (transfer function) des Filters genannt.

Mehrere Übertragungsfunktionen hintereinander (z.B. Mikrophone, Vorverstärker ...)
$\rightarrow$ Spektrum des resultierenden Signals ist Produkt aus Quellenspektrum und allen Übertragungsfunktionen.

Beispiel: Typische Sprachaufnahme

Bei einer Sprachaufnahme durchläuft das vom Sprechapparat erzeugte Sprachsignal $s(t) \circ\!-\!\bullet\, S(f)$ mehrere Systeme:

- Akustik des Raums von Mund bis Mikrophon: $A(f)$
  (Oft ist $A(f)$ wegen der kugelförmigen Abstrahlung zu höheren Frequenzen ansteigend.)
- Charakteristik des Mikrophones: $M(f)$
- Übertragungsfunktion der analogen Leitung: $L(f)$
- Übertragungsfunktion des Vorverstärkers: $V(f)$
- Nyquist-Tiefpass: $N(f)$

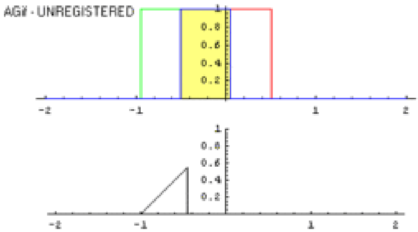
$\rightarrow$ das Signal $\check{s}(t)$, das letztlich digitalisiert wird:

$$\check{s}(t) \qquad \circ\!-\!\bullet \qquad \check{S}(f) = S(f) \cdot A(f) \cdot M(f) \cdot L(f) \cdot V(f) \cdot N(f)$$

Beispiel: Küpfmüllertiefpass

Faltung mit Rechteckfunktion

$$r(t) \quad \star \quad s(t)$$
$$\circ\!-\!\bullet$$
$$R(f) \quad \cdot \quad S(f)$$



Frage: Was bewirkt die Filterung mit einer rechteckigen Impulsantwort?

$\rightarrow$ Spektrum des gefilterten Signals $\check{S}(f)$ ist die Multiplikation des Eingangsspektrums $S(f)$ (rot) mit der Fouriertransformierten der Impulsantwort $r(t)$ (grün):

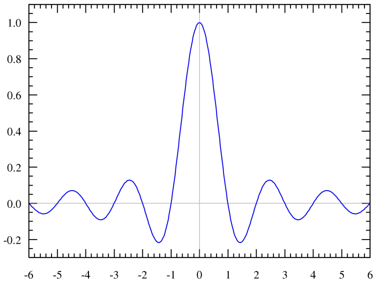
$$\check{S}(f) = R(f) \cdot S(f)$$

Beispiel: Küpfmüllertiefpass

Fouriertransformierte einer Rechteckfunktion der Breite $2b$:

$$R(f) = \mathcal{F}\{rect(\frac{t}{2b})\} = 2b\frac{sin(2\pi bf)}{2\pi bf}$$

Diese Funktion wird auch *si-Funktion* genannt: $si(x) = \frac{sin(x)}{x}$



- $si(f = 0) = 1$ d.h. tiefe Frequenzen werden nicht gedämpft
- $si(f) = 0$ das erste Mal für $b \cdot f_g = \frac{1}{2}$ (weil $sin(\pi) = 0$), also $f_g = \frac{1}{2b}$
- *Tiefpass* für $f = 0 \ldots \frac{1}{2b} Hz$

Beispiel: Küpfmüllertiefpass

Je schmaler also das Rechtecksignal (b klein) desto größer der Durchlassbereich des *Küpfmüllertiefpasses* ($\frac{1}{2b}$).

Anschaulich:
Faltung mit Rechteck ist eine Glättung des Signals, weil jeweils über die Signalwerte innerhalb des Rechtecks gemittelt (aufaddiert) wird.
Glättung heißt aber: hohe Frequenzen werden eliminiert.
Das genau macht auch ein Tiefpass.

Fragen

Welche Sprachlaute haben ein Linienspektrum? Warum?

Welche Sprachlaute haben ein stochastisches Signal, welche ein Impuls-Signal?

Hat ein Impulssignal ein diskretes (Linien) oder kontinuierliches Spektrum?

Wenn man die Amplitude eines Signals verdoppelt, was passiert dann mit dem Spektrum des Signals?

Wenn man zwei Signale addiert, was passiert dann im Spektralbereich?