

Kompressionsverfahren für Audio

Entropie=
durchschnittlicher
Informationsgehalt
pro Zeichen in einer
Zeichenkette
=untere Grenze der
zur Kodierung eines
Zeichens im Durch-
schnitt notwendigen
Bits

2 prinzipielle Verfahren:

Entropie-Kodierung

Daten werden als Folge digitaler Werte verlustfrei komprimiert.

Run-Length-Encoding, Pattern matching, Statistische Verfahren

Quellen-Komprimierung

Je nach Quelle oder Art der Daten werden Kompressionsverfahren eingesetzt, die besondere Eigenschaften der Quelldaten ausnutzen, meist auch mit (hinnehmbaren) Verlusten.

Transformationskodierung, Prädikative Verfahren

MP3
AAC
ADPCM

Entropie

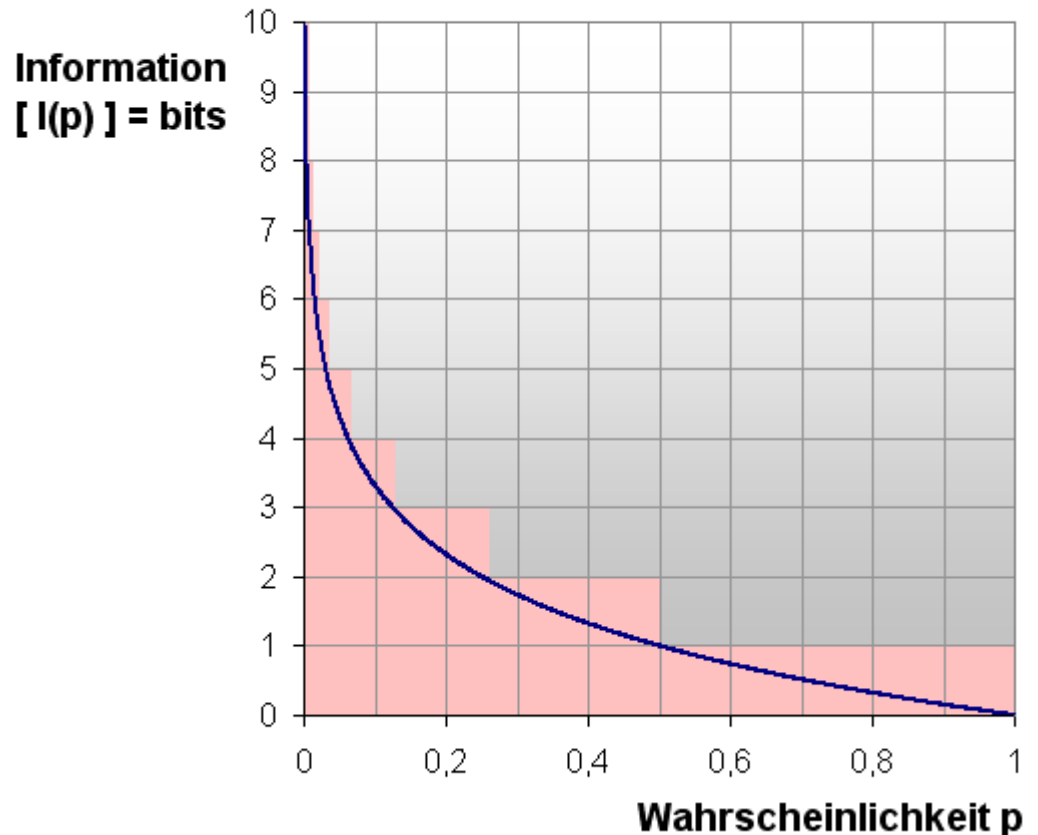
Gegeben sein eine endliche Information
(Wort über einem Alphabet von m Zeichen)

Definition;

$$I(p) = \lg\left(\frac{1}{p}\right) = -\lg(p)$$

$I(p)$ Informationsgehalt eines Zeichens, das mit Wahrscheinlichkeit p vorkommt.
 $p=0$. kommt nie vor
 $p=1$ nur das Zeichen kommt vor

Informationsgehalt eines Zeichens
mit Auftrittswahrscheinlichkeit p



Entropie

Gegeben sein eine endliche Information
(Wort über einem Alphabet von m Zeichen)

Der Erwartungswert

$$H = \sum_{i=1}^n p_i \cdot \text{ld} \left(\frac{1}{p_i} \right)$$

des Informationsgehaltes über
alle Zeichen
= mittlerer Informationsgehaltes
aller Zeichen
= Entropie

$$H = - \sum_{i=1}^m p_i \log_2 p_i$$

m Anzahl Zeichen

p_i Wahrscheinlichkeit
des Auftretens des
Zeichens i

Alternativ (normierte Entropie)

$$H(I) = - \sum_{j=1}^{|Z|} p_j \cdot \log_{|Z|}(p_j)$$

$$H(I) = - \frac{1}{\log_2(|Z|)} \cdot \sum_{j=1}^{|Z|} p_j \cdot \log_2(p_j)$$

Entropie

Beispiel: wir beobachten 2 Studenten im Hörsaal und notieren die Beobachtungen:

Nr.	Binärcode	Beobachtung	Wahrscheinlichkeit	VLC
0	00	Beide schlafen	0,5	
1	01	A schläft	0,125	
2	10	B schläft	0,125	
3	11	Beide lauschen	0,25	



16 Ereignisse, 32 Bit, 2 Bit pro Ereignis

0	0	0	1	0	1	1	0	0	1	0	0	0	1	0	1
0	1	0	0	0	1	1	0	0	1	0	1	0	0	0	1

Gesucht: Kürzeste Kodierung für diese Ereignisfolge

Entropie

Optimal
?

Beispiel: wir beobachten 2 Studenten im Hörsaal und Notieren die Beobachtungen:

Nr.	Binärcode	Beobachtung	Wahrscheinlichkeit	VLC
0	00	Beide schlafen	0,5	0
1	01	A schläft	0,125	110
2	10	B schläft	0,125	111
3	11	Beide lauschen	0,25	10

16 Ereignisse, 28 Bit, 1,75 Bitt pro Ereignis

$$L = \sum_{i=1}^4 p_i \text{Bit}(i)$$

0	1	0	1	0	1	1	0	0	1	0	1	0	1	0	1
	1		1		0	0			0		1		1		0
	0		1							0		1			

Entropie

Mathematische Definition:

Informationsgehalt:
Bit pro Zeichen, die
notwendig sind, um
die Information
darzustellen.

$$H = -\sum_{i=1}^m p_i \log_2 p_i$$

m Anzahl mögl. Ereignisse

p_i Wahrscheinlichkeit
des Auftretens des
Ereignis i

„A Mathematical
Theory of
Communication“
1948



Claude Elwood
Shannon 1916-2001

$$\begin{aligned} H &= -\left(\frac{1}{2} \log_2 \frac{1}{2} + \frac{1}{8} \log_2 \frac{1}{8} + \frac{1}{8} \log_2 \frac{1}{8} + \frac{1}{4} \log_2 \frac{1}{4}\right) \\ &= -\left(\frac{1}{2} (\log_2 1 - \log_2 2) + \frac{1}{8} \log_2 \frac{1}{8} + \frac{1}{8} \log_2 \frac{1}{8} + \frac{1}{4} \log_2 \frac{1}{4}\right) \\ &= -\left(\frac{1}{2} (0 - 1) + \frac{1}{8} \log_2 \frac{1}{8} + \frac{1}{8} \log_2 \frac{1}{8} + \frac{1}{4} \log_2 \frac{1}{4}\right) \\ &= -\left(-\frac{1}{2} - \frac{3}{8} - \frac{3}{8} - \frac{2}{4}\right) \\ &= 1,75 \end{aligned}$$

Entropie

Definition:

Wir beschränken
uns auf $m=2^n$,
Zweierpotenzen

$$H = -\sum_{i=1}^m p_i \log_2 p_i$$

m Anzahl Zeichen im Alphabet

p_i Wahrscheinlichkeit

des Auftretens des

Zeichens i

„A Mathematical
Theory of
Communication“
1948



Claude Elwood
Shannon

Falls alle Zeichen gleich häufig, so $H=n$
 $H < n$, nicht rein zufällig, gewisse Abstufung

Merke: Die Entropie einer Nachrichtenquelle W ist die untere Grenze der zur Kodierung einer Nachricht durchschnittlich benötigten Bit.

Entropie ursprünglich Begriff der Thermodynamik und Statistischen Mechanik

Komprimierung PCM-Daten

Speicherplatz für 3 min Sound

CD-Qualität, stereo: $3 \cdot 60 \cdot 176400 =$ **31,5 MB** **176.400 Byte/s**

Radio-Qualität, mono $3 \cdot 60 \cdot 22050 \cdot 2 =$ **7,9 MB** **44.100 Byte/s**

Sprachqualität, mono $3 \cdot 60 \cdot 11025 =$ **2 MB** **11.025 Byte/s**

ISDN-Telefonie, mono $3 \cdot 60 \cdot 8096 =$ **1,44 MB** **8.096 Byte/s**

Entropieverfahren wie Huffman, LZW wenig brauchbar

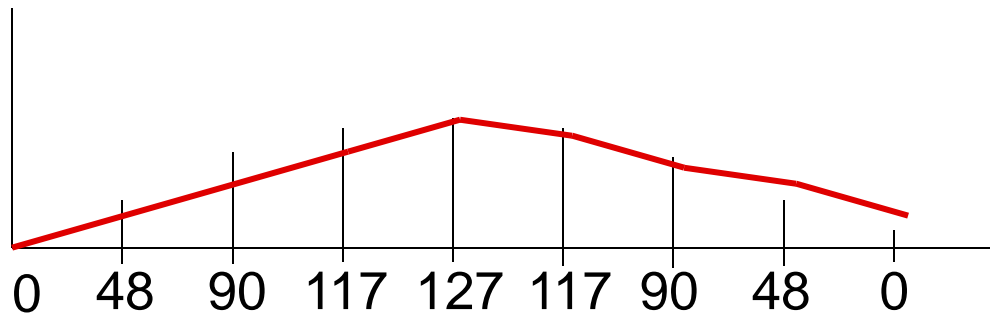
Predictive Coding:

DPCM Delta / Differential Pulse Code Modulation

ADPCM Adaptive DPCM

Komprimierung DPCM

Idee: die Differenzen zwischen den Pulswerten speichern.
In der Regel kleine Zahlen, z.B. mit 4 Bit zu kodieren



Delta 48 42 27 10 -10 -27 -42 -48

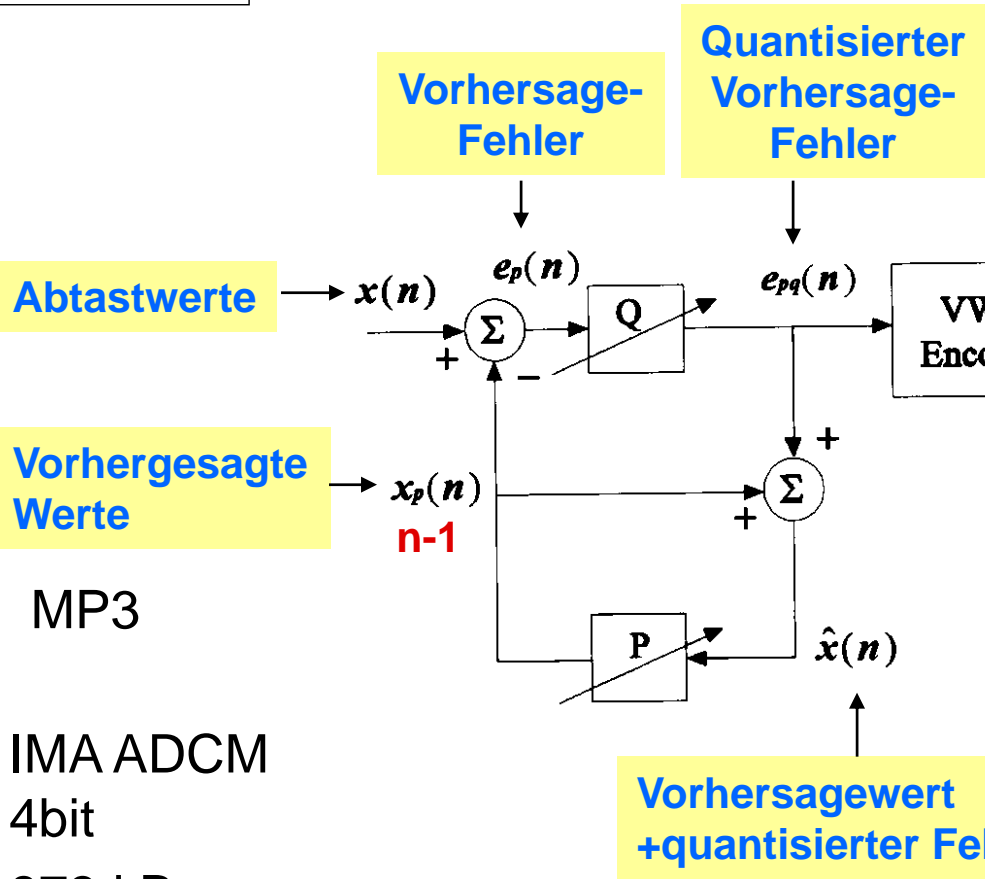
Differenzen
brauchen 7 Bit

6-Bit-Delta 32 32 32 21 -10 -27 -32 -32

0 32 64 96 127 117 90 58 26

Konstante Differenzen führen zu mäßigen Ergebnissen
Entweder wenig Komprimierung oder wenig Approximation

Komprimierung Predictive Coding



repeat

$$x_p(n) \leftarrow \text{predict}(x(n), \hat{x}(n-1), \dots)$$

$$e_p(n) \leftarrow x(n) - x_p(n)$$

$$e_{pq}(n) \leftarrow \text{quantize}(e_p)$$

output (e_{pq})

$$\hat{x}(n) \leftarrow x_p(n) + e_{pq}(n)$$

$n \leftarrow n + 1$

until eof

MP3

IMA ADPCM
4bit
873 kB

ADPCM
variables Delta
vorhersagen

Komprimierung IMA ADPCM

4:1 Komprimierung: 16Bit-Wert durch 4 Bit darstellen

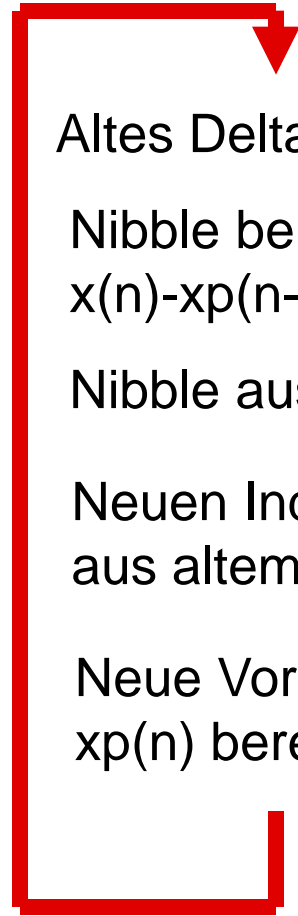
4-Bit Delta-„Nibble“

Vorzeichen	bit3	bit2	bit0
------------	------	------	------

Status des Quantisierers
 $x_p(n-1)$
index

Stepsize-Tabelle

0	7
1	8
2	9
.	
.	
88	32767



Altes Delta=Tabelle[index]

Nibble berechnen aus $x(n)-x_p(n-1)$ und altem Delta

Nibble ausgeben

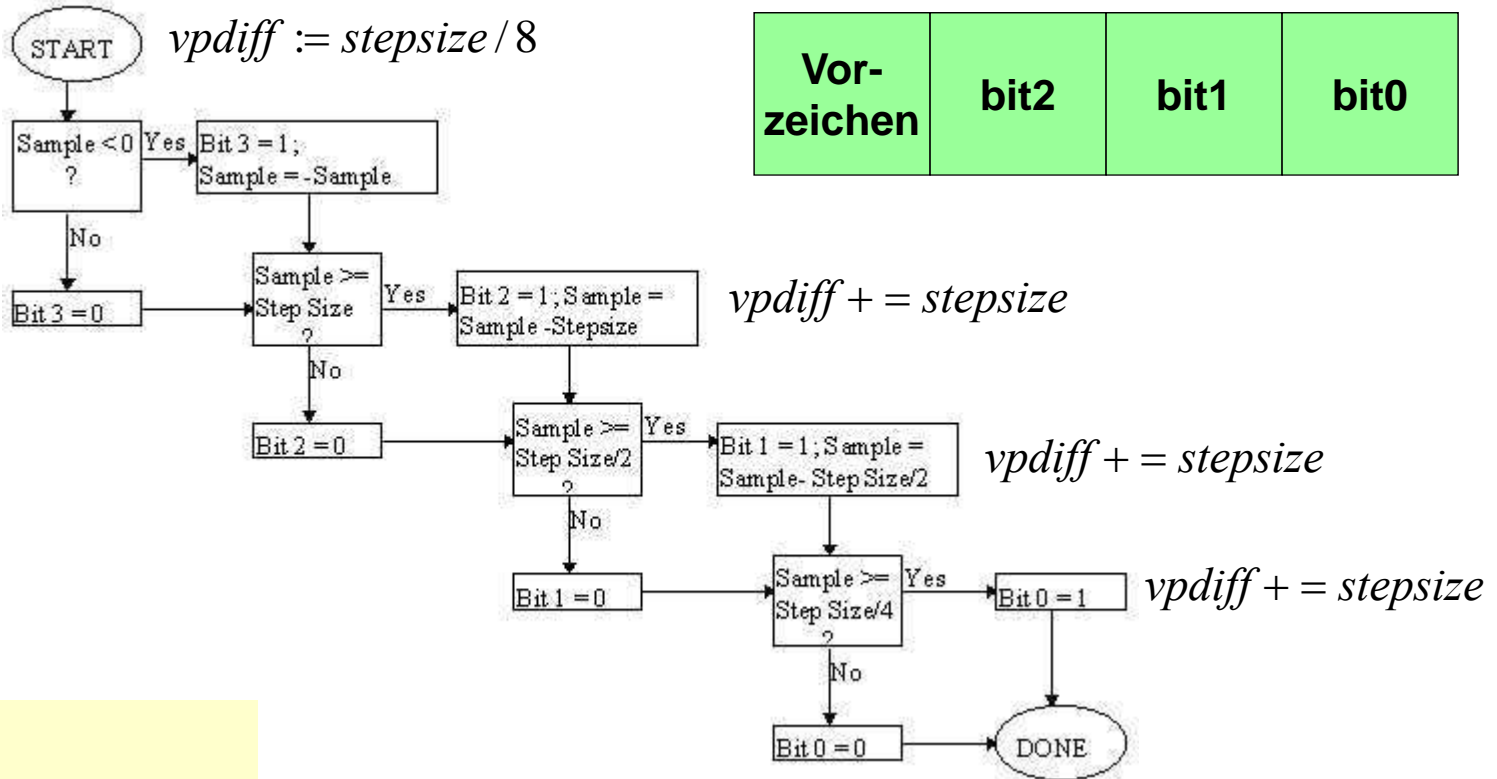
Neuen Index berechnen aus altem Index und Nibble

Neue Vorhersage $x_p(n)$ berechnen

Komprimierung IMA ADPCM

4-Bit Delta-„Nibble“

Vorzeichen	bit2	bit1	bit0
------------	------	------	------



Hilfs-
variable:

$Sample := x(n) - x_p(n-1)$

$Stepsize := StepSizeTabelle[index]$

Neue Vorhersage $x_p(n) = x_p(n-1) + vpdiff$

Fourier, Jean-Baptiste Joseph Baron de (1768-1830),
französischer Mathematiker und Physiker,
geboren in Auxerre,
ausgebildet im Mönchskloster von Saint-Benoît-sur-Loire.

$$f(t) = \dots + b_1 \sin(2\pi ft) + \dots + b_2 \sin(2 * 2\pi ft) + \dots - b_3 \sin(3 * 2\pi ft) + \dots$$



Jede „anständige“ periodische Funktion hat eine trigonometrische Reihendarstellung mit eindeutig bestimmten Koeffizienten a_i und b_i .

Fourier-Reihe

Summendarstellung

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

Alternativ: $\lambda = \frac{1}{f}$ Periodenlänge

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos\left(\frac{2\pi kt}{\lambda}\right) + b_k \sin\left(\frac{2\pi kt}{\lambda}\right) \right)$$

Beispiel:
Orgelton



$$\frac{1}{4} * (\sin(2\pi ft) + \sin(2\pi ft * 2) + \sin(2\pi ft * 4) + \sin(2\pi ft * 6))$$

$$a_k = 0 \forall k \in \mathbb{N}, k \geq 0$$

$$b_1 = b_2 = b_4 = b_6 = \frac{1}{4}$$

$$b_k = 0 \text{ sonst.}$$

Fourier-Reihe

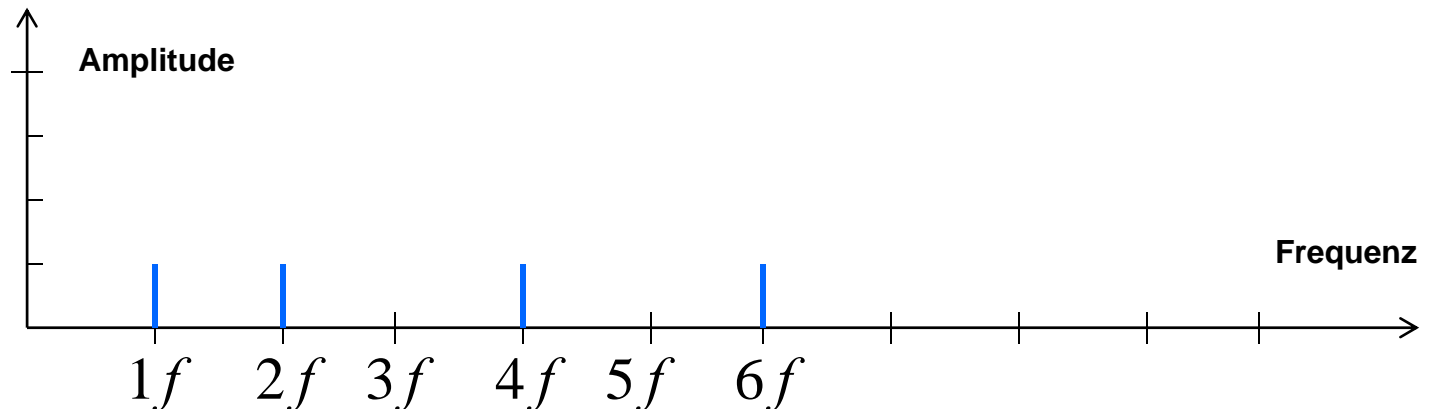
Summendarstellung

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

Die Koeffizienten a_k, b_k bestimmen, mit welcher Amplitude die zugehörige Frequenz am Klang beteiligt ist.

Periodische Funktionen haben ein diskretes **Spektrum**

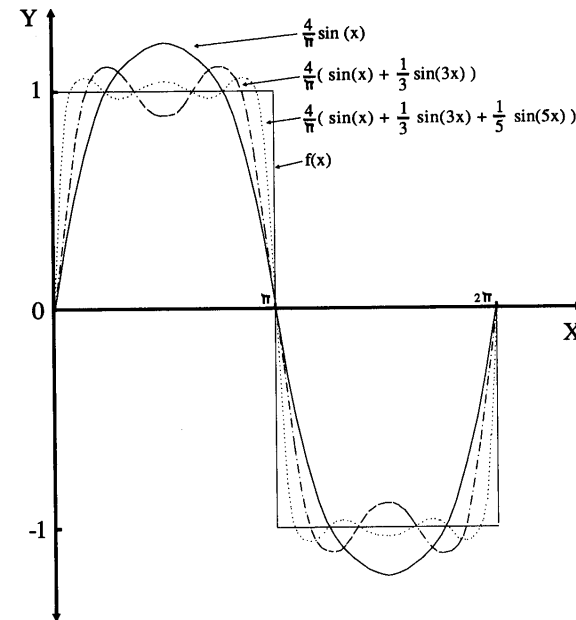
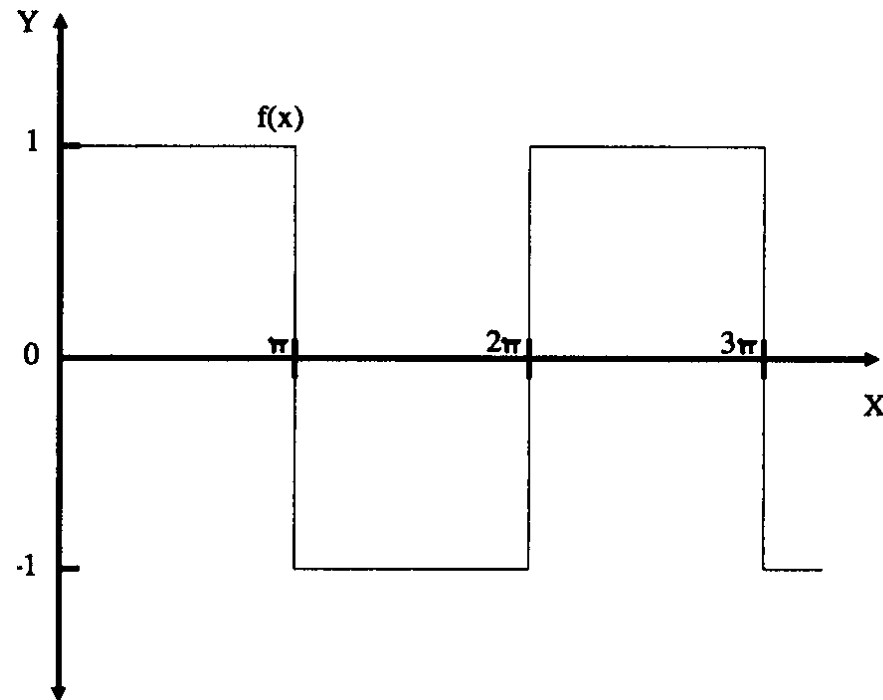
f ist die niedrigste beteiligte Frequenz.



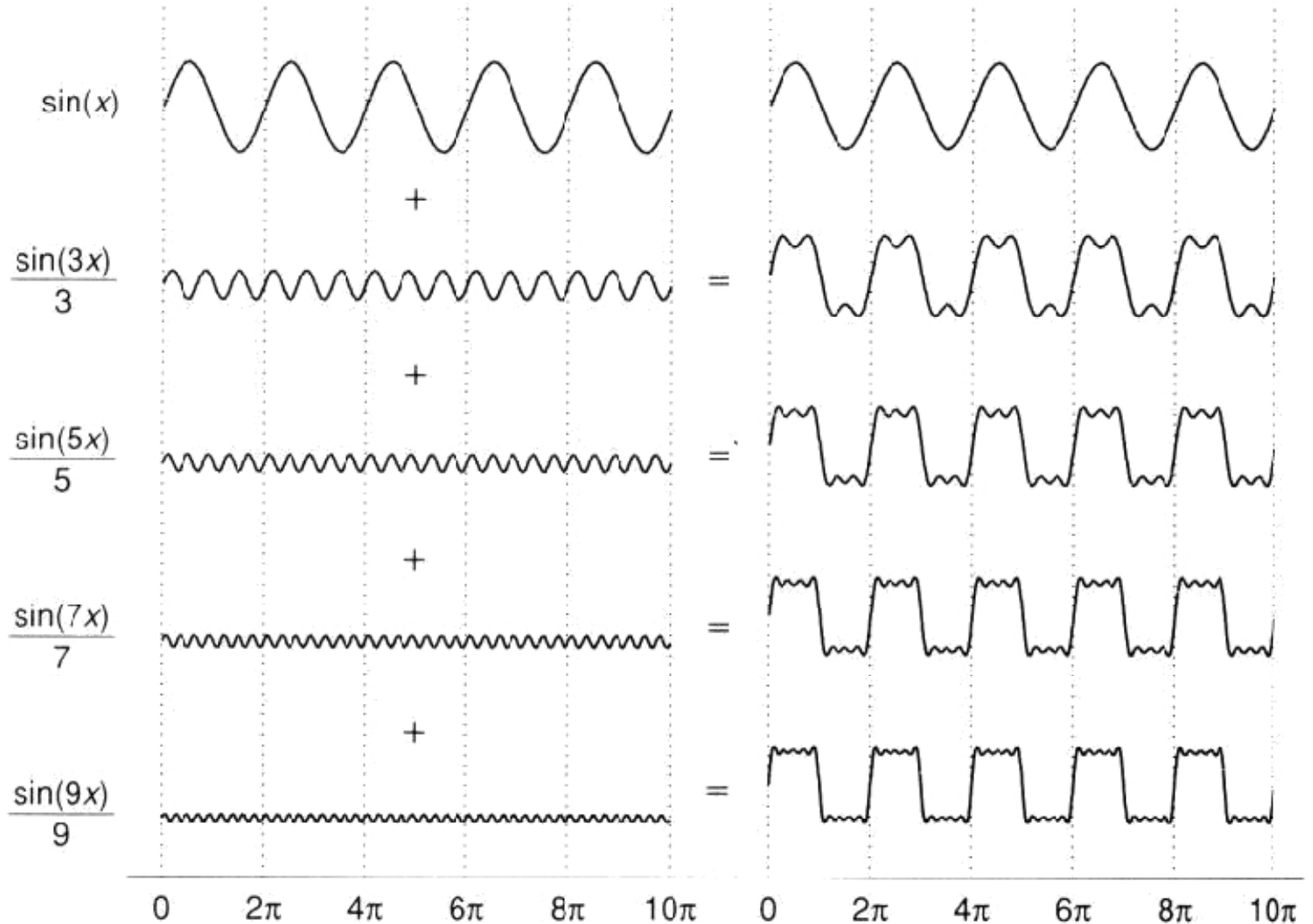
Harmonische Analyse

$$f(x) = \sum_{j=1}^{\infty} \frac{4}{(2j-1)\pi} \sin((2j-1)x)$$

Rechteckfunktion ist periodisch,
erfüllt Dirichlet'sche Bedingungen (siehe Lehrbuch Analysis),
also gibt es Fourier-Darstellung



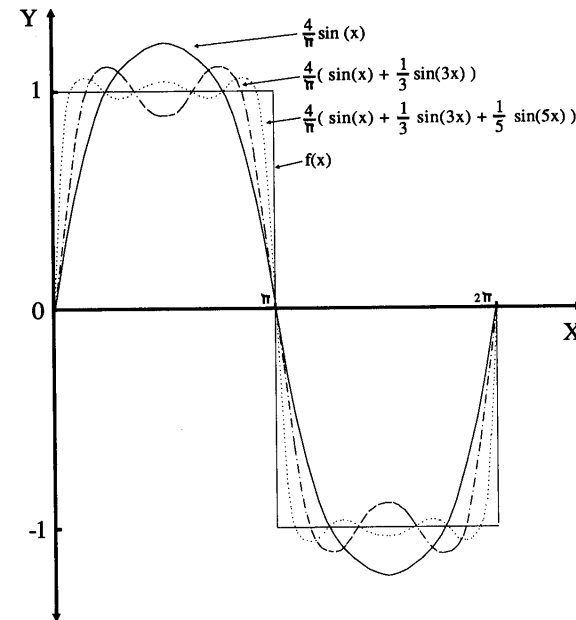
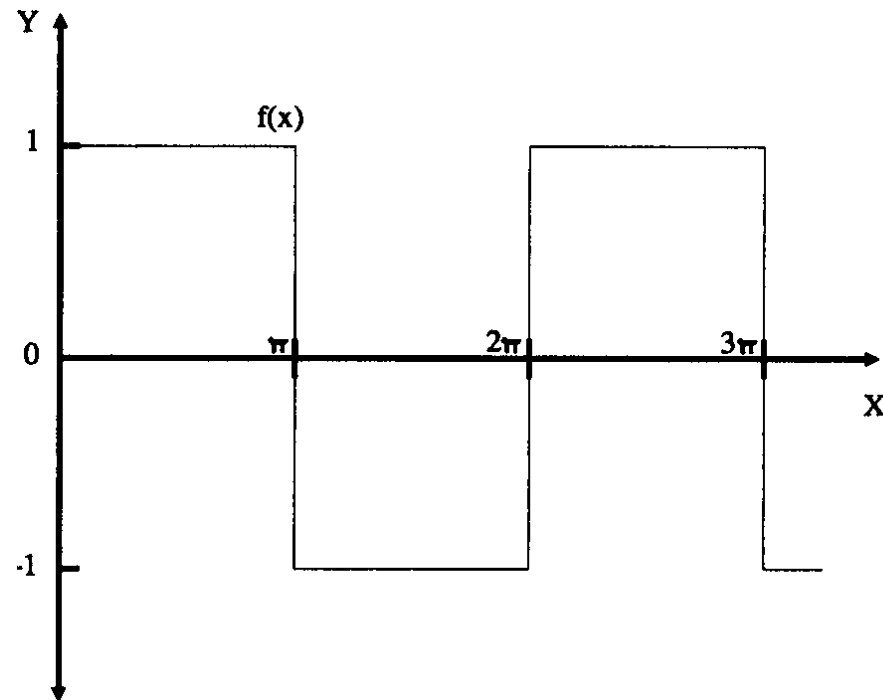
Harmonische Analyse $f(x) = \sum_{j=1}^{\infty} \frac{4}{(2j-1)\pi} \sin((2j-1)x)$



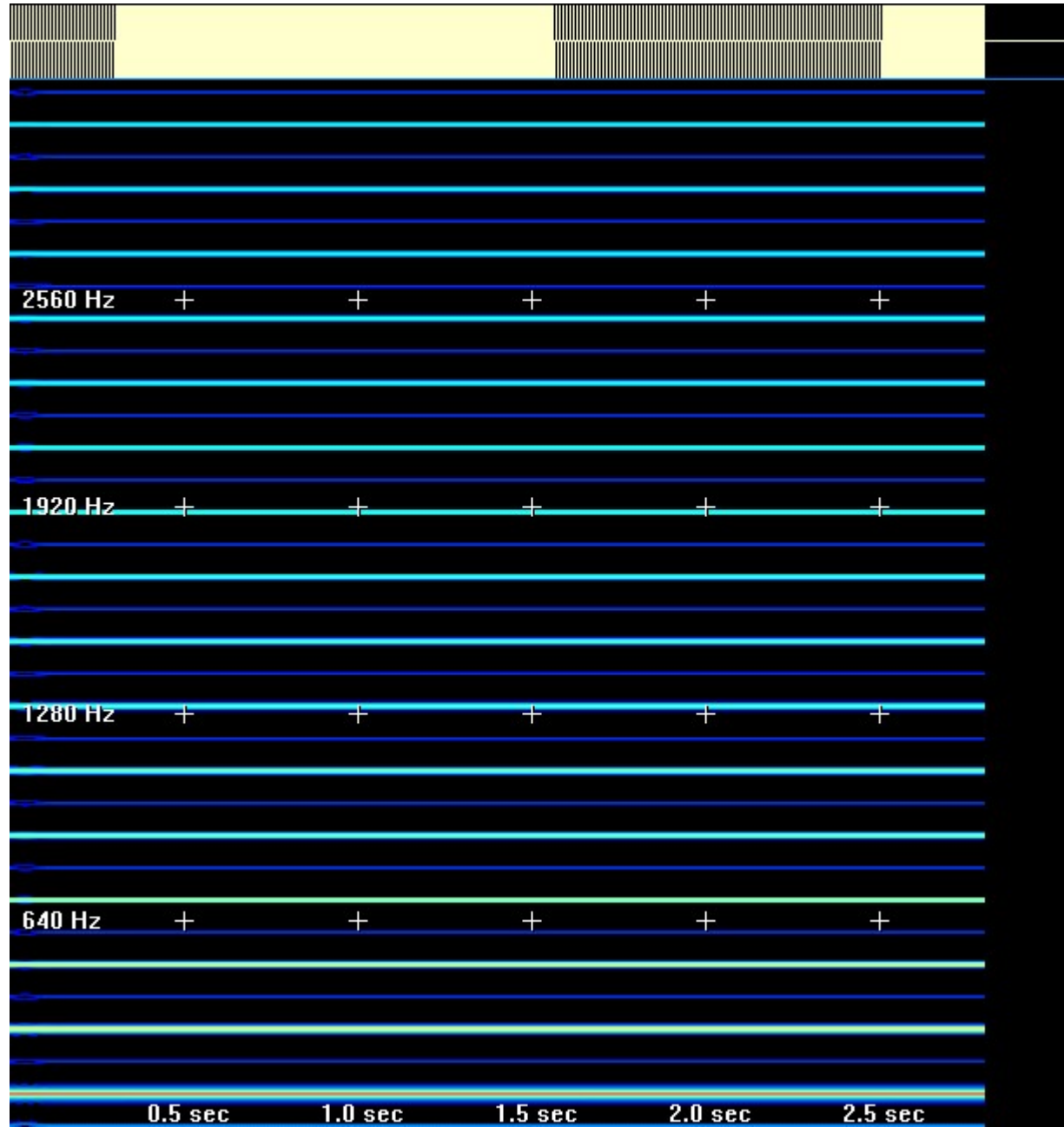
Harmonische Analyse

$$f(x) = \sum_{j=1}^{\infty} \frac{4}{(2j-1)\pi} \sin((2j-1)x)$$

Sägezahn ungerade Funktion ($f(x)=-f(-x)$), also keine Kosinus-Anteile erfüllt Dirichlet'sche Bedingungen (siehe Lehrbuch Analysis), also gibt es Fourier-Darstellung



Beispielspektrum 100 Hz Rechteck-Kurve



Berechnet mit
Spectrogram 10.0

R. S. Horne

<http://www.visualizationsoftware.com/gram.html>

Fourier-Koeffizienten berechnen

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

Mathematik: $a_0 = 2f \int_0^{1/f} g(t) dt$

$$a_k = 2f \int_0^{1/f} g(t) \cos(2\pi fkt) dt$$

$$b_k = 2f \int_0^{1/f} g(t) \sin(2\pi fkt) dt$$

Wichtige Folgerung:

ist die periodische Funktion $g(t)$ frequenzbeschränkt, so besteht das Spektrum aus endlich vielen diskreten Werten

Übergang ins Komplexe

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

Mathematik:

$$a_0 = 2f \int_0^{1/f} g(t) dt$$

$$a_k = 2f \int_0^{1/f} g(t) \cos(2\pi fkt) dt$$

$$b_k = 2f \int_0^{1/f} g(t) \sin(2\pi fkt) dt$$

Wichtige Folgerung:

ist die periodische Funktion $g(t)$ frequenzbeschränkt, so besteht das Spektrum aus endlich vielen diskreten Werten

umrechnen mit komplexer Exponentialfunktion

$$g_n(t) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(2\pi fkt) + b_k \sin(2\pi fkt))$$

Teil-
summe

Mathematik:

$$e^{ix} = \cos x + i \sin x$$

$$e^{-ix} = \cos x - i \sin x$$

Umrechnen:

$$\cos(2\pi ft) = \frac{1}{2} (e^{i2\pi ft} - e^{-i2\pi ft})$$

$$\sin(2\pi ft) = \frac{1}{2i} (e^{i2\pi ft} - e^{-i2\pi ft})$$

$$g_n(t) = \frac{a_0}{2} + \sum_{k=1}^n \left(\frac{a_k}{2} (e^{i2\pi ft} - e^{-i2\pi ft}) + \frac{b_k}{2i} (e^{i2\pi ft} - e^{-i2\pi ft}) \right)$$

Alles Geniale ist einfach

$$g_n(t) = \frac{a_0}{2} + \sum_{k=1}^n \left(\frac{a_k}{2} (e^{i2\pi ft} - e^{-i2\pi ft}) + \frac{b_k}{2i} (e^{i2\pi ft} - e^{-i2\pi ft}) \right)$$

triviale
Umrechnung

$$g_n(t) = \frac{a_0}{2} + \sum_{k=1}^n \left(\frac{1}{2} (a_k - ib_k) e^{i2\pi ft} + \frac{1}{2} (a_k + ib_k) e^{-i2\pi ft} \right)$$

geniale
Substitution

$$c_0 = \frac{a_0}{2}, c_k = \frac{1}{2} (a_k + ib_k), c_{-k} = \frac{1}{2} (a_k - ib_k)$$

$$g_n(t) = \sum_{k=-n}^n c_k e^{i2\pi ft}$$

immer noch die
Fourier-Reihe $n \rightarrow \infty$

Fourier-Koeffizienten berechnen

Wir hatten
schon:

$$a_0 = 2f \int_0^{1/f} g(t) dt$$

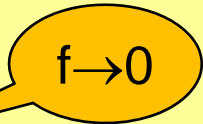
$$a_k = 2f \int_0^{1/f} g(t) \cos(2\pi fkt) dt$$

$$b_k = 2f \int_0^{1/f} g(t) \sin(2\pi fkt) dt$$

und berechnen
daraus

$$c_k = f \int_0^{1/f} g(t) e^{-i2\pi ft} dt$$

Kontinuierliche Fouriertransformation

$$\begin{aligned}
 g(t) &= \sum_{k=-\infty}^{\infty} e^{i2\pi ft} \left(f \int_0^f g(t) e^{-i2\pi ft} dt \right) \\
 &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{i2\pi ft} \left(\int_0^f g(t) e^{-i2\pi ft} dt \right) 2\pi f \quad \text{setze } \Delta u = 2\pi f \\
 &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{i\Delta ukt} \left(\int_0^f g(t) e^{-i\Delta ukt} dt \right) \Delta u \quad (\Delta u \rightarrow 0) \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} G(u) e^{iut} du
 \end{aligned}$$


Mathematische Definition

$$G(u) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} g(t) e^{-iut} dt$$

F(u) ist
Fourier-Transformierte
von f(t)

Inverse Fouriertransformation

$$g(t) = \int_{-\infty}^{+\infty} G(u) e^{iut} du$$

Eigenschaften Fouriertransformation

Transformation in den **Frequenzraum**

Fourier-Transformation berechnet das **Spektrum**

Periodische Funktionen haben ein **diskretes Spektrum**

Die Fouriertransformation lässt sich umkehren !

Die inverse Fourier-Transformation macht
aus dem Spektrum den Sound.

Anwendung der Fouriertransformation

Analyse des Spektrums, Frequenzmessung

Transpositionen

Frequenzfilter (Hoch-, Tiefpass)

Beweis des Sampling-Theorems

Schnelle, diskrete Fouriertransformation Cooley 6 Turkey

Wir betrachten die zu transformierende Funktion an $2n=2^i$ Stützstellen.

$$f_m = \sum_{k=0}^{2n-1} x_k e^{-\frac{2\pi i}{2n}mk} \quad m = 0, 1, \dots, 2n-1$$

Die Fouriertransformation (f_1, \dots, f_{2n-1}) lässt sich durch divide & conquer (gerade und ungerade Summanden) mit Aufwand $N \cdot \log(N)$ berechnen.

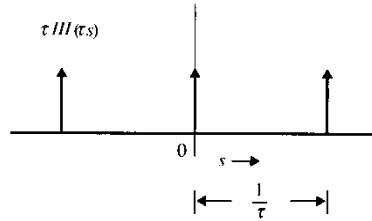
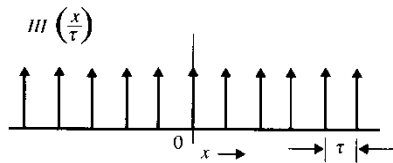
Java-Klasse: <http://www.theorie.physik.uni-goettingen.de/~honecker/wr07/fft.java>

Schnelle, diskrete Fouriertransformation Cooley 6 Turkey

$$f_m = \sum_{k=0}^{2n-1} x_k e^{-\frac{2\pi i}{2n}mk} \quad m = 0, 1, \dots, 8$$

Beispiele Abtastfunktion

$$x = (1, 1, 1, 1, 1, 1, 1, 1) \quad f = (2.8, 0, 0, 0, 0, 0, 0, 0)$$



Schnelle, diskrete Fouriertransformation Cooley 6 Turkey

Medien-
Technik

$$f_m = \sum_{k=0}^{2n-1} x_k e^{-\frac{2\pi i}{2n}mk} \quad m = 0, 1, \dots, 8$$

Beispiele Grenzfrequenz

$$x=(1,-1,1,-1,1,-1,1,-1) \quad f=(0,0,0,0,2.8,0,0,0)$$

Schnelle, diskrete Fouriertransformation Cooley 6 Turkey

$$f_m = \sum_{k=0}^{2n-1} x_k e^{-\frac{2\pi i}{2n}mk} \quad m = 0, 1, \dots, 8$$

Beispiele „Tchteck“

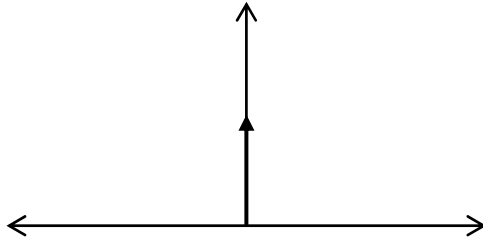
$x=(1,1,1,1,-1,-1,-1,-1)$ $f=(0,-0.7+1.7i,0,-0.7+0.3i,0,-0.7-0.3i,0,-1.7-1.7i)$

Impulsfunktion Dirac'sche Delta-Funtion

$\delta(x)$

Definition:

$$\delta(x) = 0 \quad \forall x \neq 0, \quad \int_{-\infty}^{\infty} \delta(x) dx = 1$$

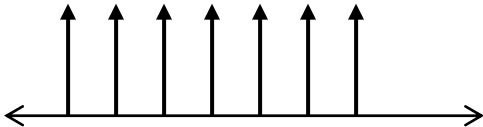


Eigenschaften:

$$\int_{-\infty}^{\infty} f(x) \delta(x) dx = f(0)$$

$$\int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx = \int_{-\infty}^{\infty} f(x - x_0) \delta(x) dx = f(x_0)$$

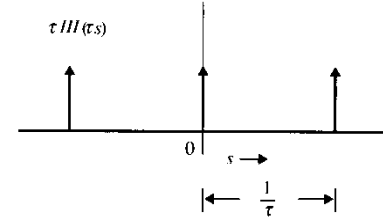
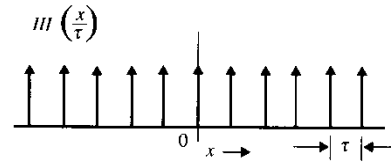
$$\delta(x) \circ f(x) = \int_{-\infty}^{\infty} \delta(t) f(x - t) dt = f(x - t)|_{t=0} = f(x)$$



Abtast-Theorem: Beweisidee

$$\text{III}\left(\frac{x}{\tau}\right) = \tau \sum_{n \in \mathbb{Z}} \delta(x - n\tau) \quad \text{Shah-Funktion mit Frequenz } \frac{1}{\tau}$$

$$\text{FT}\left(\text{III}\left(\frac{x}{\tau}\right)\right) = \tau \text{III}(\tau s) = \sum_{n \in \mathbb{Z}} \delta\left(s - \frac{n}{\tau}\right)$$



Spektrum
Ausgangs-
Signal

Spektrum
abgetastetes
Signal

f(x) zu sampelnde Funktion mit beschränktem Spektrum

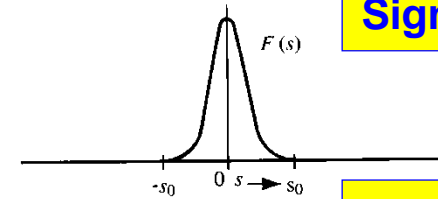
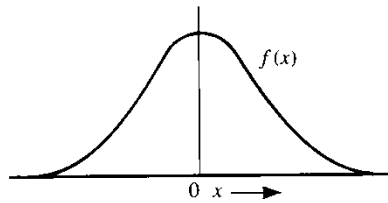
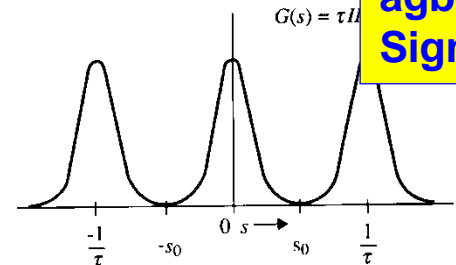
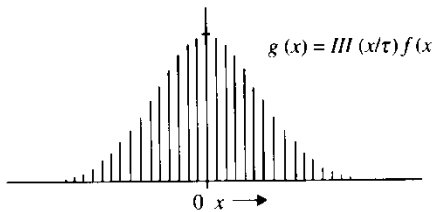


Figure 12-2 A bandlimited function

$$g(x) = \text{III}\left(\frac{x}{\tau}\right) f(x)$$



Abtast-Theorem

$$f(x) = FT^{-1}(F(x)) = FT^{-1}(G(x)\Pi_{\tau}(x))$$

$$= g(x) \circ \text{sinc}\left(\frac{x}{\tau}\right)$$

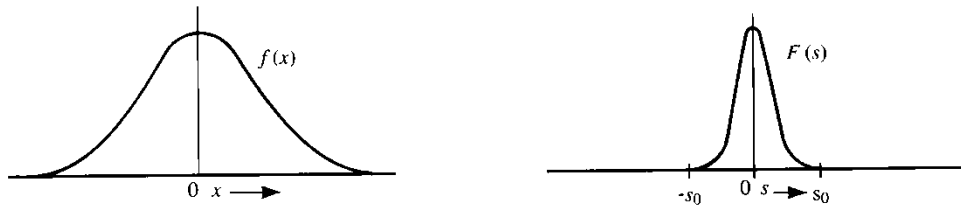
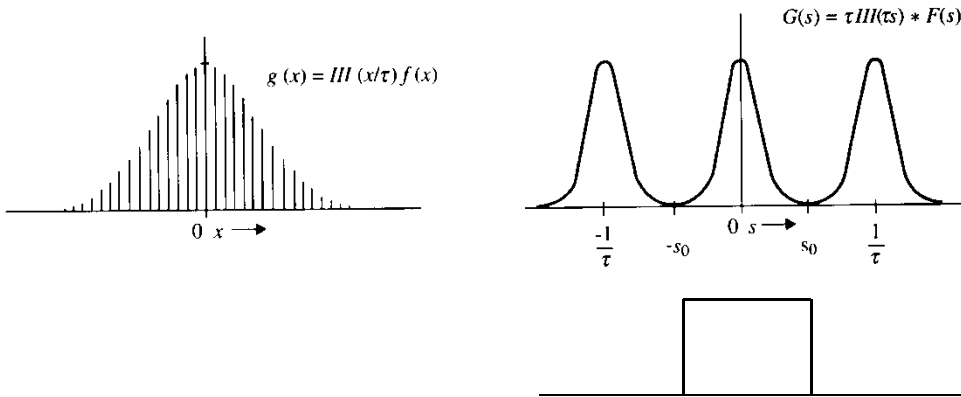


Figure 12-2 A bandlimited function

$$F(x) = FT(f(x)) \quad \text{Spektrum}$$



$$G(x) = \tau \text{III}(\tau x) \circ F(x)$$

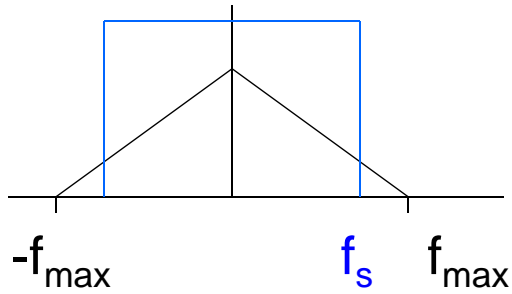
Faltung FT(Shah) mit Spektrum

Kastenfunktion $\Pi_{\tau}(x)$

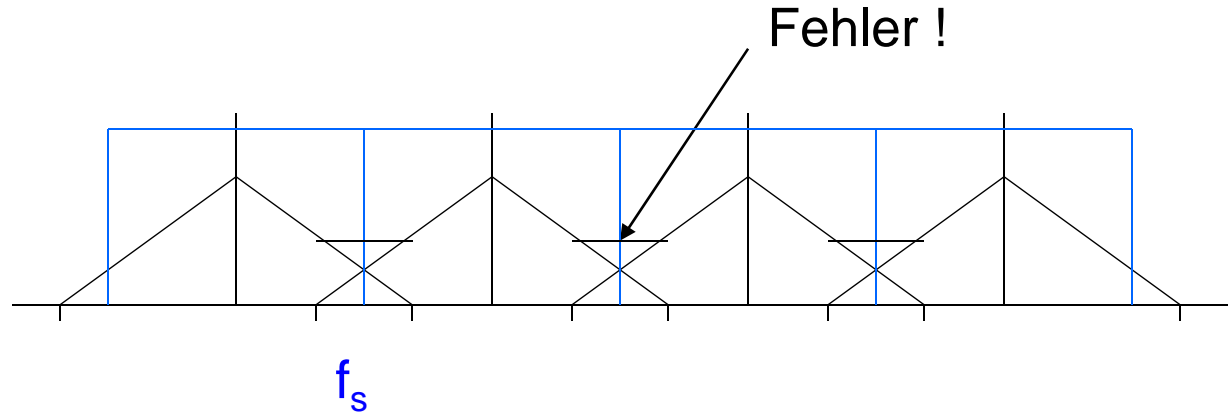
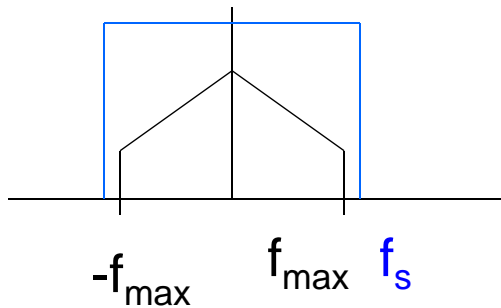
$$F(x) = \Pi_{\tau}(x)G(x)$$

Sampling

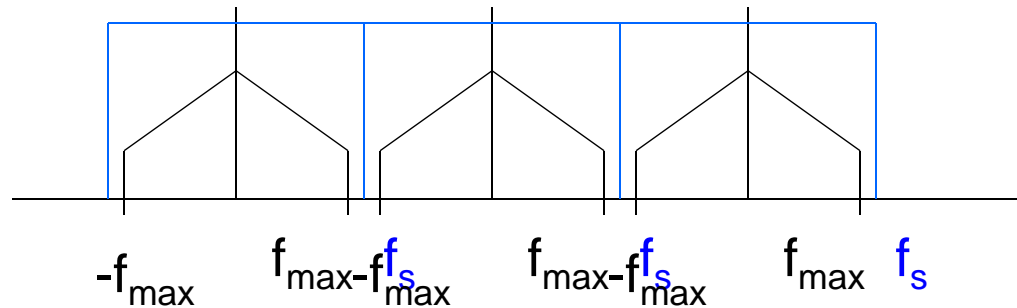
Aliasing bei falscher Abtastfrequenz



Frequenzspektrum
des Ausgangssignals
mit f_{\max}



Frequenzspektrum
des abgetasteten
Signals mit f_s



Konvolution - Faltung

Definition:

$$h(x) = \int_{-\infty}^{\infty} f(t)g(x-t)dt = f \circ g(x)$$

Faltungssatz:

$$\int_{-\infty}^{\infty} f \circ g(x)e^{-ixu} dx = \int_{-\infty}^{\infty} f(t)e^{-itu} dt \int_{-\infty}^{\infty} g(y)e^{-iyu} dy$$

kurz: $H(u) = F(u) \cdot G(u)$

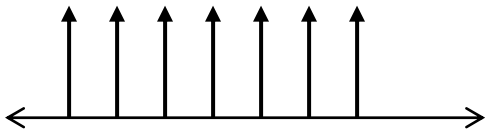
H: Fouriertransformierte von h
G: Fouriertransformierte von g
F: Fouriertransformierte von f

Abtast-Theorem Shah-Funktion

$$\delta(x) = 0 \quad \forall x \neq 0, \quad \int_{-\infty}^{\infty} \delta(x) dx = 1$$

Dirac'sche Delta-Funktion

Es gilt:
$$\delta(ax) = \frac{1}{|a|} \delta(x)$$



$$\text{III}(x) = \sum_{n=-\infty}^{n=\infty} \delta(x-n)$$

Shah-Funktion

Es gilt:
$$FT(\text{III}(x)) = \text{III}(x)$$

$$\text{III}(ax) = \sum_{n \in \mathbb{Z}} \delta(ax - n) = \sum_{n \in \mathbb{Z}} \delta\left(a\left(x - \frac{n}{a}\right)\right) = \frac{1}{|a|} \sum_{n \in \mathbb{Z}} \delta\left(x - \frac{n}{a}\right)$$

Literatur zu diesem Kapitel

Breitfeld, Peter: Fourierreihen
und Fouriertransformation

docs.s fz-
bw.de/phag/skripte/**fourier**.pdf

Hyperlinks zu diesem Kapitel

www.wikipedia.org

page.mi.fu-berlin.de/~lind/lingausarb/

[http://de.wikipedia.org/wiki/Schnelle_Fourier-
Transformation](http://de.wikipedia.org/wiki/Schnelle_Fourier-Transformation)

Grafik-Quellen

www.wikipedia.org