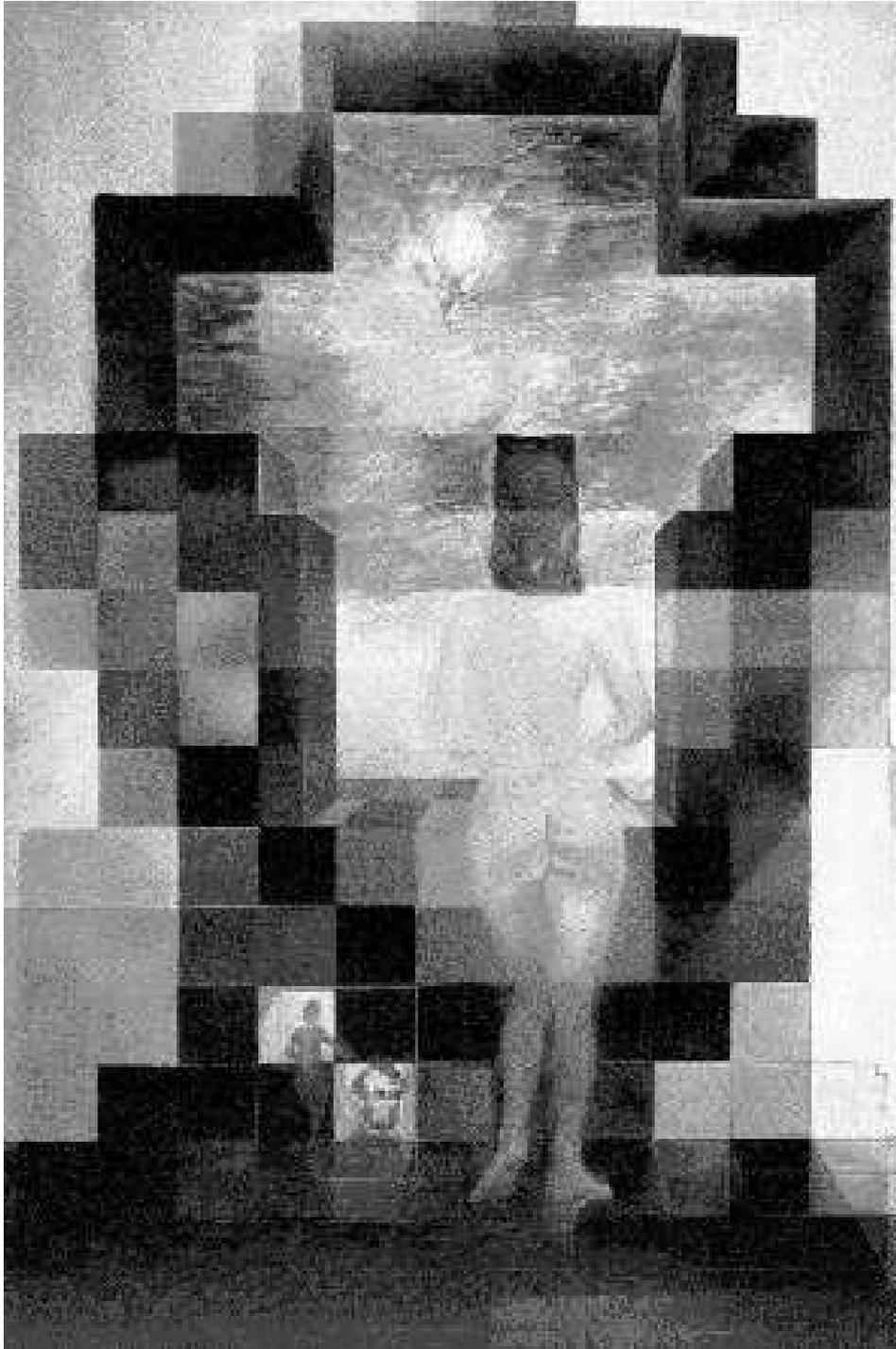


Gerald Sommer

Computer Vision I

8. Auflage



Lehrstuhl für Kognitive Systeme
Christian-Albrechts-Universität zu Kiel
2008

Vorwort zur 1. Auflage

Dieses (seit langem angekündigte – jetzt endlich erscheinende) Skript entstand aus dem Manuskript der im SS 93 gehaltenen Vorlesung „Computer Vision I“, die im Rahmen des Hauptstudiums der Informatik an der Christian–Albrechts–Universität zu Kiel angeboten wird. Es soll die Teilnehmer der Vorlesung unterstützen, den breiten Umfang dieses interdisziplinären Fachgebietes zu erfassen, ohne in jedem Fall auf Spezialliteratur zurückgreifen zu müssen.

Kapitel 5 zur Berechnung des optischen Flusses wurde von K. Daniilidis verfaßt.

Für das Setzen in Latex und die Erstellung zahlreicher Abbildungen danken wir Ulrike Garbe, Johannes Wittek und Volker Krüger.

Wir bitten den aufmerksamen Leser, Fehler jeglicher Art, die zwangsläufig in dieser Uraufgabe auftauchen werden, an uns weiterzuleiten.

August 1994

Vorwort zur 2. Auflage

In dieser 2. Auflage sind Fehler ausgebessert worden. Dazu wurde ein Index erstellt, so daß das Auffinden spezieller Fragestellungen erheblich erleichtert wurde. Darüber hinaus soll durch Querverweise mehr Klarheit über die oft verwirrende Vielfalt parallel verwendeter, aber äquivalenter Fachbegriffe geschaffen werden.

Wir bitten den aufmerksamen Leser weiterhin, eventuell noch vorkommende Fehler an uns weiterzuleiten.

April 1995

Vorwort zur dritten Auflage

Den Anlaß zu dieser Neuauflage des Skripts gaben einige Änderungen in Kapitel 3. Die Abschnitte 3.6 und 3.7 wurden grundlegend überarbeitet und um einen Abschnitt über differentialgeometrische Maße ergänzt und auch die übrigen Kapitel profitierten von einer Fehlerkorrektur. Das 5. Kapitel über Bildfolgen wurde ausgelassen und wird separat aufgelegt.

Wir bedanken uns herzlich für die unermüdliche Aufmerksamkeit der Leser beim Finden der Fehler.

Januar 1997

Vorwort zur sechsten Auflage

Gegenüber den vorigen Auflagen wurden folgende Erweiterungen vorgenommen: Einführung der quaternionwertigen Fouriertransformation und der quaternionwertigen Gaborfunktion, Hinzunahme einer Einführung in die Wavelettransformation. Gleichzeitig wurde Kapitel 4 über die stochastischen Modellierungen überarbeitet und von der Lehrveranstaltung Computer Vision I nach Computer Vision II verlegt. Kapitel 4 erscheint deshalb auch nicht mehr in diesem Skript.

September 2000

Vorwort zur siebten Auflage

In dieser Auflage wurde die Fouriertransformation ausgebaut. In den letzten Jahren stellte sich heraus, daß ihre Behandlung in der bisherigen Form zu knapp war, um ein Verstehen ihrer unterschiedlichen Facetten zu ermöglichen. Ihr wurde nun das Kapitel 2 vollständig gewidmet.

September 2004

Inhaltsverzeichnis

1	Signaltheoretische Modelle	1
1.1	Modelle der Signalanalyse	3
1.1.1	Prozeßmodelle	5
1.1.2	Signalmodelle	6
1.1.3	Systemmodelle	21
1.1.4	Strategische Modelle	27
1.2	Diskrete Signalfunktionen	32
1.2.1	Signal und Träger	32
1.2.2	Abtastfunktion und Signalbegrenzung	34
1.2.3	Zeilenweise Bildabtastung	39
1.2.4	Logischer Träger	40
1.2.5	Nachbarschaften und Umgebung	42
1.3	Synopsis der Operatoren	47
1.3.1	Lokale und globale Operatoren	47
1.3.2	Parallele und sequentielle Verarbeitung	50
1.3.3	Verschiebungsinvarianz von Operatoren	50
1.3.4	Algebraische Beziehungen von Operatoren	53
1.4	Lineare verschiebungsinvariante Operatoren	55
1.4.1	Lineare Operatoren und Impulsantwort	55
1.4.2	Matrixdarstellung eines linearen Operators	58
1.4.3	Transformation des Basissystems	59
1.4.4	Primitive Operatoren	62
1.4.5	Zyklische Faltung und Korrelation	69
1.4.6	Lineare Faltung	77
1.4.7	Implementierung der linearen Faltung	81
2	Fouriertransformation	95
2.1	Integraltransformation	96
2.2	Repräsentationen der Fouriertransformation	101
2.3	Fouriertransformationen kontinuierlicher Signale	105
2.3.1	Komplexe harmonische Funktionen	105
2.3.2	Fourierreihen periodischer Funktionen	107
2.3.3	Approximation durch Fourierreihenentwicklung	111
2.3.4	Herleitung der Fouriertransformation	117
2.3.5	Eigenschaften der Fouriertransformation	119
2.4	Fouriertransformationen diskreter Signale	124

2.4.1	Komplexe harmonische Folgen	126
2.4.2	Fourierdeskriptoren	130
2.4.3	N-te Einheitswurzel und Kreisteilung	132
2.5	Diskrete Fouriertransformation (DFT)	135
2.5.1	Herleitung der Diskreten Fouriertransformation	136
2.5.2	Eigenschaften der Diskreten Fouriertransformation	138
2.5.3	Theoreme der DFT bzw. FT	141
2.5.4	Quadratische Signalformen	149
2.5.5	Zyklische Faltung mittels DFT	151
2.5.6	Fast Fourier Transformation (FFT)	153
2.6	Hartley–Transformation	157
2.7	Quaternionwertige Fouriertransformation	161
2.7.1	Quaternionen	162
2.7.2	Quaternionwertige Fouriertransformation	164
2.7.3	Quaternionische Hermite-Symmetrie	167
2.7.4	Das Phasenkonzept der 2D–QFT	168
3	Anwendungen von LSI-Operatoren	173
3.1	Spektrale Charakterisierung von Bildern	173
3.1.1	Diskretisierung	173
3.1.2	Abtasttheorem	175
3.1.3	Bildgüteparameter	178
3.1.4	Darstellung von Spektren	180
3.2	Spektrale Charakterisierung von LSI–Operatoren	182
3.2.1	Übertragungsfunktion eines LSI–Operators	182
3.2.2	Amplitudenübertragung eines LSI–Operators	183
3.2.3	Phasenübertragung eines LSI–Operators	184
3.2.4	Filterung im Frequenzraum	185
3.3	Bildglättung mit Rechteckfiltern	187
3.4	Bildglättung mit Binomial– und Gaußfiltern	191
3.4.1	Bildung von 1D–Binomialmasken	191
3.4.2	Bildung von 2D–Binomialmasken	192
3.4.3	Eigenschaften der Gaußfunktion als LSI–Operator	193
3.5	Kantendetektion	198
3.5.1	Kanten beleuchteter 3D–Objekte	198
3.5.2	2D–Kantendetektoren	204
3.5.3	Verbesserung der lokalen Schätzung der Ableitung	205
3.5.4	Konditionierung der Ableitungsoperation	208
3.5.5	Topologische Probleme bei 2D–Strukturen	210
3.6	Hierarchien der Skala und Abstraktion	212
3.6.1	Gaußpyramide	214
3.6.2	Laplacepyramide	216
3.7	Lokale Zerlegung der Identität	218
3.7.1	Lokale stochastische Hauptachsentransformation	221
3.7.2	Unschärfe lokale Ableitungen und Jets	226
3.7.3	Differentialgeometrische Maße	230

3.8	Lokale Spektrale Repräsentation	241
3.8.1	Eindimensionale Hilberttransformation	245
3.8.2	Anwendung des analytischen Signals	248
3.9	Hilbertpaar und Quadraturfilterpaar	251
3.9.1	Detektion von Kanten und Linien in Bildern	252
3.9.2	Lokale Energie und lokale Phase	254
3.9.3	Quaternionwertiges analytisches Signal und Hilbert-Quadrupel	257
3.10	Gaborfunktion und lokale Fouriertransformation	261
3.10.1	Notation der eindimensionalen Gaborfunktion	261
3.10.2	Gaborfunktion als Bandpaßoperator	262
3.10.3	Lokale Fouriertransformation und Gabortransformation	264
3.11	Gabor-Wavelettransformation	266
3.11.1	Wavelet-Transformation	266
3.11.2	Grundzüge der Multiskalenanalyse	271
3.11.3	Gabor-Wavelets	274
3.11.4	Gaborfunktion als Quadraturpaar	275
3.12	Die 2D-Gaborfunktion	280
3.12.1	Die Pseudo-2D-Gaborfunktion	280
3.12.2	Die quaternionwertige 2D-Gaborfunktion	289

Wichtige Symbole

Mengen

\mathbb{N}	Menge der natürlichen Zahlen
\mathbb{Z}	Menge der ganzen Zahlen
\mathbb{Q}	Menge der rationalen Zahlen
\mathbb{R}	Menge der reellen Zahlen
\mathbb{C}	Menge der komplexen Zahlen
\mathbb{H}	Menge der Quaternionen

Indizes und Koordinaten

$m \in \{0 \dots M - 1\}$	Spaltenindex eines Signals
$n \in \{0 \dots N - 1\}$	Zeilenindex eines Signals
$t \in \{0 \dots T - 1\}$	Zeitindex eines Signals

(m, n)	diskrete Koordinaten eines Aufpunktes $\in \mathbb{N} \times \mathbb{N}$
$(i, j), (k, l)$	diskrete Koordinaten eines Testpunktes $\in \mathbb{N} \times \mathbb{N}$
(x, y)	kontinuierliche Ortskoordinaten $\in \mathbb{R} \times \mathbb{R}$
(u, v)	kontinuierliche/diskrete Ortsfrequenzen $\in \mathbb{R} \times \mathbb{R}$
$\mathcal{U}_{m,n}$	Umgebung des Punktes (m, n)
$\mathcal{N}_{m,n}^\nu$	ν -Nachbarschaft eines Punktes

Vektoren und Matrizen

$\mathbf{x} = (x, y)^T$	Vektor der Ortskoordinaten
A, I, Λ	Matrix, Einheitsmatrix, Diagonalmatrix

Signale

$f_m, f_{m,n}$	diskretes 1D- bzw. 2D-Signal
$f(x), f(x, y)$	kontinuierliches 1D- bzw. 2D-Signal
f, h, g	Eingangssignal, Impulsantwort, Ausgangssignal
$F_u, F_{u,v}$	diskretes 1D- bzw. 2D-Spektrum von f
$F(u), F(u, v)$	kontinuierliches 1D- bzw. 2D-Spektrum von f
F, H, G	Eingangsspektrum, Übertragungsfunktion, Ausgangsspektrum

Operatoren

\mathcal{O}	Operator (allgemein)
\mathcal{L}	linearer Operator
\mathcal{P}	Projektionsoperator
\mathcal{A}	Aperturoperator

\mathcal{T}	Translationsoperator
\mathcal{S}	Skalierungsoperator
\mathcal{R}	Rotationsoperator
Σ	Summenoperator
\mathcal{D}	Differenzoperator
Δ	Laplaceoperator
∇	Nablaoperator

Funktionskonstanten

$\operatorname{Re}\{z\}, \operatorname{Im}\{z\}$	Real- und Imaginärteil einer komplexen Zahl z
$\delta(x), \delta_m$	Dirac'sches Delta (Einheitsimpuls), Kronecker-Delta
$\operatorname{III}(x), \operatorname{III}_m$	Shahfunktion (Kammfunktion)
$\Pi(x) := \operatorname{rect}(x), \Pi_m$	Rechteckfunktion
$\operatorname{sinc}(x)$	Sincfunktion: $\operatorname{sinc}(x) = \frac{\sin \pi x}{\pi x}$
$\arg(x, y), \arg(z)$	Argument der komplexen Zahl $z = x + jy$, siehe S. 184

binäre Operatoren

*	Faltung
\otimes	zyklische Faltung
\circ	Korrelation

binäre Relationen

$f \circ \longrightarrow \bullet F$	F ist die Fouriertransformierte von f
$f_g \bullet \longrightarrow f_u$	f_u ist die Hilberttransformierte von f_g

wichtige Funktionen

$\operatorname{sgn}(x) := \begin{cases} -1; & x < 0 \\ 0; & x = 0 \\ 1; & x > 0 \end{cases}$	Signumfunktion
$\operatorname{step}(x) := \begin{cases} 0; & x < 0 \\ 1; & x \geq 0 \end{cases}$	Sprungfunktion
$\operatorname{rect}(x) := \begin{cases} 1; & x \leq 1/2 \\ 0; & x > 1/2 \end{cases}$	Rechteckfunktion
$\Lambda(x) := \begin{cases} 1 - x ; & x \leq 1 \\ 0; & x > 1 \end{cases}$	Dreieckfunktion

Kapitel 1

Signaltheoretische Modelle

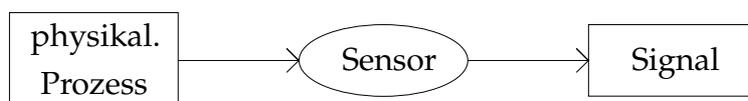
Gegenstand der *Signaltheorie* ist die *Gewinnung, Übertragung* und *Analyse* sensorisch registrierter Signale.

Die Signaltheorie beruht auf

- der *Modellierung von Signalen*,
- der *Modellierung von Problemstellungen* und
- der *Modellierung von Lösungen*.

Zielstellung der Signaltheorie ist die Identifikation von Organisationsprinzipien der Signalstrukturen. Ihre Wurzeln liegen in

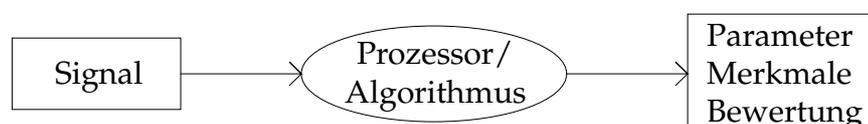
- der *Nachrichtentechnik*:
 - verlustfreie Informationsgewinnung (Sensoren)



- verlustfreie Informationsübertragung (Kanäle)



- der *Kybernetik*:
 - sekundäre Informationsverarbeitung (Rechnertechnik)



In der Kybernetik wird Informationsverlust nicht nur zugelassen, sondern gezielt eingesetzt.

Die *Systemtheorie* befaßt sich in formalisierter Weise mit den Systemen, die die obigen Abbildungen realisieren. Oft werden dazu ähnliche Mittel benutzt, wie in der Signaltheorie.

Beispiel:

- Frequenzraumdarstellung eines Operators in der Bildverarbeitung
- Frequenzraumdarstellung eines Bildes

Die gewählten Modellierungsansätze für Signale und Systeme bedingen sich gegenseitig.

Wir werden in dieser Vorlesung diese Wechselbeziehung kennenlernen: z.B. deterministisch oder stochastisch, stationär oder nicht-stationär.

Signale sind Träger der Information. Sie sind struktureller Natur mit physikalischen Eigenschaften. Ein *Bild* ist als (mehrdimensionales) Signal irgendeiner physikalischen Herkunft zu verstehen. *Muster* sind Repräsentationen der implizit im Signal enthaltenen (kodierten) Information.

Information ist als *Nachricht* aus der Umwelt an den Beobachter aufzufassen. Sie hat *strukturelle*, *semantische* und *pragmatische* Aspekte. Für die Signaltheorie relevant ist der *strukturelle Aspekt* der Information. In der Signaltheorie werden Strukturen mittels mathematischer Modelle beschrieben. Unter *Identifikation* versteht man das signaltheoretisch basierte *Erkennen* von Signalstrukturen durch Vergleich mit Modellen (*Matching*). Aber auch außerhalb der Signaltheorie liegende Zugänge (z.B.: Mustererkennung, Geometrie) werden zur Identifikation genutzt.

Der für die KI relevante *semantische Aspekt* liegt außerhalb des Modellierungsrahmens der Signaltheorie. Ebenso liegt die kontext-/ aufgabenabhängige Bewertung von Information (*pragmatische Aspekt*) außerhalb der Signaltheorie.

Alle drei Aspekte sind aber für die erfolgreiche Lösung einer CV-Aufgabe von Bedeutung.

Die Signaltheorie liefert den Zugang dafür, eine *implizite* strukturelle Information in *explizite* strukturelle Information umzuwandeln.

1.1 Modelle der Signalanalyse

Die Modelle der Signalanalyse beziehen sich auf **Prozesse** und ihre Abbilder (Ereignisse, Objekte, Muster, ...), **Signale** (sensorische Daten) und **Systeme** (Filter, Algorithmen, Operatoren, Transformationen, ...).

In vielen Aspekten ist die **Dimension eines Signals** ohne Bedeutung für die Modellierung der Signalanalyse. Eine Reihe grundsätzlicher Unterschiede führen jedoch zu prinzipiellen Problemen der signaltheoretischen Modellierung der Bildanalyse. So gilt, daß multidimensionale Signale nicht einfach logische Extrapolationen eindimensionaler Signale sind. Vielmehr stellen sie eine **neue Qualität** dar.

- **einfache dimensionale Erweiterung** → mehr Daten → technische Grenzen
Bilder in technischen Anwendungen haben oft die Dimension 512×512 Bildpunkte. In wissenschaftlichen Anwendungen werden aber auch z.B. 4000×4000 Bilder verwendet (Astronomie).

$$512 \rightarrow 512^2 = 0.262MB \rightarrow 512^3 = 0.125GB$$

Bild-Kuben werden im Ergebnis der Computer-Tomographie erzeugt.

Bild-Folgen findet man bei der Analyse dynamischer Szenen.

Video-Bildfolge: 25 Bilder / sec mit $512 \times 512 \rightarrow 6.5$ MByte/sec

Kleinbild-Film: 24×36 mm, Auflösung 0.01 mm

Digitalkamera mit vergleichbarer Auflösung zum Kleinbildfilm:

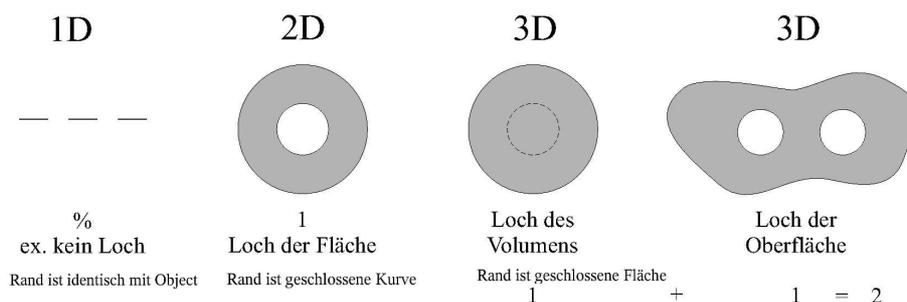
ca. 10 MB ($\approx 3000 \times 3000$) ~ S/W

30 MB ~ Farbe (3×8 bit)

Aber: Ein zweidimensionales Signal stellt gegenüber einem eindimensionalen Signal eine andere **topologische Qualität** dar! Dies hat viele Konsequenzen, die in der Signaltheorie noch nicht beherrscht werden.

- Beispiel aus der **Topologie**: Begriffe "Loch" und "Rand"

Löcher existieren in unterschiedlichen Qualitäten, je nach der Dimension eines Objektes.



- Die Theorie mehrdimensionaler Signale ist nicht gut ausgebaut:

Viele Methoden der Bildverarbeitung sind einfach übernommene 1D-Methoden.

- Beispiel: lineare Algebra (Vektoren f , Matrizen A)

$$f_2 = A \cdot f_1$$

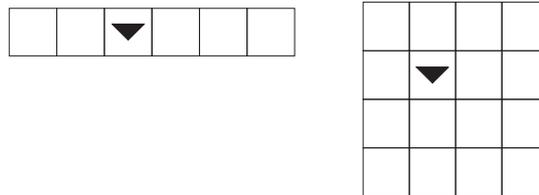
Bemerkung: Wir nehmen an, daß f eine Spaltenvektor sei.

aber: wenn f eine Matrix ist, dann ist A ein Tensor

→ Konvertierung: Matrix → Vektor

- D.E. Dudgeon, R.M. Mersereau: Multidimensional Digital Signal Processing,
Prentice Hall, Englewood Cliffs, 1984
ein gutes Standardbuch im klassischen Rahmen

- Zusätzlicher Freiheitsgrad der Abtastung



Bei zeilenweiser Abtastung: Beschränkung der Freiheitsgrade aber einfache algebraische Beschreibung der Abbildung!

- Fundamentalsatz der Algebra
Im Bereich der komplexen Zahlen existiert für ein Polynom der Ordnung n eine Darstellung als Produkt linearer Polynome

Bsp.: $n = 2$

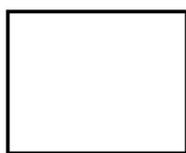
$$p(x) = a_0 + a_1x + a_2x^2 \implies p(x) = \prod_{i=0}^2 (x - x_i)$$

Für mehrdimensionale Polynome gilt der Fundamentalsatz nicht!

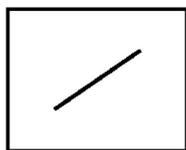
- Zusätzliche Symmetriekonzepte
 - 1D: nur gerade/ungerade Symmetrie
 - 2D: im Prinzip unendlich viele Symmetrien
(nur begrenzt durch die Gitterstruktur des Trägers)
 - unsere Arbeitsgruppe: langfristiges Forschungsprogramm
- Intrinsische Dimension eines (lokalen) Signals
Zu unterscheiden ist die globale (Einbettungs-) Dimension eines Signals und die lokale (intrinsische) Dimension einer Struktur (eines Teilsignales).
Die lokale Analyse/Interpretation von Signalen ist in der Bildverarbeitung von zentraler Bedeutung.
Die intrinsische Dimension beschreibt die Anzahl der erforderlichen Freiheitsgrade, um lokal eine Struktur zu beschreiben.

In der Mathematik wird auch der Begriff Kodimension verwendet.

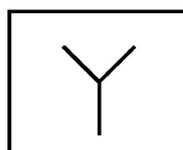
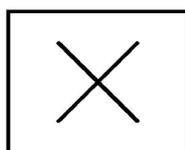
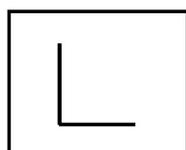
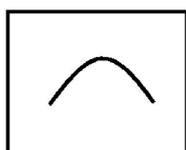
Beispiel: 2D-Signal (Einbettungsdimension)
kann intrinsische Dimensionen Null, Eins und Zwei enthalten



konstante Funktion: Null
(im Wertebereich)



lineare Funktion: Eins ~ Linien, Kanten
(in Geometrie)



nichtl. Fkt.: Zwei
Krümmung
Kreuzung
Junction

Konsequenz: nichtlineare Algorithmen/Systeme zur Verarbeitung erforderlich!

1.1.1 Prozeßmodelle

Definition (Prozeß):

Ausschnitte der objektiven Realität (Szene) bezeichnet man als *Prozesse*.

Prozesse finden in Raum und Zeit statt und besitzen eine physikalische Charakterisierung:

$$p = p(x, y, z, t; q)$$

q ist eine physikalische Variable, beispielsweise $q = \{\rho, T_1, T_2\}$ (MRI = magnetic resonance imaging: ρ ist die Protonendichte des Wasserstoffes in den Molekülen, T_1 und T_2 sind Relaxationszeiten, die etwas über die Chemie der beteiligten Moleküle aussagen).

Prozeßeigenschaften

Definition (zeitlich stationär):

Ein *zeitlich stationärer* Prozeß bedeutet

$$\frac{\partial p}{\partial t} = 0.$$

Definition (räumlich stationär):

Man bezeichnet einen Prozeß als *räumlich stationär*, wenn

$$\frac{\partial p}{\partial x} \left(= \frac{\partial p}{\partial y} = \frac{\partial p}{\partial z} \right) = 0.$$

Räumliche Szenen lassen sich meist nur planar abbilden (*Projektion*).

1.1.2 Signalmodelle

Signale sind Abbildungen von Prozessen.

Gegeben sei ein *Meßsignal* s :

$$s = h_p\{p\}$$

Der *Abbildungsoperator* h_p beschreibt geometrische Eigenschaften (Projektion des Raumes) und physikalische Eigenschaften (photometrisch: Beleuchtung, Reflektanz).

Luminanz: $L = R \cdot I$

Intensität der Beleuchtung: I

Reflektanz des Objektes: R

Physikalische Modelle werden in der 3D-Szenenanalyse verwendet.

Der Wert des Signals an der Position $d \in D$ wird mit $s(d) \in W$ bezeichnet, wobei D *Definitionsbereich* und W *Wertebereich* des Meßsignals sind:

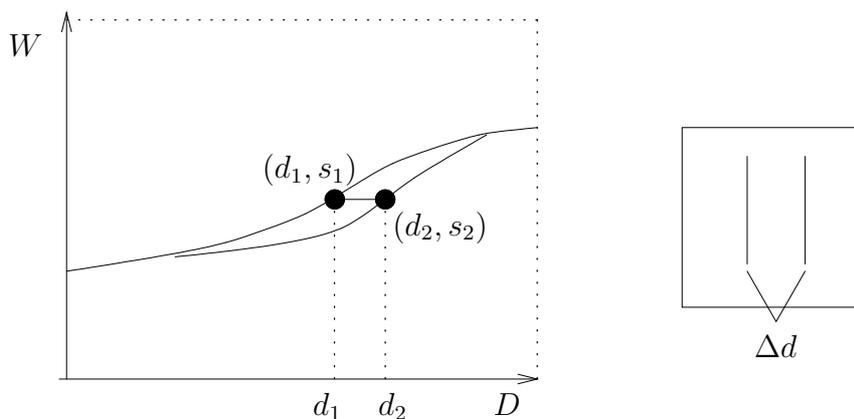
$$W \subset \mathbb{R}^q \quad D \subset \mathbb{R}^r$$

Der Definitionsbereich der Bildverarbeitung wird oft als *Ortsraum* bezeichnet. Der *Phasenraum* ist das kartesische Produkt aus dem Definitionsbereich D und dem Wertebereich W :

$$\Phi = D \times W$$

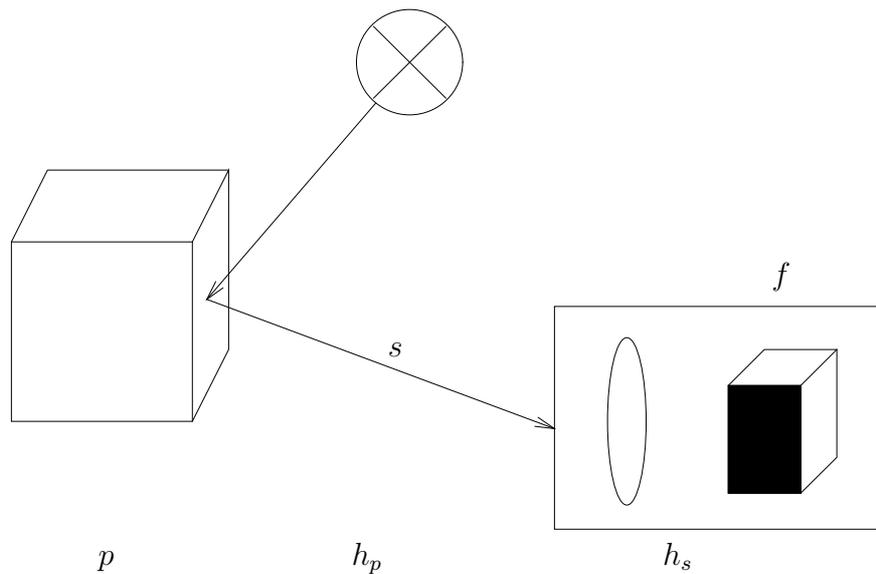
Die Elemente (Tupel (d, s)) des Phasenraums bezeichnet man als *Phasen*. Die *Phasenverschiebung* Δd zweier Phasen d_1 und d_2 ist wie folgt definiert:

$$\Delta d = |d_2 - d_1| \quad \text{für } s_1 = s_2$$



Ein *Bildsignal* f entsteht durch Abbildung eines Meßsignals s unter der Wirkung des Abbildungsoperators h_s (Kamera, Meßgerät).

$$f = h_s\{s\}$$



Digitale Bildsignale zeichnen sich durch folgende Eigenschaften aus:

1. Diskretisierung des Definitionsbereichs: $D \subset \mathbb{Z}^r$, bzw. $D \subset \mathbb{N}^r$
2. Quantisierung des Wertebereichs: $W \subset \mathbb{N}^q$

$$f(x, y) \longrightarrow f_{m,n} \quad \text{mit} \quad \begin{cases} x \in X, & y \in Y \\ m \in \{0, 1, \dots, M-1\} \\ n \in \{0, 1, \dots, N-1\} \end{cases}$$

Die Bildpunkte $(m, n) \in D = D_m \times D_n$ heißen auch *Pixel*. Die Bildpunkte $(m, n, o) \in D = D_m \times D_n \times D_o$ heißen auch *Voxel*. Die Quantisierungen $f(d)$ werden in der Bildverarbeitung als *Grauwerte* bezeichnet:

$$g \equiv f_{m,n} \quad g \in \{0, 1, \dots, G-1\}$$

Ein Bild hat eine *Dynamik* von G Grauwerten. In der digitalen Bildverarbeitung benutzt man standardmäßig eine Quantisierung $G = 2^8 = 256$ Grauwerte. Für bestimmte Anwendungen ist eine höhere Dynamik erforderlich.

Formal unterscheidet man:

- *Grauwertbild* (skalares Signal): $\{g\} \sim 8$ bit
- *Farbbild* (vektorielles Signal)

Echtfarbbild : $\{r \times g \times b\} \sim 24$ bit

Falschfarbbild : $\{g\} \longrightarrow \{r\}$

- *Tiefenbild* (z.B. Laserabtastung)

Definition (Ortsraum):

Die geordnete Menge aller Bildpunkte (Koordinaten) eines Signals, also der kanonische Definitionsbereich eines Bildsignals heißt auch Ortsraum. ■

Definition (Signalraum):

Die Menge aller möglichen Signale über einem Definitionsbereich wird durch den *Signalraum* repräsentiert. Er ist ein linearer normierter Vektorraum mit Skalarprodukt. Seine Struktur ist unabhängig von der Topologie des Signals. Das heißt, es wird nur die Vektorform eines Signals berücksichtigt. Die (diskreten) Positionen des Ortsraums bilden die den Signalraum aufspannenden Achsen. ■

Definition (Dimension des Signalraums):

Die *Dimension des Signalraumes* ist gegeben durch:

$$\dim(SR) = |D|.$$

Für ein Bild mit $D = M \times N$ mit $M = N = 2^{10}$ ist der Signalraum also 2^{20} -dimensional. Jede Achse des Signalraums spannt $|W| = G$ Signale auf. Ein Bild dieser Größe ist in diesem Fall ein 2^{20} -dimensionaler Vektor im Signalraum. ■

Definition (Volumen des Signalraums):

Das Volumen des Signalraums

$$\text{vol}(SR) = G^{|D_m \times D_n|} = G^{|D_m| \cdot |D_n|}$$

spiegelt die Zahl der durch D und G repräsentierbaren unterschiedlichen Signale wider. ■

Beispiel: Für $N = M = 2^{10} = 1024$ mit einer Quantisierung von $G = 2^8 = 256$ Grauwerten ist das Volumen des Signalraums gegeben durch $G^{M \cdot N} = 2^{8 \cdot 1024 \cdot 1024} = 2^{8388608} \approx 10^{2525222}$. ■

Bemerkung:

Ein Bild wird durch einen Punkt des Signalraumes repräsentiert. Es wird durch Angabe des Ortsvektors dieses Punktes beschrieben. Dieser Ortsvektor besitzt $\dim(D)$ Komponenten vom Betrag $f_{m,n}$. Im Signalraum der Dimension $2^{20} \approx 10^6$ stellt also ein Bild eine von ca. $10^{2525222}$ Möglichkeiten dar. Hier liegt die Ursache für die Komplexität der Bildanalyse!

Zum Vergleich: In der Kryptographie werden heute Schlüssel der Länge 128 Bit verwendet, während der Schlüssel, den ein Bild darstellt (obiges Beispiel) ≈ 8 Mio Bit Länge hat.

Glücklicherweise ist aber unsere Aufgabe nicht die vollständige Entschlüsselung. Vielmehr interessieren wir uns für die Identifikation geeigneter Unterräume in diesem hochdimensionalen Raum (das nennen wir Generalisierung \approx Zuordnung gemeinsamer Konzepte zu diesen Unterräumen; z.B. "Katze").

Veranschaulichung der Größe dieser Zahlen:

Weltalter: Erste Eddingtonsche Zahl $E_1 \approx 10^{40}$ Elementarzeiten (10^{-23} sec)

Zahl der Nukleonen des Universums: Zweite Eddingtonsche Zahl $E_2 \approx 10^{80}$

Zahl der Photonen des Universums: $N \approx 10^{90}$

Primäre Signaltypen werden unmittelbar durch den Abbildungsoperator h_s bereitgestellt. Es werden folgende primäre Signaltypen unterschieden:

1. Zeitreihe (1D-Signal) :

$$f = [f_t] \quad t \in \{0, 1, \dots, T-1\}$$

2. Bild (2D-Signal):

$$f = [f_{m,n}] \quad \begin{cases} m \in \{0, 1, \dots, M-1\} \\ n \in \{0, 1, \dots, N-1\} \end{cases}$$

3. räumliche Bildfolge (3D-Signal): Kubus

$$f = [f_{m,n,k}] \quad k \in \{0, 1, \dots, K-1\}$$

4. zeitliche Bildfolge (3D-Signal): Stapel

$$f = [f_{m,n,t}]$$

Beispiel: $f = [f_{m,n,k}]$ bezeichnet man als Kubus. Diese Darstellung entspricht einem 3D-Bild.

$f = [f_{m,n,t}]$ entspricht einem Stapel planarer Bilder, bei dem die Zeit t als freier Parameter auftritt. Eine solche Darstellung entspricht beispielsweise einem Film. ■

5. mehrkanaliges Bild:

$f = [f_{m,n}^p]$ wobei p der freie Parameter für den betrachteten Kanal ist. Während Grauwertbilder skalare Signale sind, stellen mehrkanalige Bilder vektorielle Signale dar.

Sekundäre Signaltypen entstehen im Verlauf der Bildanalyse.

1. Binärbild :

$$f_B = [f_{m,n}], \quad f_{m,n} \in \{0, 1\}$$

2. Objekt:

$$f_R = \{f_{m,n} \mid (m,n) \in R \subset D\}, f_{m,n} = 1$$

Der Definitionsbereich eines Objektes ist die *Region* $R \subset D_m \times D_n$. Man versteht unter einer *Region* R alle Bildpunkte, die unter einem bestimmten Gesichtspunkt ein *Homogenitätskriterium* erfüllen.

3. Kante:

$$f_E = \{f_{m,n} \mid (m,n) \in E, f_{m,n} = 1\}$$

Eine *Kantenregion* E bezeichnet alle Bildpunkte, die unter einem bestimmten Gesichtspunkt ein *Inhomogenitätskriterium* erfüllen.

4. Kontur, Kurve:

$$f_C = \{f_{m,n} \mid (m,n) \in C\}$$

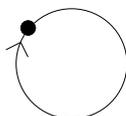
C ist die *Region* der Konturbildpunkte (dünn). Es gilt $C \subset E$. Die Menge der Konturpunkte kann auch als Liste

$$f_C = \{f_l \mid l \in \{l_0, l_1, \dots, l_{L-1}\}\}$$

repräsentiert werden (parametrische Kurvendarstellung).

Eine *Kontur* (z.B. *Kreis*) ist stets *geschlossen*, denn

$$l_0 = l_L.$$



Eine *Kurve* (z.B. *Linie*) ist stets *offen*, denn

$$l_0 \neq l_L.$$



Eigentlich wird der Begriff *Kontur* nur in Verbindung mit einem diskreten Definitionsbereich verwendet (sonst: geschlossene Kurve).

Bemerkung:

Es existieren zwei komplementäre Wege Strukturen in Bildern zu separieren: *regionenorientiert* bzw. *kantenorientiert*.

Die *Bildsegmentierung* ist ein bedeutender Schritt der *Bildanalyse*, in dessen Ergebnis sekundäre Signaltypen als explizite Repräsentation von Signalstruktur definiert werden.

Definition (Segmentierung):

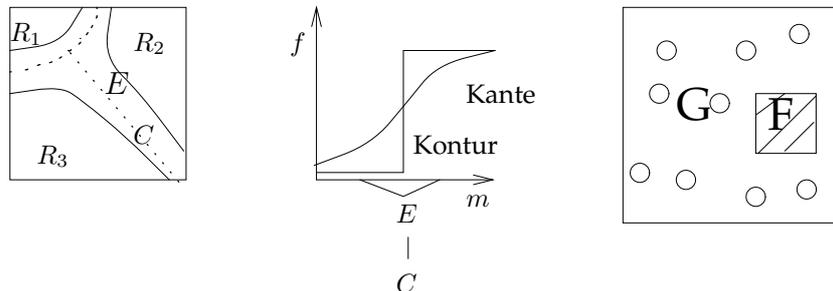
Die *Segmentierung* ist eine vollständige Zerlegung des Bildträgers $D = D_m \times D_n$ in Regionen R_s unter Wirkung eines Segmentierungsprädikates P , so daß

1. $D = \bigcup_{s=0}^{S-1} R_s$
2. $R_s, \quad s \in \{0, 1, \dots, S-1\}$ sind verbunden
3. $P(R_s) = \text{TRUE}$ für alle $s \in \{0, 1, \dots, S-1\}$
4. $P(R_s \cup R_t) = \text{FALSE}$ für $s \neq t, R_s$ und R_t sind benachbart

■

Spezialfall: Objekt-Hintergrund-Segmentierung

Die Anwendung eines einzelnen Segmentierungsprädikates wirkt wie ein syntaktisch definierter Ausblendoperator und trennt ein oder mehrere Objekte vom Hintergrund. Das Problem besteht darin, diese syntaktisch definierte Trennung mit dem semantischen und pragmatischen Konzepten der Aufgabe in Verbindung zu bringen.

**Beobachtbare und verdeckte Variable**

Wir haben als Gegenstand der Bildanalyse das Explizitmachen implizit enthaltener Information benannt. Dies soll an den Begriffen beobachtbare und verdeckte Variable erklärt werden.

Jeder Bildpunkt (m, n) trägt beobachtbare und verdeckte Attribute. Eine beobachtbare Variable repräsentiert ein Attribut, das unmittelbar beobachtbar ist (z.B. die Helligkeit als Grauwert). Deren räumliche/zeitliche Organisation führt aber zu verdeckten Attributen, die je nach dem konzipierten Abstraktionsniveau ganz unterschiedlich aussehen können. Z.B. kann $e_{m,n}$ eine Variable der Kantenzugehörigkeit oder $t_{m,n}$ eine solche der Texturiertheit sein. Auf einem anderen Niveau kann $s_{m,n}$ ein Label der Zugehörigkeit zur Region R_s darstellen und $b_{m,n}^s$ schließlich der dieser Region zugeordnete Begriff/Name.

Bildanalyse wird als (nicht unbedingt lineare) Folge von Abstraktionen bzw. Konzepten und den diesen zuzuordnenden Variablen organisiert.

Strukturen des Signalraumes

Um in der Folge des Überganges von beobachtbaren auf verdeckte Variable Zugriff auf bestimmte Unterräume des Signalraumes zu erhalten, bildet man gewöhnlich neue Basissysteme des Signalraumes.

Häufig wird als algebraischer Einbettungsraum des Signalraumes die Menge aller n -Tupel komplexer Zahlen (\mathbb{C}^n) bzw. reeller Zahlen (\mathbb{R}^n) verwendet. Beide Räume sind als linearer, normierter Vektorraum mit Skalarprodukt zu interpretieren.

Sei $\{f_0, f_1, \dots, f_{n-1}\}$ das n -Tupel des Signals f , dann stellt sich f im Signalraum dar als

$$f = \sum_{i=0}^{n-1} f_i e_i,$$

wobei $\{e_i, i = 0, \dots, n-1\}$ das n -Tupel der Basisvektoren des Signalraumes ist; $\{e_i\}$ bildet die kanonische Basis des Signalraumes.

a) Lineare Unabhängigkeit

Die Basisvektoren e_i werden als linear unabhängig betrachtet. Dann sind aber auch die ihnen zugeordneten Grauwerte f_i als linear unabhängig zu behandeln. Aus der linearen Unabhängigkeit folgt ihre Unkorreliertheit (i.A. nicht umgekehrt).

Das heißt, die Pixel haben in dieser Repräsentation keine Beziehungen zueinander. Aufgabe der Bildanalyse ist es, diese Korrelationen zu rekonstruieren.

Die Topologie des Signals wird höchstens implizit repräsentiert (durch Abtastfolge).

$$\left. \begin{array}{l} \begin{array}{|c|c|c|c|} \hline 1 & 2 & 3 & 4 \\ \hline \end{array} \quad |D_1| = 4 \\ \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array} \quad |D_1| = |D_2| = 2 \end{array} \right\} n = 4$$

Topologische und geometrische Struktureigenschaften sind nicht einfach im Signalraum zu erkennen.

Ausweg: Berechnung von m verschiedenen topol./geometrischen Eigenschaften

- durch Filter (z.B. Richtungsfiler)
- durch statistische Hauptachsentransformation ("Eigenbilder")
- Karhunen-Loeve-Transformation

Repräsentation in Kopien $\mathbb{C}_1^n, \dots, \mathbb{C}_m^n$ des ursprünglichen Signalraumes und Rekonstruktion der Zusammenhänge hieraus

b) orthonormale Basis

$$\langle e_i, e_j \rangle = \delta_{i,j} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases} \quad \text{Kroneckersymbol}$$

$$\langle \cdot, \cdot \rangle \quad \text{Skalarprodukt}$$

c) Euklidische Norm

$$\|f\| = \sqrt{|f|^2} = \sqrt{\sum_{i=0}^{n-1} |f_i|^2} = \sqrt{\langle f, f \rangle}$$

d) Skalarprodukt

$$\langle f, h \rangle = \sum_{i=0}^{n-1} f_i^* h_i \quad * \text{ konjugiert komplex}$$

Übergang zu neuer Basis $\{b_i\}$: z.B. durch System von Filtern

$$f = \sum f'_i b_i$$

Projektion des Signals auf Basis:

$$f'_i = \langle b_i, f \rangle \quad \text{Projektionskoeffizienten}$$

Forderung: $\text{span} \{ \text{Filter} \} = \text{span} \{ \text{Unterraum} \}$

Weitere wichtige Signalmodelle:

1. $\mathbb{C}^k(\mathbb{R}^n, \mathbb{C})$ Raum der **k-fach differenzierbaren Funktionen** aus \mathbb{R}^n in \mathbb{C}

- ex Modell von f als nichtdeterminierte Funktion
(Klasse der Funktion ist bekannt aber nicht deren Parameter)
z.B. partielle Ableitung k -ter Ordnung am Ort d
 - Differentialgeometrie
 - Approximation (z.B. Splines)
- gilt: $\mathbb{C}^\infty \subset \dots \subset \mathbb{C}^2 \subset \mathbb{C}^1 \subset \mathbb{C}^0$
- Bsp.: $C^0([a, b], \mathbb{C})$

$$\langle h, f \rangle = \int_a^b h^*(x) f(x) dx, \quad \|f\| = \int_a^b |f(x)|^2 dx$$

2. **Funktionsraum** $L_2(\mathbb{R}^n)$ Raum der quadratisch integrierbaren Funktionen

Dieser Raum wird in der Fouriertheorie kontinuierlicher Signale angenommen

$$\langle h, f \rangle = \int_{-\infty}^{\infty} h^*(x) f(x) dx \quad h, f \in L_2(\mathbb{R})$$

$$\|f\| = \int_{-\infty}^{\infty} |f(x)|^2 dx < \infty$$

Wichtiges Problem: Suche nach vollständigem und interessantem Basissystem dieses Raumes

~ Basisvektoren spannen Unterräume mit interessanten Struktureigenschaften auf

Bsp: Wavelets

3. **Folgenraum** l_2 Raum der quadratisch summierbaren unendlichen Folgen

$$\|c\|^2 = \sum_{n=-\infty}^{\infty} |c_n|^2 < \infty \quad \{c_n\} \in l_2$$

Fouriertheorie: Isometrie von $L_2(0, 2\pi)$ und l_2 wegen Parseval-Identität

$$\frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx = \sum_{n=-\infty}^{\infty} |c_n|^2$$

$$c_n = \langle f, e^{-2\pi i n x} \rangle \quad \text{Fourierkoeffizienten}$$

Signaltypen

Es lassen sich folgende *Signaltypen* aufgrund ihrer Eigenschaften unterscheiden:

1. Unscharfes Signal:

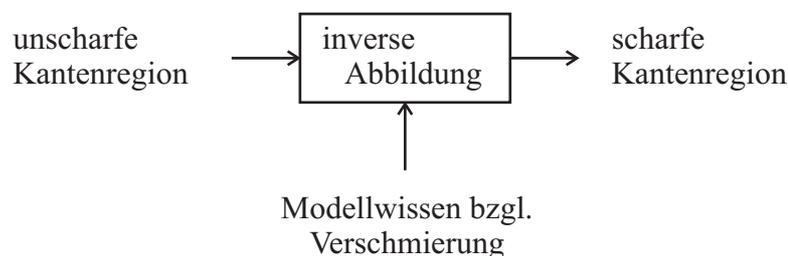
Unschärfe/scharfe Kantenkonzepte \sim unscharfe/scharfe Objektkonzepte.

Ursachen der Unschärfe:

- *inhärente Prozeßunschärfe:*
 - Mischung von Teilprozessen (z.B.: Partialvolumen - Röntgenstrahlen)
 - Zufallsprozesse (z.B.: Nuklearmedizin)
- *abbildungsseitige Unschärfen:*
 - Bewegung
 - Verschmierung
 - Rauschen

Man kann *unscharfe Kantenregionen* mit Hilfe von *zusätzlichem Wissen* auf *scharfe Konturregionen* abbilden.

"Restauration":



2. Determiniertes Signal:

- funktional eindeutig beschrieben und im Prinzip bekannt
- *Aufgaben:* Signal unter anderen Signalen zu detektieren und zu lokalisieren

Beispiel: Kreis $x^2 + y^2 = r_0^2$, r_0 bekannt

3. Nichtdeterminiertes Signal:

- ist einer (bekannten) Funktionsklasse zuzuordnen, aber Parameterspezifizierung ist unbekannt
- *Aufgaben:* Parameteridentifikation und Lokalisierung
Welche Realisierung der Klasse liegt vor?
Beispiel 1: Kreis $x^2 + y^2 = r^2$, r unbekannt
Beispiel 2: Linie: 2 Parameter \implies Hough-Transformation
siehe Einschub: Hough-Transformation zur Linienerkennung

4. Deterministisch-chaotisches Signal:

- funktional beschreibbar (DGL), aber dennoch im Wertebereich nicht vorhersagbar (Biosignale).
- Aufgabe: z.B. Bestimmung der fraktalen Dimension des seltsamen Attraktors
siehe Einschub: Analyse dynamischer Systeme

5. Stochastisches Signal:

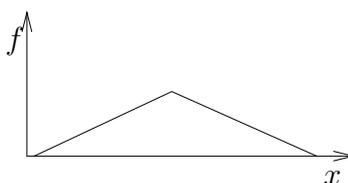
- kann prinzipiell funktional nicht beschrieben werden
- wird als *Realisierung stochastischer Ereignisse (Prozesse)* betrachtet
- Ursachen der Stochastik: Prozeß, Messung, Analyse
- Aufgaben (durch Modellierung der stochastischen Prozesse):
 - *Signalbeschreibung* durch Prozeßparameter (z.B.: Gauß-Prozeß: Mittelwert, Varianz)
 - *Approximation* durch nichtdeterminiertes Signal (z.B.: Methode der kleinsten Quadrate)

6. Stationäres Signal:

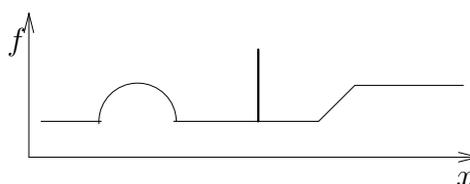
- ein einheitliches Signalmodell wird angewendet (z.B.: bei Textur)

7. Nichtstationäres Signal:

- einheitliches Signalmodell mit *unterschiedlichen Parametern*



- unterschiedliche Signalmodelle
 - regionale Stationarität
 - lokale Stationarität



Einschub: Hough-Transformation zur Linienerkennung

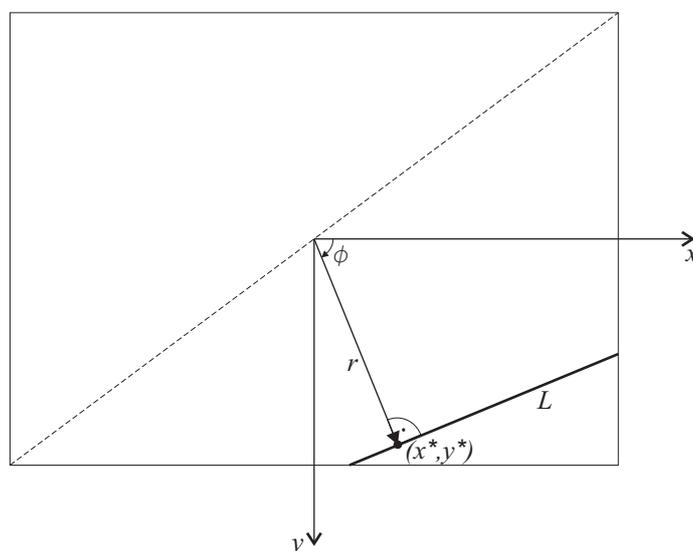
Gegeben : Bild mit Linienstrukturen (z.B. Binärbild)

Gesucht : Parameter der Linien

Lösung : Hough-Transformation

Die Hough-Transformation beruht auf der Annahme eines parametrischen Modelles geometrischer Entitäten, z.B. von Linien. Sie stellt eine Transformation aus dem Ortsraum in den Parameterraum der Entitäten dar. So wie der Ortsraum diskretisiert ist, ist auch der Parameterraum geeignet diskretisiert.

Alle Bildpunkte, die im Ortsraum auf einer Linie liegen, werden im Parameterraum auf einen Punkt projiziert. Genauer, jeder Punkt im Ortsraum erzeugt eine Kurve im Parameterraum und diese Kurven überlagern sich in einem Punkt, der eine Gerade im Ortsraum repräsentiert. Im Falle eines Binärbildes akkumulieren sich die Beiträge im Parameterraum. Dieser wird zu einem Akkumulator. Mittels Schwellenoperation im Parameterraum lassen sich alle Linien im Ortsraum auffinden.



Seien r die vertikale Distanz der Linie L zum Bildzentrum ($r \geq 0$) und ϕ die Orientierung des Distanzvektors ($0 \leq \phi \leq \pi$). Dann gilt für alle $(x, y) \in L$

$$r = x \cos \phi + y \sin \phi$$

bzw.

$$x \cos \phi + y \sin \phi - r = 0.$$

Dies ist die Hessesche Normalform der Geradengleichung. Für zwei Punkte $(x_1, y_1) \in L$ und $(x_2, y_2) \in L$ schneiden sich die entsprechenden Kurven im Punkt (r, ϕ) im Parameterraum.

Vielfältige Gründe (z.B. die Diskretisierungen beider Räume) führen dazu, daß der Schnittpunkt im Parameterraum nur ein Häufungszentrum ist, also die exakte Gleichung $x \cos \phi + y \sin \phi - r = 0$ nur unscharf erfüllt wird:

$$x \cos \phi + y \sin \phi - r \leq \varepsilon, \quad \varepsilon > 0$$

Dies entspricht einer unscharfen Definition der Linien im Ortsraum. Die Schwelle ε ist zu wählen. ■

Einschub: Analyse dynamischer Systeme

- Gegeben : Zeitreihe eines deterministisch-chaotischen Signals
 Annahme : Das System, welches als Quelle dieses Signals angesehen werden kann, möge in seiner Dynamik durch eine Differentialgleichung n -ter Ordnung beschreibbar sein.
 Gesucht : Ordnung n der Differentialgleichung als Maß der Komplexität des Systems
 Lösung : Bestimmung der fraktalen Dimension der Phasenraumtrajektorie des Systems aus der Zeitreihe. Diese Dimension wird mit der Ordnung der DGL in Beziehung gesetzt.

Hier soll nur kurz der Kerngedanke skizziert werden.

Gegeben sei ein dynamisches System f , das von n Parametern abhängt (eine DGL n -ter Ordnung)

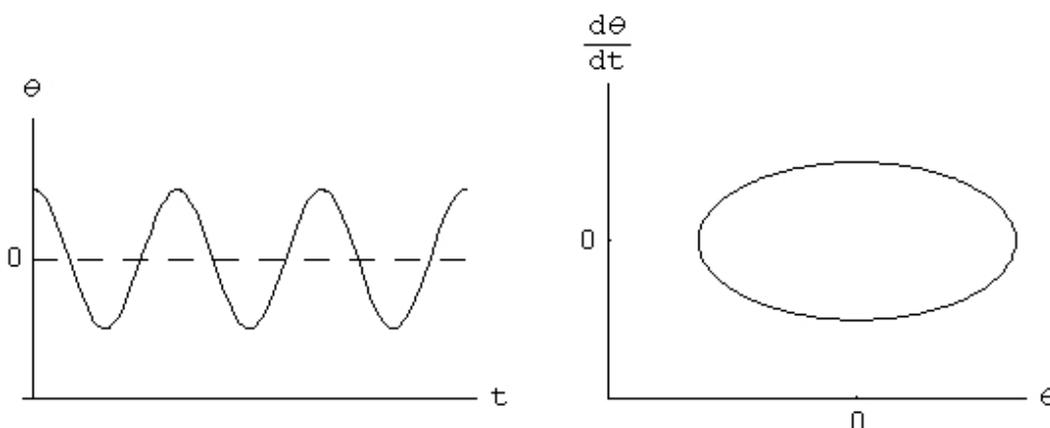
$$f := f(p_1, p_2, \dots, p_n; t) \quad \text{mit} \quad p_i = p_i(t).$$

Dieses System generiert eine Zeitreihe. Dann gilt: Zu jeder Zeitreihendarstellung kann eine äquivalente Phasenraumdarstellung (Darstellung im Parameterraum) angegeben werden.

1. triviales Beispiel: ungedämpfter harmonischer Oszillator (Pendel)

Die Zeitreihe stellt ein periodisches Signal des Auslenkwinkels θ über der Zeit dar. Das Pendel wird durch zwei Parameter beschrieben, den Auslenkwinkel $p_1 = \theta$ und die Bahngeschwindigkeit $p_2 = \frac{d\theta}{dt}$.

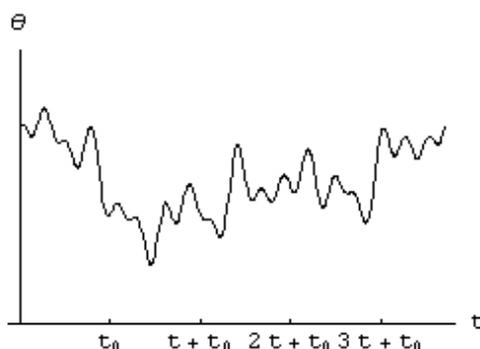
Die genaue DGL interessiert nicht, sie ist von der Ordnung $n = 2$.



Zeitreihe

Trajektorie im Phasenraum

2. komplexes Beispiel: deterministisch chaotisches Signal

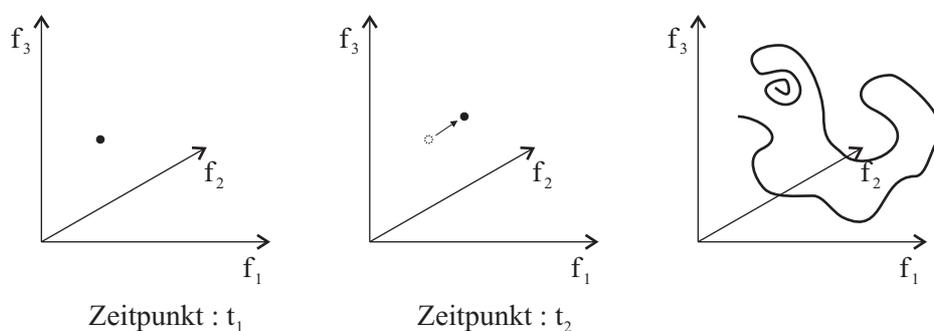


Theorem von Takens (1981):

Bestimme die Dimension der Trajektorie des der Zeitreihe zuzuordnenden seltsamen Attraktors im Phasenraum aus der Korrelationsdimension der Zeitreihe.

Hierzu wird aus der Zeitreihe $f(t)$ eine Menge von Zeitreihen $f_i(t), 1 < i \leq n$ gebildet, deren Korrelation untereinander ermittelt wird.

$$\begin{aligned} f_1(t) &= f(t) \\ f_2(t) &= f(t + \tau) \\ &\vdots \\ f_n(t) &= f(t + (n - 1)\tau) \end{aligned}$$



■

Redundanz und Relevanz

Für die Bildanalyse von grundsätzlicher Bedeutung sind die Eigenschaften *Redundanz* und *Relevanz* von Signalen.

Redundanz ist an die **syntaktische Information** gekoppelt und kann deshalb modelliert werden.

Relevanz ist an die **pragmatische Information** gekoppelt und entzieht sich daher einer Modellierung.

Definition (Redundanz):

Unter *Redundanz* R versteht man den Anteil am Signal, der mit Kopplung des Signals an den Träger (z.B. die Menge der Bildpunkte) gebunden ist, selbst aber keine Strukturinformation trägt. Der Begriff Redundanz gestattet, zwischen notwendigen (*unverzichtbaren*) und überflüssigen (*verzichtbaren*) Daten zu unterscheiden.

$$R = \frac{N_{\text{verzichtbar}}}{N_{\text{gesamt}}} = 1 - \frac{N_{\text{unverzichtbar}}}{N_{\text{gesamt}}}$$

■

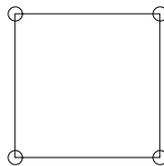
Beispiel:

1. Ein Quadrat ist hinreichend durch die Angabe seiner vier Eckpunkte definiert.

$$R = 1 - \frac{2^2}{2^2} = 0$$

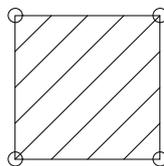
2. Ein Quadrat wird durch seine Kanten definiert (1D): Seien 2^2 Ecken durch $4 \cdot 256$ Kantenpunkte ersetzt. Damit ergibt sich für die Redundanz R :

$$R = 1 - \frac{2^2}{2^2 \cdot 2^8} = 1 - 2^{-8}$$



3. Ein Quadrat wird durch seine Fläche definiert (2D): Seien $256^2 = 2^{16}$ Bildpunkte im Quadrat und 2^2 zu seiner Definition notwendige Ecken. Damit ergibt sich für die Redundanz R :

$$R = 1 - \frac{2^2}{2^{16}} = 1 - 2^{-14}$$



■

Aus diesem Beispiel ist ersichtlich, daß 2D-Strukturen mehr Redundanz enthalten als 1D-Strukturen.

Signalkode: nicht verzichtbare Bausteine eines Signals (Basissystem).

Signalkompression bedeutet Signalkodierung mit Redundanzminderung (optimal: Redundanzminimierung). Durch zusätzliches Wissen (nicht im Code enthalten) wird die Redundanz des Codes vermindert.

Die Signalanalyse erfordert auch eine Unterscheidung von **erwünschter** (relevanter) und **unerwünschter** (nicht relevanter) Information.

Definition (Relevanz):

Unter *Relevanz* versteht man den Anteil des Signals, der in Hinblick auf die Zielstellung von Bedeutung ist.

Relevanz bezieht sich auf den Anteil an Information, der an den Beobachter gebunden ist, d.h. die pragmatische Information. ■

Definition (Irrelevanz):

Unter *Irrelevanz* versteht man den Anteil des Signals, der in Hinblick auf die Zielstellung nicht von Bedeutung ist.

Irrelevanz bezieht sich analog zum Begriff der Relevanz auf den Anteil an verzichtbarer semantischer Information, der an den Beobachter gebunden ist. ■

Triebfeder der Signalanalyse ist das Prinzip, Relevanz zu vermehren bzw. Irrelevanz zu vermindern. Relevanz ist im Gegensatz zur Redundanz nicht generell formalisierbar. *Redundanz* ist in der Bildanalyse deshalb von Bedeutung, weil sie *Stabilisierung gegenüber Störungen* bewirkt.

Beispiel: *Vermessung eines Quadrates*

- a) *Ecken sind ungenau gegeben* → *starke Auswirkungen auf Rekonstruktion des Quadrates*
 - b) *Kanten sind ungenau gegeben* → *Regressionsaufgabe für Rekonstruktion der Kanten aus ungenauen Punkten mit Fehlerminimierung!*
-

Einschub: Signale als Systeme

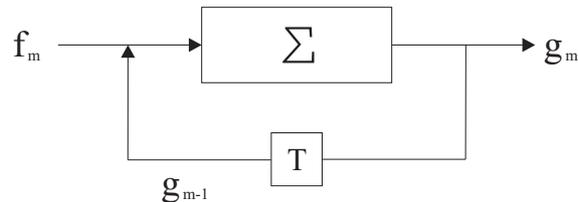
Wir erwähnten die enge Verbindung von Signaltheorie und Systemtheorie. Hier soll eine spezielle aber bedeutende Methode der Signalanalyse angesprochen werden, die auf einer operationalen Definition von Signaleigenschaften beruht. Diese Methode heißt

Analyse durch Synthese

Wir werden im Kapitel 4 (stochastische Modellierungen) hierauf genauer eingehen.

Aufgabe: Bestimme die Parameter eines stochastischen Systemes, welches das vorliegende Signal synthetisieren kann. (Erinnerung: Analyse eines dynamischen Systems, das ein deterministisch-chaotisches Signal erzeugt.) Dies geschieht durch Minimierung eines Fehlermaßes in einem Optimierungsverfahren. Das so parametrisierte System steht für das Signal.

Annahme: Ein rekursives System filtert den Input (Rauschen) in der Weise, daß diesem die gewünschten Korrelationen aufgeprägt werden.



Das System wird durch eine Differenzgleichung (Analogon zu DGL) beschrieben. Der momentane Output wird aus dem momentanen Input und zurückliegendem Output beschrieben.

Der zurückliegende Output beschreibt den Zustand des Systems in zurückliegenden Phasen. Es genügen wenige Zustände, um die Textur eines Bildes zu beschreiben, wenn das Modell gut genug ist.

■

1.1.3 Systemmodelle

Seien $f \in F$ und $g \in G$ Signalrepräsentationen und F, G Signorräume.

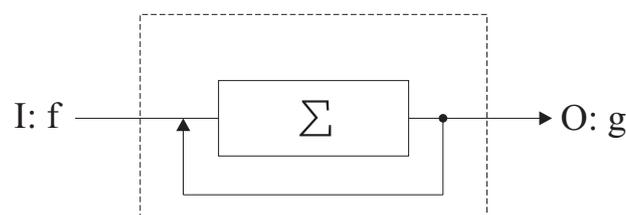
Definition (System):

Ein System Σ beschreibt die Relation zwischen dem Input f und dem Output g bzw. eine Abbildung von f auf g .

- $\Sigma = (f, g)$: System
- $g = h\{f\}$: Abbildung der Signalrepräsentation

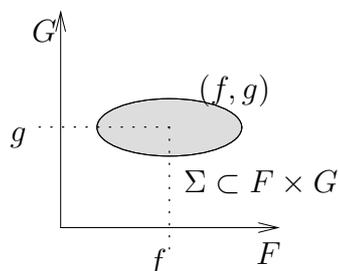
■

Blockschema: Input-Output-Relation



Jedes rekursive System kann auf abstrakterer Ebene als vorwärts gerichtetes System betrachtet werden.

Der Begriff System steht für die in konkreten Anwendungen verwendeten Begriffe Operator, Filter, Transformation, Algorithmus, u.s.w. Ein System ist eine Untermenge des Realisierungsraumes $F \times G$, der von den Signalräumen F und G aufgespannt wird.



Daraus folgt, daß Signale und Systeme bzw. Signalmodelle und Systemmodelle sich gegenseitig bedingen!

Signaltheorie \iff Systemtheorie

(Wahl des Begriffs hängt von Fragestellung ab!)

Signalanalyse:

Mit Hilfe eines bekannten Systems werden die Parameter eines unbekanntes Inputsignals ermittelt.

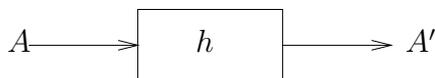
Systemanalyse:

Mit Hilfe eines bekannten Inputsignals werden die Parameter eines unbekanntes Systems (*black box*) ermittelt.

Die Eigenschaften der Systeme prägen sich dem Output auf:

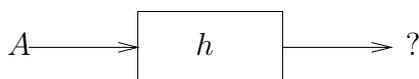
- **deterministisches System**

Abbildung Input-Output eindeutig und vorhersagbar (\sim Verschmierung)

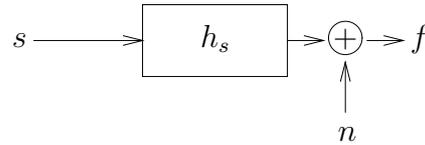


- **stochastisches System**

Abbildung Input-Output mehrdeutig und nicht vorhersagbar (\sim Rauschen)



Beispiel: Beobachtungsgleichung eines Abbildungssystems:

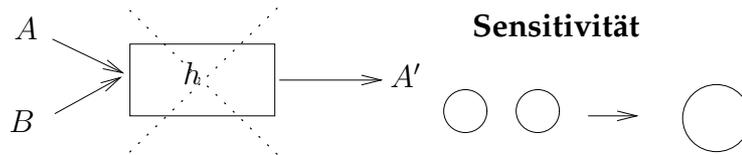


- $f = h_s * s + n$, wobei h_s eine deterministische Störung des Signals s bewirkt und n eine stochastische Störung ist. Man bezeichnet $*$ als Faltungsoperator.



Die folgenden Systemeigenschaften sind von Bedeutung.

1. Vollständigkeit des Systems:



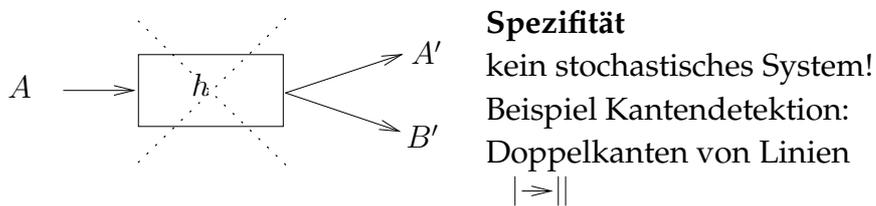
Keine (im Sinne der Modellbildung) erwünschten Strukturen werden unterdrückt.

Sensitivität : Grad der Vollständigkeit;

Beispiel: Die Verschmelzung benachbarter Strukturen bei der Kantendetektion

2. Eindeutigkeit des Systems:

Keine (im Sinne der Modellbildung) zusätzlichen Strukturen werden durch das System erzeugt.



Spezifität: Grad der Eindeutigkeit

Beispiel: Entstehen von Doppelkanten bei der Kantendetektion

Einschub: Spezifität und Sensitivität

Die Begriffe Spezifität und Sensitivität spielen in der statistischen Entscheidungstheorie eine tragende Rolle. Da sie aber auch den eingeführten Begriffen Eindeutigkeit und Vollständigkeit eines Systems entsprechen, sollen sie hier kurz beleuchtet werden.

Annahme: Es existiere ein Operator zur Kantendetektion, der parametrisch an die eine oder andere oben genannte Eigenschaft angepaßt werden kann.

Gesucht: Strategien, die diese Parameteranpassung ermöglichen.

Lösung: Experimente mit unterschiedlichen Parametern und ihre statistische Auswertung.

Folgende Situationen werden in der *Entscheidungsmatrix* in Beziehung gesetzt:

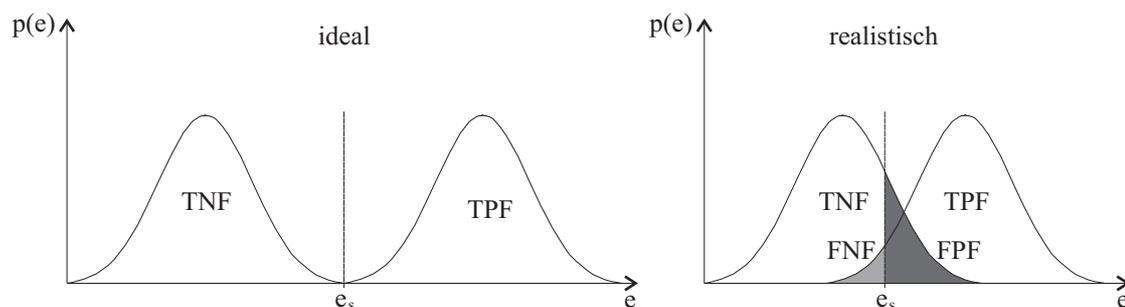
- Eine Struktur (Kante) ist vorhanden
- Eine Struktur (Kante) wird erkannt

Entscheidungsmatrix:

		Struktur wird erkannt	
		ja	nein
Struktur ist vorhanden	ja	TP	FN
	nein	FP	TN

TP : true positive
 TN: true negative
 FP : false positive
 FN: false negative

Die Häufigkeiten $p(e)$ der (negativen bzw. positiven) Antworten über dem Entscheidungskriterium (Merkmal e) lassen sich nur im Idealfall gut mittels einer Entscheidungsschwelle e_s separieren. In der Realität überlagern sich beide Häufigkeiten. Eine Verschiebung der Entscheidungsschwelle führt zu Änderungen der Fehlerhäufigkeiten. Um hier richtig zu handeln, gibt es keine allgemeine Lösung. Vielmehr muß eine Strategie gewählt werden, welche die Konsequenzen der zugelassenen Fehler berücksichtigt.



Normierung der Häufigkeitsfunktionen:

$$\begin{aligned}
 TPF + FNF &= 1 & TPF &= \frac{N_{TP}}{N_P} & , & & TNF &= \frac{N_{TN}}{N_N} \\
 TNF + FPF &= 1 & FPF &= \frac{N_{FP}}{N_P} & , & & FNF &= \frac{N_{FN}}{N_N}
 \end{aligned}$$

Dann gilt

Spezifität: $TNF = 1 - FPF$

Um eine hohe Spezifität zu erreichen, muß der Anteil der falsch-positiven Antworten minimiert werden. Dies erhöht den Anteil der falsch-negativen Antworten.

Beispiel: Kantensuche - das Hochsetzen der Schwelle führt dazu, daß nur noch die stärksten Gradienten im Bild erfaßt werden. Nachteil: Es entstehen Lücken → Nachbearbeitung ■

Beispiel: Differentialdiagnostik - die negativen Antworten werden in weiteren Schritten spezifisch untersucht. Es besteht hoher Grad an Sicherheit, daß $PF \approx TPF$. ■

Sensitivität: $TPF = 1 - FNF$

Um eine hohe Sensitivität zu erreichen, muß der Anteil der falsch-negativen Antworten minimiert werden. Dies erhöht den Anteil der falsch-positiven Antworten.

Beispiel: Kantensuche - das Herabsetzen der Schwelle führt dazu, daß auch schwache Gradienten im Bild erfaßt werden. Nachteil: Es werden Nicht-Kanten-Punkte erfaßt → Nachbearbeitung. ■

Beispiel: Screening - die positiven Antworten werden in weiteren Schritten spezifisch untersucht. Es besteht hoher Grad an Sicherheit, daß $NF \approx TNF$. ■

3. Robustheit:

Empfindlichkeit eines Systems bei Änderung der Voraussetzungen (Modellannahmen).

⇒ Modellannahmen für System sollten *robust* sein.

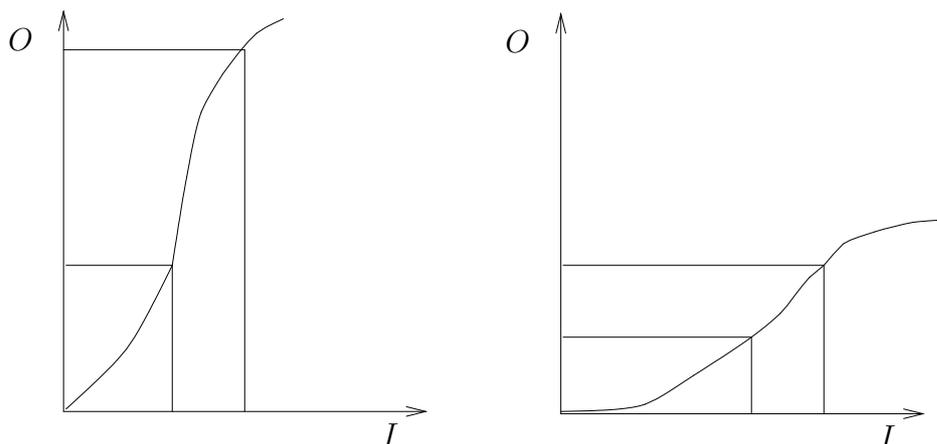
Beispiel: relative Unabhängigkeit eines Kantenoperators von Bildkontrast.



4. Stabilität:

Grad der *Toleranz* gegenüber kleinen Schwankungen der Inputwerte oder bezüglich des Ausfalls von Systemkomponenten.

- instabiles System reagiert auf kleine Störungen des Inputs mit großen Amplitudenänderungen der Outputwerte.



5. Optimalität:

Verhalten eines Systems, dessen Parameter so gewählt sind, daß sie eine *Kostenfunktion* minimieren

Kosten:

- Abweichung Output vom gewünschten Output $C = |g - g_{opt}|$
- Aufwand zum Erreichen des Zieles

Beispiel:

- Maximierung einer Wahrscheinlichkeit (MAP-Verfahren; Bayes-Strategie),
- Minimierung des mittleren quadratischen Fehlers (MMSE-Verfahren: $|g - g_{opt}|^2 \rightarrow \min$).



- Optimalität, Robustheit und Stabilität hängen eng miteinander zusammen.
- Suboptimale Systeme sind oft robuster oder stabil.

6. Adaptivität:

Adaptivität ist die Fähigkeit eines Systems, sich an Änderungen der Parameter eines Signals anzupassen. Das System ist nicht mehr unabhängig vom Signal.

Adaptive Systeme sind bez. der Systemparameter nichtstationär;
 ~ die Systemparameter sind vom Inputsignal abhängig.

- Kosten eines Optimalverfahrens werden signalabhängig verändert (lokale Optimalverfahren).
 - bessere Approximation einer gewünschten Outputfunktion
 - bessere Sensitivität (Vollständigkeit) des Systems (oft *nichtlinear* wegen der Wechselwirkung von Signal und System).

7. Linearität:

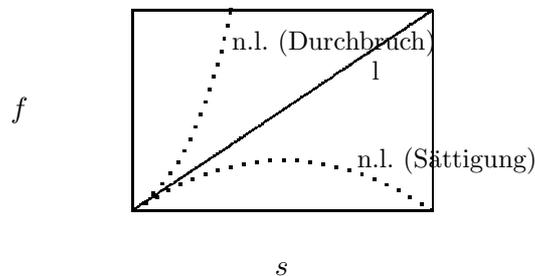
- Lineare Abbildung von Input und Output:

$$h\{a_1 f_1 + a_2 f_2\} = a_1 h\{f_1\} + a_2 h\{f_2\}$$

- gewünschte Eigenschaft technischer Systeme (Kamera, Verstärker, etc.)
- aber: biologische visuelle Systeme sind nichtlinear!
- Linearität bedeutet oft Unabhängigkeit der Systemparameter vom Eingangssignal.
- Nichtlinearität ist oft durch Abhängigkeit der Systemparameter vom Eingangssignal bedingt.

Beispiel:

$$\begin{aligned} h_s(s) &\sim \text{nichtlinear}, \\ h_s &\sim \text{linear}. \end{aligned}$$



■

1.1.4 Strategische Modelle

Strategische Modelle werden herangezogen, um die Art und Weise des Erreichens der Zielstellung zu spezifizieren.

- Reduzierung der Komplexität der Systeme und/oder Problemstellung
- Reduzierung der Freiheitsgrade der Signale

Transformationen:

Signale werden in andere Darstellungsformen (anderes Basissystem) zu folgenden Zwecken transformiert:

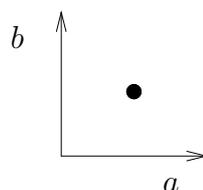
1. Reduzierung der Komplexität von Operationen

- *logarithmische Darstellung:*
Multiplikation \rightarrow Addition

<i>Transformation</i>	<i>ursprüngliche Operation</i>	<i>transformierte Operation</i>
Fouriertransformation	Faltung	Multiplikation
zahlentheoretische Transformation	Faltung	Addition
Laplacetransformation	Differentiation	Multiplikation
konforme Transformation	Skalierung, Rotation	Verschiebung
Hough-Transformation	Auffinden und Identifizieren geometrischer Objekte	Maximumsuche im Parameterraum

Beispiel: *Houghtransformation*

$y = ax + b$, wobei $a, b \sim$ Parameter
Gerade \rightarrow *Punkt*



■

2. Hervorheben invarianter Eigenschaften:

- Fouriertransformation: Verschiebungsinvarianz
- Mellintransformation [Bracewell, 1984]: Skalierungs-/Rotationsinvarianz

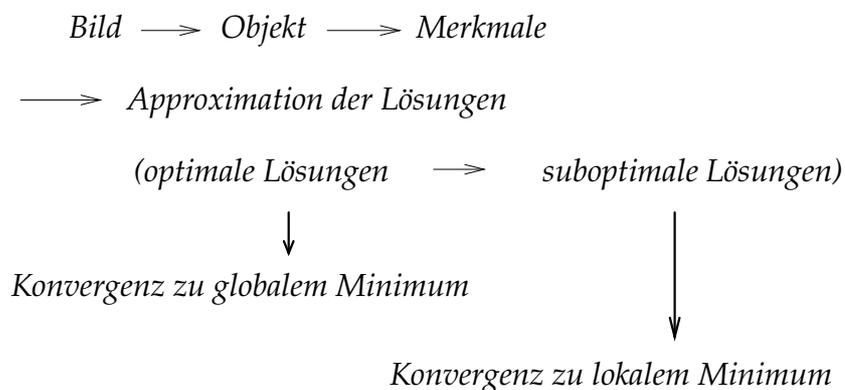
Teile- und Herrsche-Strategie: (divide and conquer)

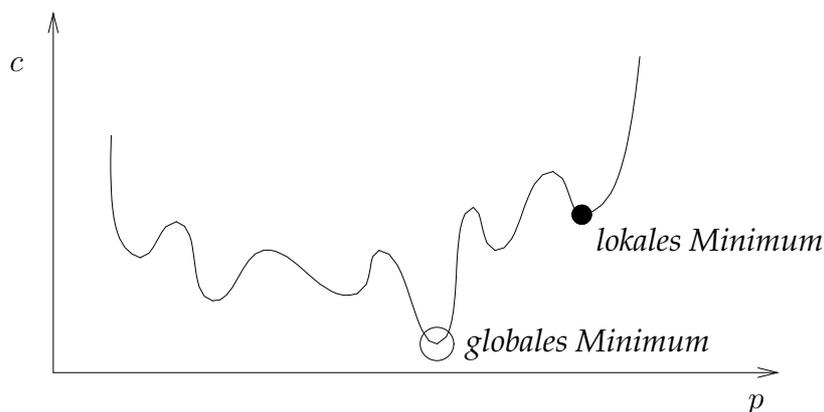
Man versteht darunter die Abbildung eines Problems auf eine Menge kleiner Probleme, die leichter lösbar sind.

1. Reduziere alle Probleme auf Teilprobleme, die beherrschbar sind.
2. Löse Teilprobleme.
3. Rekonstruiere die Lösung des Problems aus Lösungen der Teilprobleme.

Dadurch können aber Wechselwirkung und Korrelation verloren gehen.

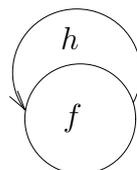
Beispiel:





1. **Rekursion:** Abbildung der Problemlösung auf zyklischen Graphen:

$$f^{q+1} = h\{f^q\} \quad g = f^t$$



Abbruch an Terminalpunkt, für den Ziel- oder Kostenfunktion vorgegebene Schwelle erreicht.

2. **Iteration:**

Abbildung der Problemlösung auf linearen Graphen:

$$g = h\{f\} = h^q\{h^{q-1}\{\dots\{h^1\{f\}\}\dots\}\}$$

Zum Beispiel: Kaskadierung von Elementaroperationen:

$$f \longrightarrow \boxed{h^1} \longrightarrow g^1 \longrightarrow \boxed{h^2} \longrightarrow g^2 \longrightarrow \dots \longrightarrow g^{q-1} \longrightarrow \boxed{h^{q-1}} \longrightarrow g$$

Äquivalenz von Iteration und Rekursion bei identischen Elementaroperationen.

3. **Separierung:**

Unabhängige Behandlung von Teilproblemen (bei Nichtbeachten anderer Probleme), z.B.: Das Einfrieren eines Freiheitsgrades des Problems reduziert die Komplexität:

$$2D \longrightarrow 1D$$

Beispiel:

- $3D \rightarrow 2D$ (Projektion der Welt auf Bildebene)
- Tomographie ($2D \rightarrow 1D$) \rightarrow Radontransformation

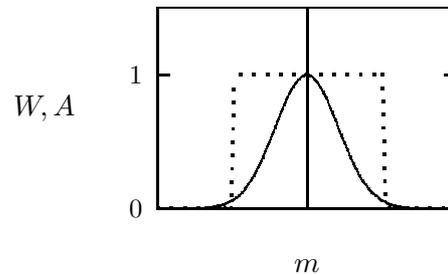
■

5. Ausblendung:

Anwendung des Aperturoperators (Fensteroperator) \mathcal{A} , blendet einen Teil des Signals aus.

$$\mathcal{A}f_m = \begin{cases} f_m & \text{für } m \in D_A \subset D_m \\ 0 & \text{sonst} \end{cases}$$

- Aperturoperator ist Spezialfall eines Wichtungsoperators
- Wichtungsfunktion: $\mathcal{W}f = [w_m f_m]$



1.2 Diskrete Signalfunktionen

In diesem Abschnitt werden grundlegende Begriffe zur Beschreibung diskreter Signalfunktionen eingeführt. Insbesondere wird die Beziehung zwischen dem Signal und seinem Definitionsbereich diskutiert. Von besonderer Bedeutung für die im Abschnitt 1.3 einzuführenden Operatoren sind die Begriffe *Nachbarschaft* und *Umgebung* eines Bildpunktes.

1.2.1 Signal und Träger

Der Begriff *Träger* steht für den Definitionsbereich des Signals. Man unterscheidet:

- *physikalischer* Träger: Art und Weise der Verknüpfung der Zellen des Definitionsbereiches (inhärente Struktur).
- *logischer* Träger: Art und Weise des Zugriffs auf das Signal (Abtastung des Signals).

physikalischer Träger:

Repräsentiert meist Ortsraum oder Zeit, aber auch Frequenz (MRI Magnetic Resonance Imaging, SAR Synthetic Apertur Radar).

Forderung an den physikalischen Träger: Struktur eines *Gitters* (topologischer Begriff):

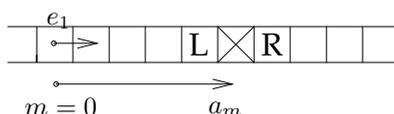
1. Die Struktur ist *zusammenhängend*.
2. Die Struktur ist *homogen*, d.h.
 - jeder Bildpunkt hat die gleiche Anzahl von Nachbarn und
 - jeder Bildpunkt hat gleiche Gestalt und Orientierung.

Nur *quadratische* und *hexagonale Gitter* erfüllen diese Forderungen.

Die Forderungen sind Zugeständnisse an

- die Grenzen der technischen Realisierung (Kameras, Bildspeicher) und an
- die Grenzen der Modellierung von Verarbeitungsstrategien

1D-Träger: lineares Gitter



topologische Eindeutigkeit: jeder Bildpunkt hat linken und rechten Nachbarn

- Gitter \cong topologischer Raum \rightarrow Vektorraum!
- Vektorraum:
 1. Träger hat Koordinatenursprung ($m = 0$)

2. Träger hat Basis (Einheitsvektor e_1)

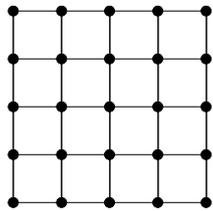
Einheitsvektor definiert *elementare Verschiebung*, um von einem Punkt des Trägers zum nächsten zu gelangen:

$$a_m = m \cdot e_1 \quad a_m \in D_m = \{0, 1, 2, \dots\}$$

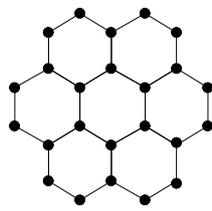
2D-Träger: planares Gitter

Es besteht topologische Mehrdeutigkeit:

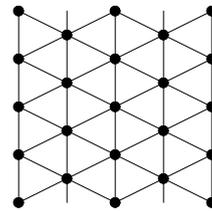
- | | | |
|---|---|---|
| <ol style="list-style-type: none"> 1. Viereckgitter 2. Sechseckgitter 3. Dreieckgitter | } | → liefern vollständige und topologisch
homogene Zerlegung der Ebene. |
|---|---|---|



Viereckgitter



Sechseckgitter



Dreieckgitter

Basis des Viereckgitters: orthogonale Basis $e = [e_1, e_2]^T$ von Einheitsvektoren.

- Bildpunkt (m, n) ist durch Vektor $a_{mn} = me_1 + ne_2$ erreichbar
- Einheitsvektoren definieren Größe *und* Richtung einer elementaren Verschiebung.
Übergang zu schiefwinkligem Gitter für: $\alpha = \angle(e_1, e_2) \neq 90^\circ$
→ Basis: $e^d = [e_1^d, e_2^d]^T$. In Vektornotation ergibt sich:

$$e^d = De \quad \text{wobei} \quad D = \begin{bmatrix} d_{11} & d_{21} \\ d_{12} & d_{22} \end{bmatrix}$$

$$\alpha = 90^\circ$$

$$\alpha = 60^\circ$$

orthogonales Viereckgitter: hexagonales Gitter:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$$

Beim Dreieckgitter ($\alpha = 120^\circ$) existieren drei linear abhängige (Basis-) Vektoren.

- Es ist nicht metrisch homogen (nicht geeignet, den topologischen Raum in einen Vektorraum eindeutig abzubilden).

→ Es existieren keine metrisch eindeutigen Verschiebungen!

1.2.2 Abtastfunktion und Signalbegrenzung

Diracimpuls

- diskret: Einheitsimpuls

$$\delta_m := \begin{cases} 1 & \text{für } m = 0 \\ 0 & \text{sonst} \end{cases}$$

- kontinuierlich: Der *Diracimpuls* (*Deltaimpuls*) $\delta(x)$ repräsentiert als ein Vertreter der Distributionen¹ eine Klasse von Funktionen, die allein über ihre Eigenschaften bei Integration beschrieben sind. Der Diracimpuls ist hier ausreichend definiert durch die Eigenschaft:

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

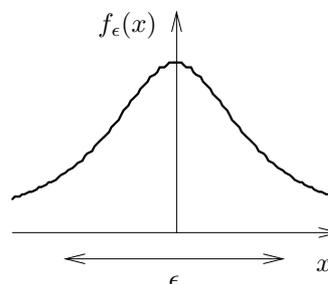
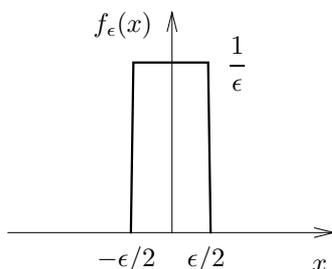
Insbesondere gilt dann:

$$\begin{aligned} \delta(x) &= 0 & \text{für } x \neq 0 \\ \int_{-\infty}^{\infty} \delta(x) dx &= 1 \end{aligned}$$

Der Diracimpuls geht durch Grenzübergang aus der Funktionen $f_\epsilon(x)$ hervor:

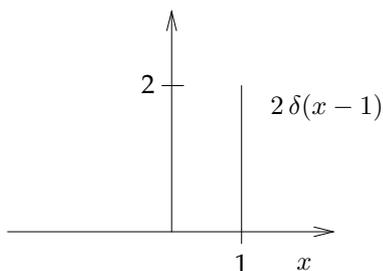
$$\delta(x) := \lim_{\epsilon \rightarrow 0} f_\epsilon(x)$$

$$f_\epsilon(x) = \begin{cases} 1/\epsilon & \text{für } |x| < \epsilon/2 \\ 0 & \text{sonst} \end{cases} \quad f_\epsilon(x) = \frac{2\epsilon}{\epsilon^2 + 4\pi^2 x^2}$$



Die grafische Notation für eine Deltafunktion soll eine Linie sein, die die Länge des Impulsintegrals der Deltafunktion besitzt, also z. B.:

¹Damit ist $\delta(x)$ eigentlich keine Funktion, wird aber trotzdem meist als Funktion bezeichnet.



Folgende Sätze gelten:

$\int_{-\infty}^{\infty} \delta(x - x_0) f(x) dx = f(x_0)$	Verschiebung, Ausblenden
$\int_{-\infty}^{\infty} \delta(ax) f(x) dx = \frac{1}{a} f(0)$	Ähnlichkeit
$\int_{-\infty}^{\infty} \delta^{(n)}(x) f(x) dx = (-1)^n f^{(n)}(0)$	n -fache Differentiation

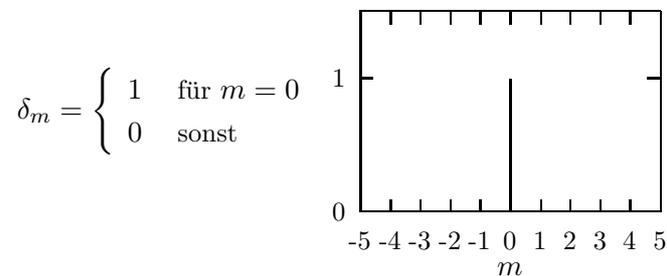
diskrete Version der Ausblendungseigenschaft:

$$f_m = \sum_k f_k \delta_{m-k}$$

unbegrenzte Signal:

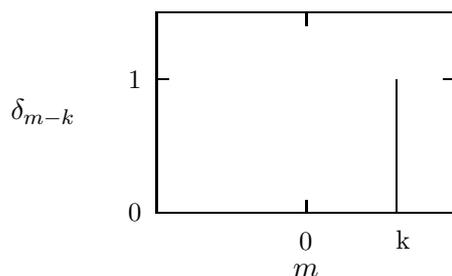
$f = [f_m]$, wobei $m \in D \equiv \mathbb{Z}$ ist.

- Einheitsimpuls am Koordinatenursprung:



- Verschiebung an Position k :

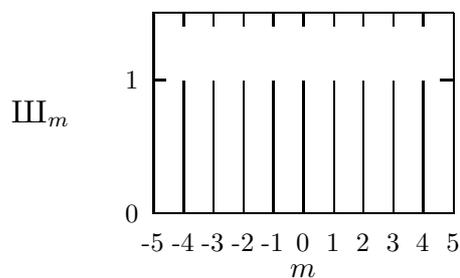
$$\delta_{m-k} := \begin{cases} 1 & \text{für } m = k \\ 0 & \text{sonst} \end{cases}$$



- Überlagerung aller möglichen Verschiebungen:

Wert am Ort m wegen Ausblendeigenschaft für $f_k = 1$

$$\text{III}_m = \sum_{k=-\infty}^{\infty} \delta_{m-k}$$



$\text{III} \sim$ unendliche Folge von Einheitsimpulsen \cong ideale Abtastfunktion, Kammfunktion oder Shahfunktion.

$$\text{III} = [\text{III}_m] = \left[\sum_{k=-\infty}^{\infty} \delta_{m-k} \right]$$

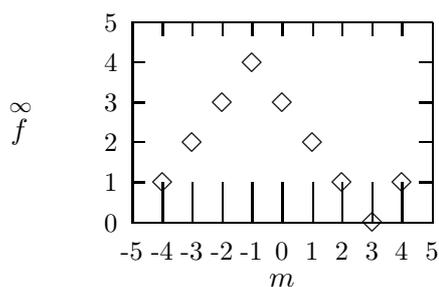
\Rightarrow

- Signalfunktion f ist als Wichtung der Abtastfunktion mit Signalwerten f_m zu interpretieren:

$$f = \left[\sum_{k=-\infty}^{\infty} f_k \delta_{m-k} \right] = [f_m]$$

Dies folgt aus der Ausblendeigenschaft des Deltaimpulses

$$f_m = \sum_{k=-\infty}^{\infty} f_k \delta_{m-k}$$



- Erweiterung auf $r = 2$:

$$\text{III}_{m,n} = \sum_k \sum_l \delta_{m-k,n-l}$$

begrenzttes Signal:

- unbegrenzttes Signal: nützliches Modell, um Probleme des Signalrandes zu vermeiden.
- begrenzttes Signal: durch Ausblenden mittels Aperturoperator \mathcal{A} , der auf die Abtastfunktion wirkt, so daß

$$\mathcal{A} f = \begin{cases} f & \text{für } m \in \{0, 1, 2, \dots, M-1\} \\ 0 & \text{sonst} \end{cases}$$

$$f = \left[\sum_{k=0}^{M-1} f_k \delta_{m-k} \right] = [f_m] \quad \text{mit } m \in \{0, 1, 2, \dots, M-1\}$$

- Der Signalrand erfordert besondere Beachtung!
- Die Abtastfunktion entspricht der *kanonischen Basis des Signalraumes*, denn es gilt:

$$\delta_{m-k} = e_k^m = e_m$$

$$f = [f_0, f_1, f_2, \dots, f_{M-1}]^T \quad \text{Signal im Ortsraum}$$

$$= f_0 \underbrace{[1, 0, 0, 0, \dots, 0]^T}_{\delta_{0-k}=e_k^0} + f_1 \underbrace{[0, 1, 0, 0, \dots, 0]^T}_{\delta_{1-k}=e_k^1} + f_2 \underbrace{[0, 0, 1, 0, \dots, 0]^T}_{\delta_{2-k}=e_k^2} \dots$$

Signal im Signalraum

Sind die e_k die Basisvektoren des Signalraums, so ist die Notation für das Signal

$$f = \sum_{m=0}^{M-1} f_m e_m.$$

periodisches Signal:

- ist modulo der Kardinalität M des Definitionsbereichs erklärt.
- Nutzen: Randprobleme werden umgangen.
- Zwangsgleichung:

$$\tilde{f}_m = \tilde{f}_{m+M} = \tilde{f}_{m-M}$$

- *Periode* D_m : Menge unabhängiger Signalwerte.

$$D_m = \{0, 1, \dots, M-1\} \quad \longrightarrow \quad M = |D|$$

- Ein begrenzttes Signal erhält man aus einem periodischen Signal, durch die Anwendung des Aperturoperators.

- Erweiterung auf $r = 2$: Periodizität allgemeiner definiert

$$\left. \begin{aligned} \tilde{f}_{m,n} &= f_{(m+p_{11}, n+p_{21})} \\ \tilde{f}_{m,n} &= f_{(m+p_{12}, n+p_{22})} \end{aligned} \right\} \text{gemeinsame Zwangsgleichungen}$$

- Periodizitätsmatrix

$$P = \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix}, \quad \det P \neq 0$$

- Allgemeiner Fall für beliebige Perioden: $p_{ij} \neq 0$ bei $D = D_m \times D_n$:

$$p_{11} = M, \quad p_{22} = N, \quad p_{21} = J, \quad p_{12} = K \quad \longrightarrow \quad P = \begin{bmatrix} M & J \\ K & N \end{bmatrix}$$

$$\implies \left. \begin{aligned} \tilde{f}_{m,n} &= \tilde{f}_{m+M, n+J} = \tilde{f}_{m-M, n-J} \\ \tilde{f}_{m,n} &= \tilde{f}_{m+K, n+N} = \tilde{f}_{m-K, n-N} \end{aligned} \right\} \text{Zwangsgleichungen}$$

- Spezialfall: separable horizontale/vertikale periodische Fortsetzung bei $D = D_m \times D_n$:

$$p_{11} = M, \quad p_{22} = N, \quad p_{21} = p_{12} = 0 \quad \longrightarrow \quad P = \begin{bmatrix} M & 0 \\ 0 & N \end{bmatrix}$$

$$\implies \left. \begin{aligned} \tilde{f}_{m,n} &= \tilde{f}_{m+M, n} = \tilde{f}_{m-M, n} \\ \tilde{f}_{m,n} &= \tilde{f}_{m, n+N} = \tilde{f}_{m, n-N} \end{aligned} \right\} \text{Zwangsgleichungen}$$

- Zahl der Bildpunkte einer Periode von \tilde{f} ist $|\det P|$

$$\det P = p_{11}p_{22} - p_{12}p_{21}$$

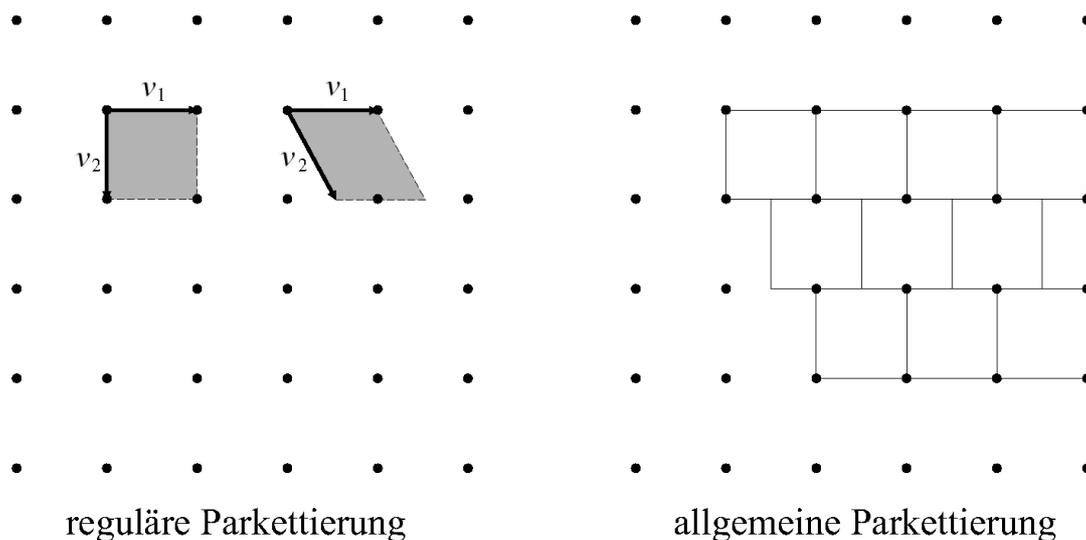
Auf die Tilde zur Kennzeichnung der periodischen Signalfunktion wird in den folgenden Abschnitten verzichtet.

Offensichtlich beschreibt die hier interessierende Periodizität das Erzeugen eines Supergitters, dessen elementare Zellen durch D gegeben sind mit jeweils $|D|$ Bildpunkten. Ist $v = [v_1, v_2]^T$ ein Vektor, der einen ausgewählten Punkt aus D , z.B. den Ursprung $(m, n) = (0, 0)$, in denselben Punkt einer benachbarten Zelle verschiebt, gilt $v_1 = |D_m|e_1$ und $v_2 = |D_n|e_2$.

Das erzeugte Gitter repräsentiert eine reguläre Parkettierung von $\mathbb{Z} \times \mathbb{Z}$ mit Zellen, deren Gestalt durch P gegeben ist und deren Fläche sich durch $|\det P|$ ergibt. Es gilt

$$v = Pe.$$

Beispiel:



$$P = \begin{bmatrix} M & 0 \\ 0 & N \end{bmatrix}$$

$$P = \begin{bmatrix} M & 0 \\ \frac{M}{2} & N \end{bmatrix}$$

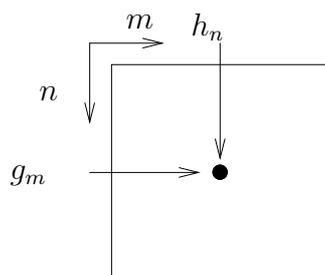
In der Bildverarbeitung interessieren nur reguläre Parkettierungen. In der Computergrafik interessiert man sich auch für allgemeine Parkettierungen (siehe S.G. Hoggar: *Mathematics for Computer Graphics*, Cambridge University Press, 1994).



1.2.3 Zeilenweise Bildabtastung

Die gebräuchlichste Art und Weise der Bildabtastung bei sequentieller Verarbeitung ist die zeilenweise Abtastung. Sie ist modellierbar mittels Separabilität und Linienimpuls. Für $r = 2$: *Separabilität* \rightarrow Signal läßt sich als Produkt zweier orthogonaler 1D-Signale darstellen.

Für den Bildpunkt (m, n) ist die Zerlegung $f_{m,n} = g_m \cdot h_n$ nur im Ausnahmefall möglich.



Aber Verallgemeinerung ist i. allg. möglich:

Linearkombination von Produkten separabler Komponenten:

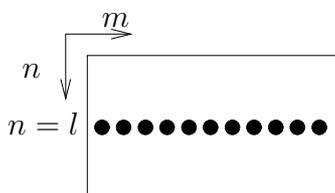
$$f_{m,n} = \sum_{i=0}^{I-1} g_{m,i} h_{i;n}$$

Linienimpuls:

Impulsfolge, die in einer Richtung homogen ist und in dazu senkrechter Richtung impulsförmig:

$$\mathbb{I} \cdot \delta_{n-l} = \left[\sum_{k=0}^{M-1} \delta_{m-k} \right] \cdot \delta_{n-l} \quad \text{für } m \in D_m, \quad n \in D_n$$

$$f_{m,n} = \sum_{l=0}^{N-1} f_{m;l} \mathbb{I}_m \cdot \delta_{n-l}$$



δ_{n-l} selektiert Zeile $n = l$. Die Zeile $n = l$ wird dann wie ein 1D-Signal durch die Shahfunktion $\mathbb{I} = [\mathbb{I}_m]$ abgetastet.

Vektordarstellung der Bildfunktion: Für $D = M \times N$ wird ein Bild als Vektor der Länge $P = MN$ dargestellt.

$$\Rightarrow f = [f_p \mid 0 \leq p \leq P-1, \quad P = M \cdot N, \quad p = nM + m]$$

1.2.4 Logischer Träger

physikalischer Träger : Art und Weise der Einbettung des Signals in den Definitionsbereich

logischer Träger : Art und Weise des Zugriffs auf Signal

Beispiel: MVS (Menschliches Visuelles System) \sim selektives (aufgabenbezogenes) irreguläres Abtasten. ■

- *irreguläre Abtastung:*

Schrittweite und Abtastung orts- bzw. zeitabhängig (\sim hierfür Signaltheorie bisher kaum entwickelt).

- inhomogenes Gitter \sim irreguläre Abtastung

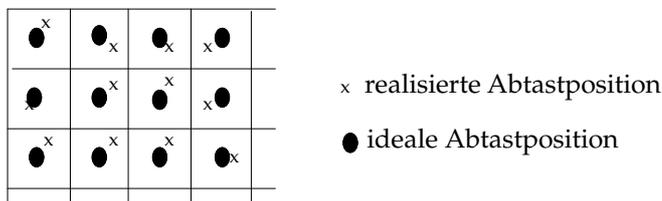
- *reguläre Abtastung:*

Schrittweite und Abtastrichtung verhalten sich homogen.

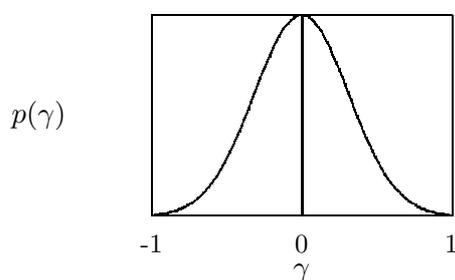
- homogenes Gitter \sim reguläre Abtastung

Jittering:

Zugelassene kleine Störung einer regulären Abtastposition:



z.B.: Abtastung gestört um zufälligen Wert $|\gamma| < 1$ mit Verteilungsdichte $p(\gamma)$ um Wert Null.

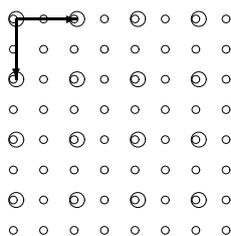


Anwendung: "natürliche" Darstellung von Computergraphiken

Resampling:

Reguläres Abtasten des Signals in Schrittweiten, die Vielfachen oder Bruchteilen der Einheitsvektoren des physischen Trägers ($\Delta = 1$) entspricht

Beispiel: $\Delta = 2$



Abtastfunktion bei Resampling:

$$\begin{aligned} \text{III}_{\Delta m} &= \sum_{k=-\infty}^{\infty} \delta_{m-\Delta k} \\ \hat{f}'_m &= f_m \cdot \text{III}_{\Delta m} \end{aligned}$$

\Rightarrow Abbildung des Definitionsbereichs $D_m \rightarrow D_m^\Delta$ mit $M^\Delta = \frac{M}{\Delta}$

$\Delta > 1$: Kontraktion (Stauchung) durch Unterabtastung:



logischer Träger wird kleiner als physikalischer Träger

$\Delta < 1$: Expansion (Streckung) durch Überabtastung:



logischer Träger wird größer als physikalischer Träger

Falls A eine zweidimensionale Matrix ist mit $|\det A| \neq 0$, so gilt für das abgetastete Signal \hat{f} :

$$\hat{f}_{\vec{m}l} = f_{A\vec{m}} \implies f_{m,n,l} = f_{m,n} \cdot \prod_{a_{11}m+a_{12}n, a_{21}m+a_{22}n}$$

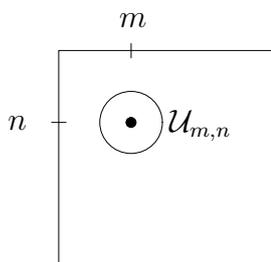
A steht für Streckung, falls $|\det A| > 1$. A steht für Stauchung, falls $|\det A| < 1$.

1.2.5 Nachbarschaften und Umgebung

Nachbarschaften und Umgebungen sind *Fensterfunktionen*, die einen Bildausschnitt bezüglich eines *Aufpunktes* (m, n) definieren:

$$\mathcal{A}_{m,n}f = \{ f_{i,j} | (i, j) \in \mathcal{U}_{m,n} \subset D \}$$

$\mathcal{U}_{m,n}$ ist der Einzugsbereich des Aperturoperators \mathcal{A} an der Position (m, n) .



Distanz: $p \sim$ Metrik-Parameter

- *Distanzfunktion:*

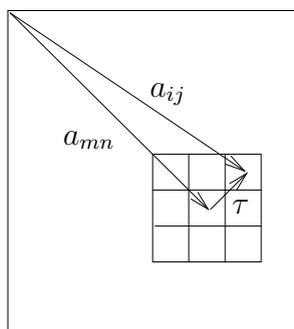
$$d_p(a_{mn}, a_{ij}) = \|a_{mn} - a_{ij}\|_p = \|\tau\|_p$$

- *Distanzvektor:*

$$\tau = [\xi, \eta]^T$$

- Norm des Distanzvektors:

$$\|\tau\|_p = (|\xi|^p + |\eta|^p)^{\frac{1}{p}}$$



wichtige Metriken:

- absolute (Minkowski- oder City-Block-) Metrik: $p = 1$

$$d_1((m, n), (i, j)) = |i - m| + |j - n|$$

- Euklidische Metrik: $p = 2$

$$d_2((m, n), (i, j)) = ((i - m)^2 + (j - n)^2)^{\frac{1}{2}}$$

- Maximum- (Tschebyscheff-) Metrik: $p = \infty$

$$d_\infty((m, n), (i, j)) = \max\{|i - m|, |j - n|\}$$

Definition (Nachbarschaft \mathcal{N}):

Die Nachbarschaft \mathcal{N} eines Bildpunktes (m, n) ist durch die Menge aller Bildpunkte (i, j) gegeben, deren Abstand $d_p = 1$ ist.

- $(m, n) \sim$ Aufpunkt
- $(i, j) \sim$ Testpunkt



absolute Metrik ($p = 1$):

	N	
N	A	N
	N	

$$A \sim (m, n) \quad \sim 4\text{-Nachbarschaft}$$

$$N \in \mathcal{N} \quad \mathcal{N}_4 = \{(i, j) \mid d_1((i, j), (m, n)) = 1\}$$

$\tau = |\xi|e_1 + |\eta|e_2$ entspricht elementarem Verschiebungsvektor in Richtung Basis der Signalfunktion

\implies Ecken sind nicht erreichbar, denn $\tau = [1, 1]^T \implies \|\tau\|_1 = 2$.

\sim City-Block-Metrik

Euklidische Metrik ($p = 2$):

N	N	N
N	A	N
N	N	N

$$\tau = [1, 1]^T \implies \|\tau\|_2 = \sqrt{2} \implies 1 \quad !$$

- Euklidische Metrik nicht sehr geeignet für Nachbarschaftsdefinition im kartesischen Koordinatensystem

→ Wichtung der Nachbarschaftsbildpunkte (Komponenten des Distanzvektors) erzeugt eine bessere Approximation der Euklidischen Metrik:

z.B.: Chamfer-Metrik:

$$\tau = [3, 3]^T \mapsto \|\tau\|_{CM} = \frac{3}{4}\sqrt{2}$$

4	3	4
3	0	3
4	3	4

- Euklidische Metrik ist ideal in polar separablem Koordinatensystem, da dieses gegenüber Rotation invariant ist, falls der Aufpunkt um Ursprung liegt.

Maximum-Metrik ($p = \infty$):

N	N	N
N	A	N
N	N	N

~ 8-Nachbarschaft: $\mathcal{N}_8 = \{(i, j) \mid d_\infty((i, j), (m, n)) = 1\}$

$$\tau = [1, 1]^T \mapsto \|\tau\|_\infty = \max(1, 1) = 1$$

Definition (Markovmasche \mathcal{M}):

Eine Markovmasche \mathcal{M} der Ordnung q ist durch alle Bildpunkte (i, j) gegeben, für deren Distanz $\|\tau\|$ zum Aufpunkt (m, n) gilt: $0 < \|\tau\| \leq q$.



Bemerkung: Eine Nachbarschaft ist eine Markovmasche der Ordnung 1.

diskrete "Kreise" vom Radius 2:

\mathcal{M}_2 absolute Metrik
(City-Block-Metrik)

4	3	2	3	4
3	2	1	2	3
2	1	0	1	2
3	2	1	2	3
4	3	2	3	4

\mathcal{M}_2 Euklidische Metrik

3	2	2	2	3
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
3	2	2	2	3

\mathcal{M}_2 Maximum-Metrik
(Schachbrett-Metrik)

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

Definition (kausales Signal):

Ein Signal $g(x)$ wird als *kausal* bezeichnet, wenn gilt:

$$g(x) = 0 \quad \text{für} \quad x < 0$$

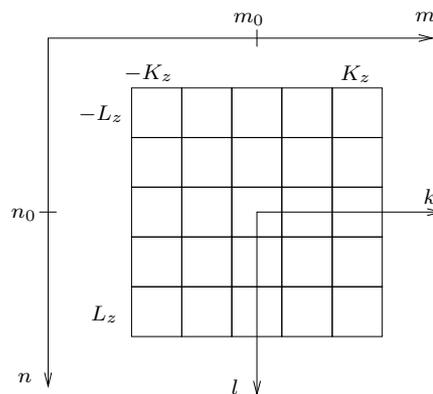


In einem physikalisch realisierbaren System folgt die Wirkung $g(x)$ immer der Ursache (zeitliche Kausalität). Die Antwort eines solchen Systems auf eine Ursache bei $x = 0$ beginnt daher frühestens bei $x = 0$ und ist somit kausal. Die Pixel eines Bildes besitzen — außer durch die

Verarbeitung in einem sequentiellen Rechner — keine bindende zeitliche Ordnung. Bildsignale sind daher oft als akausal zu betrachten.

Anwendungen:

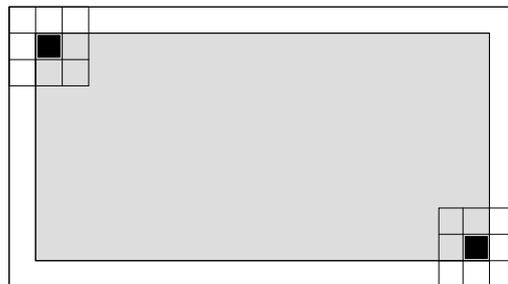
1. *Distanztransformation* (Matching): Bestimmung des Abstandes von Strukturen.
 2. Definition des Einflußbereiches auf den Grauwert am Aufpunkt (Operatorfenster). Die Grauwerte im Einflußbereich werden verrechnet und dem Aufpunkt zugeordnet.
- symmetrischer (akausaler) Einzugsbereich



Für einen Einzugsbereich der Größe $K \times L$, K, L ungerade, eines beliebigen Operators gilt dann mit $K_z = \frac{(K-1)}{2}$ und $L_z = \frac{(L-1)}{2}$:

$$U = \{(k, l) \mid -K_z \leq k \leq K_z, -L_z \leq l \leq L_z\}$$

Bildrand: symmetrisch



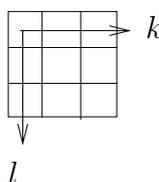
Definitionsbereich der Aufpunktkoordinaten:

$$D_m = \{K_z, K_z + 1, \dots, M - K_z - 1\}$$

$$D_n = \{L_z, L_z + 1, \dots, N - L_z - 1\}$$

$$D = D_m \times D_n$$

- unsymmetrischer (kausaler) Einzugsbereich



Für einen Einzugsbereich der Größe $K \times L$ eines beliebigen Operators gilt dann

$$\mathcal{U} = \{(k, l) \mid 0 \leq k \leq K - 1, 0 \leq l \leq L - 1\}$$

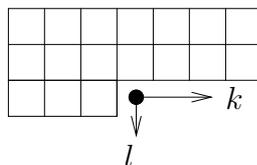
Bildrand: unsymmetrisch



$$\begin{aligned} D_m &= \{0, 1, \dots, M - K\} \\ \text{Definitionsbereich der Aufpunktkoordinaten: } D_n &= \{0, 1, \dots, N - L\} \\ D &= D_m \times D_n \end{aligned}$$

- nichtsymmetrische Halbebenen - Kausalität (NSHP)

“Vergangenheit”



$$\mathcal{U} = \{(k, l) \mid l < 0 : \forall k, \text{ und } l = 0 : k < 0\}$$

1.3 Synopsis der Operatoren

In diesem Abschnitt werden grundlegende Eigenschaften von Operatoren eingeführt. In den folgenden Abschnitten wird die besonders wichtige Klasse der linearen Operatoren behandelt.

Definition (Operator):

Ein Operator \mathcal{O} erzeugt eine Abbildung aus einem Signalraum \mathbb{F} in einen Signalraum \mathbb{F}' :

$$\mathcal{O} : \mathbb{F} \rightarrow \mathbb{F}', g = \mathcal{O}f$$

■

Operatoren in der Bildverarbeitung sind i. allg. Funktionen, die Bilder in die Menge der Bilder abbilden, also $\mathbb{F}' \equiv \mathbb{F}$. Häufig gilt aber auch $\mathbb{F}' \subset \mathbb{F}$ (Bsp: Binarisierungsoperator, Konturfolgeoperator).

Die Signaltheorie ist eng an die Klasse der *linearen Operatoren* geknüpft. Allgemein gilt, daß Operatoren der Struktur des Signalraumes angepaßt sein müssen (metrisch, topologisch).

Wir werden die Darstellung von Operatoren im Ortsraum, im Signalraum und im Frequenzraum behandeln. Die Darstellung im Ortsraum ist an keinerlei Einschränkungen gebunden. Auch topologische/gestaltsändernde Operatoren sind darstellbar. Im Signalraum lassen sich nur lineare Operatoren behandeln. Die Eigenschaften Linearität und Verschiebungsinvarianz sind für die Repräsentation von Operatoren im Frequenzraum erforderlich.

In diesem Abschnitt werden wir eine Synopsis der Operatoren für die Bildverarbeitung einführen und die algebraischen Regeln der Verbindung von Operatoren behandeln.

1.3.1 Lokale und globale Operatoren

Definition (Support eines Operators):

Sei \mathcal{O} ein Operator und f_m, g_m Signalwerte am Ort m mit $g_m = \mathcal{O}f_m$. Dann heißt die Menge $\mathcal{U}(m) \subset D_m$ aller Bildpunkte, die bei der Berechnung von g_m von \mathcal{O} betrachtet werden, *Einzugsbereich* oder *Support des lokalen Operators* \mathcal{O} .

■

Die Definition besagt, daß die Arbeitsweise eines Operators \mathcal{O} von der Menge \mathcal{U} abhängt. Die Koordinaten der Punkte in \mathcal{U} können auf zwei Arten angegeben werden:

1. Transformation in globalen Koordinaten:

$$g_m = \mathcal{O}\{f_i | i \in \mathcal{U}(m)\}$$

m, i heißen dann *globale Koordinaten*.

2. Transformation in lokalen Koordinaten:

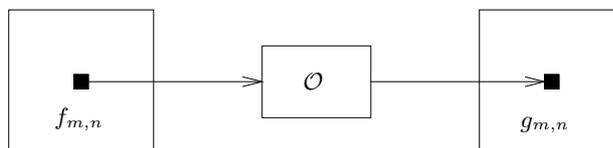
$$g_m = \mathcal{O}\{f_i | k \in \mathcal{U}(m), k = m - i\}$$

k heißt hier *lokale Koordinate*.

Es gibt drei wichtige Arten von Einzugsbereichen \mathcal{U} , die zu einer Typeneinteilung der Operatoren führen:

1. Punktoperator: $\mathcal{U}_{m,n} = (m, n)$

Der Output eines Punktoperators \mathcal{O} für einen Bildpunkt (m, n) hängt nur von der Koordinate und/oder dem Signalwert dieses Bildpunktes ab.



Beispiele:

1. Geometrische Operatoren:

$$g_{m,n} = \mathcal{O}f_{m,n} = f_{m',n'} \text{ mit } \mathcal{O} \in \{\mathcal{S}, \mathcal{T}, \mathcal{M}, \mathcal{R}\}$$

1.1 Skalierungsoperator \mathcal{S}

$$\mathcal{S}f_m = f_{sm}$$

i. Stauchung: $|s| > 1$

ii. Streckung: $|s| < 1$

1.2 Bewegungsoperatoren $\mathcal{T}, \mathcal{M}, \mathcal{R}$

i. Translation $\mathcal{T}f_m = f_{m-1}$

ii. Spiegelung $\mathcal{M}f_m = f_{-m}$

iii. Rotation $\mathcal{R}f_{m,n} = \mathcal{R}f_{\vec{m}} = f_{A\vec{m}} = f_{m',n'}$

A. \mathcal{R} nur für mehrdimensionales Signal sinnvoll

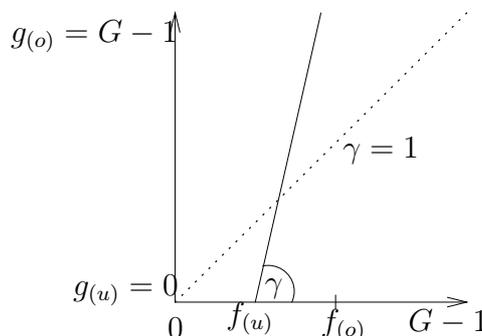
B. \mathcal{R} Spezialfälle auf diskretem Träger:

$$\varphi = n \cdot \frac{\pi}{2} \text{ (Viereckgitter)}$$

$$\varphi = n \cdot \frac{\pi}{3} \text{ (Sechseckgitter)}$$

2. Grauwerttransformation, Amplitudenskalierung oder Quantisierung

2.1 Funktion Tonwertanpassung (z.B. in Photoshop)



Der Operator ergibt sich aus ursprünglicher und gewünschter Spanne des Grauwertbereiches mit der linearen Übertragungsfunktion

$$g_{mn} = (f_{mn} - f_{(u)}) \frac{g_{(o)} - g_{(u)}}{f_{(o)} - f_{(u)}} + g_{(u)}$$

2.2 Im Gegensatz zu dieser linearen Grauwerttransformation beruht die Methode der Histogramm-Egalisierung auf der Anwendung einer nichtlinearen Transformation, die durch die Form des kumulativen Histogramms gegeben ist. Im Idealfall zeigt ein egalisiertes Histogramm für alle Grauwerte die gleiche Häufigkeit.

Für ein Bild der Ausdehnung $M \times N$ Bildpunkte sei $h(f_i)$, $0 \leq i \leq G - 1$, ein normiertes Histogramm,

$$h(f_i) = \frac{\text{card}(f_i)}{M N}.$$

Das kumulative Histogramm $h_c(f_j)$,

$$h_c(f_j) = \sum_{f_i=0}^{f_j} h(f_i)$$

stellt eine monoton wachsende nichtlineare Funktion dar. Diese Funktion kann als Kennlinie einer nichtlinearen Grauwerttransformation dienen. Diese Grauwerttransformation entsprechend

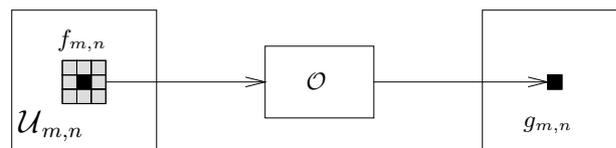
$$g_i = (G - 1) h_c(f_i)$$

führt dazu, daß alle Grauwerte g_i etwa mit gleicher Häufigkeit im Bild auftreten (entspricht etwa der Maximierung der Entropie).

■

2. Lokaler Operator: $\mathcal{U}_{m,n} \subset D$

Der Output eines lokalen Operators \mathcal{O} für einen Bildpunkt (m, n) hängt von den Koordinaten und/oder den Signalwerten an anderen Bildpunkten in einer begrenzten Umgebung $\mathcal{U}_{m,n}$ ab. Ein einzelner Bildpunkt hat nur im lokalen Kontext Bedeutung.



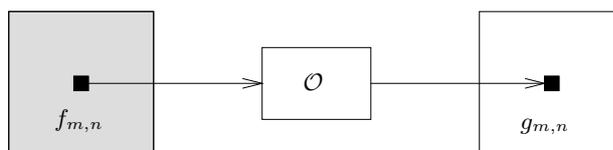
Beispiel:

- lokale Filter
- lokale Interpolationsoperatoren
- morphologische Operatoren

■

3. Globaler Operator: $\mathcal{U}_{m,n} = D$

Der Output eines globalen Operators \mathcal{O} für einen Bildpunkt (m, n) hängt vom gesamten Inputsignal ab. Jeder einzelne Bildpunkt hat nur im globalen Kontext eine Bedeutung, d.h. jeder Bildpunkt des Outputs von \mathcal{O} trägt globale Information des Inputs.

**Beispiel:**

- Integraltransformation
- Frequenzraumfilter
- ideale Interpolation

■

1.3.2 Parallele und sequentielle Verarbeitung

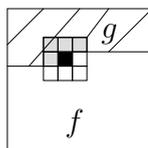
Für lokale Operatoren werden zwei verschiedene Verarbeitungsstrategien unterschieden:

1. Die parallele Verarbeitung:

Die parallele Verarbeitungsstrategie zeichnet sich dadurch aus, daß die Outputmatrix verschieden von der Inputmatrix ist. Das bedeutet, daß die Verarbeitung eines jeden Bildpunktes der Inputmatrix völlig unabhängig von der Verarbeitung der übrigen Bildpunkte geschieht. Es bestünde also die Möglichkeit, alle BP gleichzeitig, bzw *parallel*, zu transformieren.

2. Die sequentielle Verarbeitung:

In diesem Fall entspricht die Inputmatrix der Outputmatrix. Es werden also bei der Transformation der einzelnen BP i.allg. nicht nur die ursprünglichen Inputdaten sondern auch die Outputdaten betrachtet. Dies hat zur Folge, daß u.a. die Reihenfolge, in der die BP verarbeitet werden, für das Transformationsergebnis entscheidend ist.



1.3.3 Verschiebungsinvarianz von Operatoren

Lokale Operatoren können über dem gesamten Definitionsbereich eines Signals die gleiche Abbildungsvorschrift realisieren oder diese regional bzw. bildpunktweise variieren.

Definition (Verschiebungsinvarianz):

Seien \mathcal{O} ein Operator mit der Abbildungsvorschrift $g_m = \mathcal{O}f_m$ und \mathcal{T} ein Verschiebungsoperator mit der Abbildungsvorschrift $g_m = \mathcal{T}f_m = f_{m-\tau}$. Gilt

$$\mathcal{T}\mathcal{O}f_m = \mathcal{O}\mathcal{T}f_m$$

bzw.

$$\mathcal{T}\mathcal{O} = \mathcal{O}\mathcal{T}$$

so heißt \mathcal{O} *verschiebungsinvariant* oder *homogen*. ■

Ein verschiebungsinvarianter Operator kommutiert mit dem Verschiebungsoperator. Dies ist einfach zu zeigen:

$$\mathcal{T}\mathcal{O}f_m = \mathcal{T}g_m = g_{m-\tau}$$

$$\mathcal{O}\mathcal{T}f_m = \mathcal{O}f_{m-\tau} = g_{m-\tau}.$$

Demzufolge gilt die Abbildungsvorschrift $g_m = \mathcal{O}f_m$ für alle $m \in D$.

Die Verschiebungsinvarianz eines Operators besagt, daß jeder BP des Definitionsbereiches mit dem gleichen Operator abgebildet wird. Seine Arbeitsweise hängt nicht von den Koordinaten oder von sonstigen Eigenschaften des jeweiligen Aufpunktes ab. Eine Verschiebung des Inputs hat also eine gleiche Verschiebung des Outputs zur Folge.

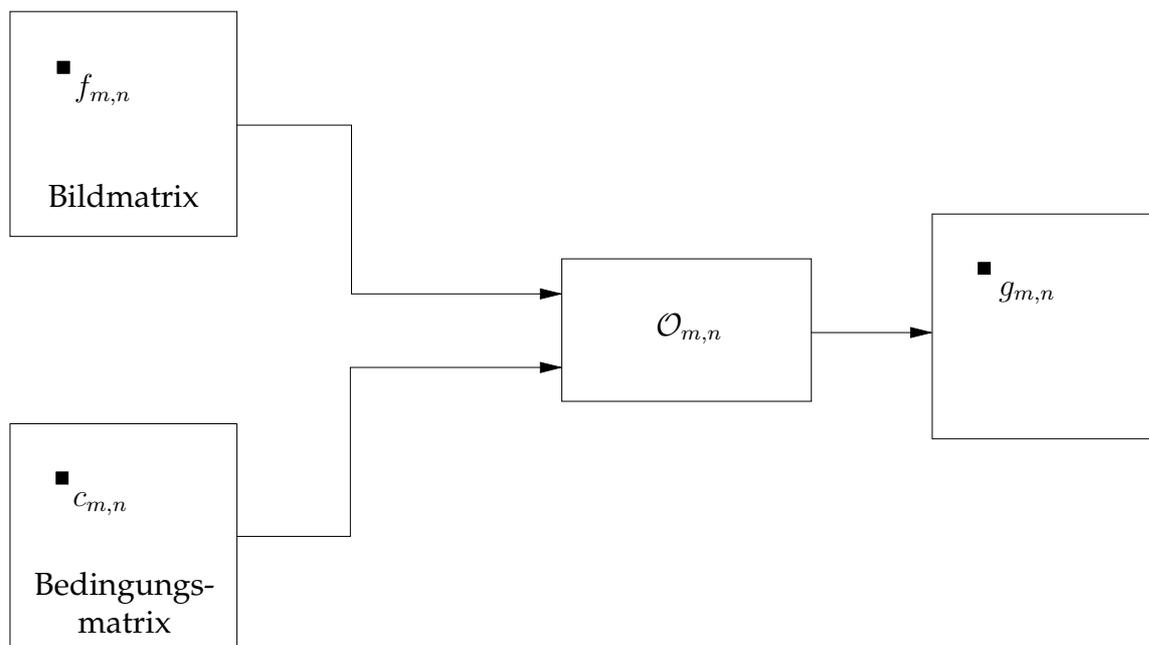
Die Verschiebungsinvarianz eines Operators ist für die Definition der Faltung und für die spektrale Charakterisierung von Operatoren von grundlegender Bedeutung (siehe auch Kapitel 2).

Definition (Verschiebungsvarianz):

Ein nicht verschiebungsinvarianter Operator heißt *verschiebungsvariant* oder *inhomogen*. ■

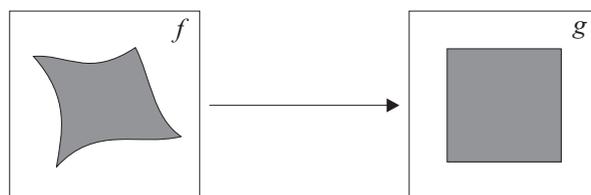
Die Verschiebungsvarianz eines Operators besagt, daß seine Wirkungsweise von den Koordinaten seines jeweiligen Aufpunktes abhängt. Das bedeutet, daß die Wirkung des Operators auf einen verschobenen Input nicht das gleiche Ergebnis liefert wie seine Anwendung auf das nicht verschobene Signal und nachträgliche Verschiebung des Outputs. Im Extremfall kann es sogar sein, daß jeder BP durch einen eigenen Operator abgebildet wird: $g_m = \mathcal{O}_m f_m$.

Für einen verschiebungsvarianten Operator \mathcal{O}_m kann man eine (u.U. virtuelle) Bedingungsma-
trix c angeben, die für jeden Bildpunkt m die Erfüllung der Bedingung c_m verwaltet.
Also gilt $\mathcal{O}_m = \mathcal{O}(f_m)$.

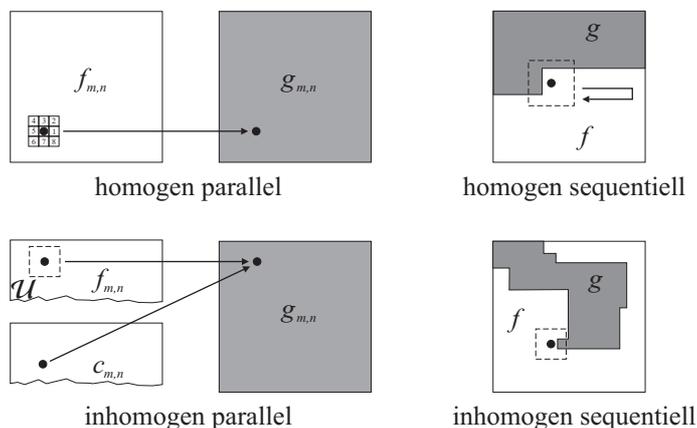


Beispiel:

1. **Filterbank-Methode:** Je nach den lokalen Bedingungen (z.B. „Glattheit“ der Signalfunktion oder Rauschen) geschieht ein Zugriff auf einen entsprechenden Operator aus einer Menge von Operatoren.
2. **Inhomogene geometrische Entzerrung (Warping):**



Zusammenfassend erhalten wir folgendes Schema der Anwendung lokaler Operatoren:



1.3.4 Algebraische Beziehungen von Operatoren

Für Operatoren gelten die folgenden drei wichtigen Eigenschaften:

1. **Operatorprodukt:** $\mathcal{O} = \mathcal{O}_2\mathcal{O}_1$ (Serienschaltung)

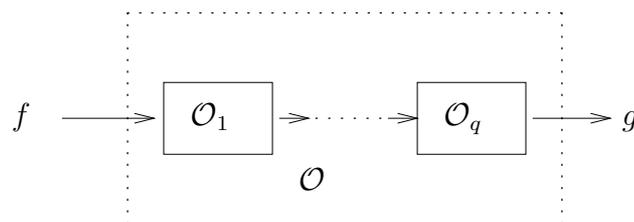
Das Operatorprodukt (Verkettung) ist i.allg. *nicht kommutativ* ! Es gilt:

$$\mathcal{O}_2\mathcal{O}_1f = \mathcal{O}_2\{\mathcal{O}_1f\} \neq \mathcal{O}_1\mathcal{O}_2f = \mathcal{O}_1\{\mathcal{O}_2f\}$$

Zum Beispiel sind die Operatoren \mathcal{T} , \mathcal{R} (Translation / Rotation) nicht kommutativ.

Die Klasse der linearen Operatoren ist kommutativ (Abelsche Operatoren). Aus der Seriendarstellung ergibt sich die *Kaskadenform* eines Operators: Die sequentielle (abhängige) Anwendung von Teiloperatoren \mathcal{O}_i , $1 \leq i \leq q$, hat die gleiche Wirkung wie die Anwendung des Operatorprodukts: Die Teiloperationen \mathcal{O}_i werden sukzessive auf die Zwischenergebnisse angewendet.

$$g = \mathcal{O}f = \mathcal{O}_q\mathcal{O}_{q-1}\cdots\mathcal{O}_2\mathcal{O}_1f$$



- Reihenfolge ist i. allg. von Bedeutung
- Teiloperatoren wirken auf Zwischenergebnisse (Originaldaten können verloren gehen)

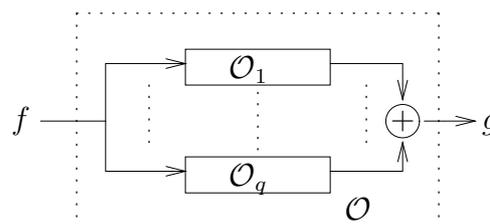
2. **Operatorsumme:** $\mathcal{O} = \mathcal{O}_1 + \mathcal{O}_2$ (Parallelverarbeitung)

Aus der Kommutativität der Addition folgt:

$$(\mathcal{O}_1 + \mathcal{O}_2)f = \mathcal{O}_1f + \mathcal{O}_2f = f(1) + f(2) = f(2) + f(1) = (\mathcal{O}_2 + \mathcal{O}_1)f$$

Die *Operatorsumme* ist kommutativ. Analog zur Kaskadenform des Operatorproduktes wird hier auch von der *Parallelfarm* gesprochen.

$$g = \mathcal{O}f = \mathcal{O}_1f + \mathcal{O}_2f + \cdots + \mathcal{O}_qf$$

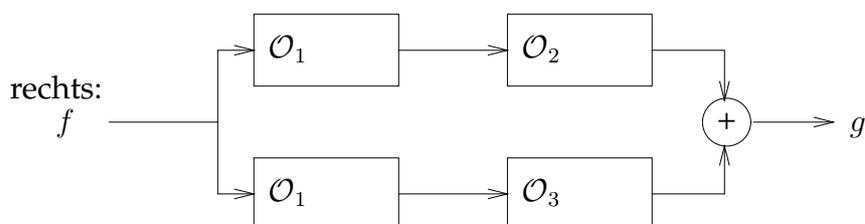
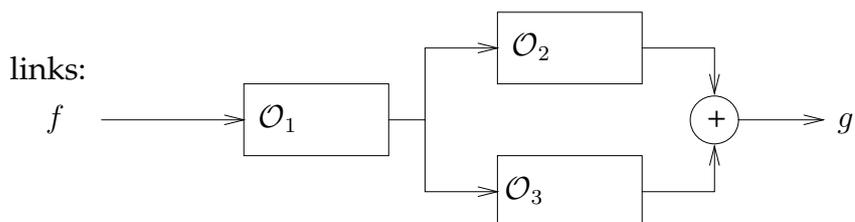


- Reihenfolge ist ohne Bedeutung
- Teiloperationen wirken auf Originaldaten (müssen erhalten bleiben)

3. Distributivität:

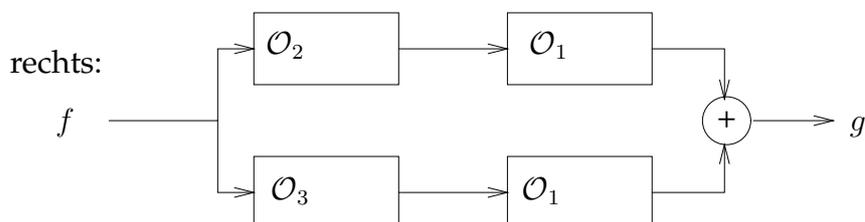
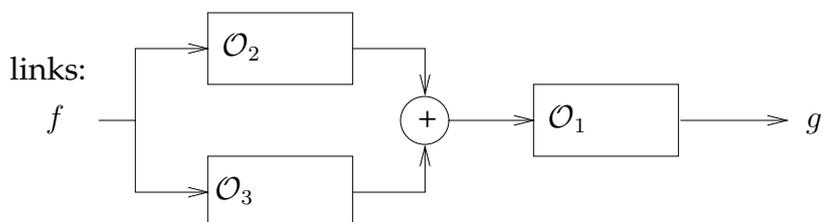
Beim Distributivgesetz ist eine wichtige Eigenart zu bemerken, die direkt aus 1. und 2. folgt: Es gilt nämlich

$$(\mathcal{O}_2 + \mathcal{O}_3)\mathcal{O}_1 = \mathcal{O}_2\mathcal{O}_1 + \mathcal{O}_3\mathcal{O}_1$$



wohingegen

$$\mathcal{O}_1(\mathcal{O}_2 + \mathcal{O}_3) \neq \mathcal{O}_1\mathcal{O}_2 + \mathcal{O}_1\mathcal{O}_3$$



zu beachten ist.

Nur wenn \mathcal{O}_1 linear ist, gilt auch das Distributivgesetz. Wenn $\mathcal{O}_1, \mathcal{O}_2$ und \mathcal{O}_3 linear sind, fallen beide Distributivgesetze zusammen.

1.4 Lineare verschiebungsinvariante Operatoren

Die linearen verschiebungsinvarianten Operatoren (LSI-Operatoren) spielen in der Bildverarbeitung eine besonders wichtige Rolle: *Faltung* und *Korrelation* lassen sich mit diesen Operatoren beschreiben. Durch Fouriertransformation kann man eine Frequenzraumdarstellung von LSI-Operatoren angeben. Außerdem bietet sich die effiziente Ausführung von Faltung und Korrelation im Frequenzraum an.

1.4.1 Lineare Operatoren und Impulsantwort

Für lineare Operatoren gilt das Überlagerungsprinzip .

Definition (lineare Operatoren):

Sei \mathcal{L} ein Operator. Falls gilt:

$$\mathcal{L}(af + bg) = a\mathcal{L}f + b\mathcal{L}g \quad a, b \in \mathbb{R}$$

so heißt \mathcal{L} linear. ■

Bemerkung:

1. Ein linearer Operator bildet Bruchteile eines Signals in gleicher Weise wie das gesamte Signal ab. Dies kann man durch die Operatorgleichung ausdrücken $\Sigma\mathcal{L} = \mathcal{L}\Sigma$, welche die Invarianz bezüglich Summenbildung darstellt!
2. Die hier interessierenden linearen Operatoren können beliebig additiv und multiplikativ vertauscht werden.

Für lineare Operatoren der Signalverarbeitung gilt die Kommutativität des Produktes linearer Operatoren:

$$\mathcal{L} = \mathcal{L}_1\mathcal{L}_2 = \mathcal{L}_2\mathcal{L}_1$$

Insbesondere gilt auch die Distributivität:

$$\mathcal{L}_1(\mathcal{L}_2 + \mathcal{L}_3) = \mathcal{L}_1\mathcal{L}_2 + \mathcal{L}_1\mathcal{L}_3$$

Betrachten wir ein unbegrenztes Signal f als gewichtete Impulsfolge $f_m = \sum_{k=-\infty}^{\infty} f_k\delta_{m-k}$ im Ortsraum. Ferner sei \mathcal{L} ein linearer Operator mit $g = \mathcal{L}f$:

$$g_m = \mathcal{L} \sum_{k=-\infty}^{\infty} f_k\delta_{m-k} = \sum_{k=-\infty}^{\infty} f_k\mathcal{L}\delta_{m-k} = \sum_{k=-\infty}^{\infty} f_k h_k^m$$

Wir zeigen die Kommutativität der linearen Operatoren der Signalverarbeitung am Beispiel von Punktoperatoren und einem Signal, das nur aus einem Impuls an der Position $k = m$ besteht:

$$g_m = \mathcal{L}f_m = \mathcal{L}f_k\delta_{m-k} = f_k\mathcal{L}\delta_{m-k}$$

Der lineare Punktoperator \mathcal{L}_1 skaliert den Einheitsimpuls in folgender Weise:

$$\mathcal{L}_1\delta_{m-k} = a\delta_{m-k} + b$$

Für einen anderen Punktoperator gelte $\mathcal{L}_2\delta_{m-k} = c\delta_{m-k} + d$. Dann folgt

$$\mathcal{L}_1\mathcal{L}_2\delta_{m-k} = a(c\delta_{m-k} + d) + b = ac\delta_{m-k} + ad + b$$

$$\mathcal{L}_2\mathcal{L}_1\delta_{m-k} = c(a\delta_{m-k} + b) + d = ca\delta_{m-k} + bc + d$$

Also gilt die Kommutativität der Produkts linearer Operatoren nur für $b = d = 0$.

Da aber $|\delta_{m-k}| = 1$ für $k = m$, gilt auch

$$\mathcal{L}_1\delta_{m-k} = a\delta_{m-k} + b = (a + b)\delta_{m-k} = \alpha\delta_{m-k}$$

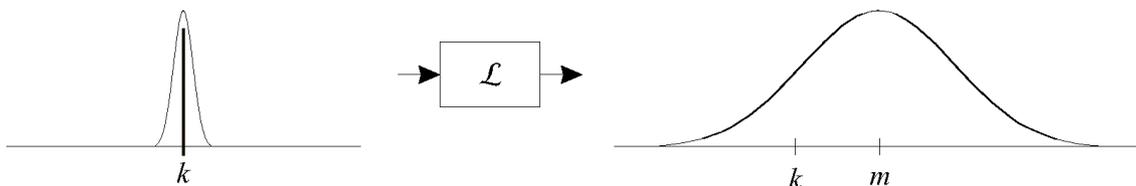
und entsprechend $\mathcal{L}_2\delta_{m-k} = \beta\delta_{m-k}$. Also gilt stets die Kommutativität $\mathcal{L}_1\mathcal{L}_2 = \mathcal{L}_2\mathcal{L}_1$.

Wir behandeln nun die Darstellung eines linearen Operators im Ortsraum.

Definition (verallgemeinerte Impulsantwort):

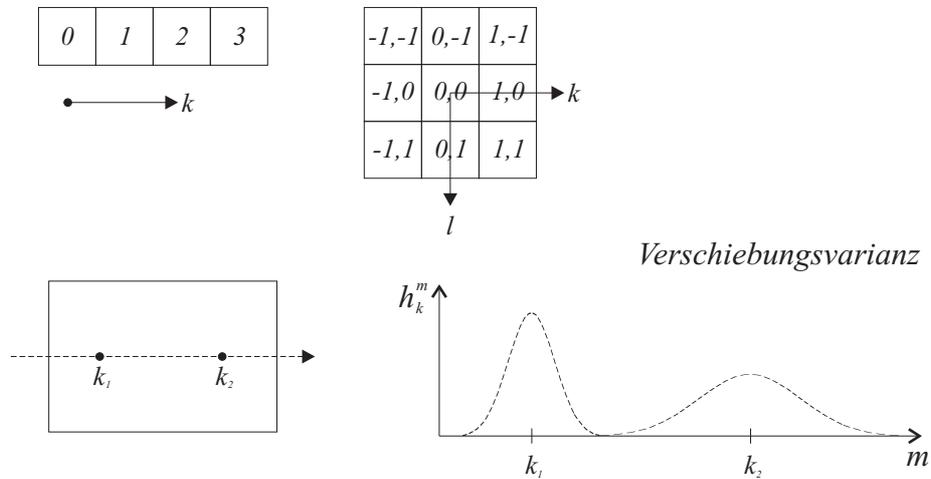
Die Beschreibungsform eines linearen Operators \mathcal{L} im Ortsraum $h_k^m = \mathcal{L}\delta_{m-k}$ heißt *verallg. Impulsantwort von \mathcal{L}* . ■

Die Impulsantwort h_k^m von \mathcal{L} ist der Output des linearen Operators \mathcal{L} am Ort m für einen Einheitsimpuls als Input am Ort k .



Ein allgemeiner linear Operator verändert die Form des Eingangssignals und kann auch eine Verschiebung des Signals bewirken. Die Reichweite der Wirkung eines Operators auf einen Einheitsimpuls hängt von seiner Ausdehnung im Ortsraum ab. Die Veränderung der Form des Signals hängt von den Koeffizienten des Operators im Einzugsbereich ab.

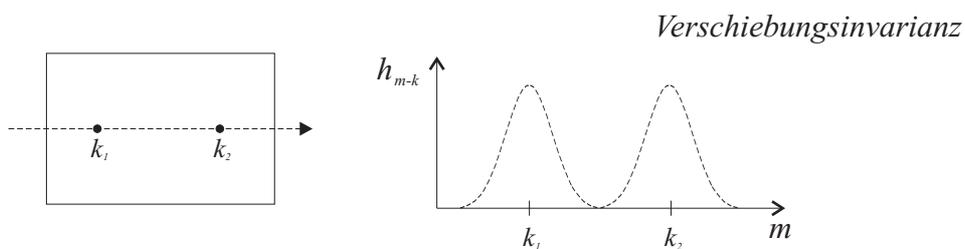
Ist der betrachtete Ortsraum eindimensional, so ist h ein Vektor, liegt 2-Dimensionalität vor, so ist h eine Matrix. Ist der Operator verschiebungsvariant, so hängt die (verallgemeinerte) Impulsantwort vom Wirkungsort des Operators ab.



Definition (Impulsantwort):

Ist ein Operator linear und verschiebungsinvariant (LSI-Operator), dann ist seine Impulsantwort eine verschobene Kopie einer räumlich invarianten Impulsantwort, d.h.

$$h_k^m = h_{m-k}$$



Die Impulsantwort hängt also nur von der **Differenz zwischen Aufpunkt (k) und Testpunkt (m)** ab. Hieraus folgt die Darstellung der **Faltungssumme**:

$$g_m = \sum_{k=-\infty}^{\infty} f_k h_{m-k}$$

oder in symbolischer Schreibweise:

$$\begin{aligned}
 g_m &= (f * h)_m \\
 g &= f * h.
 \end{aligned}$$

1.4.2 Matrixdarstellung eines linearen Operators

Wir behandeln nun die Darstellung eines linearen Operators im Signalraum.

Betrachten wir ein endliches Signal mit den Werten f_m , $m \in D_m = \{0, 1, \dots, M-1\}$, dann gilt im Signalraum:

$$f = \sum_{m=0}^{M-1} f_m e_m$$

Im Signalraum stellt sich das Signal als Vektoraddition seiner Komponenten dar. Wenden wir nun einen linearen Operator \mathcal{L} auf f an, so ergibt dies

$$g = \mathcal{L}f = \mathcal{L} \sum_{m=0}^{M-1} f_m e_m = \sum_{m=0}^{M-1} f_m \mathcal{L}e_m$$

mit

$$\mathcal{L}e_m = \sum_{k=0}^{M-1} l_{k,m} e_k \quad l_{k,m} = \langle \mathcal{L}e_m, e_k \rangle$$

Der Operator wirkt somit nur auf die Basis des Signalraumes. Man erkennt, daß unter Umständen alle BP betrachtet werden müssen, um nur einen BP zu ändern. Der Koeffizient $l_{k,m}$ beschreibt hierbei den durch den Operator bewirkten Einfluß des Eingangssignals am Ort k auf das Ausgangssignal am Ort m . Es gilt also

$$g = \sum_{k=0}^{M-1} g_k e_k$$

mit

$$g_k = \sum_{m=0}^{M-1} l_{k,m} f_m$$

bzw.

$$g = Lf,$$

wenn die Signalvektoren als Spaltenvektoren vereinbart sind.

Aus obiger Gleichung folgt die Matrixdarstellung des linearen Operators \mathcal{L} :

$$L = \begin{array}{c} \begin{array}{c} \overrightarrow{k} \\ \downarrow \\ m \end{array} \left[\begin{array}{cccc} l_{0,0} & l_{1,0} & \cdots & l_{M-1,0} \\ l_{0,1} & l_{1,1} & \cdots & l_{M-1,1} \\ \dots & \dots & \dots & \dots \\ l_{0,M-1} & l_{1,M-1} & \cdots & l_{M-1,M-1} \end{array} \right] \end{array}$$

Der 1D-Fall läßt sich auf den 2D-Fall übertragen. Für das 2D-Signal $f_{j,k}$ entsteht das transformierte Signal $g_{m,n}$, $m, j \in D_m = \{0, 1, \dots, M-1\}$, $n, k \in D_n = \{0, 1, \dots, N-1\}$

durch die Transformation

$$g_{m,n} = \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} \lambda_{j,m}^{k,n} f_{j,k}$$

Stellt man die Matrix $L = [l_{k,n}]$ durch eine Matrix aus Matrizen dar mit $l_{k,n} = \left[\lambda_{j,m}^{k,n} \right]$, so liefert dies die Matrixschreibweise:

$$g = Lf$$

Das Element $l_{k,n}$ beschreibt dann den Anteil der Wirkung der Zeile k in f auf die Zeile n in g und $\lambda_{j,m}^{k,n}$ bezeichnet den Anteil der Wirkung des Bildpunktes (j, k) in f am Bildpunkt (m, n) im Signal g .

Diese Blockmatrix-Struktur von L ist eine Folge der zeilenweisen Abtastung. Also gilt: $l_{k,n} \sim N \times N$ und $L \sim MN \times MN$.

Beispiel: $f \sim (3 \times 3)$ -Signal

	f		g
	00	10	20
	01	11	21
	02	12	22
		\xrightarrow{L}	
	00	10	20
	01	11	21
	02	12	22

Wirkung von Zeile 1 in f auf Zeile 0 in g :

$$L = \begin{matrix} \begin{matrix} \downarrow \\ n \end{matrix} & \begin{matrix} \rightarrow \\ k \end{matrix} \\ \begin{bmatrix} l_{0,0} & l_{1,0} & l_{2,0} \\ l_{0,1} & l_{1,1} & l_{2,1} \\ l_{0,2} & l_{1,2} & l_{2,2} \end{bmatrix} \end{matrix}$$

Wirkung von Bildpunkt (1,1) in f auf Bildpunkt (2,0) in g :

$$l_{1,0} = \begin{matrix} \begin{matrix} \rightarrow \\ j \end{matrix} \\ \begin{matrix} \downarrow \\ m \end{matrix} \end{matrix} \begin{bmatrix} \lambda_{00}^{10} & \lambda_{10}^{10} & \lambda_{20}^{10} \\ \lambda_{01}^{10} & \lambda_{11}^{10} & \lambda_{21}^{10} \\ \lambda_{02}^{10} & \lambda_{12}^{10} & \lambda_{22}^{10} \end{bmatrix}$$

■

1.4.3 Transformation des Basissystems des Signalraumes – Eigenwerttransformation

Die Matrixdarstellung eines linearen Operators hängt vom Basissystem des Ortsraumes ab. Wird ein Operator in einem diagonalen statt einem kartesischen Koordinatensystem angewendet, muß seine Signalraumdarstellung die Transformation des Basissystems berücksichtigen. Wie die Transformation eines Operators aussieht, wenn das

Basissystems $B = \{e_0, e_1, \dots, e_{M-1}\}$ in das Basissystem $B' = \{e'_0, e'_1, \dots, e'_{M-1}\}$ linear transformiert wird, liefert die lineare Algebra. Sei $P = [p_{km}]$ die Transformationsmatrix des linearen, invertierbaren Operators \mathcal{P} für die Transformation des Basissystems und sei \mathcal{L} ein anzuwendender linearer Operator mit der Matrixdarstellung L .

$$e'_m = \sum_{k=0}^{M-1} p_{km} e_k \quad \text{d.h.} \quad f' = P f$$

Wegen

$$g' = L' f' = L' P f$$

und

$$g' = P g = P L f$$

folgt

$$L' P = P L$$

oder

$$L' = P L P^{-1}$$

Die Operatorgleichung beschreibt die Vorgehensweise zur Anwendung von \mathcal{L} in B' :

1. Rücktransformation der Basis $\mathcal{P}^{-1} : B' \rightarrow B$
2. Anwendung des Operators $\mathcal{L} : f \rightarrow g$
3. Hintransformation der Basis $\mathcal{P} : B \rightarrow B'$

Beispiel: Eigenwerttransformation:

Frage 1: Welche Strukturen bleiben unter der Wirkung eines linearen Operators \mathcal{L} erhalten (invariant) - bis auf Skalierungsfaktor?

Antwort: Die Eigenvektoren des Operators \mathcal{L} sind gegenüber diesem invariant.

Frage 2: Welche Matrixdarstellung ist die einfachste für einen linearen Operator?

Antwort: Die Darstellung als Diagonalmatrix Λ mit den Diagonalelementen λ_{ii} , die man als Eigenwerte bezeichnet, wird durch die Eigenwerttransformation erzeugt.

■

Die Eigenvektoren x_i und die Eigenwerte λ_i , $i \in \{0, 1, \dots, M-1\}$ sind Lösungen der Eigenwertgleichung des linearen Operators \mathcal{L} :

$$\mathcal{L}x_i = \lambda_i x_i \quad i \in \{0, 1, \dots, M-1\}$$

Die Säkulargleichung

$$\det(L - \lambda I) = 0$$

stellt die Lösungsgleichung der Eigenwerttransformation dar. Aus ihr erhält man das charakteristische Polynom, dessen Nullstellen die Eigenwerte λ_i sind. Hier ist λ der Vektor der Eigenwerte (bzw. das Eigenwertspektrum) und I ist die Identitäts-Diagonalmatrix.

Die Matrixdarstellung des linearen Operators \mathcal{L} wird vereinfacht, indem die Basis in eine neue Basis, bestehend aus den Eigenvektoren des Operators, transformiert wird. Dann hat die Operatormatrix die Gestalt einer Diagonalmatrix Λ , auf deren Diagonale die Eigenwerte des Operators stehen.

Ist darüber hinaus der Operator symmetrisch, so bilden die Eigenvektoren sogar eine *Orthonormalbasis*. In diesem Fall ist die transformierte Matrix des Operators \mathcal{L} durch

$$\Lambda = C^{-1}LC$$

darstellbar. Hierbei sind L die Ursprungsmatrix von \mathcal{L} und C die (orthogonale) Matrix der *Hauptachsentransformation*.

Allgemein gilt also: Wird einem Operator \mathcal{L} bezüglich einer Basis $B = \{e_0, e_1, \dots, e_{M-1}\}$ eine Diagonalmatrix Λ zugeordnet, so sind die Basisvektoren sämtlich Eigenvektoren von \mathcal{L} , d.h. bezüglich dieser Basis ist die Wirkung des Operators \mathcal{L} auf ein Element $f \in \mathbb{F}$ invariant.

Anwendungen:

1. Komplexe harmonische Funktionen (das Basissystem der Fourier-Transformation) stellen die Eigenvektoren von LSI-Operatoren dar.
Die Ausführung einer Faltung, siehe Abschnitt 1.4.5, im Frequenzraum (nach Fourier-Transformation) erfolgt durch punktweise Multiplikation, weil die Faltungsmatrix diagonalisiert ist. Das ist die Grundlage des Faltungstheorems der Fourier-Transformation (Abschnitt 2.5.3). Die Eigenwerte der Faltungsoperation sind operatorspezifisch. Sie heißen Koeffizienten der Frequenzübertragungsfunktion des LSI-Operators (Abschnitt 3.2).
2. Die Karhunen-Loeve-Transformation (oder Hauptkomponenten-Transformation, Hotelling-Transformation) ist eine statistische Eigenwerttransformation der Kovarianzmatrix (siehe Abschnitt 3.7.1). Während die Merkmale der Kovarianzmatrix korreliert sind, sind die Eigenvektoren (Hauptkomponenten) entkorreliert.

Beispiel: Lösung der Eigenwertgleichung eines LSI-Operators

$$L = \begin{pmatrix} 2 & -5 \\ 1 & -4 \end{pmatrix}$$

1. Berechnung des Eigenwertspektrums

$$\det(L - \lambda I) = 0$$

$$\begin{vmatrix} (2 - \lambda) & -5 \\ 1 & (-4 - \lambda) \end{vmatrix} = \lambda^2 + 2\lambda - 8 + 5 = 0$$

$$\lambda_{1/2} = -1 \pm \sqrt{1 + 3} = -1 \pm 2$$

$$\lambda = [1, -3]^T$$

2. Berechnung der Eigenvektoren

$$(L - \lambda_1 I)u = \begin{pmatrix} 1 & -5 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{aligned} u_1 - 5u_2 = 0 &\longrightarrow u_1 = 5u_2 \\ u_1 - 5u_2 = 0 &\longrightarrow u_1 = 5u_2 \end{aligned}$$

Folglich ist ein möglicher Eigenraum $L_1 = \{\mu [5, 1], \mu \in \mathbb{R}\}$.

$$(L - \lambda_2 I)u = u = \begin{pmatrix} 5 & -5 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{aligned} 5u_1 - 5u_2 = 0 &\longrightarrow u_1 = u_2 \\ u_1 - u_2 = 0 &\longrightarrow u_1 = u_2 \end{aligned}$$

Folglich ist ein anderer möglicher Eigenraum $L_2 = \{\mu [1, 1], \mu \in \mathbb{R}\}$. ■

1.4.4 Primitive Operatoren

In diesem Kapitel werden einfachste Operatoren betrachtet, die als Basis aller komplexeren Operatoren dienen. Angenommen sei eine periodische Signalfunktion \tilde{f} .

Der Translationsoperator \mathcal{T}

Der *Translationsoperator* \mathcal{T} verschiebt den Input zyklisch um eine Position nach rechts:

$$\tilde{g}_m = \mathcal{T}\tilde{f}_m = \tilde{f}_{m-1}$$

Entsprechend verschiebt der Operator \mathcal{T}^k den Input um k Positionen nach rechts. Der Translationsoperator \mathcal{T}^{-1} verschiebt den Input zyklisch um eine Position nach links:

$$\tilde{g}_m = \mathcal{T}^{-1}\tilde{f}_m = \tilde{f}_{m+1}$$

Beispiel: \tilde{f} sei eindimensional, $D_m = \{0, 1, 2\}$. Mit \mathcal{T} gilt dann

$$\tilde{f} : \begin{array}{|c|c|c|} \hline \tilde{f}_0 & \tilde{f}_1 & \tilde{f}_2 \\ \hline \end{array} \xrightarrow{\mathcal{T}} \tilde{g} : \begin{array}{|c|c|c|} \hline \tilde{f}_2 & \tilde{f}_0 & \tilde{f}_1 \\ \hline \end{array}$$

bzw mit \mathcal{T}^{-1} :

$$\tilde{f} : \begin{array}{|c|c|c|} \hline \tilde{f}_0 & \tilde{f}_1 & \tilde{f}_2 \\ \hline \end{array} \xrightarrow{\mathcal{T}^{-1}} \tilde{g} : \begin{array}{|c|c|c|} \hline \tilde{f}_1 & \tilde{f}_2 & \tilde{f}_0 \\ \hline \end{array}$$

Bezeichnungen:

- \mathcal{T} heißt der Operator der Rechtsverschiebung.
- \mathcal{T}^{-1} heißt der Operator der Linksverschiebung.
- \mathcal{I} heißt der Identitätsoperator.

Bemerkung: Der Identitätsoperator ist das neutrale Element, das am Signal nichts ändert, also $g_m = \mathcal{I}f_m = f_m$. Er wird durch die Einheits-Diagonalmatrix repräsentiert:

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

Offensichtlich ist \mathcal{T} zu \mathcal{T}^{-1} invers (und umgekehrt):

$$\mathcal{T}\mathcal{T}^{-1} = \mathcal{T}^{-1}\mathcal{T} = \mathcal{I}$$

bzw.

$$\mathcal{T}^M = \mathcal{T}^{-M} = \mathcal{I}$$

Der Operator \mathcal{T} ist nicht involutorisch, da $\mathcal{T}^2 \neq \mathcal{I}$, und nicht idempotent, da $\mathcal{T}^2 \neq \mathcal{T}$. Ferner ist \mathcal{T} linear und verschiebungsinvariant, woraus die Existenz der Matrizen und der Impulsantworten folgt.

$$T = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \quad T^{-1} = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Bemerkung: Wegen der Linearität von \mathcal{T} gilt für den skalierten Operator $a\mathcal{T}$

$$g_m = a\mathcal{T}f_m = \mathcal{T}af_m.$$

Der Skalierungsfaktor wirkt also nur auf die Signalkoeffizienten.

Für die Impulsantwort gilt:

$$\begin{aligned} h_T &= \mathcal{T}\delta_0 = \mathcal{T}[1, 0, \cdots, 0]^T = [0, 1, 0, \cdots, 0]^T \\ h_{T^{-1}} &= \mathcal{T}^{-1}\delta_0 = \mathcal{T}^{-1}[1, 0, \cdots, 0]^T = [0, \cdots, 0, 1]^T \end{aligned}$$

Wir betrachten nun ein 2D-Signal f . Unter der Annahme der Separabilität des Verschiebungsoperators gilt (wegen der Linearität gilt auch die Kommutativität der Operatoren)

$$\mathcal{T} = \mathcal{T}_h \mathcal{T}_v = \mathcal{T}_v \mathcal{T}_h$$

wobei

$$\begin{aligned} \tilde{g}_{m,n} &= \mathcal{T}_h \tilde{f}_{m,n} = \tilde{f}_{m-1,n} \\ \tilde{g}_{m,n} &= \mathcal{T}_v \tilde{f}_{m,n} = \tilde{f}_{m,n-1} \end{aligned}$$

Für die Matrix T gilt $T = [t_{kn}]$ mit $t_{kn} = [\lambda_{jm}^{kn}]$.

Beispiel: Gegeben sei ein periodisches 2D-Signal \tilde{f} :

$$\tilde{f} : \begin{array}{|c|c|} \hline \tilde{f}_{00} & \tilde{f}_{10} \\ \hline \tilde{f}_{01} & \tilde{f}_{11} \\ \hline \end{array} \quad \begin{array}{l} \mathcal{T} = \mathcal{T}_h \mathcal{T}_v \\ \rightarrow \\ \mathcal{T} = \mathcal{T}_v \mathcal{T}_h \end{array} \quad \tilde{g} : \begin{array}{|c|c|} \hline \tilde{g}_{00} & \tilde{g}_{10} \\ \hline \tilde{g}_{01} & \tilde{g}_{11} \\ \hline \end{array} = \begin{array}{|c|c|} \hline \tilde{f}_{11} & \tilde{f}_{01} \\ \hline \tilde{f}_{10} & \tilde{f}_{00} \\ \hline \end{array}$$

Es soll folgende Transformation mit Hilfe von verketteten horizontalen und vertikalen Einzeltransformationen durchgeführt werden: Aus \tilde{f} soll \tilde{g} entstehen.

Aus \tilde{f} soll $\tilde{g} : \begin{array}{|c|c|} \hline \tilde{f}_{11} & \tilde{f}_{01} \\ \hline \tilde{f}_{10} & \tilde{f}_{00} \\ \hline \end{array}$ entstehen.

$$T_h : \left. \begin{array}{l} (0,0) \rightarrow (1,0) : \lambda_{01}^{00} = 1 \\ (1,0) \rightarrow (0,0) : \lambda_{10}^{00} = 1 \\ (0,1) \rightarrow (1,1) : \lambda_{01}^{11} = 1 \\ (1,1) \rightarrow (0,1) : \lambda_{10}^{11} = 1 \end{array} \right\} \begin{array}{l} t_{00} \\ \\ t_{11} \end{array} \quad T_v : \left. \begin{array}{l} (0,0) \rightarrow (0,1) : \lambda_{00}^{01} = 1 \\ (1,0) \rightarrow (1,1) : \lambda_{11}^{01} = 1 \\ (0,1) \rightarrow (0,0) : \lambda_{00}^{10} = 1 \\ (1,1) \rightarrow (1,0) : \lambda_{11}^{10} = 1 \end{array} \right\} \begin{array}{l} t_{01} \\ \\ t_{10} \end{array}$$

Daraus folgt:

$$T_h = \left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array} \right] \quad T_v = \left[\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array} \right]$$

So gilt für das Matrizenprodukt:

$$T = T_h T_v = T_v T_h = \left[\begin{array}{cccc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array} \right] = \left[\begin{array}{cccc} 0 & 0 & 0 & \lambda_{10}^{10} \\ 0 & 0 & \lambda_{01}^{10} & 0 \\ 0 & \lambda_{10}^{01} & 0 & 0 \\ \lambda_{01}^{01} & 0 & 0 & 0 \end{array} \right]$$

■

In diesem Beispiel werden Zeilen- und Spaltindizes vertauscht. Dies entspricht einer Punktspiegelung um das Bildzentrum (aber nur in diesem Spezialfall einer 2×2 -Matrix!).

Bemerkungen:

1. Ein Operator \mathcal{O} ist verschiebungsinvariant, falls er mit dem Translationsoperator \mathcal{T} kommutiert, d.h.: $\mathcal{O} \mathcal{T} = \mathcal{T} \mathcal{O}$.
2. Ist ein linearer Operator \mathcal{L} verschiebungsinvariant, so kann er als Linearkombination von Translationsoperatoren dargestellt werden.

3. Der Verschiebungsoperator kann als spezieller Summenoperator interpretiert werden. Er ist also sowohl als Punktoperator als auch als lokaler Operator interpretierbar:

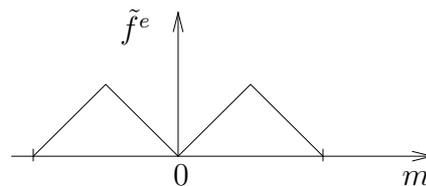
$$\begin{aligned}\mathcal{T} : \tilde{g}_m &= \mathcal{T}\tilde{f}_m = \tilde{f}_{m-1} \\ &= 0 \cdot \tilde{f}_m + 1 \cdot \tilde{f}_{m-1}\end{aligned}$$

Der Spiegeloperator \mathcal{M}

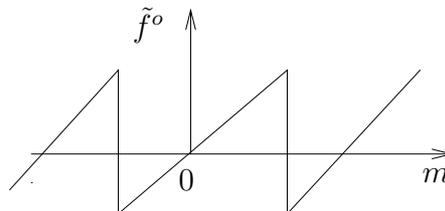
Sei \tilde{f} wieder ein 1D-Signal. Der Spiegeloperator \mathcal{M} spiegelt den Input an einem bez. \mathcal{M} festgelegten Symmetriezentrum im Definitionsbereich der Signalfunktion (z.B. Koordinatenursprung). Da also der Spiegeloperator an ein Symmetriezentrum gebunden ist, ist er *nicht* verschiebungsinvariant. I.a. gilt für \mathcal{M} :

$$\tilde{g}_m = \mathcal{M}\tilde{f}_m = \tilde{f}_{-m}$$

Ist \tilde{f}^e eine gerade/symmetrische Funktion, so gilt offensichtlich $\mathcal{M}\tilde{f}_m^e = \tilde{f}_{-m}^e = \tilde{f}_m^e$.



Für eine ungerade/antisymmetrische Funktion \tilde{f}^o gilt $\mathcal{M}\tilde{f}_m^o = \tilde{f}_{-m}^o = -\tilde{f}_m^o$.



Offensichtlich wirkt \mathcal{M} wie \mathcal{T}^{2m} im Falle \tilde{f}^e und wie $-\mathcal{T}^{2m}$ im Falle \tilde{f}^o . Aus der Koordinatenabhängigkeit des Verschiebungsoperators erkennt man die Verschiebungsvarianz des Spiegeloperators.

Jede beliebige Funktion \tilde{f} lässt sich bezüglich eines gewählten Symmetriezentrums lokal zerlegen in einen geraden Teil \tilde{f}_m^e und einen ungeraden Teil \tilde{f}_m^o , $\tilde{f}_m = \tilde{f}_m^e + \tilde{f}_m^o$, wobei gilt:

$$\begin{aligned}\tilde{f}_m^e &= \frac{1}{2} (\mathcal{I} + \mathcal{M}) \tilde{f}_m \\ \tilde{f}_m^o &= \frac{1}{2} (\mathcal{I} - \mathcal{M}) \tilde{f}_m\end{aligned}$$

Für ein Signal der Länge $|D| = M$ gibt es aber M derartige Zerlegungen. Diese integrale Zerlegung über alle M Spiegelungen wird im Frequenzraum sichtbar. Sie wird

durch die Fouriertransformation realisiert.

Ferner ist eine gespiegelte Funktion \tilde{f} gleich der Differenz von \tilde{f}^e und \tilde{f}^o :

$$\mathcal{M}\tilde{f}_m = \mathcal{M}(\tilde{f}_m^e + \tilde{f}_m^o) = \tilde{f}_m^e - \tilde{f}_m^o$$

Zu bemerken bleibt noch, daß \mathcal{M} involutorisch, d.h. $\mathcal{M}^2 = \mathcal{I}$ (wegen $\mathcal{M}^{-1} = \mathcal{M}$), aber nicht idempotent ist (da $\mathcal{M}^2 \neq \mathcal{M}$) und linear, jedoch nicht verschiebungsinvariant ist. Der Spiegeloperator kommutiert nicht mit dem Translationsoperator:

$$\mathcal{T}\mathcal{M} = \mathcal{M}\mathcal{T}^{-1}.$$

Beweis:

$$\mathcal{T}\tilde{g}_m = \mathcal{T}\mathcal{M}\tilde{f}_m = \tilde{f}_{-m-1} = \mathcal{M}\tilde{f}_{m+1} = \mathcal{M}\mathcal{T}^{-1}\tilde{f}_m$$

Der Skalierungsoperator \mathcal{S}

Der Skalierungsoperator \mathcal{S} bewirkt eine Skalierung des Definitionsbereichs einer Fkt. \tilde{f} :

$$\tilde{g}_m = \mathcal{S}\tilde{f}_m = \tilde{f}_{sm},$$

$s \in \mathbb{Z}$. Die Skalierung ist nicht verschiebungsinvariant, da \mathcal{S} bezüglich des Koordinatensprungs skaliert. \mathcal{S} ist jedoch linear, besitzt aber i. allg. kein Inverses.

Beispiel: Sei $\tilde{f} : \boxed{a \mid b \mid c \mid d}$ ein periodisches Signal. Es soll mit $s = 2$ skaliert werden, also $\tilde{g}_m = \tilde{f}_{2m}$ mit $m \in D_m = \{0, 1, 2, 3\}$

$$\tilde{g}_m = \mathcal{S}\tilde{f}_m : \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} a \\ c \\ a \\ c \end{pmatrix}$$

Beim periodischen Signal hängt die Struktur der Matrix des Skalierungsoperators von $|D_m|$ ab. ■

Beispiel: $\tilde{f} : \boxed{a \mid b \mid c}$ $s = 2$

$$\tilde{g}_m = \mathcal{S}\tilde{f}_m : \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ c \\ b \end{pmatrix}$$

Achtung: periodisches Signalmodell (keine Änderung von D_m) ■

Summen- und Differenzoperatoren Σ, \mathcal{D}

Sei \tilde{f} ein 1D-Signal. Dann ist der Differenzoperator \mathcal{D} definiert durch:

$$\begin{aligned}\tilde{g}_m^{r,d} &= \mathcal{D}_r \tilde{f}_m = \tilde{f}_m - \tilde{f}_{m-1} & \mathcal{D}_r &= \mathcal{I} - \mathcal{T} \\ \tilde{g}_m^{l,d} &= \mathcal{D}_l \tilde{f}_m = \tilde{f}_m - \tilde{f}_{m+1} & \mathcal{D}_l &= \mathcal{I} - \mathcal{T}^{-1}\end{aligned}$$

Analog ist der Summenoperator Σ f.a. $m \in D_m$ definiert durch:

$$\begin{aligned}\tilde{g}_m^{r,s} &= \Sigma_r \tilde{f}_m = \tilde{f}_m + \tilde{f}_{m-1} & \Sigma_r &= \mathcal{I} + \mathcal{T} \\ \tilde{g}_m^{l,s} &= \Sigma_l \tilde{f}_m = \tilde{f}_m + \tilde{f}_{m+1} & \Sigma_l &= \mathcal{I} + \mathcal{T}^{-1}\end{aligned}$$

Beide Operatoren sind linear und verschiebungsinvariant.

Beispiel: $\mathcal{D}_r = \mathcal{I} - \mathcal{T}$ für $M = 3$

$$\begin{aligned}g_0 &= f_0 - f_2 \\ g_1 &= f_1 - f_0 \\ g_2 &= f_2 - f_1\end{aligned}$$

Hieraus folgt

$$D_r = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

und für $h_{D_r} = \mathcal{D}_r \delta_0$

$$h_{D_r} = [1, -1, 0]^T$$

■

Der Differenzoperator ist ungerade symmetrisch, denn es gilt $\mathcal{D}_r = -\mathcal{T}\mathcal{D}_l$ und $\mathcal{D}_l = -\mathcal{T}^{-1}\mathcal{D}_r$. Ferner besitzen \mathcal{D} und Σ keine Inversen.

Beispiel:

$$\begin{aligned}\Sigma_l \mathcal{D}_l &= (\mathcal{I} + \mathcal{T}^{-1})(\mathcal{I} - \mathcal{T}^{-1}) \\ &= \mathcal{I}\mathcal{I} - \mathcal{I}\mathcal{T}^{-1} + \mathcal{I}\mathcal{T}^{-1} - \mathcal{T}^{-1}\mathcal{T}^{-1} \\ &= \mathcal{I} - \mathcal{T}^{-2} \stackrel{!}{=} \mathcal{D}_{l_2} \\ \Sigma_l \mathcal{D}_r &= \mathcal{T}^{-1} - \mathcal{T} \neq \mathcal{I}\end{aligned}$$

■

Die Differenz und die Summe im Diskreten entsprechen der Differentiation und Integration im Stetigen. Die zweifache Differenz ist dann also ein Analogon zur zweiten

Ableitung:

$$\begin{aligned}
 \mathcal{D}^2 &= \mathcal{D}_r \mathcal{D}_l \\
 &= (\mathcal{I} - \mathcal{T}) (\mathcal{I} - \mathcal{T}^{-1}) \\
 &= \mathcal{I}^2 - \mathcal{I}\mathcal{T} - \mathcal{I}\mathcal{T}^{-1} + \mathcal{T}\mathcal{T}^{-1} \\
 &= 2\mathcal{I} - \mathcal{T} - \mathcal{T}^{-1}
 \end{aligned}$$

Es gilt auch

$$\mathcal{D}^2 = \mathcal{D}_r \mathcal{D}_l = \mathcal{D}_l \mathcal{D}_r = \mathcal{D}_r + \mathcal{D}_l = \mathcal{D}_l + \mathcal{D}_r.$$

Definition (Laplaceoperator):

$\mathcal{D}^2 = 2\mathcal{I} - \mathcal{T} - \mathcal{T}^{-1}$ heißt *Laplace-Operator*, in der Folge auch mit \mathcal{L} bezeichnet.

$$\mathcal{D}^2 = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 & -1 \\ -1 & 2 & -1 & & & 0 \\ 0 & -1 & 2 & & & \vdots \\ \vdots & 0 & -1 & & & \vdots \\ 0 & \vdots & \vdots & & & -1 \\ -1 & 0 & \vdots & & -1 & 2 \end{bmatrix}$$

■

Dieser Operator erbt alle Eigenschaften des Differenzoperators, ist jedoch symmetrisch. Seine Impulsantwort ist $h_{\mathcal{D}^2} = [2, -1, 0, \dots, 0, -1]^T$. Sie läßt sich aus der Matrix von \mathcal{D}^2 direkt ablesen, jedoch auch aus der Impulsantwort von \mathcal{D}_l und \mathcal{D}_r herleiten.

Für ein 2D-Signal ist der Laplace-Operator definiert durch

$$\begin{aligned}
 \mathcal{D}^2 &= \mathcal{D}_v^2 \mathcal{D}_h^2 \\
 &= 4\mathcal{I} - \mathcal{T}_h - \mathcal{T}_h^{-1} - \mathcal{T}_v - \mathcal{T}_v^{-1}
 \end{aligned}$$

Das 2D-Faltungsintegral lautet:

$$(\tilde{f} * \tilde{h})(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{f}(\xi, \eta) \tilde{h}(x - \xi, y - \eta) d\xi d\eta$$

Die 2D-Faltungssumme hat folgende Gestalt für eine periodische und diskrete Funktion \tilde{f} :

$$(\tilde{f} \circledast \tilde{h})_{m,n} := \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} \tilde{f}_{k,l} \tilde{h}_{m-k,n-l}$$

Bemerkung:

Die Funktionen \tilde{f} und \tilde{g} , sowie die Impulsantwort \tilde{h} besitzen den gleichen Definitionsbereich. Wir werden in der Folge nur die Bezeichnungen f, g, h verwenden, wenn aus dem Kontext (z.B. der Benennung der Art der Faltung) klar ist, wie der Definitionsbereich zu verstehen ist.

Algebraische Eigenschaften der Faltung

Diese Eigenschaften gelten auch für die nicht-zyklische Faltung.

1. **Kommutativität:** Es gilt

$$\begin{aligned} g &= f \circledast h & &= h \circledast f \\ g_m &= \sum_{k=0}^{M-1} f_k h_{m-k} & &= \sum_{k=0}^{M-1} h_k f_{m-k} \end{aligned}$$

2. **Assoziativität:** Es gilt

$$g = f \circledast (h_{(1)} \circledast h_{(2)}) = (f \circledast h_{(1)}) \circledast h_{(2)}$$

3. **Distributivität:** Es gilt

$$g = f \circledast (h_{(1)} + h_{(2)}) = (f \circledast h_{(1)}) + (f \circledast h_{(2)})$$

4. **Neutrales Element:**

$$g = f \circledast \delta = f$$

Beweis von 1: Indexsubstitution $m - k = l$ bzw. $k = m - l$

Beispiel für 2: $(\Sigma \mathcal{L})f = \Sigma(\mathcal{L}f) = \mathcal{L}(\Sigma f)$

In diesem Beispiel ist \mathcal{L} ein Laplaceoperator (berechnet ein Krümmungsmaß) und Σ ein Operator, der auf irgend eine Weise einen lokalen Mittelwert berechnet. Es ist gleichwertig, ob ein gemitteltes Krümmungssignal oder die Krümmung eines gemittelten Signals berechnet wird. ■

Beispiel für 3: $\mathcal{L} \approx \Sigma_1 - \Sigma_2$

Wir werden später zeigen, daß sich der Laplaceoperator durch die Differenz zweier Operatoren mit schwacher (Σ_1) und stärkerer (Σ_2) Mittelung approximieren läßt. ■

Geometrische Deutung der Faltungssumme

Versucht man, die Faltung geometrisch zu deuten, so ergeben sich drei elementare Schritte:

1. Spiegelung der Impulsantwort am Ursprung des lokalen Koordinatensystems $h_k \rightsquigarrow h_{-k}$.
2. Verschieben der gespiegelten Impulsantwort unter Beachtung der Periodizität um m Stellen im globalen Koordinatensystem nach rechts: $h_{-k} \rightsquigarrow h_{m-k}$.
3. Bilden der Summe der Produktfolge.

$$\begin{aligned}
 g_m &= (h \circledast f)_m \\
 &= \sum_{k=0}^{M-1} f_k \mathcal{T}^m \mathcal{M} h_k \\
 &= \sum_{k=0}^{M-1} f_k \mathcal{T}^m h_{-k} \\
 &= \sum_{k=0}^{M-1} f_k h_{m-k} \\
 &\stackrel{\mathcal{T}^m \mathcal{M} = \mathcal{M} \mathcal{T}^{-m}}{=} \sum_{k=0}^{M-1} f_k \mathcal{M} h_{k-m} \\
 &= \sum_{k=0}^{M-1} f_k \mathcal{M} \mathcal{T}^{-m} h_k
 \end{aligned}$$

Beispiel: 1D-Fall $M = |D_m| = 3$, $f = [3, 4, 5]^T$

Operator $\Sigma_l = \mathcal{I} + \mathcal{T}^{-1}$ für $M = 3$ gilt auch $\Sigma_{r_2} = \mathcal{I} + \mathcal{T}^2$

$$\Sigma_l : g_m = f_m + f_{m+1} \quad \Sigma_{r_2} : g_m = f_m + f_{m-2}$$

f	3	4	5
m	0	1	2

h_{Σ_l}	1	0	1
k	0	1	2

1. Spiegelung:

Bemerkung: Wegen der Periodizität M gilt $h_{-k} = h_{0-k} = h_{M-k}$.

$\mathcal{M}h_{\Sigma_l}$	1	1	0	1	1	0
k	-3	-2	-1	0	1	2

2.,3. Verschieben und Bilden der Produkte:

1. $m = 0$:

$\mathcal{M}h$	1	1	0
f	3	4	5
k	0	1	2

$$g_0 = 3 + 4 + 0 = 7$$

2. $m = 1$:

$\mathcal{M}h$	0	1	1
f	3	4	5
k	0	1	2

$$g_1 = 0 + 4 + 5 = 9$$

3. $m = 2$:

$\mathcal{M}h$	1	0	1
f	3	4	5
k	0	1	2

$$g_2 = 3 + 0 + 5 = 8$$

Ergebnis: $g = [7, 9, 8]^T$

■

Beachte die Äquivalenz der Ausführung eines Operators
im Ortsraum: Faltung mit der Impulsantwort
im Signalraum: Multiplikation mit der Operatormatrix

Beispiel: gewichtete Summe $\sum_r^w = a\mathcal{I} + b\mathcal{T}$
periodisches Signal f , $M = 3$

Aus der Operatorgleichung folgt unmittelbar

$$\begin{aligned} g_0 &= af_0 + bf_{-1} && = af_0 + bf_2 \\ g_1 &= af_1 + bf_0 && = bf_0 + af_1 \\ g_2 &= af_2 + bf_1 = bf_1 + af_2 \end{aligned}$$

Hieraus ergibt sich

$$S_\Sigma = \begin{bmatrix} a & 0 & b \\ b & a & 0 \\ 0 & b & a \end{bmatrix} \quad h_\Sigma = [a, b, 0]^T, \quad \mathcal{M}h_\Sigma = [a, 0, b]^T$$

Faltung $g_m = \sum_{k=0}^2 f_k h_{m-k}$ liefert

$$\begin{aligned} g_0 &= af_0 + 0f_1 + bf_2 \\ g_1 &= bf_0 + af_1 + 0f_2 \\ g_2 &= 0f_0 + bf_1 + af_2 \end{aligned}$$

Dies entspricht auch

$$\begin{pmatrix} g_0 \\ g_1 \\ g_2 \end{pmatrix} = \begin{bmatrix} a & 0 & b \\ b & a & 0 \\ 0 & b & a \end{bmatrix} \begin{pmatrix} f_0 \\ f_1 \\ f_2 \end{pmatrix}$$

■

Beispiel: 2D-Faltung

$$\begin{array}{ccc}
 \begin{array}{|c|c|c|} \hline \overset{m}{\rightarrow} & & \\ \hline \underset{n}{\downarrow} & 3 & 5 & 3 \\ \hline & 3 & 5 & 3 \\ \hline & 3 & 5 & 3 \\ \hline \end{array} & \otimes & \begin{array}{|c|c|c|} \hline \overset{k}{\rightarrow} & & \\ \hline \underset{l}{\downarrow} & -1 & 2 & -1 \\ \hline & -1 & 2 & -1 \\ \hline & -1 & 2 & -1 \\ \hline \end{array} & = & \begin{array}{|c|c|c|} \hline \overset{m}{\rightarrow} & & \\ \hline \underset{n}{\downarrow} & -6 & -6 & 12 \\ \hline & -6 & -6 & 12 \\ \hline & -6 & -6 & 12 \\ \hline \end{array} \\
 f & & h & & g
 \end{array}$$

Berechnung der linken oberen Ecke:

$$\begin{array}{cccccc}
 3 & 5 & 3 & 3 & 5 & 3 \\
 3 & \begin{array}{|c|c|c|} \hline -1 & 2 & -1 \\ \hline -1 & 2 & -1 \\ \hline -1 & 2 & -1 \\ \hline \end{array} & 5 & 3 & \\
 3 & \begin{array}{|c|c|c|} \hline -1 & 2 & -1 \\ \hline -1 & 2 & -1 \\ \hline -1 & 2 & -1 \\ \hline \end{array} & 5 & 3 & \\
 3 & \begin{array}{|c|c|c|} \hline -1 & 2 & -1 \\ \hline -1 & 2 & -1 \\ \hline -1 & 2 & -1 \\ \hline \end{array} & -1 & 5 & 2 \\
 3 & 5 & 3 & -1 & -1 & 2 \\
 3 & 5 & 3 & -1 & -1 & 2
 \end{array}$$



Operatordarstellung der Faltung

Die Faltung einer Funktion f mit der Impulsantwort eines LSI-Operators kann als Anwendung des komplexeren Faltungsoperators C_h , der von der Art des LSI-Operators abhängt, aufgefaßt werden

$$g = h \otimes f = C_h f,$$

denn die Summendarstellung der Faltung ist nichts anderes als die Multiplikation eines Vektors f mit einer Matrix, deren Spaltenvektoren durch h bestimmt werden:

$$C_h = \begin{bmatrix} h_0 & h_{M-1} & \cdots & h_2 & h_1 \\ h_1 & h_0 & \cdots & h_3 & h_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ h_{M-1} & h_{M-2} & \cdots & h_1 & h_0 \end{bmatrix}$$

Da die Matrizen aller LSI-Operatoren die gleiche Gestalt besitzen, lassen sich diese, wie bereits früher erwähnt, als Faltung ausführen. Auch das Verschieben kann z.B. als Faltung mit der Impulsantwort des Verschiebungsoperators aufgefaßt werden:

Beispiel: $M = 4$

$$g_m = \mathcal{T}^2 f_m = f_{m-2}$$

$$f : \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix} \quad h_{\mathcal{T}^2} : \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}$$

$$g_m = \sum_{k=0}^3 f_k h_{m-k}$$

Achtung: Hier gilt $h_{\mathcal{T}^2} = \mathcal{M}h_{\mathcal{T}^2}$.

f	1	2	3	4	g
$m = 0$	0	0	1	0	3
$m = 1$	0	0	0	1	4
$m = 2$	1	0	0	0	1
$m = 3$	0	1	0	0	2

also $g : \begin{bmatrix} 3 & 4 & 1 & 2 \end{bmatrix}$

■

Eigenschaften des Faltungsoperators

1. Die Matrixdarstellung jedes LSI-Operators hat die Struktur eines Faltungsoperators.
2. Eine quadratische Matrix mit gleichen Diagonalelementen (auch der Nebendiagonalen) heißt *Töplitzmatrix* und spielt in der Signaltheorie eine besondere Rolle.

Beispiel: Die Matrix

$$M = \begin{bmatrix} 1 & 6 & 7 & 8 & 9 \\ 2 & 1 & 6 & 7 & 8 \\ 3 & 2 & 1 & 6 & 7 \\ 4 & 3 & 2 & 1 & 6 \\ 5 & 4 & 3 & 2 & 1 \end{bmatrix}$$

ist eine Töplitzmatrix, gebildet nach der Regel

$$m_{i,j} = \begin{cases} c_{i-j+1} & \text{für } i-j \geq 0 \\ r_{j-i+1} & \text{für } i-j < 0 \end{cases}$$

aus dem Spaltenvektor $c = [1, 2, 3, 4, 5]^T$ und dem Zeilenvektor $r = [6, 7, 8, 9]$.

■

3. Ist jede Zeile einer Töplitzmatrix die rechtsverschobene Variante der vorhergehenden Zeile, bezeichnet man die Matrix als *Zirkularmatrix*. Zirkularmatrizen lassen sich durch Fouriertransformation diagonalisieren.
Offensichtlich ist die Matrix M des obigen Beispiels keine Zirkularmatrix.
4. Ist die Determinante der Zirkularmatrix ungleich Null, so ex. der zum Faltungsoperator inverse *Entfaltungsoperator* C_h^{-1} .

$$g = h \otimes f \sim g = C_h f \rightsquigarrow f = C_h^{-1} g = C_h^{-1} C_h f = \mathcal{I} f$$

5. Es gilt:

$$\begin{aligned} \mathcal{T} C_h &= C_h \mathcal{T} \\ \mathcal{M} C_h &= C_h^T \mathcal{M} \end{aligned}$$

6. Der Faltungsoperator läßt sich als Linearkombination von Verschiebungsoperatoren darstellen:

$$C_h = \sum_{m=0}^{M-1} h_m \mathcal{T}^m$$

Beispiel: D^2 -Operator: $D^2 = 2\mathcal{I} - \mathcal{T} - \mathcal{T}^{-1}$, $M = 3$

$$h_{D^2} : \begin{bmatrix} 2 & -1 & -1 \end{bmatrix}$$

$$D^2 = 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - 1 \cdot \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} - 1 \cdot \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

$$D^2 = 2 \cdot \mathcal{T}^0 - 1 \cdot \mathcal{T}^1 - 1 \cdot \mathcal{T}^2$$



Struktur des 2D-Faltungsoperators:

Der 2D-Faltungsoperator hat eine Repräsentation als Blockzirkularmatrix der Dimension $MN \times MN$. Sie besteht aus N^2 Blöcken, die jeweils eine Zirkularmatrix der Dimension $M \times M$ darstellen.

$$C_h = \begin{bmatrix} C_{h_0} & C_{h_{N-1}} & \cdots & C_{h_1} \\ C_{h_1} & C_{h_0} & & C_{h_2} \\ \vdots & & & \vdots \\ C_{h_{N-1}} & C_{h_{N-2}} & & C_{h_0} \end{bmatrix}$$

$$C_{h_l} = \begin{bmatrix} h_{0,l} & h_{M-1,l} & \cdots & h_{1,l} \\ h_{1,l} & h_{0,l} & & h_{2,l} \\ \vdots & & & \vdots \\ h_{M-1,l} & h_{M-2,l} & & h_{0,l} \end{bmatrix}$$

Der Korrelationsoperator

Der *zyklische Korrelationsoperator* ist definiert durch eine Faltung mit einer gespiegelten Impulsantwort:

$$g_m = \sum_{k=0}^{M-1} f_k h_{m+k}$$

bzw:

$$g = h \circ f = \mathcal{M}h \circledast f$$

d.h. die Wichtungsfaktoren der verschobenen Impulsantwort werden direkt mit den darunter liegenden Funktionswerten multipliziert.

Ist h von gerader Symmetrie, d.h. $\mathcal{M}h = h$, dann ist die Faltung mit der Korrelation identisch. Dies ist ein besonders wichtiger Fall. Filter mit derartiger Symmetrie heißen *Nullphasenfilter* (Abschnitt 3.2.3).

Operatordarstellung der Korrelation:

$$g = \mathcal{K}_h f$$

Ortsraum : $g = h \circ f$

Signalraum : $g = \mathcal{K}_h f$

Die Matrix des Korrelationsoperators hat die Gestalt

$$K_h = \begin{bmatrix} h_0 & h_1 & \cdots & h_{M-1} \\ h_1 & h_2 & \cdots & h_0 \\ \dots & \dots & \dots & \dots \\ h_{M-1} & h_0 & \cdots & h_{M-2} \end{bmatrix}$$

Es gelten u.a. folgende wichtige Eigenschaften:

1. Quadratische Matrizen mit gleichen Antidiagonalelementen heißen *Hankelmatrizen*. Der Spiegeloperator besitzt die Darstellung einer Hankelmatrix.
2. Jede Zeile ist gegenüber der vorherigen linksverschoben.
3. Der Korrelationsoperator ist gegenüber Translation und Spiegelung invariant.

$$\begin{aligned} \mathcal{T}\mathcal{K}_h &= \mathcal{K}_h\mathcal{T} \\ \mathcal{M}\mathcal{K}_h &\stackrel{K_h=K_h^T}{=} \mathcal{K}_h\mathcal{M} \end{aligned}$$

Beispiel: Spiegelung um Mittelpunkt des Signals ($M = 3$)

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} c \\ b \\ a \end{pmatrix}$$

Achtung: Spiegeloperatoren sind nicht verschiebungsinvariant. Spiegelungen sind bedeutsame Operationen zur Darstellung von Geometrie. Z.B. jede Rotation läßt sich durch Hinterein-

anderausführung von zwei Spiegelungen erzeugen. ■

1.4.6 Lineare Faltung

Bisher wurde die Berechnung der Faltung für folgende Fälle behandelt:

- a) unendlich ausgedehntes Signal \tilde{f}
Über die Ausdehnung der Impulsantwort werden keine Aussagen gemacht. Aber offensichtlich ist sie maximal (unendlich).
- b) periodisches Signal \tilde{f}
Die Ausdehnung der Impulsantwort ist ebenfalls maximal (eine Periode) und selbst periodisch.

Hier wird der praktisch wichtige Fall der linearen Faltung eingeführt:

- c) endlich ausgedehntes Signal f
Die Impulsantwort hat minimale Ausdehnung.

Vorteil: Reduktion der Komplexität der Faltung um die Anteile der zyklischen Faltung, die mit Koeffizienten der Impulsantwort \tilde{h} vom Wert Null berechnet werden.

Nachteile:

1. Es muß der Rand des Signals beachtet werden.
2. Die Wahl des Ursprungs des lokalen Koordinatensystems der Impulsantwort muß beachtet werden.
3. Einige schöne Eigenschaften gehen verloren, die für die zyklische Faltung gelten, wie
 - Wenn $i + j = M$, $M = |D_m|$, dann $\mathcal{T}^i = \mathcal{T}^{-j}$.

Unter der Annahme, daß ein endliches Signal f über den Rand des Definitionsbereiches mit Nullen fortgesetzt wird, lautet die lineare Faltungssumme:

$$g_m = \sum_{k=-\infty}^{\infty} f_k h_{m-k} \quad m \in \{\dots, 0, 1, \dots, M-1, \dots\}$$

Die tatsächliche endliche Begrenzung der Faltungssumme hängt vom Definitionsbereich des Signals und der Impulsantwort, sowie von der Lage des lokalen Koordinatensystems in der Impulsantwort ab.

Eigenschaften der linearen Faltung:

1. Je nach Wahl des lokalen Koordinatensystems und nach Ausdehnung der Impulsantwort des Operators verbreitert sich die Ergebnisfunktion links/rechts gegenüber der Eingabefunktion.

$$M_g = M_f + M_h - 1$$

2. Ohne Erweiterung des Signals f auf M_g Stützstellen (z.B. durch Auffüllen mit Nullen), bleibt ein unbearbeiteter Rand.

Beispiel: h_l : rechts 1 BP, h_r : links 1 BP

3. Das Ergebnis der zyklischen Faltung entspricht dem der linearen Faltung eines endlichen Signals, wenn Input \tilde{f} und Impulsantwort \tilde{h} eine Länge M_g haben. Es ist also erforderlich, die zyklische Faltung auf einem erweiterten Definitionsbereich zu berechnen. Der sonst entstehende Fehler ist ein Alias-Fehler. Er heißt "Wraparound".

Beispiel: $f : [1, 1, 1]^T$

a) nicht erweitert zyklisch: $h : [1, 1, 1]^T$

$$M_f = M_h = M_g = 3$$

$$\begin{array}{r} \tilde{f} : \underline{1 \quad 1 \quad 1} \\ \tilde{h} : \mathbf{1 \quad 1 \quad 1} \\ \quad \quad \quad 1 \quad \mathbf{1} \quad 1 \\ \quad \quad \quad \underline{1 \quad 1 \quad 1} \\ \tilde{g} : 3 \quad 3 \quad 3 \end{array}$$

b) erweitert zyklisch: $h : [1, 1, 1, 0, 0]^T$

$$M_f = M_h = M_g = 5$$

$$\begin{array}{r} \tilde{f} : \underline{1 \quad 1 \quad 1 \quad 0 \quad 0} \\ \tilde{h} : \mathbf{1 \quad 0 \quad 0 \quad 1 \quad 1} \\ \quad \quad \quad 1 \quad \mathbf{1} \quad 0 \quad 0 \quad 1 \\ \quad \quad \quad 1 \quad 1 \quad \mathbf{1} \quad 0 \quad 0 \\ \quad \quad \quad 0 \quad 1 \quad 1 \quad \mathbf{1} \quad 0 \\ \quad \quad \quad \underline{0 \quad 0 \quad 1 \quad 1 \quad 1} \\ \tilde{g} : 1 \quad 2 \quad 3 \quad 2 \quad 1 \end{array}$$

c) linear: $h : [1, \mathbf{1}, 1]^T$

$$\begin{array}{r}
 f: \quad 0 \quad 0 \quad 1 \quad 1 \quad 1 \quad 0 \quad 0 \\
 \hline
 h: \quad 1 \quad \mathbf{1} \quad 1 \\
 \quad \quad \quad 1 \quad \mathbf{1} \quad 1 \\
 \quad \quad \quad \quad 1 \quad \mathbf{1} \quad 1 \\
 \quad \quad \quad \quad \quad 1 \quad \mathbf{1} \quad 1 \\
 \quad \quad \quad \quad \quad \quad 1 \quad \mathbf{1} \quad 1 \\
 \hline
 g: \quad 0 \quad 1 \quad 2 \quad 3 \quad 2 \quad 1 \quad 0
 \end{array}$$

■

Beispiel:

$$\tilde{f} = \begin{bmatrix} 3 & 5 & 3 \\ 3 & 5 & 3 \\ 3 & 5 & 3 \end{bmatrix} \quad \tilde{h} = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}$$

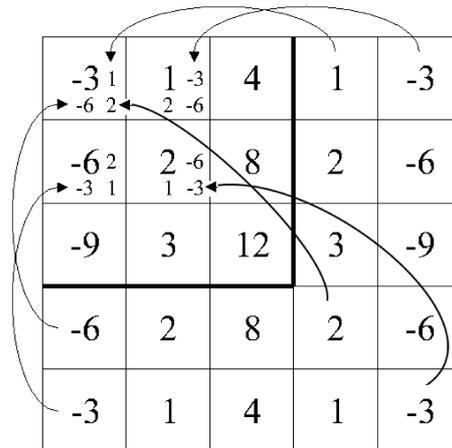
a) nicht erweitert zyklisch: Koordinatenursprung linke obere Ecke

$$\tilde{g} = \begin{bmatrix} -6 & -6 & 12 \\ -6 & -6 & 12 \\ -6 & -6 & 12 \end{bmatrix}$$

b) erweitert zyklisch (gleich linear):

$$g = \begin{bmatrix} -3 & 1 & 4 & 1 & -3 \\ -6 & 2 & 8 & 2 & -6 \\ -9 & 3 & 12 & 3 & -9 \\ -6 & 2 & 8 & 2 & -6 \\ -3 & 1 & 4 & 1 & -3 \end{bmatrix}$$

Beitrag des Wraparound:



■

1.4.7 Implementierung der linearen Faltung

Die Komplexität pro Bildpunkt der Faltung mit einem Operator der Ausdehnung $K \times L$ ist $O(K \cdot L)$ in der Anzahl der Multiplikationen und Additionen.

Da auch auf modernen Computern die Faltung noch immer einen Kostenfaktor darstellt, ist das Ziel, die Anzahl der Operatoren zu reduzieren.

Möglichkeiten:

1. Beachtung der Symmetrie der Impulsantwort
 - Reihenform (Kaskade)
 - Parallelforn
 - gerade/ungerade Symmetrie
4. Rekursive Implementierung
5. Realisierung im Frequenzraum (Faltungstheorem)
6. Tschebyscheff-Transformation der Impulsantwort
7. Zahlentheoretische/Polynom-Transformation der Impulsantwort

Beispiel: $K = 13, K \times K$

	A-N	E-N	E-S	K-B	K-E	T-E	AC-CRT	N-SN
M	169	28	14	6	4	7	13.65	3.95
A	168	168	24	24	16	57	94.15	37.13
Restr.	ohne	oktale Symm.	separ./ okt. S.	Gauß	Gauß	oktale Symm.	$N \times N$ $N=72$	$N \times N$ $N=72$

A-N *unsymmetrisch, normale Faltungssumme*

E-N *oktale Symmetrie, nicht separabel*

E-S *oktale Symmetrie, separabel*

K-B *Kaskadierung von 3×3 - Binomialfiltern*

K-E *Kaskadierung von 3×3 - Binomialfiltern und Expansion*

T-E *Tschebyscheff-Transformation (oktale Symmetrie)*

AC-CRT *Transformation nach chinesischem Restklassentheorem
nach Aggarwal und Cooley*

N-SN *Polynom-Transformation nach Nussbaumer*



Symmetrie der Impulsantwort

Normalerweise besitzen Impulsantworten in der Bildverarbeitung ein hohes Maß an Symmetrie. Diese kann genutzt werden, die Faltungsgleichung zu modifizieren.

Folgende Symmetrieklassen sollen unterschieden werden (keine genormten Klassenbezeichnungen).

A: keine Symmetrie
ohne praktische Relevanz

B: Halbebenen - Symmetrie
Beschreibung durch folgende Zwangsgleichungen:

$$h_{k,l} = h_{-k,l}$$

oder

$$h_{k,l} = h_{k,-l}$$

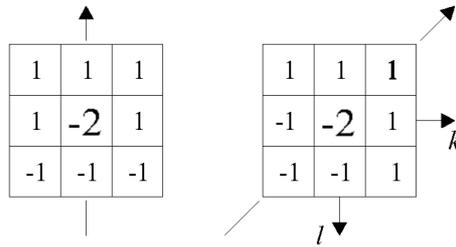
oder

$$h_{k,l} = h_{l,k}$$

oder

$$h_{k,l} = h_{-l,-k}$$

Beispiel: *Richtungsgradient (modulo 8)*

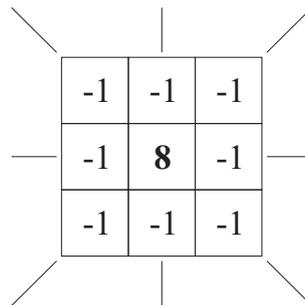


■

C: Punkt-Symmetrie

$$h_{k,l} = h_{-k,-l}$$

Beispiel: *modifizierter Laplace-Operator*



■

Bemerkung: Punkt-Symmetrie entspricht Isotropie. Auf dem Gitter fällt Punkt-Symmetrie mit gewissen Linien-Symmetrien zusammen (Isotropie kann nur approximiert werden).

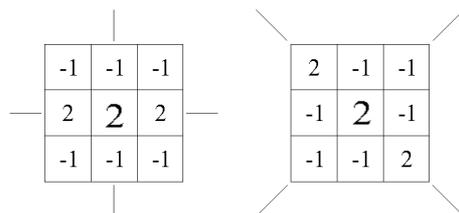
Filter mit Punkt-Symmetrie sind sogenannte Nullphasenfilter, siehe Abschnitt 3.2.3. Nullphasenfilter stellen sicher, daß der Output gegenüber dem Input nicht verschoben wird.

Bei Punkt-Symmetrie sind Faltung und Korrelation identisch.

D: Quadranten-Symmetrie

$$h_{k,l} = h_{|k|,|l|}$$

Beispiel: *Linien-Detektoren (modulo 4)*



■

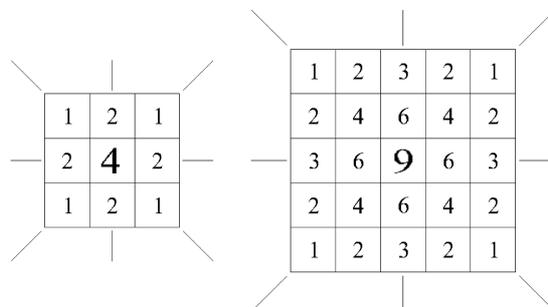
Bemerkung: Richtung versus Orientierung

Einer Richtung kann ein Vektor zugeordnet werden. Eine Orientierung ist ein Merkmal eines eindimensionalen Unterraumes, der in einem Raum höherer Dimension eingebettet ist. Dieses Merkmal ist ein Winkel bzw. Vektor von Winkeln, der gegenüber Rotation um 180° invariant ist. Linien besitzen eine Orientierung, da sie einen 1D-Unterraum bilden.

E: Oktal-Symmetrie

$$h_{k,l} = h_{|l|,|k|}$$

Beispiel: *Binomialfilter niedriger Ordnung (für höhere Ordnung approximieren sie Punktsymmetrie)*



Binomialfilter (Abschnitt 3.4) approximieren gaußförmige Wichtungen für lokale Glättung.

■

Implementierung einer oktalsymmetrischen Faltung:

Die Impulsantwort kann für den Fall eines 5×5 -Operators aus vier Komponenten synthetisiert werden. Hieraus ergibt sich folgende Zerlegung des Outputs:

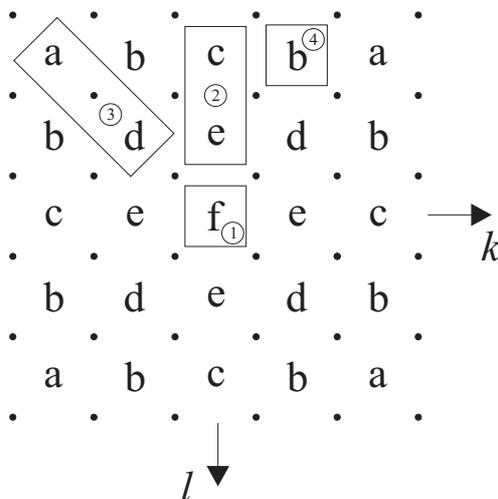
$$g_{m,n} = g_{m,n}^1 + g_{m,n}^2 + g_{m,n}^3 + g_{m,n}^4$$

$$g_{m,n}^1 = f_{m,n} h_{0,0}$$

$$g_{m,n}^2 = \sum_{l=1}^{K_z} (f_{m,n+l} + f_{m,n,-l} + f_{m+l,n} + f_{m-l,n})h_{0,l}$$

$$g_{m,n}^3 = \sum_{l=1}^{K_z} (f_{m+l,n+l} + f_{m-l,n+l} + f_{m+l,n-l} + f_{m-l,n-l})h_{l,l}$$

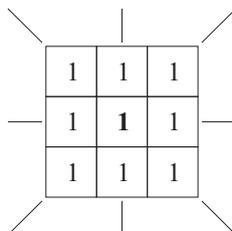
$$g_{m,n}^4 = \sum_{l=2}^{K_z} \sum_{k=1}^{l-1} (f_{m+k,n+l} + f_{m-k,n-l} + f_{m+l,n+k} + f_{m-l,n-k} + f_{m-k,n+l} + f_{m+k,n-l} + f_{m-l,n+k} + f_{m+l,n-k})h_{k,l}$$



F: Identität

$$h_{k,l} = 1$$

Beispiel: ungewichtetes Glättungsfilter



Komplexität der Faltung (asymptotische Zeitkomplexität): $K \times K$, $K = 2K_z + 1$

	M	A
A	K^2	$K^2 - 1$
B	$K(K - K_z)$	$K^2 - 1$
C	$K + 2K_z^2$	$K^2 - 1$
D	$K + K_z^2$	$K^2 - 1$
E	$K + \frac{1}{2}K_z(K_z - 1)$	$K^2 - 1$
F	1	4

Die Berücksichtigung der Symmetrie linearisiert in gewissem Maße die Anzahl der Multiplikationen, ändert aber nichts an der Anzahl der Additionen.

Für Symmetrieklasse F wurde die konstante Komplexität des Boxfilter-Algorithmus angegeben.

Separabilität der Impulsantwort

Separabilität führt zu linearer asymptotischer Zeitkomplexität $O(K)$ der Faltung mit einer 2D-Impulsantwort der Dimension $K \times K$. Es gilt für die Anzahl der arithmetischen Operationen pro BP:

$$M = 2K + 1$$

$$A = 2(K - 1)$$

Unter Separabilität der Impulsantwort ist eine Symmetrieeigenschaft zu verstehen, die zu einer Synthese einer 2D-Impulsantwort $h = [h_{kl}]$ aus dem äußeren (Tensor-) Produkt zweier senkrecht zueinander orientierter 1D-Impulsantworten $h^v = [h_l^v]$ und $h^h = [h_k^h]^T$ führt.

$$h = h^v \otimes h^h$$

mit

$$h = \begin{bmatrix} h_0^v h_0^h & \cdots & h_0^v h_{M-1}^h \\ \vdots & & \\ h_{M-1}^v h_0^h & \cdots & h_{M-1}^v h_{M-1}^h \end{bmatrix}$$

Tatsächlich wird aber der Zusammenhang

$$h = h^v * h^h$$

für die Implementierung von h genutzt. Um zu zeigen, daß

$$h^v \otimes h^h = h^v * h^h,$$

müssen h^v und h^h in Matrixform gefaltet werden.

Beispiel: Faltung zweier senkrecht orientierter 1D-Impulsantworten

0	0	1	0	0
0	0	4	0	0
0	0	6	0	0
0	0	4	0	0
0	0	1	0	0

*

0	0	0	0	0
0	0	0	0	0
1	4	6	4	1
0	0	0	0	0
0	0	0	0	0

=

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	1	4	6	4	1	0	0	0
0	0	4	16	24	16	4	0	0	0
0	0	6	24	36	24	6	0	0	0
0	0	4	16	24	16	4	0	0	0
0	0	1	4	6	4	1	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0



Implementierung separabler Faltung:

1. nichtrekursives 2-Paß-Verfahren

$$g = h^v * (h^h * f) = h^v * (h^v * f)$$

$$g_{m,n} = \sum_{k=-K_z}^{K_z} h_k^h \sum_{l=-K_z}^{K_z} h_l^v f_{m-k,n-l}$$

2. rekursives 1-Paß-Verfahren Mit

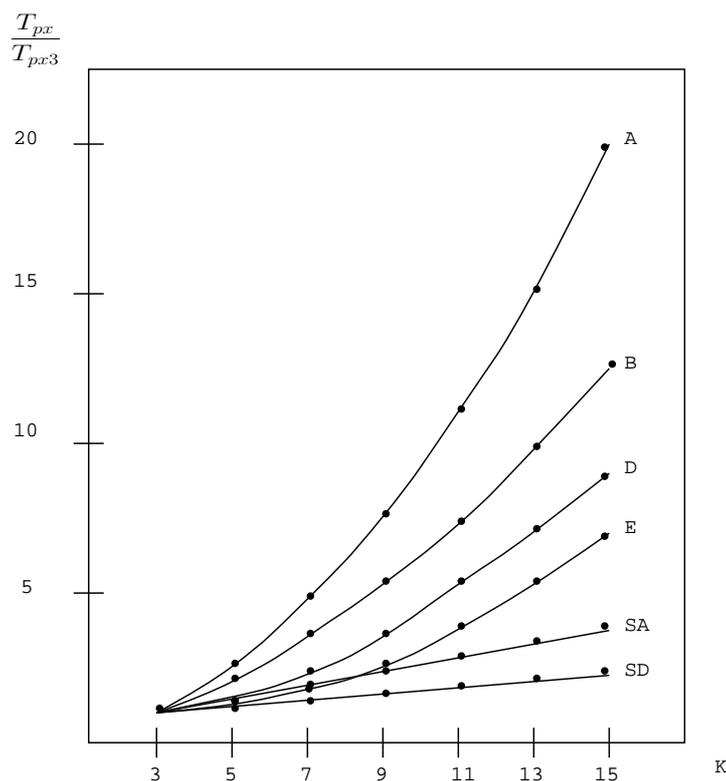
$$g_{m-k,n}^1 = \sum_{l=-K_z}^{K_z} h_l^v f_{m-k,n-l}$$

folgt

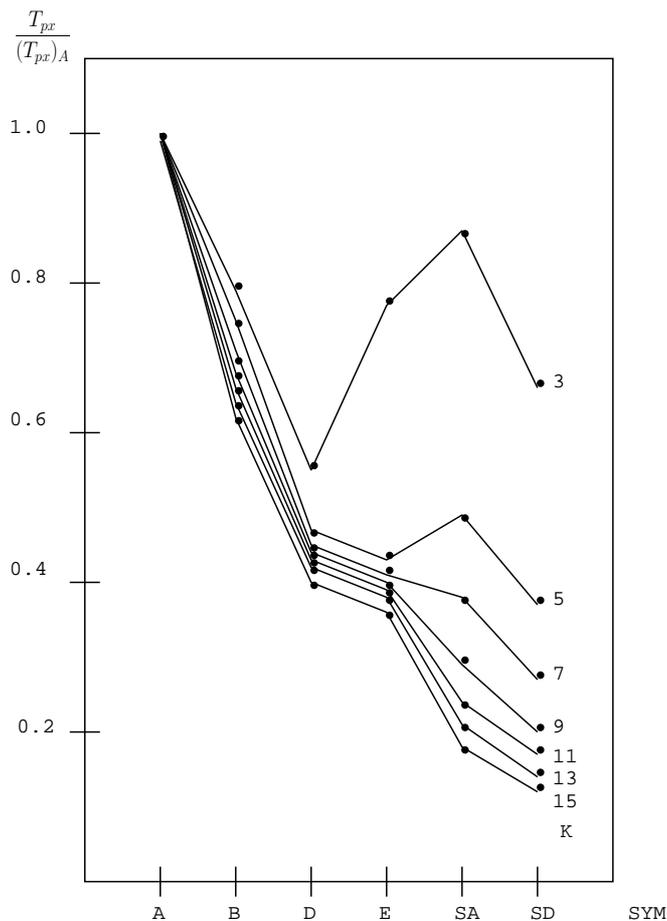
$$g_{m,n} = \sum_{k=-(K_z-1)}^{K_z} h_k^h g_{m-1-(k-1),n}^1 + h_{-K_z}^h g_{m+K_z,n}^1$$

Vergleich der Komplexität:

In der folgenden Abbildung werden gemessene Ausführungszeiten der Faltung für die Symmetrieklassen A,B,D,E, SA und SD verglichen. Die Breite K des zweidimensionalen quadratischen Operatorfensters variiert zwischen $K = 3$ und $K = 15$. Aufgetragen ist die auf die Ausführungszeit für $K = 3$ normierte Zeit T_{px} pro Bildpunkt. Der Nutzen der Beachtung der Symmetrie ist klar erkennbar.



In der folgenden Abbildung ist der gleiche Sachverhalt in Abhängigkeit von der Symmetrie aufgetragen. Die Abszisse ist also nicht metrisch korrekt! Dargestellt ist die auf die Ausführungszeit bei Nichtbeachtung der Symmetrie normierte Zeit T_{px} pro Bildpunkt. Es ist erkennbar, daß bei kleinen Operatorfenstern die Ausnutzung der Symmetrie (auch Separabilität) zu einem Mehraufwand führt, der den gewünschten Effekt reduziert.



Box-Filter-Algorithmus:

McDonnell (1980) ; Burr (1984)

Der Box-Filter-Algorithmus ist von Interesse, wenn alle Koeffizienten der Impulsantwort von gleichem Wert sind ($h_{k,l} = const$). Der Algorithmus nutzt Separabilität und Identität und muß als Spezialfall der rekursiven 1-Paß-Implementierung separabler Filter angesehen werden.

Unabhängig von der Ausdehnung K der Impulsantwort hat der Algorithmus die konstante Komplexität $M = 1$ und $A = 4$. Der Algorithmus benötigt einen Puffer der Länge M für ein $M \times M$ -Bild als Histogramm-Puffer (s.u.).

0. Initialisierung

$$g_{m,K_z}^1 = \sum_{l=-K_z}^{K_z} f_{m,K_z-l}$$

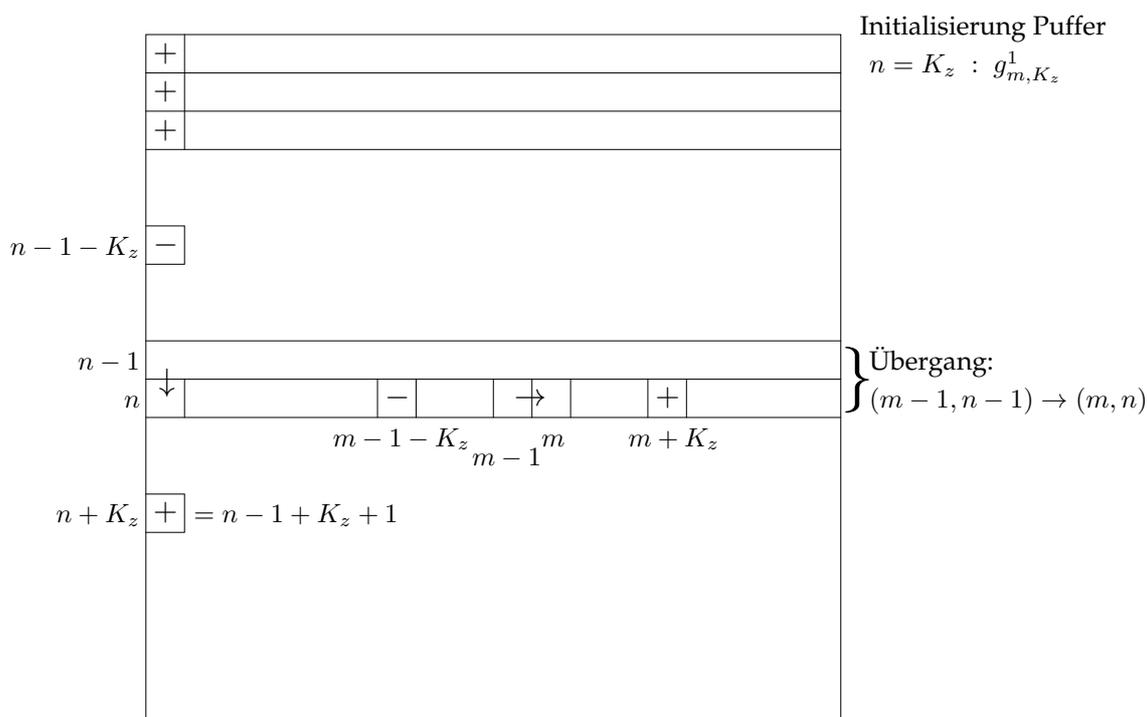
1. Zeilen-Übergang

$$g_{m-1,n}^1 = g_{m-1,n-1}^1 + f_{m-1,n-1+K_z+1} - f_{m-1,n-1-K_z}$$

2. Spalten-Übergang

$$g_{m,n} = \frac{1}{K^2} (g_{m-1,n}^1 + g_{m+K_z,n}^1 - g_{m-1-K_z,n}^1)$$

Pro Übergang $(m - 1, n - 1) \rightarrow (m, n)$ sind 4 Additionen und eine Multiplikation erforderlich.



Zerlegung der Impulsantwort:

Wie bei separabler Implementierung besteht das Ziel, eine große Impulsantwort durch die Verschachtelung kleiner Impulsantworten zu realisieren. Verschachtelung bedeutet Reihenschaltung oder Parallelschaltung.

A) Reihenschaltung

Die Kaskadierung der Faltung nutzt deren Assoziativität.

$$g = h * f = (h_{(2)} * h_{(1)}) * f = h_{(2)} * g_{(1)}$$

Ist $K \times L$ die Ausdehnung von h und gilt $K^i \times L^i$ für die $h_{(i)}$, so folgt

$$K = \sum_{i=1}^n K^i - (n - 1) \text{ und } L = \sum_{i=1}^n L^i - (n - 1)$$

für die Realisierung

$$h = h_{(1)} * h_{(2)} * \dots * h_{(n)}.$$

Die Impulsantworten $h_{(i)}$ werden Wurzeln von h genannt. Die Wurzeln können verschieden oder gleich sein.

Da $\sum_{i=1}^n K^i = K + (n-1)$, erhöht die Kaskadierung von Wurzelfiltern $h_{(i)}$ den Aufwand der Implementierung einer 1D-Faltung im Verhältnis zur direkten Implementierung von h .

Im 2D-Fall hingegen führt die Kaskadierung zur Reduzierung der Komplexität wegen

$$\sum_{i=1}^n K^i L^i < KL \text{ bzw. } \sum_{i=1}^n (K^i)^2 < K^2.$$

Beispiel: Einheitliche quadratische Wurzel $K^i = \kappa$ (z.B. $\kappa = 3$)

$$h = (h_{(i)})^{*n} \text{ mit } K = n(\kappa - 1) + 1, \quad n = \frac{K - 1}{\kappa - 1}$$

$\kappa = 3$	$K = 5$	$K = 7$	$K = 9$
$n = 2$	$M = 18$	–	–
$n = 3$	–	$M = 27$	–
$n = 4$	–	–	$M = 36$
direkt	$M = 25$	$M = 49$	$M = 81$

■

Beispiel: Sei Π_κ ein $\kappa \times \kappa$ -Boxfilter. Im Moment lassen wir den speziellen Box-Filter-Algorithmus außer Betracht. Dann erzeugt die Faltung zweier Boxfilter der Ausdehnung 2×2

$$\Lambda_3 = \Pi_2 * \Pi_2$$

einen Dreiecksfilter

$$\Lambda_3 = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Allgemein gilt für $\kappa \times \kappa$ -Boxfilter

$$\Lambda_{2\kappa-1} = \Pi_\kappa^{*2}$$

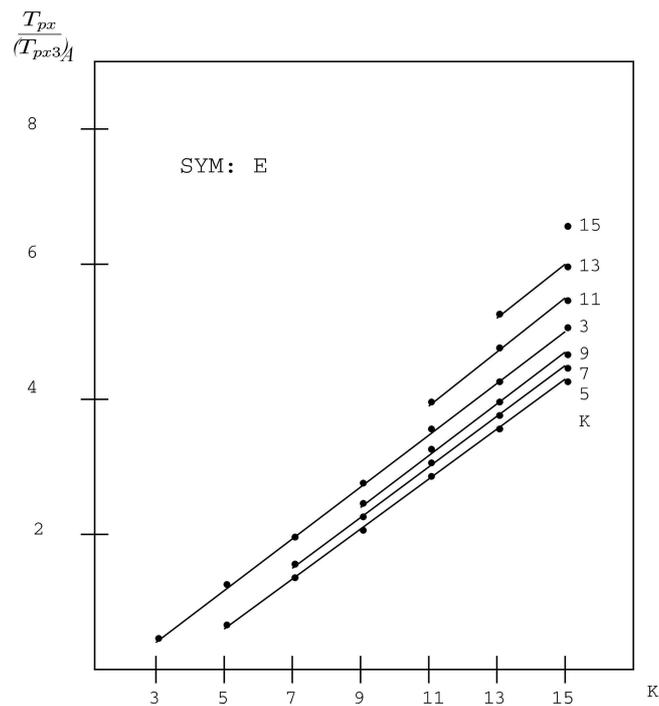
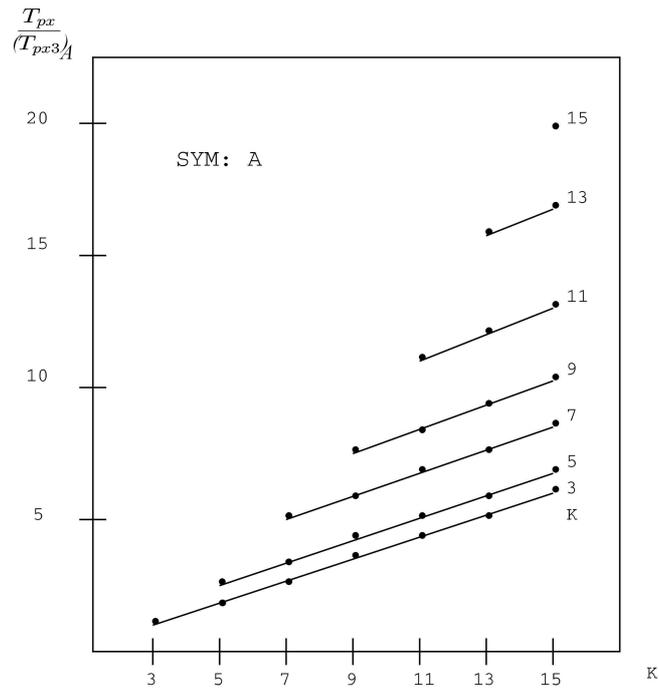
■

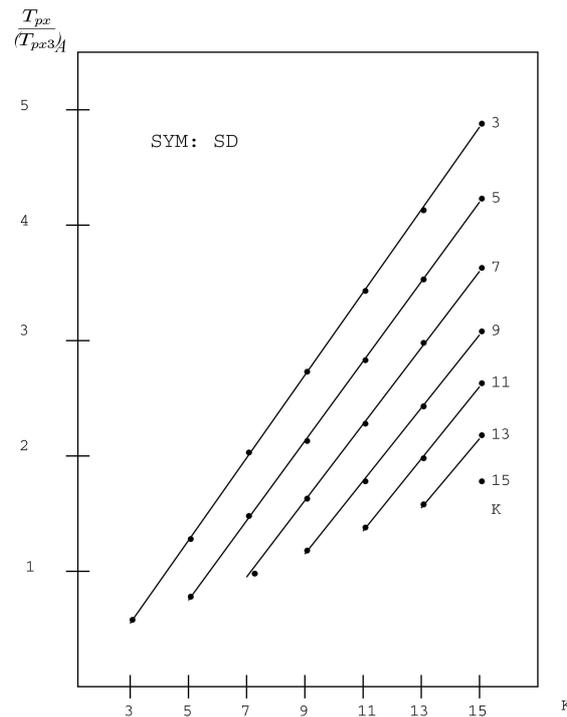
Hinweise:

Aus den folgenden Abbildungen lassen sich drei Regeln für die Verkopplung von Symmetrie und Kaskadierung in der Faltung ableiten.

1. Ist die Symmetrie kleiner als oktal, dann optimale Kaskadierung durch minimale Wurzeln (z.B. 3×3)

2. Oktale Symmetrie:

 $K < 7$: nicht kaskadiert $K \geq 7$: zwei etwa gleich große Wurzeln3. Separable Implementierung schließt Kaskadierung aus, da $n\kappa > K$ 



B) Parallelschaltung

Die Realisierung der Faltung durch Parallelschaltung kleinerer Operatoren entsprechend

$$g = h * f = (h_{(1)} + \dots + h_{(n)}) * f$$

mit

$$h = \sum_{i=1}^n h_{(i)}$$

wird durch die Distributivität der Faltungsoperation ermöglicht. Ein offensichtlicher Nutzen entsteht, wenn parallele Ausführung möglich ist. Sind die $h_{(i)}$ im Vergleich zu h bezüglich ihrer Symmetrie oder Kaskadierung besser, kann dies auch ohne Parallelverarbeitung vorteilhaft sein.

Dabei können die $h_{(i)}$ von anderer Gestalt/Dimension als h sein.

Beispiele:

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} = 9 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

A: $M = 9, A = 8$ $M = 1, A = 0$ $M = 1, A = 4$

C: $M = 5, A = 8$

Kapitel 2

Fouriertransformation

In Abschnitt 1.4.3 wurde behandelt, wie mit Hilfe einer linearen Transformation des Basissystems eines Signals eine für gewisse Problemstellung günstigere Signalrepräsentation erreicht wird. Dabei handelt es sich um das durch einen Vektor von Einheitsimpulsen aufgespannte Basissystem des Ortsraumes. Diesem Basissystem konnte eindeutig ein Basissystem des Signalraumes zugeordnet werden.

In diesem Kapitel wird die Fouriertransformation behandelt, die eine lineare Integraltransformation eines Signals bewirkt. Ihre Wirkung besteht darin, eine Abbildung des Signalraumes auf sich selbst zu erzeugen, in der Eigenschaften der Ähnlichkeit oder Unterschiedlichkeit von Signalwerten an verschiedenen Orten des Signals hervorgehoben werden. Dies erreicht man durch einen Wechsel des Basissystems. Das neue Basissystem wird durch komplexe harmonische Funktionen mit unterschiedlichen Frequenzen gebildet. Diese Frequenzen spannen einen Raum auf, der Frequenzraum genannt wird.

Während also in Kapitel 1 der Ortsraum und der Signalraum zur Formulierung signaltheoretischer Modelle benutzt werden, werden wir in diesem Kapitel die Frequenzraum-Repräsentation (auch spektrale Repräsentation genannt) zu Modellbildung für Signale und Operatoren heranziehen.

Je nach Art der Repräsentation des ursprünglichen Signals werden verschiedene Fouriertransformationen unterschieden. Wir behandeln fünf Repräsentationen der Fouriertransformation. Welche hiervon in späteren Kapiteln verwendet wird, hängt vom spezifischen Problem ab.

2.1 Integraltransformation

Die Fouriertransformation ist eine Integraltransformation. Eine Integraltransformation ist eine spezielle Funktionaltransformation. In der Informationstechnik und speziell in der Signalverarbeitung spielen Funktionaltransformationen eine bedeutende Rolle, siehe [Oppenheim und Schafer, 1999], [Posthoff und Woschni, 1984] und [Preuß, 2002]

Beispiele: (* wird in Vorlesung behandelt)

Fouriertransformation (*), Hartleytransformation (*)

Hilberttransformation (*), Gabortransformation (*)

Wavelettransformation (*)

Laplacetransformation, Z-Transformation

Haartransformation, Walsh-Hadamard-Transformation

Fermattransformation, Mersennetransformation

Wir wollen hier Signale als Funktionen betrachten. Auch Impulsantworten von Operatoren können derartige Funktionen sein. Zunächst seien diese eindimensional und kontinuierlich im Definitionsbereich, also gegeben durch $f(x)$.

Definition (Funktionaltransformation):

Eine *Funktionaltransformation* ist eine Abbildung \mathcal{T} , die jeder Funktion $f(x)$ eines gewissen Funktionenraumes eine Funktion

$$F(u) = \mathcal{T}\{f(x)\}$$

eines anderen Funktionenraumes zuordnet. Die Funktion $f(x)$ heißt Originalfunktion, die Funktion $F(u)$ heißt Bildfunktion. Läßt sich umgekehrt die Originalfunktion aus der Bildfunktion bestimmen,

$$f(x) = \mathcal{T}^{-1}\{F(u)\},$$

so nennt man \mathcal{T}^{-1} die zu \mathcal{T} inverse Funktionaltransformation. ■

Eines "gewissen Funktionenraumes" heißt hier, daß an die zu transformierenden Funktionen gewisse Eigenschaften geknüpft werden (z.B. im Falle der Fourier-Reihenentwicklung die Dirichlet-Bedingungen).

Definition (Integraltransformation):

Eine *Integraltransformation* ist eine Funktionaltransformation \mathcal{T}_K , die mit Hilfe einer Kernfunktion $K(x, u)$ durch ein Integral

$$F(u) = \mathcal{T}_K\{f(x)\} = \int_C K(x, u)f(x)dx$$

definiert ist. In der komplexen Analysis ist das Integral im Allgemeinen ein Linienintegral in der komplexen x -Ebene. Der Integrationsbereich C kann aber auch die reelle Achse oder ein endliches Intervall auf ihr sein.

Die inverse Integraltransformation \mathcal{T}_K^{-1} wird definiert durch

$$f(x) = \mathcal{T}_K^{-1}\{f(u)\} = \int_C K^{-1}(x, u)F(u)du$$

■

Genauer handelt es sich bei dieser Integraltransformation um eine Fredholmsche Integralgleichung 1. Art. Der erzeugende Kern dieser Integraltransformation definiert die Beziehung der beiden Funktionenräume untereinander, bzw. den Wechsel des Basissystems der betrachteten Funktionen.

Die Integraltransformation ist linear, d.h.

$$\mathcal{T}_K\{g(x) + h(x)\} = \mathcal{T}_K\{g(x)\} + \mathcal{T}_K\{h(x)\}$$

$$\mathcal{T}_K\{\alpha g(x)\} = \alpha \mathcal{T}_K\{g(x)\}.$$

Im Fall der Fouriertransformation ist der Kern K gegeben durch

$$K(x, u) = e^{-j2\pi ux}.$$

Nicht jede Funktion $f(x)$ kann mittels Fouriertransformation nach $F(u)$ transformiert werden. Hinreichende Bedingungen für $f(x)$ sind gegeben durch (*Dirichlet-Bedingungen*):

1. $f(x)$ ist absolut integrierbar im Integrationsbereich C , d.h.

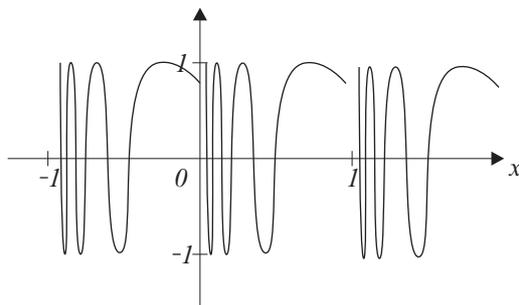
$$\int_C |f(x)|dx \leq M < \infty,$$

2. $f(x)$ hat eine endliche Zahl von Diskontinuitäten und eine endliche Anzahl von Extrema im Integrationsbereich.

Gegenbeispiele:

- 1) $f(x) = \frac{1}{1-x}$: nicht absolut integrierbar über $[0, 1]$. Bei $x = 1$ tritt eine Polstelle auf.

- 2) $f(x) = \sin \frac{1}{x}$: bei $x \rightarrow 0$ geht die Zahl der Oszillationen gegen Unendlich (die Funktion hat unendlich viele Maxima und Minima)



■

Nehmen wir an, daß die Funktion f als (unendliche) Folge f_m gegeben ist, also als diskrete Funktion vorliegt. Dann lautet die diskrete Formulierung der Integraltransformation

$$F_u = \sum_{m \in D_m} K_{mu} f_m, \quad u \in D_u.$$

Offensichtlich ist der Transformationskern hier durch eine (unendliche) Matrix $K = [K_{mu}]$ repräsentiert; f und F sind Vektoren. Für festes u wird aus dem Transformationskern ein Vektor b_u (ein *Basisvektor*) mit den Komponenten b_m^u . Dann stellt

$$F_u = \sum_{m \in D_m} b_m^u f_m = \langle b_u, f \rangle$$

ein Skalarprodukt der Vektoren b_u und f dar.

Mit anderen Worten, für ein festes u ist F_u der Wert der Projektion von f auf b_u . Er repräsentiert die durch b_u spezifizierte Eigenschaft von f .

Ist K der Transformationskern der Fouriertransformation, dann wird u als *Frequenz* oder *Raumfrequenz* bezeichnet und D_u spannt den *Frequenzraum* auf. In diesem Fall repräsentiert F_u den Anteil der Frequenz u an der Synthese des Signals f in einem integralen Sinn. Dies ist der Zugang zur Frequenzanalyse des Signals f .

Die Basisfunktion der Fouriertransformation b_u heißt *komplex-harmonische Funktion* der Frequenz u . Im Fall der kontinuierlichen Fouriertransformation ist

$$b_u(x) = e^{-j2\pi ux} = \cos 2\pi ux - j \sin 2\pi ux$$

die Basisfunktion. Sie wird durch ein Paar trigonometrischer Funktionen repräsentiert, von denen die \cos -Funktion von gerader Symmetrie ist ($\cos(-x) = \cos(x)$) und die \sin -Funktion von ungerader Symmetrie ist ($\sin(-x) = -\sin(x)$). Also projiziert die Fouriertransformation frequenzabhängig eine Funktion f auf diese beiden Symmetriekonzepte in einem globalen Sinn. Das heißt, diese Symmetrieanteile werden über D_x integriert.

Im Fall der diskreten Fouriertransformation stellt b_u eine diskretisierte Version solcher trigonometrischer Funktionen dar.

Die aus den Basisvektoren b_u für alle Frequenzen u gebildete Basis ist orthogonal. Deshalb kann eine Rücktransformation (inverse Integraltransformation) formuliert werden:

$$f_m = \sum_{u \in D_u} K_{mu}^{-1} F_u, \quad m \in D_m.$$

Die Orthogonalitätsbedingung für komplexe Harmonische werden wir im nächsten Abschnitt zeigen.

Im Falle der *Orthogonalität* gilt also $K K^{-1} = I$. Dann ist K eine reguläre Matrix, d.h. es gilt $\det K \neq 0$. K hat also D_m linear unabhängige Zeilen. Von Orthogonalität spricht man allerdings nur im Falle reellwertiger Transformationskerne. Im Falle komplexwertiger Transformationskerne spricht man von *Unitarität*.

Eine Integraltransformation \mathcal{T}_K heißt unitär, wenn gilt

$$K K^{*T} = I, \text{ d.h. } K^{*T} = K^{-1},$$

wobei K^* den komplex-konjugierten Kern und K^T den transponierten Kern bedeuten. Der Kern K^{*T} heißt *adjungiert*.

Die komplexwertige Fouriertransformation ist eine unitäre Transformation. In Abschnitt 2.6 lernen wir auch eine reellwertige Fouriertransformation, die Hartleytransformation, und in Abschnitt 2.7 die quaternionwertige Fouriertransformation kennen.

Nun werden die Konzepte der Integraltransformation auf den 2D-Fall erweitert. Eine 2D-Integraltransformation ist im kontinuierlichen Fall definiert durch

$$F(u, v) = \mathcal{T}_K\{f(x, y)\} = \iint_C K(x, y, u, v) f(x, y) dx dy.$$

und die inverse Integraltransformation ist definiert durch

$$f(x, y) = \mathcal{T}_K^{-1}\{F(u, v)\} = \iint_C K^{-1}(x, y, u, v) F(u, v) du dv$$

Im Fall der Fouriertransformation ist der Transformationskern

$$\begin{aligned} K(\mathbf{x}, \mathbf{u}) &= e^{-j2\pi(xu+yv)} \\ &= \cos 2\pi(xu + yv) - j \sin 2\pi(xu + yv), \end{aligned}$$

wobei $\mathbf{x} = (x, y)$ und $\mathbf{u} = (u, v)$.

Für konstantes \mathbf{u} ist die zugehörige komplexe Harmonische (die Basisfunktion) ein sinusförmiger Wellenzug in der durch $D_{\mathbf{x}}$ aufgespannten Ebene mit der Frequenz $\nu = \sqrt{u^2 + v^2}$ und der Orientierung θ , $\tan \theta = \frac{v}{u}$.

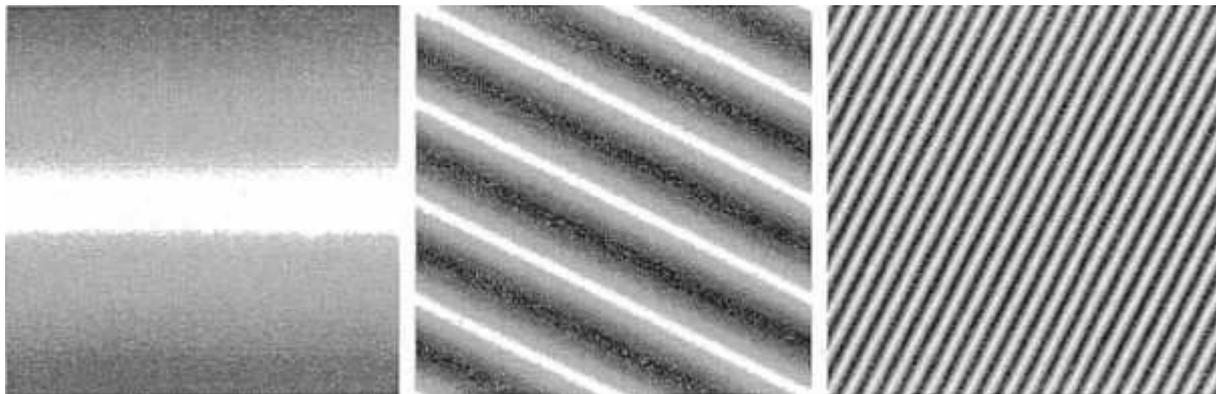
In der folgenden Abbildung sind Beispiele für diese Basisfunktionen dargestellt.

Bei genauerer Betrachtung des Transformationskernes erkennt man, daß er auch geschrieben werden kann

$$K(\mathbf{x}, \mathbf{u}) = e^{-j2\pi\langle \mathbf{x}, \mathbf{u} \rangle} = e^{-j2\pi \mathbf{x} \cdot \mathbf{u}}$$

Das Skalarprodukt, $\langle \mathbf{x}, \mathbf{u} \rangle$, ist invariant gegenüber isometrischen Transformationen (lassen den Abstand zwischen zwei Punkten unbeeinflusst), also auch der Rotation. Die Fouriertransformation hängt deshalb nicht von der Wahl der orthogonalen Basisvektoren der Ebene ab. Deshalb ist $K(\mathbf{x}, \mathbf{u})$ gleichzeitig als isotrop in einem polaren Koordinatensystem zu interpretieren. Die komplexen Harmonischen sind in einem polaren Koordinatensystem also als isotrope sinusförmige Wellenzüge vorstellbar. Sie heißen dann *sphärische* (oder polare) *Harmonische*.

Beispiele: $D_x = [-1, 1] \times [-1, 1]$, Darstellung von $\sin 2\pi \mathbf{xu}$



$$(u, v) = (0, 0.4)$$

$$(u, v) = (1, 2)$$

$$(u, v) = (10, -5)$$

■

Im Rahmen der komplexen Algebra gilt aber auch, daß der 2D-Fouriertransformationskern separabel ist:

$$K(\mathbf{x}, \mathbf{u}) = e^{-2\pi j u x} e^{-2\pi j v y}.$$

Dies hat Konsequenzen für die Berechnung der 2D-Fouriertransformation, wie wir am Beispiel der diskreten Formulierung zeigen:

$$F_{uv} = \sum_{m \in D_m} \sum_{n \in D_n} K_{mnuv} f_{mn} \quad , u \in D_u, v \in D_v.$$

Die 2D-Transformationsmatrix $K = [K_{mnuv}]$ hat die Dimension $D_m D_n \times D_u D_v$, für ein Bild der Dimension $M \times N$ also $MN \times MN$. Betrachten wir nun die Konsequenzen der Separabilität des Transformationskernes. Dann folgt hieraus, daß Zeilen und Spalten eines Bildes mittels zweier 1D-Transformationen unabhängig transformiert werden können.

Sei $S = [S_{nv}]$ die $N \times N$ -Transformationsmatrix in Spaltenrichtung und $Z = [Z_{mu}]$ die $(M \times M)$ -Transformationsmatrix in Zeilenrichtung. Dann gilt

$$F = Z f S,$$

Die Transformationsvorschrift

$$F_{uv} = \sum_{m \in D_m} \left(\sum_{n \in D_n} f_{mn} S_{nv} \right) Z_{mu}$$

entspricht der Ausführung folgender Berechnungsschritte:

1. Schritt: Spaltentransformation

$$F_{mv} = \sum_{n \in D_n} f_{mn} S_{nv}$$

2. Schritt: Zeilentransformation

$$F_{uv} = \sum_{m \in D_m} F_{mv} Z_{mu}.$$

Die Reihenfolge der Fouriertransformationen kann auch vertauscht werden.

Zusammenfassend stellen wir fest, daß das komplexwertige Basissystem der Fouriertransformation im mehrdimensionalen Fall aus algebraischen Gründen einige Möglichkeiten der Interpretation zuläßt. Dies hat nicht nur Konsequenzen für die Berechnung der Fouriertransformation. Vielmehr ist die Wahl des geeigneten Koordinatensystems in Abhängigkeit von der Problemstellung zu treffen.

2.2 Repräsentationen der Fouriertransformation

Wir geben zunächst eine Übersicht über fünf zu unterscheidende Repräsentationen der Fouriertransformation:

1. *Fourierreihenentwicklung*
2. *Fouriertransformation (kontinuierlicher Signale)*
3. *Fouriertransformation diskreter Signale*
4. *diskrete Fourierreihenentwicklung*
5. *diskrete Fouriertransformation*

Die Unterschiede resultieren aus dem Definitionsbereich der Signale. Dieser kann sein:

- endlich periodisch, endlich oder unendlich ausgedehnt
- kontinuierlich oder diskret

Bemerkung:

Bei allen Repräsentationen der Fouriertransformation nehmen wir reellwertige Funktionen $f(x)$ an. Im Allgemeinen ist die komplexwertige Fouriertransformation über komplexen Funktionen definiert.

Fourierreihenentwicklung (FR)

$f(x)$	○ — ●	$\gamma_u \in \mathbb{C}$
endlich periodisch		
Grundintervall $[-\pi, \pi]$		unendlich ausgedehnt
kontinuierlich		diskret
reell		komplex

$$\gamma_u = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-j2\pi ux} dx$$

sind die *Fourierkoeffizienten* für jedes $u \in \mathbb{Z}$. Die Funktionen $b = \{b_u(x) = \frac{1}{2\pi}e^{-j2\pi ux}\}$ heißen *komplexe Harmonische* und bilden das orthogonale Basissystem der FR-Entwicklung:

$$f(x) = \sum_{u=-\infty}^{\infty} \gamma_u e^{j2\pi ux}$$

Bemerkung:

Bei der Fourierreihenentwicklung und allen anderen Repräsentationen der Fouriertransformation treten negative Frequenzen auf.

Die Fourierreihenentwicklung kann für jedes beliebige Grundintervall $I' = [a, b]$ ausgeführt werden. Seien $x \in I = [-\pi, \pi]$ und $x' \in I' = [a, b]$, dann

$$\begin{aligned} x &= \frac{\pi}{b-a} (2x' - a - b) \\ x' &= \frac{1}{2} \left(\frac{b-a}{\pi} x + a + b \right). \end{aligned}$$

Fouriertransformation (FT)

$f(x)$	○ —●	$F(u)$
unendlich ausgedehnt		unendlich ausgedehnt
kontinuierlich		kontinuierlich
reell		komplex

Die Herleitung der FT aus der FR erfolgt durch

1. $x \in [-\frac{L}{2}; \frac{L}{2}]$
2. periodische Fortsetzung
3. Grenzübergang $L \rightarrow \infty$

$$f(x) = \int_{-\infty}^{\infty} F(u) e^{j2\pi ux} du$$
$$F(u) = \int_{-\infty}^{\infty} f(x) e^{-j2\pi ux} dx$$

Die Konvergenz des Fourierintegrals bei nicht *quadratintegrierbaren* (energiebegrenzten) Funktionen $f(x)$ bzw. $F(u)$ kann mit Hilfe eines „Konvergenzfaktors“ $e^{-\epsilon|x|}$ im Grenzübergang $\epsilon \rightarrow 0$ erzwungen werden:

$$f(x) = \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} F(u) e^{j2\pi ux - \epsilon|u|} du$$
$$F(u) = \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} f(x) e^{-j2\pi ux - \epsilon|x|} dx$$

Beispiel: $f(x) = 1$: unendlich ausgedehnte konstante Funktion

$$1 \quad \text{○} \text{---} \text{●} \quad \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} e^{-j2\pi ux - \epsilon|x|} dx$$

$$= \lim_{\epsilon \rightarrow 0} \frac{2\epsilon}{\epsilon^2 + 4\pi^2 u^2}$$

$$=: \delta(0) \quad (\text{Dirac-Impuls an der Stelle } 0)$$



Fouriertransformation diskreter Signale (FTD)

f_m ○—●	$F(e^{j\omega})$
unendlich ausgedehnt	periodisch in 2π
diskret	kontinuierlich
reell	komplex

$\approx z$ -Transformation, wenn $z = e^{-j\omega m}$ mit der Kreisfrequenz $\omega = 2\pi u$

$$f_m = \int_{-\pi}^{\pi} F(e^{j\omega}) e^{j\omega m} d\omega \quad \text{○—●} \quad F(e^{j\omega}) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} f_m e^{-j\omega m}$$

Diskrete Fourierreihenentwicklung (DFR)

\tilde{f}_m ○—●	\tilde{F}_u
periodisch in M	periodisch in M
diskret	diskret
reell	komplex

$$\tilde{f}_m = \sum_{u=0}^{M-1} \tilde{F}_u e^{\frac{j2\pi um}{M}} \quad m = 0, 1, 2, \dots, M-1$$
$$\tilde{F}_u = \frac{1}{M} \sum_{m=0}^{M-1} \tilde{f}_m e^{-\frac{j2\pi um}{M}} \quad u = 0, 1, 2, \dots, M-1$$

Die DFR benötigt im Unterschied zu FR nur endlich viele Koeffizienten.

Diskrete Fouriertransformation (DFT)

f_m	○ — ●	F_u
endlich		endlich
diskret		diskret
reell		komplex

Die DFT entsteht aus der DFR durch Ausblenden der Grundperiode M von Signal und Spektrum mit dem Ausblendoperator

$$\mathcal{A}_m^M = \begin{cases} 1 & 0 \leq m \leq M-1 \\ 0 & \text{sonst} \end{cases}$$

bzw.

$$\mathcal{A}_u^M = \begin{cases} 1 & 0 \leq u \leq M-1 \\ 0 & \text{sonst} \end{cases}$$

$$\begin{aligned} F_u &= M \mathcal{A}_u^M \{ \tilde{F}_u \} \\ &= \begin{cases} \sum_{m=0}^{M-1} f_m W_M^{mu} & 0 \leq u \leq M-1 \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

$$\begin{aligned} f_m &= \mathcal{A}_m^M \{ \tilde{f}_m \} \\ &= \begin{cases} \frac{1}{M} \sum_{u=0}^{M-1} F_u W_M^{-mu} & 0 \leq m \leq M-1 \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

mit

$$\begin{aligned} W_M^{mu} &= e^{\frac{-j2\pi mu}{M}} \\ W_M^u &= e^{\frac{-j2\pi u}{M}} \end{aligned}$$

2.3 Fouriertransformationen kontinuierlicher Signale

Für die Fouriertransformation kontinuierlicher Signale muß unterschieden werden zwischen

- periodischen Signalen \rightarrow Fourierreihenentwicklung
- aperiodischen unendlich ausgedehnten Signalen \rightarrow Fouriertransformation

Wir werden hier einige Grundkonzepte betrachten, die wir in Abschnitt 2.4 auf den Fall diskreter Signale übertragen.

2.3.1 Komplexe harmonische Funktionen

Die komplexe harmonische Funktion mit der Kreisfrequenz $\omega = 2\pi u$,

$$\phi_u(x) = e^{j\omega x} = e^{j2\pi ux} \quad , \quad u = 0, \pm 1, \pm 2, \dots$$

ist periodisch mit der Periode $P = \frac{2\pi}{\omega}$. Es gilt für alle $x \in (-\infty, \infty)$

$$\phi_u(x) = \phi_u(x + P) = \phi_u(x + 2P) = \dots = \phi_u(x + nP)$$

für alle positiven ganzen Zahlen n . Da die Funktion auch periodisch mit der Periode $2P, 3P, \dots$ ist, heißt die kleinste dieser Perioden, P , die **Grundperiode**. Als **Grundfrequenz** bezeichnet man $\omega_0 = \frac{2\pi}{P}$ bzw. $u_0 = \frac{1}{P}$.

Daß die Funktion $e^{j\omega x}$ komplexwertig ist, erkennt man unmittelbar an der **Eulerschen Formel**

$$e^{j\omega x} = \cos \omega x + j \sin \omega x.$$

Aus der Reihenentwicklung der Exponentialfunktion folgt allgemein, daß sich die algebraische Natur des Exponenten auf die Exponentialform überträgt.

In der Fourieranalyse von Signalen treten positive und negative Frequenzen auf. In diesem Fall gilt

$$\cos \omega x = \frac{e^{j\omega x} + e^{-j\omega x}}{2} \quad , \quad \sin \omega x = \frac{e^{j\omega x} - e^{-j\omega x}}{2j}.$$

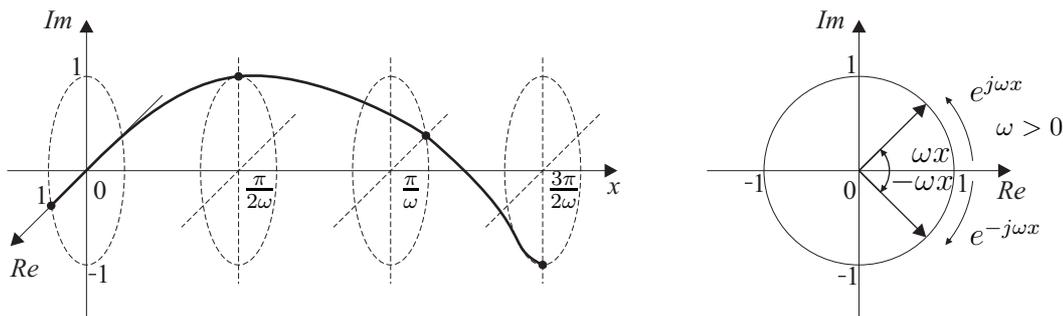
Dies impliziert auch die Eulersche Formel der **komplexen konjugierten Harmonischen**

$$e^{-j\omega x} = \cos \omega x - j \sin \omega x.$$

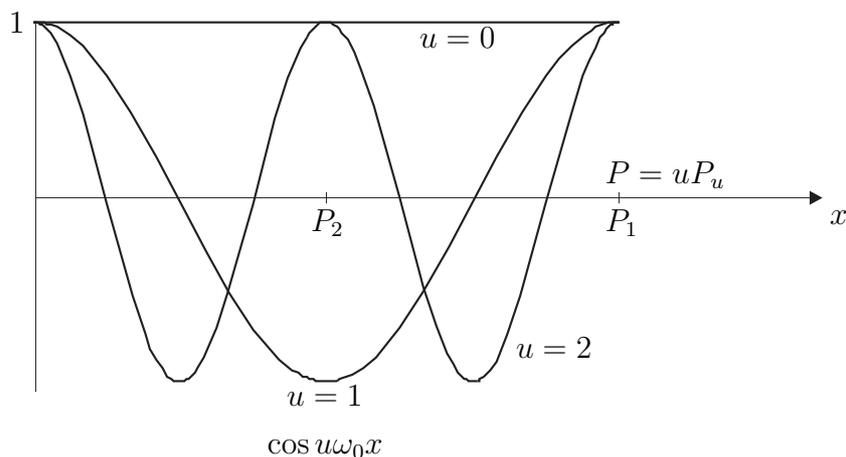
Da

$$|e^{j\omega x}| = \sqrt{\cos^2 \omega x + \sin^2 \omega x} = 1,$$

verläuft $e^{j\omega x}$ in der komplexen Ebene auf dem Einheitskreis für alle ω und x . Für $\omega > 0$ rotiert $e^{j\omega x}$ entgegen dem Uhrzeigersinn, für $\omega < 0$ im Uhrzeigersinn mit wachsendem x .



Wir schreiben nur die komplexe Harmonische etwas anders, um eine Interpretation der Frequenz u zu zeigen. Für jeden Wert $u, u = 0, \pm 1, \pm 2, \dots$, ist die Funktion $\phi_u(x) = e^{ju\omega_0 x}$ eine periodische Funktion mit der Frequenz $\omega = u\omega_0$ und der Periode $P_u = \frac{2\pi}{\omega}$. Wenn die Grundperiode definiert ist durch $P = \frac{2\pi}{\omega_0}$, dann $P = uP_u$ und irgendein Intervall der Länge P enthält u vollständige Zyklen von $\phi_u(x)$. Die Frequenz u gibt also die Anzahl der Perioden P_u in der Grundperiode P an.



Für jedes x_0 gilt

$$\begin{aligned} \int_{x_0}^{x_0+P} e^{ju\omega_0 x} dx &= \int_{x_0}^{x_0+P} \cos u\omega_0 x + j \sin u\omega_0 x dx \\ &= \begin{cases} P & \text{für } u = 0 \\ 0 & \text{für } u \neq 0 \end{cases} \end{aligned}$$

Beweis:

$u = 0$: $e^{ju\omega_0 x} = 1$ für alle x , folglich $\int_{x_0}^{x_0+P} dx = P$

$u \neq 0$: Fläche unter \cos/\sin ist Null

Sei die komplexe konjugierte harmonische Funktion

$$\phi_u^*(x) = (e^{ju\omega_0 x})^* = e^{-ju\omega_0 x}.$$

Dann impliziert das obige Integral die **Orthogonalitätseigenschaft** für die Menge aller komplexen Harmonischen $\{\phi_u(x) | u = 0, \pm 1, \pm 2, \dots\}$:

$$\begin{aligned} \int_P \phi_m(x) \phi_n^*(x) dx &= \int_{x_0}^{x_0+P} e^{j(m-n)\omega_0 x} dx \\ &= \begin{cases} P & \text{für } m = n \\ 0 & \text{für } m \neq n \end{cases} \end{aligned}$$

für jedes x_0 .

Die Orthogonalität der komplexen Harmonischen ist die Grundlage der Fourierreihenentwicklung einer Funktion $f(x)$.

2.3.2 Fourierreihen periodischer Funktionen

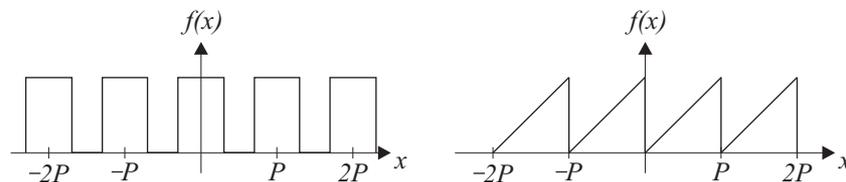
Gegeben sei eine periodische Funktion $f(x)$ mit der Grundperiode P und der Grundfrequenz $\omega_0 = \frac{2\pi}{P}$. Dann ist die Fourierreihe (FR) der periodischen Funktion $f(x)$ gegeben durch

$$f(x) = \sum_{u=-\infty}^{\infty} \gamma_u e^{ju\omega_0 x}$$

mit den *Fourierkoeffizienten*

$$\gamma_u = \frac{1}{P} \int_P f(x) e^{-ju\omega_0 x} dx = \frac{1}{P} \int_{x_0}^{x_0+P} f(x) e^{-ju\omega_0 x} dx$$

für $u = 0, \pm 1, \pm 2, \dots$ und für jedes x_0 .



Die Grundfrequenz ω_0 ist der größte gemeinsame Teiler (ggT) aller Frequenzen $\omega = u\omega_0$. Folglich ist $f(x)$ periodisch mit der Grundperiode P für jeden Koeffizienten γ_u . Umgekehrt kann jede periodische Funktion mit der Grundperiode P repräsentiert werden durch eine Linearkombination komplexer Harmonischer mit unterschiedlichen Frequenzen.

Zwei periodische Funktionen $f_i(x)$ mit den Perioden $P_i, i = 1, 2$, addieren sich zu einer Funktion $f(x) = f_1(x) + f_2(x)$. Die Funktion $f(x)$ ist nur dann ebenfalls periodisch, wenn P_1/P_2 eine rationale Zahl ist oder ein Bruch. Andernfalls haben die beitragenden Funktionen keine gemeinsame Periode und ihre Summe ist nicht periodisch.

Beispiele:

$$f_1(x) = \sin 2x \quad P_1 = \frac{2\pi}{2} = \pi$$

$$f_2(x) = \sin 2\pi x \quad P_2 = \frac{2\pi}{2\pi} = 1$$

$$f_3(x) = \sin 21x \quad P_3 = \frac{2\pi}{21}$$

$$a) f(x) = f_1(x) + f_2(x) : P_1/P_2 = \frac{\pi}{1} = \pi \quad (\text{ggT}\{\pi, 1\} \text{ existiert nicht})$$

$f(x)$ ist nicht periodisch, da $f_1(x)$ und $f_2(x)$ keine gemeinsame Periode haben.

$$b) f(x) = f_1(x) + f_3(x) : P_1/P_3 = \frac{\pi}{2\pi/21} = \frac{21}{2} \quad (\text{ggT}\{2, 21\} = 1)$$

$f(x)$ ist periodisch mit Grundperiode $P = 2P_1 = 21P_3 = 2\pi$ und Grundfrequenz $\omega_0 = \frac{2\pi}{2\pi} = 1$.

$$f_4(x) = \pi \sin 0.6x - 3 \cos 1.5x + 11 \sin 3x$$

$$\text{ggT}\{0.6, 1.5, 3\} = 0.3 \rightarrow f_4 \text{ ist periodisch}$$

Hieraus folgt für f_4 die Grundfrequenz $\omega_0 = 0.3$ und die Grundperiode $P = \frac{2\pi}{0.3}$.

■

Herleitung der Entwicklungskoeffizienten der Fourierreihe:

$$\begin{aligned} \int_{x_0}^{x_0+P} f(x) e^{-jn\omega_0 x} dx &= \sum_{m=-\infty}^{\infty} \int_{x_0}^{x_0+P} \gamma_m e^{jm\omega_0 x} e^{-jn\omega_0 x} dx \\ &= \sum_{m=-\infty}^{\infty} \gamma_m \int_{x_0}^{x_0+P} e^{j(m-n)\omega_0 x} dx \end{aligned}$$

Wegen der Orthogonalitätseigenschaft verschwindet jedes Integral für $m \neq n$ und wird gleich P für $m = n$. Folglich gilt für jedes beliebige x_0

$$\int_{x_0}^{x_0+P} f(x) e^{-jn\omega_0 x} dx = P\gamma_n.$$

Hieraus folgt unmittelbar der Ausdruck für die Fourierkoeffizienten γ_n .

$$\gamma_n = \frac{1}{P} \int_0^P f(x) e^{-jn\omega_0 x} dx = \frac{1}{P} \int_{-P/2}^{P/2} f(x) e^{-jn\omega_0 x} dx$$

Die Verschiebung der Integrationsgrenzen ist möglich, weil das Integral für den geraden/ungeraden Signalanteil erhalten bleibt. Die Fourierkoeffizienten können reellwertig oder komplexwertig sein.

Beispiel 1: Fourierreihenentwicklung

$$f(x) = 1 - 3 \cos 0.6\pi x + 2 \sin 1.2\pi x + \cos 2.1\pi x \text{ für alle } x$$

$\text{ggT}\{0.6\pi, 1.2\pi, 2.1\pi\} = 0.3\pi \rightarrow f(x)$ ist periodisch, weil alle $\frac{P_i}{P_j}$ rational sind, mit

$$\omega_0 = 0.3\pi, \quad P = \frac{2\pi}{0.3\pi} = \frac{20}{3}$$

Also kann $f(x)$ geschrieben werden als

$$f(x) = 1 - 3 \cos 2\omega_0 x + 2 \sin 4\omega_0 x + \cos 7\omega_0 x \text{ für alle } x.$$

Berechnung der Fourierkoeffizienten (nicht nach Formel):

$$\sin \omega x = \frac{e^{j\omega x} - e^{-j\omega x}}{2j} \quad \cos \omega x = \frac{e^{j\omega x} + e^{-j\omega x}}{2}$$

$$\begin{aligned} f(x) &= \frac{1}{2} e^{-j2.1\pi x} - \frac{2}{2j} e^{-j1.2\pi x} - \frac{3}{2} e^{-j0.6\pi x} \\ &+ 1 - \frac{3}{2} e^{j0.6\pi x} + \frac{2}{2j} e^{j1.2\pi x} + \frac{1}{2} e^{j2.1\pi x} \\ &= 0.5 e^{-j7\omega_0 x} + j e^{-j4\omega_0 x} - 1.5 e^{-j2\omega_0 x} \\ &+ 1 - 1.5 e^{j2\omega_0 x} - j e^{j4\omega_0 x} + 0.5 e^{j7\omega_0 x} \end{aligned}$$

Bemerkung: Da $j = \sqrt{-1}$, $j^2 = -1$ ist $-\frac{1}{j} = \frac{j^2}{j} = j$.

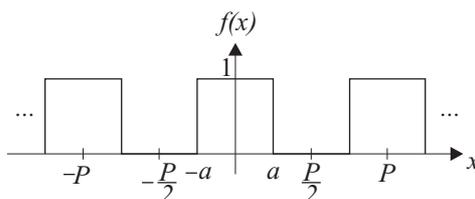
Folglich sind die Fourierkoeffizienten γ_u der Frequenzen $\omega = u\omega_0$:

$$\gamma_{-7} = 0.5, \gamma_{-4} = j, \gamma_{-2} = -1.5, \gamma_0 = 1, \gamma_2 = -1.5, \gamma_4 = -j, \gamma_7 = 0.5$$

Alle anderen Fourierkoeffizienten sind Null. ■

Beispiel 2: Fourierreihenentwicklung

Rechteckimpulse mit Amplitude 1, Breite $2a$, Periode P , Grundfrequenz $\omega_0 = \frac{2\pi}{P}$



$$\gamma_0 = \frac{1}{P} \int_{-\frac{P}{2}}^{\frac{P}{2}} f(x) dx = \frac{1}{P} \int_{-a}^a 1 dx = \frac{2a}{P}$$

$$\begin{aligned} \gamma_u &= \frac{1}{P} \int_{-\frac{P}{2}}^{\frac{P}{2}} f(x) e^{-ju\omega_0 x} dx = \frac{1}{P} \int_{-a}^a e^{-ju\omega_0 x} dx \\ &= \frac{1}{-ju\omega_0 P} [e^{-ju\omega_0 a} - e^{ju\omega_0 a}] = \frac{2 \sin u\omega_0 a}{u\omega_0 P} \end{aligned}$$

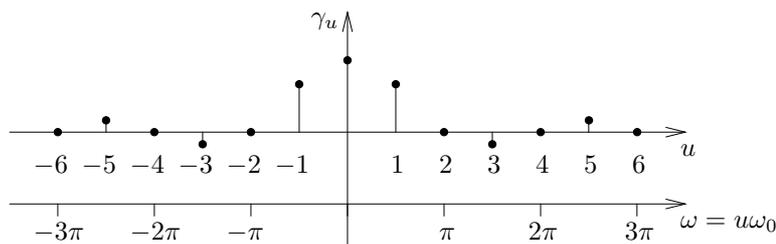
Es werden also alle Fourierkoeffizienten γ_u , $u \in (-\infty, \infty)$ benötigt, um die Funktion darzustellen. Dies ist typisch für Funktionen, die unendlich steile Kanten (Sprungstellen) haben.

Annahme: $P = 4$, $a = 1$

$$\begin{aligned} \omega_0 &= \frac{2\pi}{P} = \frac{\pi}{2}, \quad \gamma_0 = \frac{2a}{P} = 0.5 \\ \gamma_u &= \frac{2 \sin u\omega_0 a}{u\omega_0 P} = \frac{2 \sin(\frac{u\pi}{2})}{2\pi u} = \frac{\sin(\frac{u\pi}{2})}{u\pi} \end{aligned}$$

oder $\gamma_1 = \gamma_{-1} = \frac{1}{\pi}$, $\gamma_2 = \gamma_{-2} = 0$, $\gamma_3 = \gamma_{-3} = -\frac{1}{3\pi}$, $\gamma_4 = \gamma_{-4} = 0$, $\gamma_5 = \gamma_{-5} = \frac{1}{5\pi}$, ...

Obwohl die Sinusfunktion von ungerader Symmetrie ist, führt die Division durch die vorzeichenbehaftete Frequenz zu gerader Symmetrie der Fourierkoeffizienten.



Folglich

$$f(x) = 0.5 + \frac{1}{\pi} (e^{j\omega_0 x} + e^{-j\omega_0 x}) - \frac{1}{3\pi} (e^{j3\omega_0 x} + e^{-j3\omega_0 x}) + \dots$$

Da die Fourierkoeffizienten von gerader Symmetrie sind, sind sie auch reellwertig. ■

Bemerkung: Aus der Eulerschen Formel folgt die **trigonometrische Fourierreihe**, die nur positive Frequenzen nutzt.

$$\begin{aligned} f(x) &= \gamma_0 + \sum_{u=1}^{\infty} ((\gamma_u + \gamma_{-u}) \cos u\omega_0 x + j(\gamma_u - \gamma_{-u}) \sin u\omega_0 x) \\ &= \alpha_0 + \sum_{u=1}^{\infty} (\alpha_u \cos u\omega_0 x + \beta_u \sin u\omega_0 x) \end{aligned}$$

mit

$$\alpha_0 = \gamma_0, \quad \alpha_u = \gamma_u + \gamma_{-u}, \quad \beta_u = j(\gamma_u - \gamma_{-u})$$

bzw.

$$\begin{aligned} \alpha_0 &= \frac{1}{P} \int_{x_0}^{x_0+P} f(x) dx \\ \alpha_u &= \frac{2}{P} \int_{x_0}^{x_0+P} f(x) \cos u\omega_0 x dx \\ \beta_u &= \frac{2}{P} \int_{x_0}^{x_0+P} f(x) \sin u\omega_0 x dx \end{aligned}$$

Die Erweiterung zu nichtperiodischen Funktionen (Fouriertransformation) ist einfacher mit der Exponentialform der Fourierreihe.

Diskrete Frequenzspektrum:

Die Fourierreihenentwicklung liefert eine umkehrbar eindeutige Zuordnung von $f(x)$ und $\{\gamma_u | u = 0, \pm 1, \pm 2, \dots\}$. Die Menge der Fourierkoeffizienten $\{\gamma_u\}$ heißt diskretes *Frequenzspektrum*.

Da γ_u eine komplexe Zahl ist, können wir schreiben $\gamma_u = a_u + jb_u$ und es folgen

$$\begin{aligned} \text{Amplitude von } \gamma_u &: |\gamma_u| = \sqrt{a_u^2 + b_u^2} \\ \text{Phase von } \gamma_u &: \angle \gamma_u = \arctan \frac{b_u}{a_u} \equiv \varphi_u \end{aligned}$$

Polarform von γ_u : $\gamma_u = |\gamma_u| e^{j\varphi_u}$

komplex Konjugierte: $\gamma_u^* = a_u - jb_u = |\gamma_u| e^{-j\varphi_u}$

• **Konjugierte Symmetrie (Hermiteische Symmetrie):** Für reellwertiges $f(x)$, d.h. $f^*(x) = f(x)$, folgt

$$\gamma_u^* = \gamma_{-u}$$

bzw.

$$|\gamma_u| = |\gamma_{-u}|, \quad \varphi_u = -\varphi_{-u}$$

Das **Amplitudenspektrum** $|\gamma_u|$ ist folglich von gerader Symmetrie und das **Phasenspektrum** φ_u ist von ungerader Symmetrie. Wir sprechen in diesem Fall von Hermite-Symmetrie des Amplituden- und Phasenspektrums.

Ist $f(x)$ reell und gerade, dann ist γ_u eine reelle und gerade Funktion von u :

$$\gamma_u = a_u = \frac{1}{P} \int_{-\frac{P}{2}}^{\frac{P}{2}} f(x) \cos u\omega_0 x \, dx$$

Ist $f(x)$ reell und ungerade, dann ist γ_u eine imaginäre und ungerade Funktion von u :

$$\gamma_u = jb_u = -\frac{1}{P} \int_{-\frac{P}{2}}^{\frac{P}{2}} f(x) \sin u\omega_0 x \, dx$$

Kosinus und Sinus sind die geraden/ungeraden Beiträge zu den Fourierkoeffizienten.

2.3.3 Approximation durch Fourierreihenentwicklung

Es ist oft unpraktikabel, eine Funktion $f(x)$ mittels unendlich vieler Fourierkoeffizienten zu repräsentieren. Vielmehr ist eine Approximation $\hat{f}(x)$ gesucht, die irgend ein Fehlerkriterium minimiert, wenn die Fourierreihenentwicklung mit endlich vielen Koeffizienten erfolgt. Generell gilt: Es werden für eine gute Approximation umso mehr Koeffizienten benötigt, je steiler lokal die Funktion verläuft. Für $\frac{d}{dx}f(x) \rightarrow \infty$ folgt $u \rightarrow \infty$.

Sei

$$e(x) = f(x) - \hat{f}(x)$$

der **lokale Fehler** am Ort x . Für die Definition des Gesamtfehlers im Intervall $[x_0, x_1]$ gibt es verschiedene Möglichkeiten.

Der **totale Fehler**

$$E_t = \int_{x_0}^{x_1} e(x) dx = \int_{x_0}^{x_1} (f(x) - \hat{f}(x)) dx$$

besitzt den Nachteil, daß kleines E_t nicht auch kleines $e(x)$ impliziert, da positive und negative lokale Fehler sich auslöschen können. Diese Auslöschung wird vermieden mit dem **totalen absoluten Fehler** E_1 ,

$$E_1 = \int_{x_0}^{x_1} |e(x)| dx = \int_{x_0}^{x_1} |f(x) - \hat{f}(x)| dx.$$

Die Minimierung dieses Fehlers mit L_1 -Norm ist analytisch nicht möglich.

Hingegen ist die Minimierung des **quadratischen Fehlers** (L_2 -Norm)

$$E_2 = \int_{x_0}^{x_1} e^2(x) dx = \int_{x_0}^{x_1} (f(x) - \hat{f}(x))^2 dx$$

einfach analytisch zu realisieren. Dieses Kriterium berücksichtigt in gleicher Weise positive und negative Fehler und berücksichtigt große Fehler stärker als kleine Fehler.

Existiert für die approximierte Funktion \hat{f} die Darstellung mit Hilfe von Basisfunktionen $b_i(x)$ als Modell der Funktion $f(x)$,

$$\hat{f}(x) = \sum_{i=1}^n \alpha_i b_i(x),$$

dann erlaubt das **Gaußsche Fehlerminimierungsverfahren** diejenigen α_i analytisch zu berechnen, die E_2 minimieren.

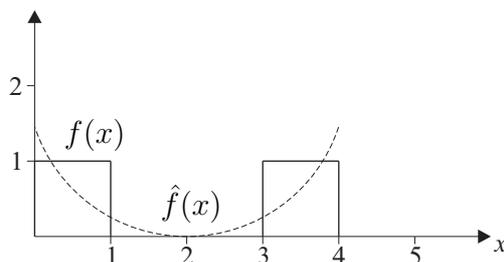
Wird anstelle des quadratischen Fehlers E_2 der **mittlere quadratische Fehler** E verwendet,

$$E = \frac{1}{x_1 - x_0} \int_{x_0}^{x_1} e^2(x) dx,$$

heißt das Optimalverfahren auch **MMSE-Verfahren** (Minimierung des mean square error). Ist $e(x)$ komplexwertig, wird $|e(x)|^2$ unter dem Integral verwendet.

Das Verfahren führt auf ein Gleichungssystem, die **Normalgleichungen**, deren Lösung algebraisch möglich ist, allerdings am einfachsten nur für eine kleine Anzahl von Basisfunktionen.

Beispiel 1: Minimierung des quadratischen Fehlers



Basisfunktionen: $b_1(x) = 1$, $b_2(x) = x$, $b_3(x) = x^2$

Modell: $\hat{f}(x) = \alpha_1 b_1(x) + \alpha_2 b_2(x) + \alpha_3 b_3(x)$

lokaler Fehler: $e(x) = f(x) - \hat{f}(x)$

Finde im Intervall $[0, 4]$ die α_i , $i = 1, 2, 3$, die den quadratischen Fehler minimieren

$$E_2 = \int_0^4 (f(x) - \alpha_1 b_1(x) - \alpha_2 b_2(x) - \alpha_3 b_3(x))^2 dx$$

Notwendige Bedingung, um E_2 zu minimieren:

$$\frac{\partial E_2}{\partial \alpha_i} = 0 \quad \text{für } i = 1, 2, 3$$

Hieraus folgen für $i = 1, 2, 3$ die Normalgleichungen der Art

$$\alpha_1 \int b_1(x) b_i(x) dx + \alpha_2 \int b_2(x) b_i(x) dx + \alpha_3 \int b_3(x) b_i(x) dx = \int f(x) b_i(x) dx$$

mit folgenden beispielhaften Nebenbedingungen

$$\begin{aligned}\int_0^4 b_1^2(x) dx &= 4 \\ \int_0^4 b_1(x)b_2(x) dx &= \int_0^4 x dx = \frac{1}{2}x^2 \Big|_0^4 = 8 \\ \int_0^4 f(x)b_1(x) dx &= \int_0^1 dx + \int_3^4 dx = 2 \quad (\text{da } f(x) = 0 \text{ in } [1, 3]) \\ \int_0^4 f(x)b_2(x) dx &= \int_0^1 x dx + \int_3^4 x dx = \frac{1}{2}x^2 \Big|_0^1 + \frac{1}{2}x^2 \Big|_3^4 = \frac{1}{2} + \frac{7}{2} = 4\end{aligned}$$

...

Das System der Normalgleichungen kann in folgender Art geschrieben werden:

$$\begin{bmatrix} \int b_1^2(x) dx & \int b_2(x)b_1(x) dx & \int b_3(x)b_1(x) dx \\ \int b_1(x)b_2(x) dx & \int b_2^2(x) dx & \int b_3(x)b_2(x) dx \\ \int b_1(x)b_3(x) dx & \int b_2(x)b_3(x) dx & \int b_3^2(x) dx \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} \int f(x)b_1(x) dx \\ \int f(x)b_2(x) dx \\ \int f(x)b_3(x) dx \end{bmatrix}$$

Die Lösung dieses Gleichungssystems existiert, wenn die $b_i(x)$ linear unabhängig sind. Dann läßt sich folgende Gleichung durch Invertierung der Matrix nach dem Parametervektor $\alpha = [\alpha_1, \alpha_2, \alpha_3]^T$ auflösen.

$$\begin{bmatrix} 4 & 8 & \frac{64}{3} \\ 8 & \frac{64}{3} & 64 \\ \frac{64}{3} & 64 & \frac{1024}{5} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ \frac{38}{3} \end{bmatrix}$$

Lösung: $\alpha_1 = \frac{23}{16} = 1.43$, $\alpha_2 = -\frac{45}{32} = -1.4$, $\alpha_3 = \frac{45}{128} = 0.35$

$$\hat{f}(x) = 1.43 - 1.4x + 0.35x^2$$

■

Bemerkung: Das Basissystem $\{1, x, x^2\}$ in diesem Beispiel ist nicht orthogonal. Der Nachteil dieses Umstandes besteht in der aufwendigen Erweiterung der Ordnung des Funktionsmodells. Es müssen hierbei alle Koeffizienten neu berechnet werden.

Die **Legendreschen Polynome**

$$L_n(x) = \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} (1 - x^2)^n$$

bilden für $n = 0, 1, 2, \dots$ ein Orthogonalsystem $\{L_n(x)\}$. Die ersten Basisfunktionen lauten

$$L_0(x) = 1, \quad L_1(x) = x, \quad L_2(x) = \frac{1}{2}(3x^2 - 1), \dots$$

Für praktische Anwendungen ist folgende Rekursionsvorschrift bedeutsam:

$$(n+1)L_{n+1}(x) - (2n+1)xL_n(x) + nL_{n-1}(x) = 0$$

für $n = 1, 2, \dots$. Ausgehend von $L_0(x) = 1$ und $L_1(x) = x$ kann hieraus bequem $L_2(x), L_3(x), \dots$ berechnet werden.

Im Fall eines orthogonalen Basissystems ist die Erweiterung der Modellordnung möglich, ohne die Koeffizienten der bisherigen Modellordnung neu zu berechnen. Die Matrix der Integrale über Produkten der Basisfunktionen ist eine Diagonalmatrix. Deshalb entartet die Matrixgleichung in n entkoppelte Gleichungen. Dies zeigen wir an folgendem Beispiel.

Beispiel 2: Minimierung des quadratischen Felders

*Annahme: gleiche Funktion wie in Beispiel 1
aber orthogonales Basissystem*

$$b_1(x) = 1, \quad b_2(x) = \cos\left(\frac{\pi x}{2}\right), \quad b_3(x) = \sin\left(\frac{\pi x}{2}\right)$$

$$\int_0^4 b_i(x)b_j(x)dx = 0 \quad \text{für } i \neq j$$

$$\int_0^4 b_1^2(x)dx = \int_0^4 dx = 4$$

$$\int_0^4 b_2^2(x)dx = \int_0^4 \cos^2\left(\frac{\pi x}{2}\right)dx = \int_0^4 \frac{1}{2}(1 + \cos \pi x)dx = 2$$

$$\int_0^4 b_3^2(x)dx = \int_0^4 \sin^2\left(\frac{\pi x}{2}\right)dx = \int_0^4 \frac{1}{2}(1 - \cos \pi x)dx = 2$$

Gleichungssystem mit Diagonalmatrix:

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} \int_0^4 f(x)dx \\ \int_0^4 f(x) \cos\left(\frac{\pi x}{2}\right)dx \\ \int_0^4 f(x) \sin\left(\frac{\pi x}{2}\right)dx \end{bmatrix}$$

Die Matrixgleichung reduziert sich zu drei entkoppelten Gleichungen:

$$\begin{aligned}\alpha_1 &= \frac{1}{4} \int_0^4 f(x) dx \\ \alpha_2 &= \frac{1}{2} \int_0^4 f(x) \cos\left(\frac{\pi x}{2}\right) dx \\ \alpha_3 &= \frac{1}{2} \int_0^4 f(x) \sin\left(\frac{\pi x}{2}\right) dx\end{aligned}$$

■

Wir werden nun die Fourierreihenentwicklung als Minimierungsproblem im obigen Sinne behandeln:

- quadratisches Fehlermaß
- linear unabhängige und orthogonale Basisfunktionen.

Gegeben: reell- oder komplexwertige Funktion $f(x)$ über Intervall $[0, P]$
komplexe Harmonische (orthonormal !)

$$\begin{aligned}\phi_u(x) &= \frac{1}{\sqrt{P}} e^{ju\omega_0 x} \quad , \quad u = 0, \pm 1, \pm 2, \dots, \pm N \\ & \quad x \in [0, P] \\ & \quad \omega_0 = \frac{2\pi}{P}\end{aligned}$$

Gesucht: Fourierkoeffizienten γ_u , so daß

$$\hat{f}_N(x) = \sum_{u=-N}^N \gamma_u \phi_u(x)$$

den quadratischen Fehler

$$E_2 = \int_0^P \left(f(x) - \hat{f}_N(x) \right) \left(f(x) - \hat{f}_N(x) \right)^* dx$$

minimiert. Hier wird $f(x)$ (und $\hat{f}_N(x)$) als komplexwertig zugelassen.

$$\begin{aligned}\{\phi_u(x) \mid u = 0, \pm 1, \dots, \pm N\} & \text{ ist orthonormale Basis:} \\ \int_0^P \phi_n(x) \phi_m^*(x) dx &= 0 \quad \text{für } n, m = 0, \pm 1, \dots, \pm N \text{ und } n \neq m \\ \int_0^P \phi_m(x) \phi_m^*(x) dx &= 1 \quad \text{für } m = 0, \pm 1, \dots, \pm N\end{aligned}$$

Notation (Skalarprodukt):

$$\langle x, y \rangle = \int_0^P x(t)y^*(t)dt = \langle y, x \rangle^*$$

$$|x|^2 = \langle x, x \rangle = \int_0^P x(t)x^*(t)dt = \int_0^P |x(t)|^2 dt$$

Entwicklung des quadratischen Fehlermaßes:

$$E_2 = \int_0^P f(x)f^*(x)dx - \sum_{u=-N}^N \gamma_u \int_0^P \phi_u(x)f^*(x)dx$$

$$- \sum_{u=-N}^N \gamma_u^* \int_0^P f(x)\phi_u^*(x)dx + \sum_{u=-N}^N \sum_{v=-N}^N \gamma_u \gamma_v^* \int_0^P \phi_u(x)\phi_v^*(x)dx$$

Verkürzte Schreibweise:

$$E_2 = |f|^2 - \sum_{u=-N}^N (\gamma_u \langle \phi_u, f \rangle + \gamma_u^* \langle f, \phi_u \rangle) + \sum_{u=-N}^N \gamma_u \gamma_u^*$$

Da γ_u komplex ist, soll die Berechnung von $\frac{\partial E_2}{\partial \gamma_u}$ vermieden werden. Stattdessen wird zum Ausdruck von E_2 der Term

$$\sum_{u=-N}^N |\langle f, \phi_u \rangle|^2 = \sum_{u=-N}^N \langle f, \phi_u \rangle \langle f, \phi_u \rangle^* = \sum_{u=-N}^N \langle f, \phi_u \rangle \langle \phi_u, f \rangle$$

addiert und subtrahiert :

$$E_2 = |f|^2 - \sum_{u=-N}^N |\langle f, \phi_u \rangle|^2 + \sum_{u=-N}^N (\langle f, \phi_u \rangle \langle \phi_u, f \rangle - \gamma_u \langle \phi_u, f \rangle - \gamma_u^* \langle f, \phi_u \rangle + \gamma_u \gamma_u^*)$$

$$= |f|^2 - \sum_{u=-N}^N |\langle f, \phi_u \rangle|^2 + \sum_{u=-N}^N (\gamma_u - \langle f, \phi_u \rangle)(\gamma_u^* - \langle \phi_u, f \rangle)$$

$$= |f|^2 - \sum_{u=-N}^N |\langle f, \phi_u \rangle|^2 + \sum_{u=-N}^N |\gamma_u - \langle f, \phi_u \rangle|^2$$

Da die ersten zwei Terme unabhängig von γ_u sind und der dritte Term stets nicht-negativ ist, wird E_2 minimiert, wenn

$$|\gamma_u - \langle f, \phi_u \rangle| = 0$$

oder

$$\gamma_u = \int_0^P f(x)\phi_u^*(x)dx = \frac{1}{\sqrt{P}} \int_0^P f(x)e^{-ju\omega_0 x} dx$$

für alle u .

Folglich ist die optimale Approximation mittels $2N + 1$ Termen möglich:

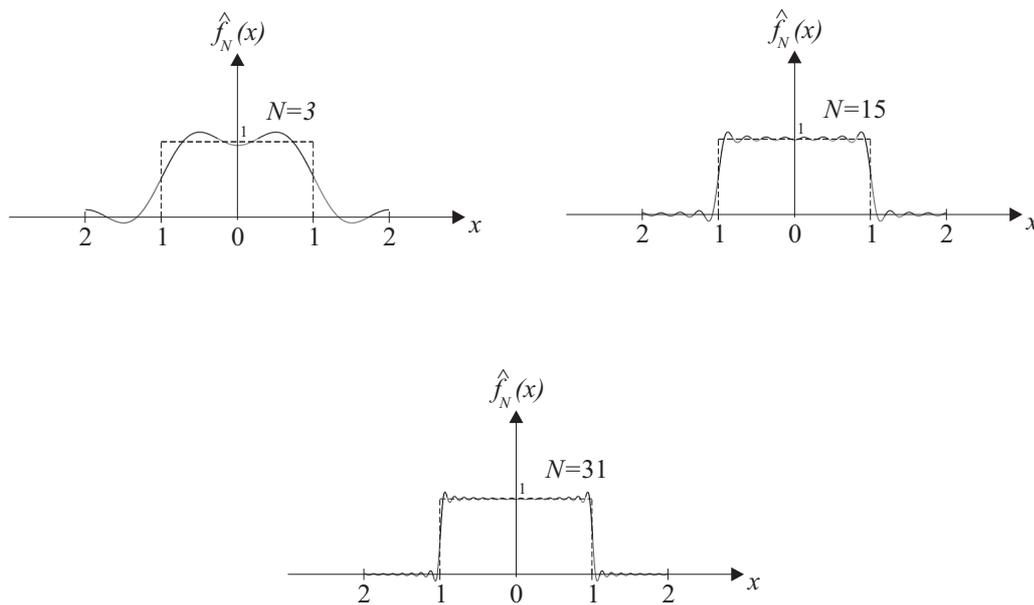
$$\hat{f}_N(x) = \sum_{u=-N}^N \gamma_u \phi_u(x) = \sum_{u=-N}^N \frac{1}{P} \left(\int_0^P f(x) e^{-ju\omega_0 x} dx \right) e^{ju\omega_0 x}$$

Dies ist aber die Fourier-Reihenentwicklung mit endlich vielen Fourierkoeffizienten. Wir stellen deshalb fest:

Eine optimale Näherung im Sinne der Minimierung des quadratischen Fehlers läßt sich durch Fourierentwicklung mit endlich vielen Termen realisieren. Für $N \rightarrow \infty$ folgt $E_2 \rightarrow 0$.

Beispiel: Abgeschnittene Fourierentwicklung für $N = 3, 15$ und 31 .

Die Funktionen zeigen ein für endliches N nicht zu beseitigendes Überschwingen an Diskontinuitäten. Diese Erscheinung heißt **Gibbssches Phänomen**.



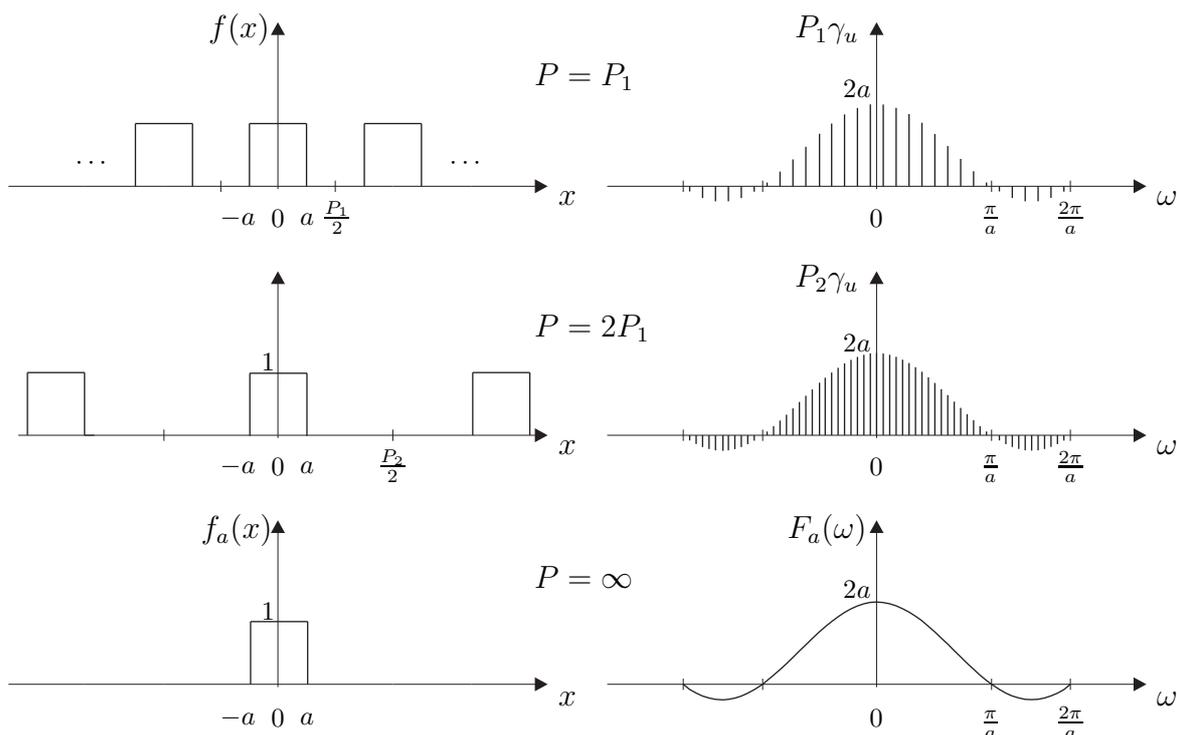
2.3.4 Herleitung der Fouriertransformation

Die Fouriertransformation (FT) ist anwendbar auf nichtperiodische, unendlich ausgedehnte, kontinuierliche Signale. Wir leiten die Fouriertransformation aus der Fourierreihenentwicklung her, indem wir für die Grundperiode den Grenzübergang $P \rightarrow \infty$ betrachten. Dazu gehen wir von Beispiel 2 (Fourierreihenentwicklung) aus.

Dort galt für die Fourierkoeffizienten

$$\gamma_0 = \frac{2a}{P}, \quad \gamma_u = \frac{2 \sin u\omega_0 a}{u\omega_0 P}$$

Wenn P wächst, nimmt $|\gamma_u|$ ab und kann u.U. nicht mehr den Frequenzgehalt eines aperiodischen Signals wiedergeben. Stattdessen wird die Größe $P\gamma_u = \frac{2 \sin u\omega_0 a}{u\omega_0}$ betrachtet, die unabhängig von P ist. Diese Unabhängigkeit ist in der folgenden Abbildung für die Fälle $P = P_1$, $P = 2P_1$ und $P = \infty$ zu erkennen.



Wenn P wächst, nimmt die Grundfrequenz ab wegen $\omega_0 = \frac{2\pi}{P}$. Deshalb nimmt die Dichte der Frequenzen $\omega = u\omega_0$ zu, um schließlich zu einem kontinuierlichen Frequenzspektrum überzugehen.

Sei die Fourierreihenentwicklung gegeben durch

$$f(x) = \frac{1}{P} \sum_{u=-\infty}^{\infty} \gamma_u P e^{ju\omega_0 x}$$

wobei der Vorfaktor die Skalierung der Fourierkoeffizienten

$$\gamma_u = \frac{1}{P} \int_{-P/2}^{P/2} f(x) e^{-ju\omega_0 x} dx$$

aufhebt.

Mit $\omega = u\omega_0$, $P = \frac{2\pi}{\omega_0}$ und $F(\omega) = P\gamma_u$ folgt

$$f(x) = \frac{1}{P} \sum_{u=-\infty}^{\infty} \gamma_u P e^{ju\omega_0 x} = \frac{1}{2\pi} \sum_{u=-\infty}^{\infty} F(u\omega_0) e^{ju\omega_0 x} \omega_0$$

Im Grenzfall $P \rightarrow \infty$ gilt $\omega_0 \rightarrow 0$, also wird ω ein Element des Kontinuums und ω_0 wird

ersetzt durch das Differenzial $d\omega$. Hieraus folgt das Paar der *Fouriertransformation*

$$F(\omega) = \int_{-\infty}^{\infty} f(x)e^{-j\omega x} dx \quad \text{für alle } \omega$$



$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{j\omega x} d\omega \quad \text{für alle } x.$$

Die symbolische Notation der Fouriertransformation als linearer Operator lautet

$$F(\omega) = \mathcal{F}\{f(x)\}$$

bzw.

$$f(x) = \mathcal{F}^{-1}\{F(\omega)\}$$

$F(\omega)$ heißt das *Frequenzspektrum* von $f(x)$.

In Abschnitt 2.2 führten wir die Fouriertransformation als $F(u)$ ein. Dies ist gleichwertig zu der hier abgeleiteten Schreibweise. Lediglich die Normierung muß beachtet werden.

2.3.5 Eigenschaften der Fouriertransformation

Ebenso wie bei der Fourierentwicklung gilt auch hier, daß eine nicht periodische Funktion nur dann eine Fouriertransformierte hat, wenn sie absolut integrierbar ist oder allgemeiner, wenn die Dirichlet-Bedingungen erfüllt sind.

Beispiele:

1) $f(x) = e^{-2x}$, $-\infty \leq x \leq \infty$, ist nicht absolut integrierbar, da für $x \rightarrow -\infty$ die Fläche unter der Kurve unendlich wird.

2) $f(x) = \begin{cases} e^{-2x} & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$ ist absolut integrierbar:

$$\int_{-\infty}^{\infty} |f(x)| dx = \int_0^{\infty} e^{-2x} dx = \frac{1}{-2} e^{-2x} \Big|_0^{\infty} = -\frac{1}{2} [0 - 1] = 0.5$$

3) $f(x) = e^{-2|x|}$ ist absolut integrierbar für alle x .



Die meisten in der Praxis relevanten Funktionen sind absolut integrierbar, weil sie keine unendlich großen Werte annehmen, und sind von endlicher Ausdehnung. Sie sind deshalb transformierbar.

Im Fall nichtstetiger Funktionen $f(x)$ mit einer Unstetigkeitsstelle bei $x = x_0$ liefert die inverse Fouriertransformation das Mittel von $f(x_{0+})$ und $f(x_{0-})$,

$$\frac{f(x_{0+}) + f(x_{0-})}{2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega x_0} d\omega,$$

wobei $f(x_{0+})$ und $f(x_{0-})$ die rechts- bzw. linksseitigen Funktionswerte an der Stelle x_0 sind.

Eine besondere Rolle spielt die *Dirac-Deltafunktion* $\delta(x - x_0)$ im Kontinuierlichen, da für $x \rightarrow x_0$ folgt $\delta(x - x_0) \rightarrow \infty$, obwohl

$$\int_{-\infty}^{\infty} \delta(x - x_0) dx = 1.$$

Obwohl das Fouriertransformationspaar $f(x) \circ \text{---} \bullet F(\omega)$ hierfür keine Bedeutung hat, kann die Ausblendeigenschaft

$$\int_a^b \delta(x - x_0) f(x) dx = \begin{cases} f(x_0) & \text{für } x_0 \in [a, b] \\ 0 & \text{für } x_0 \notin [a, b] \end{cases}$$

formal angewendet werden. Dies liefert mit

$$\int_{-\infty}^{\infty} \delta(x - x_0) e^{-j\omega x} dx = e^{-j\omega x_0}$$

die komplexe Funktion $e^{-j\omega x}$ an der Stelle $x = x_0$. Deshalb wird auch die Dirac-Deltafunktion als Fourier-transformierbar angesehen:

$$F(\omega) = \mathcal{F} \{ \delta(x - x_0) \} = e^{-j\omega x_0}.$$

Allerdings läßt sich die Dirac-Deltafunktion nicht durch inverse Fouriertransformation aus der Spektralfunktion $F(\omega) = e^{-j\omega x_0}$ zurückgewinnen. Man kann aber zu den Spektralfunktionen auch die Funktion $F(\omega) = \delta(\omega - \omega_0)$ für beliebiges aber fest gewähltes ω_0 hinzunehmen. Dann erhält man durch formale Anwendung der Ausblendeigenschaft die komplexe harmonische Funktion

$$f(x) = \mathcal{F}^{-1} \{ \delta(\omega - \omega_0) \} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega x} \delta(\omega - \omega_0) d\omega = \frac{1}{2\pi} e^{j\omega_0 x},$$

die allerdings nicht absolut integrierbar ist, weil sie im Intervall $[-\infty, \infty]$ unendlich viele Extrema besitzt.

Die inverse Fouriertransformation von $F(\omega) = \delta(\omega - \omega_0)$ und speziell für $\omega_0 = 0$ liefert $f(x) = 1$, denn

$$\mathcal{F} \{ f(x) = 1 | x \in (-\infty, \infty) \} = 2\pi \delta(\omega).$$

Dies ist ein Impuls bei $\omega = 0$ mit dem Gewicht 2π .

Interpretation: $f(x) = 1$ ist ein konstantes Signal, das keine Frequenzen ungleich Null enthält.

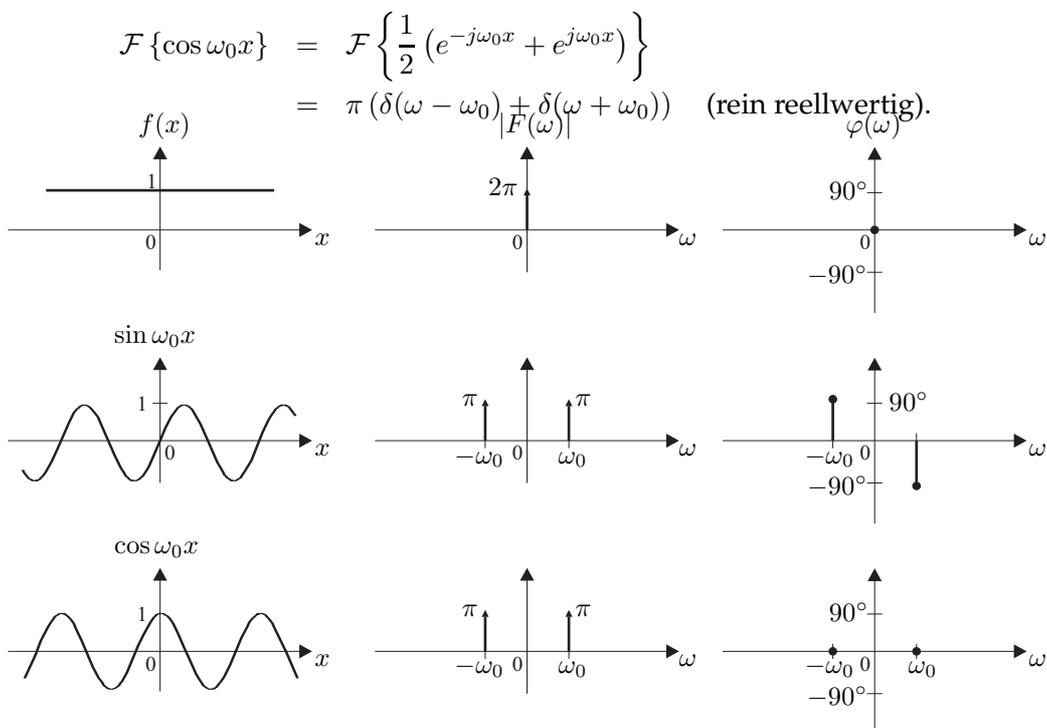
Absolute Integrierbarkeit ist eine hinreichende Bedingung der Fourier-Transformierbarkeit, aber keine notwendige. Wir zeigen nun, daß für periodische Funktionen, die nicht absolut integrierbar sind (wie die eben angegebene komplexe Harmonische), die Fouriertransformation mittels Dirac-Impulsen angegeben werden kann.

Deshalb stellen wir fest: Ist $f(x)$ absolut integrierbar **oder** periodisch, so ist die Funktion Fouriertransformierbar.

Beispielsweise erhalten wir wegen der Linearität der Fouriertransformation

$$\begin{aligned} \mathcal{F}\{\sin \omega_0 x\} &= \mathcal{F}\left\{\frac{1}{2j}(e^{j\omega_0 x} - e^{-j\omega_0 x})\right\} \\ &= \frac{\pi}{j}(\delta(\omega - \omega_0) - \delta(\omega + \omega_0)) \\ &= -\pi j\delta(\omega - \omega_0) + \pi j\delta(\omega + \omega_0) \quad (\text{rein imaginärwertig}). \end{aligned}$$

und



Eine allgemeine periodische, kontinuierliche Funktion $f(x)$ mit der Periode P und der Grundfrequenz $\omega_0 = \frac{2\pi}{P}$ besitzt die Fourierreihendarstellung

$$f(x) = \sum_{u=-\infty}^{\infty} \gamma_u e^{ju\omega_0 x}$$

Die entsprechende Fouriertransformation von $f(x)$ ist

$$F(\omega) = \mathcal{F}\{f(x)\} = \sum_{u=-\infty}^{\infty} 2\pi\gamma_u \delta(\omega - u\omega_0).$$

Das Fourierspektrum besteht also aus einer Folge von Impulsen mit dem Gewicht $2\pi\gamma_u$ bei den Frequenzen $\omega = u\omega_0$.

Mit Hilfe der Fouriertransformation von Dirac-Impulsen ist es also möglich, auch für periodische Funktionen, die wegen der unendlichen Anzahl von Extremwerten die Dirichletbedingung nicht erfüllen, die Fouriertransformation zu realisieren.

Wir bemerken:

Periodizität einer kontinuierlichen Funktion bewirkt ein diskretes Frequenzspektrum.

Im Extremfall kann die periodische (kontinuierliche) Funktion auch in Gestalt einer gewichteten Folge von Dirac-Impulsen auftreten. Damit schlägt die Verwendung des Dirac-Impulses eine Brücke zwischen kontinuierlichen und diskreten Funktionen. Es sei an Abschnitt 1.2.2 erinnert, wo diskrete Funktionen als gewichtete Folge von Dirac-Impulsen eingeführt werden.

Beispiel 1: Impulsfolge (Abtastfunktion)

Impulsfolge mit Periode $P = x_0$

$$f(x) = \sum_{k=-\infty}^{\infty} \delta(x - kx_0)$$

Fourierreihe mit $\omega_0 = \frac{2\pi}{x_0}$ (Grundfrequenz ist invers proportional zum Impulsabstand im Ortsraum)

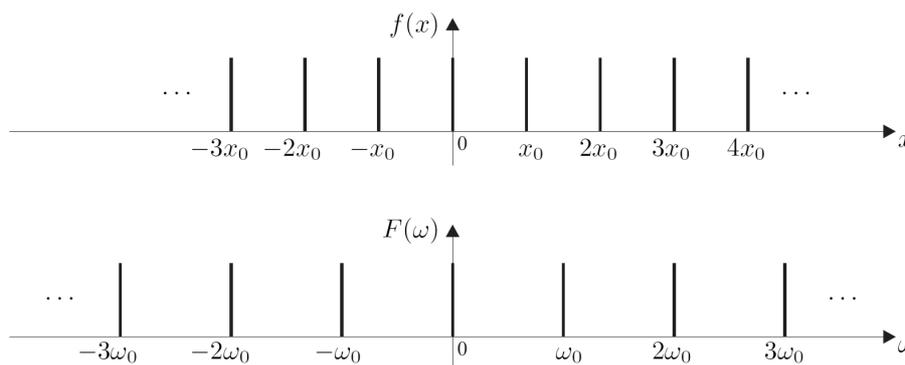
$$f(x) = \sum_{u=-\infty}^{\infty} \gamma_u e^{ju\omega_0 x} = \sum_{u=-\infty}^{\infty} \frac{1}{x_0} e^{ju\omega_0 x}$$

mit den (frequenzunabhängigen) Fourierkoeffizienten wegen Integration über P

$$\gamma_u = \frac{1}{x_0} \int_{-\frac{x_0}{2}}^{\frac{x_0}{2}} \delta(x) e^{-ju\omega_0 x} dx = \frac{1}{x_0} e^{-ju\omega_0 x} \Big|_{x=0} = \frac{1}{x_0}$$

Fouriertransformation als Folge von Impulsen

$$F(\omega) = \sum_{u=-\infty}^{\infty} \frac{2\pi}{x_0} \delta(\omega - u\omega_0)$$



Beispiel 2: Signumfunktion

$$\text{sgn}(x) = \begin{cases} 1 & \text{wenn } x > 0 \\ 0 & \text{wenn } x = 0 \\ -1 & \text{wenn } x < 0 \end{cases}$$

Ist nicht absolut integrierbar im Intervall $(-\infty, \infty)$, aber die Approximation $f(x) = e^{-a|x|}\text{sgn}(x)$ mit a kleine positive Konstante ist absolut integrierbar.

$$\begin{aligned}\mathcal{F}\{\text{sgn}(x)\} &= \lim_{a \rightarrow 0} \left(- \int_{-\infty}^0 e^{ax} e^{-j\omega x} dx + \int_0^{\infty} e^{-ax} e^{-j\omega x} dx \right) \\ &= \lim_{a \rightarrow 0} \left(\frac{-1}{a - j\omega} + \frac{1}{a + j\omega} \right)\end{aligned}$$

$$\omega \neq 0 : \quad \mathcal{F}\{\text{sgn}(x)\} = \frac{2}{j\omega}$$

$$\omega = 0 : \quad \lim_{a \rightarrow 0} \left(\frac{-1}{a} + \frac{1}{a} \right) = 0$$

$$\mathcal{F}\{\text{sgn}(x)\} = \begin{cases} \frac{2}{j\omega} & \text{wenn } \omega \neq 0 \\ 0 & \text{wenn } \omega = 0 \end{cases}$$

Die Signumfunktion ist von ungerader Symmetrie. Deshalb hat sie ein rein imaginäres Spektrum. Die Signumfunktion werden wir in Abschnitt 3.8 benötigen. ■

Beispiel 3: Sinc-Funktion und Rechteckfunktion

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$$

mit $\text{sinc}(0) = 1$. Die sinc-Funktion hat Nullstellen bei $x = \pm 1, \pm 2, \dots$ und für $x \rightarrow \pm\infty$ folgt $\text{sinc}(x) \rightarrow 0$.

Die sinc-Funktion ist von Bedeutung beim Filterentwurf (siehe Abschnitt 3.2.3).

Rechteckfunktion f_a und sinc-Funktion bilden ein Fouriertransformations-Paar.

$$f_a(x) = \begin{cases} 1 & \text{für } |x| \leq a \\ 0 & \text{sonst} \end{cases}$$

$$\begin{aligned}F_a(\omega) &= \int_{-\infty}^{\infty} f_a(x) e^{-j\omega x} dx = \int_{-a}^a e^{-j\omega x} dx = \frac{-1}{j\omega} (e^{-j\omega a} - e^{j\omega a}) \\ &= \frac{2 \sin(a\omega)}{\omega} \\ &= \frac{2a \sin\left(\pi \frac{a\omega}{\pi}\right)}{\pi \cdot \frac{a\omega}{\pi}} = 2a \text{sinc}\left(\frac{a\omega}{\pi}\right)\end{aligned}$$

Die sinc-Funktion und die Rechteckfunktion sind von gerader Symmetrie. Deshalb haben sie ein reellwertiges Spektrum. ■

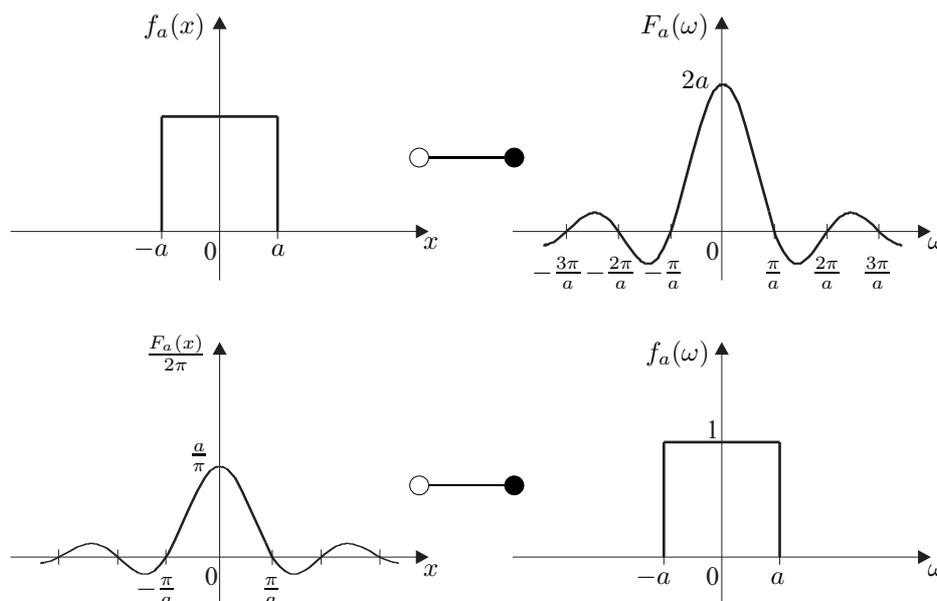
Wegen der **Dualitätseigenschaft** der Fouriertransformation gilt:

$$\text{Wenn } F(\omega) = \mathcal{F}\{f(x)\}, \text{ dann } f(-\omega) = \frac{1}{2\pi} \mathcal{F}\{F(x)\}$$

Hieraus folgt

$$\frac{1}{2\pi} \mathcal{F}\{F_a(x)\} = \mathcal{F}\left\{\frac{a}{\pi} \operatorname{sinc}\left(\frac{ax}{\pi}\right)\right\} = f_a(-\omega) = f_a(\omega)$$

Sinc-Funktion und seine inverse Fouriertransformation:



2.4 Fouriertransformationen diskreter Signale

Wir wollen in diesem Abschnitt diskrete Signale $f = [f_m]$ einer Fouriertransformation unterziehen. Dabei sind, wie in 2.2 zusammengefaßt, wieder unterschiedliche Annahmen für den Definitionsbereich dieser Signale zu machen. Hieraus resultieren verschiedene Formen der Fouriertransformation diskreter Signale:

- unendlich ausgedehnt \rightarrow Fouriertransformation diskreter Signale (FTD)
- periodisch mit endlicher Periode \rightarrow Fourierreihenentwicklung diskreter Signale (DFR)
- endlich ausgedehnt \rightarrow Diskrete Fouriertransformation (DFT).

Wie Abschnitt 2.2 verdeutlicht, haben diskrete Originalfunktionen auch andere Bildfunktionen zur Folge. Vieles, was wir in Abschnitt 2.3 kennenlernten ist natürlich einfach übertragbar.

Sei x_0 der Abtastabstand der diskreten Funktion f , so daß für alle x gilt $x = mx_0$. Ist f periodisch, dann

$$f_m = f_{m+M} = f_{m+2M} = \dots \quad \text{für } m \in \mathbb{Z}.$$

Die kleinste Periode, M , heißt wieder Grundperiode.

Während aber für die kontinuierliche komplexe Harmonische $\phi_u(x) = e^{j\omega x}$ gilt $\omega = u\omega_0 \in (-\infty, \infty)$, ist der Frequenzbereich der *diskreten komplexen Harmonischen* $\phi_{m,u} = e^{j\omega m x_0}$ begrenzt,

$$-\frac{\pi}{x_0} < \omega \leq \frac{\pi}{x_0}.$$

Diese, für die Eindeutigkeit der Repräsentation wichtige Einschränkung stellt den wichtigsten Unterschied zwischen den Fällen kontinuierlich/diskret dar.

Die Einschränkung folgt aus der Forderung, daß für eine periodische komplexe Folge $e^{j\omega_0 m x_0}$ mit der Grundfrequenz $\omega_0 = \frac{2\pi}{M}$ gelten muß, daß $\frac{\omega_0 x_0}{\pi}$ eine rationale Zahl ist. Wir zeigen dies am Beispiel der Sinus-Funktion. Für die Kosinus-Funktion gilt entsprechendes. Offensichtlich ist die kontinuierliche Funktion $\sin \omega_0 x$ periodisch für jedes ω_0 . Aber die abgetastete Funktion $\sin \omega_0 m x_0$ ist nicht periodisch für jedes ω_0 und jedes x_0 . Wenn die diskrete Folge $\sin \omega_0 m x_0$ periodisch ist, dann existiert eine ganze Zahl N , so daß

$$\begin{aligned} \sin \omega_0 m x_0 &= \sin \omega_0 (m + N) x_0 \\ &= \sin(\omega_0 m x_0 + \omega_0 N x_0) \\ &= \sin \omega_0 m x_0 \cos \omega_0 N x_0 + \cos \omega_0 m x_0 \sin \omega_0 N x_0 \end{aligned}$$

für alle m . Dies gilt nur, wenn

$$\cos \omega_0 N x_0 = 1 \quad \text{und} \quad \sin \omega_0 N x_0 = 0,$$

also wenn für die Argumente dieser Funktionen gilt

$$\omega_0 N x_0 = 2\pi n \quad \text{oder} \quad \frac{\omega_0 x_0}{\pi} = \frac{2n}{N} \quad \text{für } n \in \mathbb{Z}.$$

Also ist $\sin \omega_0 m x_0$ periodisch nur dann, wenn eine ganze Zahl n existiert, so daß $N = \frac{2n\pi}{\omega_0 x_0}$ eine ganze Zahl ist bzw., wenn $\frac{\omega_0 x_0}{\pi}$ eine rationale Zahl ist. Die kleinste Zahl N entspricht der Grundperiode von $\sin \omega_0 m x_0$.

Beispiele:

1. Für die diskrete Folge $\sin 2k$, $k \in \mathbb{Z}$ gilt $\omega_0 x_0 = 2$. Da $\frac{\omega_0 x_0}{\pi} = \frac{2}{\pi}$ keine rationale Zahl ist, ist die Folge nicht periodisch.
2. Die Folge $\sin 0.1\pi k$ ist periodisch, da $\frac{\omega_0 x_0}{\pi} = 0.1$ eine rationale Zahl ist.

■

Wie erklärt man die Frequenz einer diskreten Folge wie $\sin \omega_0 m x_0$? Zur Erinnerung: Im Fall einer kontinuierlichen Funktion $\sin \omega_0 x$ ist $\omega_0 = \frac{2\pi}{P}$ mit P als Grundperiode die Frequenz. Dies kann nicht direkt auf eine diskrete Folge übertragen werden. Vielmehr gilt:

Definition (Frequenz einer diskreten Folge):

Die Frequenz einer diskreten Folge $\sin \omega_0 m x_0$, egal ob periodisch oder nicht periodisch, ist definiert als die Frequenz der kontinuierlichen Funktion $\sin \omega x$ mit

$$-\pi < \omega x_0 \leq \pi \quad \text{bzw.} \quad -\frac{\pi}{x_0} < \omega \leq \frac{\pi}{x_0},$$

die nach Abtastung mit der Schrittweite x_0 (entspricht $\sin \omega x|_{x=m x_0} = \sin \omega_0 m x_0$) die diskrete Folge $\sin \omega_0 m x_0$ ergibt.



Man bezeichnet die Funktion $\sin \omega x$ mit der nach dieser Definition begrenzten Frequenz als die primäre einhüllende Funktion der Folge $\sin \omega_0 m x_0$.

2.4.1 Komplexe harmonische Folgen

Von nun an sei angenommen, daß $x_0 = 1$ gilt. Folglich gilt für $\omega = u\omega_0$

$$-\pi < \omega \leq \pi \quad \text{bzw.} \quad -\frac{1}{2} < u \leq \frac{1}{2}.$$

Die komplexe exponentielle Folge $e^{j\omega_0 m}$ mit

$$\omega_0 = \frac{2\pi}{M}$$

ist periodisch mit der Grundperiode M :

$$e^{\frac{j2\pi m}{M}} = e^{\frac{j2\pi}{M}(m+M)} = e^{\frac{j2\pi m}{M}} e^{j2\pi}$$

Die komplexe harmonische Folge

$$\phi_{m,u} = e^{ju\left(\frac{2\pi}{M}\right)m} = e^{ju\omega_0 m}, \quad u \in \mathbb{Z}$$

besteht aus unendlich vielen komplexen Exponentialfunktionen, von denen nur M verschieden sein können. Hieraus folgt, daß alle diese Funktionen periodisch sind mit der Periode M :

$$\begin{aligned} \phi_{m+M,u} &= e^{ju\left(\frac{2\pi}{M}\right)(m+M)} = e^{ju\left(\frac{2\pi}{M}\right)m} e^{ju2\pi} \\ &= e^{ju\left(\frac{2\pi}{M}\right)m} = \phi_{m,u} \end{aligned}$$

und

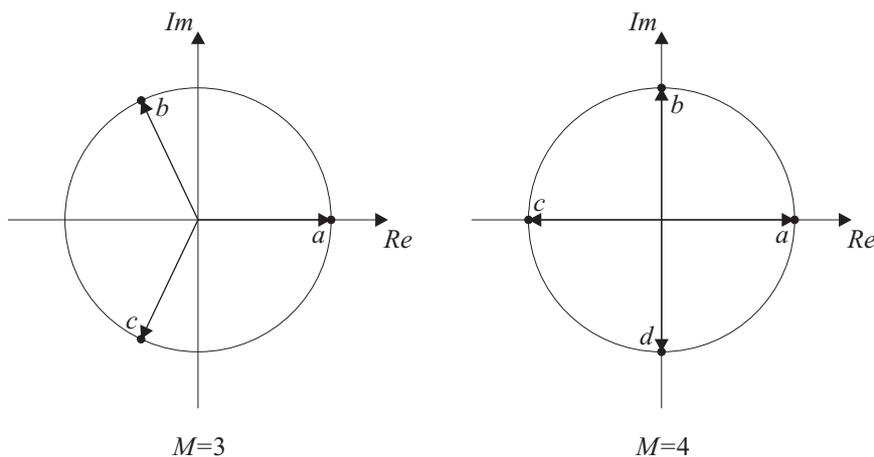
$$\phi_{m,u+M} = e^{j(u+M)\left(\frac{2\pi}{M}\right)m} = e^{ju\left(\frac{2\pi}{M}\right)m} = \phi_{m,u}.$$

Für gerades M und ungerades M ergeben sich unterschiedliche Schreibweisen der diskreten (positiven und negativen) Frequenzen u , wenn der Frequenzraum positive und negative Frequenzen überspannt. Deshalb wird vereinbart

$$u = 0, 1, \dots, M - 1.$$

Die diskreten komplexen Harmonischen können wieder auf dem Einheitskreis visualisiert werden, da $|\phi_{m,u}| = 1$ für alle m und alle u .

Im Unterschied zum kontinuierlichen Fall kommen aber nur M verschiedene Positionen vor.



Beispiel: $M = 3$ für $m = \dots, -3, -2, -1, 0, 1, 2, 3, 4, 5, \dots$

$$\phi_{m,0} = e^{j0\left(\frac{2\pi}{M}\right)m} : \{\dots, a, \dots\}$$

$$\phi_{m,1} = e^{j1\left(\frac{2\pi}{M}\right)m} : \{\dots, a, b, c, a, b, c, a, b, c, \dots\}$$

$$\phi_{m,2} = e^{j2\left(\frac{2\pi}{M}\right)m} : \{\dots, a, c, b, a, c, b, a, c, b, \dots\}$$

In Klammern stehen die Positionen der Projektionen der Harmonischen auf den Einheitskreis. Wenn $u = 1$ und m alle möglichen Positionen annimmt, erhält man

$$m = 0 : a = e^{j0\frac{2\pi}{3}} = 1$$

$$m = 1 : b = e^{j1\frac{2\pi}{3}} = e^{j\frac{2\pi}{3}} = \cos \frac{2\pi}{3} + j \sin \frac{2\pi}{3} = -0,5 + j0.866$$

$$m = 2 : c = e^{j2\frac{2\pi}{3}} = e^{j\frac{4\pi}{3}} = -0.5 - j0.866$$

Offensichtlich gilt:

$\phi_{m,0}$ hat Periode 1 und Frequenz 0

$\phi_{m,1}$ hat Periode 3 und Frequenz $\omega = \frac{2\pi}{3}$

$\phi_{m,2} = \phi_{m,-1}$ hat Periode 3 und Frequenz $\omega = -\frac{2\pi}{3}$

Es gilt

$$\sum_{m=0}^{M-1} \phi_{m,u} = \sum_{m=0}^{M-1} e^{ju\left(\frac{2\pi}{M}\right)m} = \begin{cases} M & \text{für } u = 0 \text{ (modulo } M) \\ 0 & \text{sonst} \end{cases}$$

Bemerkung: $u = 0$ (modulo M) bedeutet $u = 0, \pm M, \pm 2M, \dots$ Wir zeigen, dass für $u \neq 0$ die Summe Null ergibt. Wegen $\sum_{k=0}^{K-1} r^k = 1 + r + r^2 \dots + r^{K-1} = \frac{1-r^K}{1-r}$ folgt

$$\sum_{m=0}^{M-1} \phi_{m,u} = \frac{1 - e^{ju\left(\frac{2\pi}{M}\right)M}}{1 - e^{ju\left(\frac{2\pi}{M}\right)}}$$

Wenn $u \neq 0$ (modulo M), dann $e^{ju\left(\frac{2\pi}{M}\right)} \neq 1$, $e^{ju\left(\frac{2\pi}{M}\right)M} = 1$ und der obige Bruch wird Null.

Hieraus folgt die **Orthogonalitätsbeziehung**

$$\sum_{m=0}^{M-1} \phi_{m,u} \phi_{m,v}^* = \sum_{m=0}^{M-1} e^{j(u-v)(\frac{2\pi}{M})m} = \begin{cases} M & \text{für } u = v \text{ (modulo } M) \\ 0 & \text{für } u \neq v \text{ (modulo } M) \end{cases}$$

Die diskrete Fourierreihenentwicklung (DFR) für diskrete periodische Funktionen wurde in Abschnitt 2.2 notiert. Ähnlich wie im kontinuierlichen Fall kann hieraus die Fouriertransformation diskreter, unendlich ausgedehnter Funktionen (FTD) abgeleitet werden. Anstelle der absoluten Integrierbarkeit tritt hier der Begriff absolute Summierbarkeit.

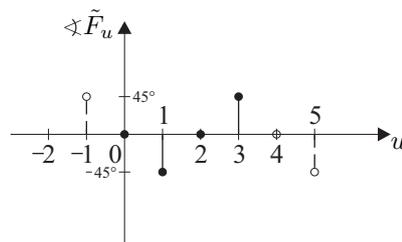
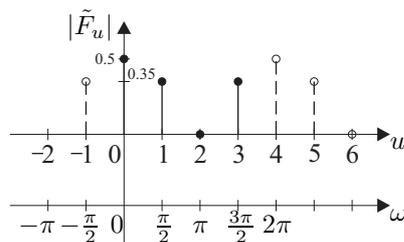
Beispiel: Diskrete Fourierreihenentwicklung

Die Funktion \tilde{f} sei periodisch mit M als Grundperiode. Hier gelte $M = 4$.

$$\begin{aligned} \tilde{f}_0 &= \tilde{f}_1 = 1, \quad \tilde{f}_2 = \tilde{f}_3 = 0 \\ \tilde{F}_0 &= \frac{1}{4} \sum_{m=0}^3 \tilde{f}_m e^{-j \cdot 0 \cdot \frac{2\pi}{4} m} = \frac{1}{4} (1 \cdot 1 + 1 \cdot 1 + 0 \cdot 1 + 0 \cdot 1) \\ &= 0.5 = 0.5 e^{j0} \\ \tilde{F}_1 &= \frac{1}{4} \sum_{m=0}^3 \tilde{f}_m e^{-j \frac{2\pi}{4} m} = \frac{1}{4} (1 \cdot 1 + 1 \cdot e^{-j \frac{\pi}{2}} + 0 \cdot e^{-j \pi} + 0 \cdot e^{-j \frac{3\pi}{2}}) \\ &= \frac{1}{4} (1 - j) = 0.35 e^{-j \frac{\pi}{4}} \\ \tilde{F}_2 &= \frac{1}{4} \sum_{m=0}^3 \tilde{f}_m e^{-j \pi m} = \frac{1}{4} (1 \cdot 1 + 1 \cdot (-1) + 0 \cdot 1 + 0 \cdot (-1)) \\ &= \frac{1}{4} (1 - 1) = 0 \\ \tilde{F}_3 &= \frac{1}{4} \sum_{m=0}^3 \tilde{f}_m e^{-j \frac{3\pi}{2} m} = \frac{1}{4} (1 \cdot 1 + 1 \cdot j) \\ &= \frac{1}{4} (1 + j) = 0.35 e^{j \frac{\pi}{4}} \end{aligned}$$

Offensichtlich ist das Spektrum diskret und periodisch. Hieraus erhält man die diskrete Fourierreihenentwicklung der Funktion:

$$\begin{aligned} \tilde{f}_m &= \sum_{u=0}^3 \tilde{F}_u e^{ju \frac{2\pi}{4} m} = 0.5 + 0.25(1 - j) e^{j \frac{\pi}{2} m} + 0 e^{j \pi m} + 0.25(1 + j) e^{j \frac{3\pi}{2} m} \\ &= 0.5 + 0.35 e^{j(\frac{\pi}{2} m - \frac{\pi}{4})} + 0.35 e^{j(\frac{3\pi}{2} m + \frac{\pi}{4})} \end{aligned}$$





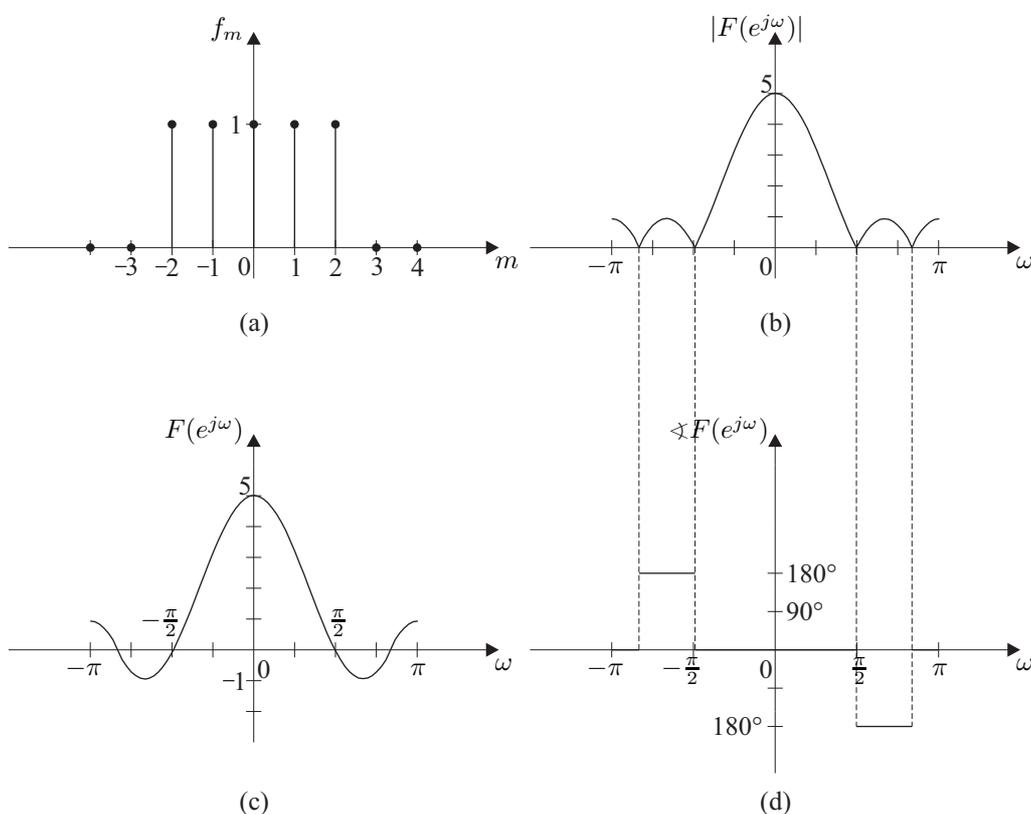
Beispiel: Fouriertransformation diskreter Signale

Es sei f eine unendlich ausgedehnte Funktion, die aber nur endlich viele Werte ungleich Null habe.

Beachte hierbei die Periodizität des Spektrums.

$$f_{-2} = f_{-1} = f_0 = f_1 = f_2 = 1$$

$$\begin{aligned} \mathcal{F}_d \{f_m\} = F(e^{j\omega}) &= \sum_{m=-2}^2 e^{-j\omega m} = e^{j2\omega} + e^{j\omega} + 1 + e^{-j\omega} + e^{-j2\omega} \\ &= 1 + 2 \cos \omega + 2 \cos 2\omega \end{aligned}$$



Da f symmetrisch um den Ursprung ist, ist das Spektrum eine kontinuierliche, gerade, reellwertige und periodische Funktion.



Beziehung zwischen FT und FTD:

Die diskrete Folge $f_m, m = 0, \pm 1, \pm 2, \dots$, wird für $x_0 = 1$ mittels

$$f_s(x) = \sum_{m=-\infty}^{\infty} f_m \delta(x - m) \quad \text{für alle } x \in \mathbb{R}$$

in eine kontinuierliche Repräsentation überführt. Hieraus folgt

$$\mathcal{F}\{f_s(x)\} = \sum_{m=-\infty}^{\infty} f_m \mathcal{F}\{\delta(x-m)\} \quad \text{für alle } x \in \mathbb{R}$$

wegen der Linearität der Fouriertransformation. Wegen der Ausblendeigenschaft des Dirac-Impulses folgt

$$\mathcal{F}\{\delta(x-m)\} = \int_{-\infty}^{\infty} \delta(x-m) e^{-j\omega x} dx = e^{-j\omega x}|_{x=m} = e^{-j\omega m}$$

Die Fouriertransformation von $f_s(x)$ lautet also

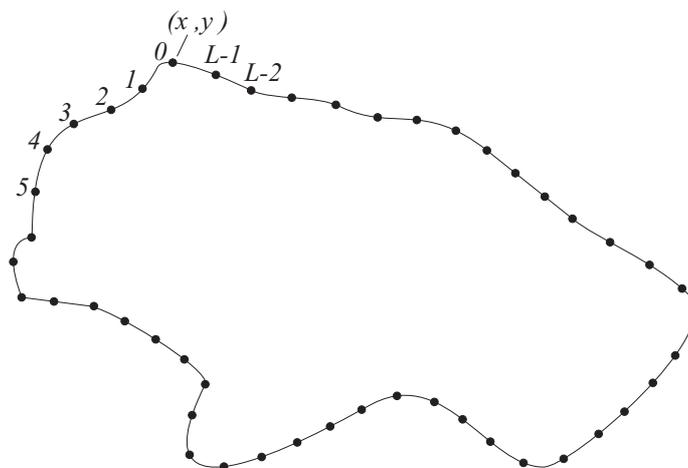
$$\mathcal{F}\{f_s(x)\} = \sum_{m=-\infty}^{\infty} f_m e^{-j\omega m}$$

und ist identisch mit der FTD:

$$\mathcal{F}\{f_s(x)\} = \mathcal{F}_d\{f_m\}.$$

2.4.2 Fourierdeskriptoren

Die Gestaltsbeschreibung von Objekten ist eine wichtige Problemstellung der Bildanalyse. Die Gestalt wird durch den Objektrand wiedergegeben. Dieser Objektrand wird in der Bildverarbeitung durch eine diskrete geschlossene Kurve repräsentiert, die man Objektkontur nennt. Da die benachbarten Konturpunkte in Achsenrichtung und in Richtung der Diagonalen unterschiedliche Abstände aufweisen (in 8-Nachbarschaft), ist eine genaue Bestimmung der Gestaltsparameter problematisch. In der 4-Nachbarschaft treten ähnliche Störungen auf. Wir wollen von diesen Problemen absehen und betrachten eine diskrete Randkurve der Länge L eines Objektes in Parameterdarstellung, $z_l = (x_l, y_l)$ mit $0 \leq l < L$.



Offensichtlich ist die Randkurve periodisch mit dem Umfang L :

$$z_{l+nL} = z_l \quad , \quad n \in \mathbb{Z}$$

Diese periodische Kurve wird in eine Fourierreihe entwickelt, wobei z_l als komplexwertige eindimensionale Funktion modelliert wird,

$$z_l = x_l + jy_l.$$

Die Fourierkoeffizienten

$$c_u = \frac{1}{L} \sum_{l=0}^{L-1} z_l e^{-\frac{2\pi j u l}{L}}, \quad u \in \mathbb{Z}$$

heißen **Fourierdeskriptoren** der Kurve.

Eine Objektkontur ist tatsächlich eine diskrete periodische Funktion. Offensichtlich ist die diskrete Fourierreihendarstellung (DFR) die adäquate Repräsentation der Fouriertransformation.

Aus den Fourierdeskriptoren kann die Kurve selbst wieder rekonstruiert werden.

$$z_l = \sum_{u=0}^{L-1} c_u e^{\frac{2\pi j u l}{L}} \quad \text{für alle } 0 \leq l < L$$

Geometrische Bedeutung der Fourierdeskriptoren:

Der Fourierdeskriptor

$$c_0 = \frac{1}{L} \sum_{l=0}^{L-1} z_l$$

steht unmittelbar für den Mittelpunkt (Schwerpunkt) der Kontur. Der Fourierdeskriptor

$$c_1 = \frac{1}{L} \sum_{l=0}^{L-1} z_l e^{-\frac{2\pi j l}{L}}$$

kann für eine Approximation der Kontur durch einen Kreis verwandt werden,

$$z_l^1 = c_1 e^{\frac{2\pi j l}{L}} \Rightarrow z_l^1 = r_1 e^{j(\varphi_1 + \frac{2\pi l}{L})}.$$

Die Gleichung beschreibt eine Kreisbewegung eines Punktes entgegen der Uhrzeigerrichtung, die beim Winkel φ_1 (gemessen von der reellen Achse) beginnt. Entsprechend steht der Fourierdeskriptor c_{-1} für einen Kreis mit entgegengesetztem Umlaufsinn (in Uhrzeigerrichtung) aber gleichem Radius,

$$z_l^{-1} = r_{-1} e^{j(\varphi_{-1} - \frac{2\pi l}{L})}, \quad r_{-1} = r_1.$$

Beide Koeffizienten gemeinsam (4 Parameter): Ellipse mit Halbachsen a und b , beliebiger Orientierung θ und beliebigem Startwinkel φ_0 auf der Ellipse. Für $\varphi_1 = \varphi_{-1} = 0$ folgt

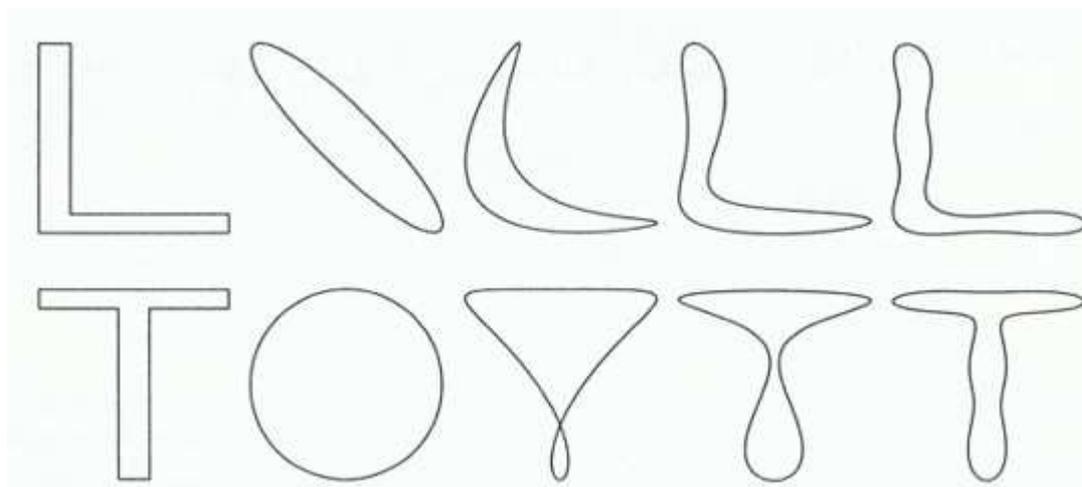
$$z_l^{|1|} = z_l^1 + z_l^{-1} = (r_1 + r_{-1}) \cos\left(\frac{2\pi l}{L}\right) + j(r_1 - r_{-1}) \sin\left(\frac{2\pi l}{L}\right)$$

als die Parameterform der Ellipse, deren Halbachsen parallel zum Koordinatensystem liegen und deren Startpunkt auf der x -Achse liegt.

Paarweise Fourierdeskriptoren höherer Frequenz ergeben ebenfalls Ellipsen, die $|u|$ mal durchlaufen werden. Diese tragen zunehmend mehr Details zur Beschreibung der Form bei.

Beispiel:

Rekonstruktion der Gestalt mit $N = 2, 3, 4, 8$ Deskriptoren:



Invarianzeigenschaften der Fourierdeskriptoren:

- Translationsinvarianz: Alle Fourierdeskriptoren, bis auf c_0 , sind translationsinvariant.
- Skalierungsinvarianz: Skalierung der Kurve mit dem Faktor Δ bewirkt eine Skalierung der Fourierdeskriptoren mit dem Faktor Δ . Normierung der Deskriptoren auf $|c_1|$ führt zu skaleninvarianten Fourierdeskriptoren.
- Rotationsinvarianz: Drehung einer Kurve um den Winkel θ bewirkt Multiplikation der Fourierdeskriptoren mit dem frequenzabhängigen Phasenfaktor $e^{ju\theta}$. Es ist möglich, die Phasen aller Fourierdeskriptoren zu der von c_1 in Beziehung zu setzen. Dazu subtrahiert man von allen Phasenwinkeln φ_u die Phasenverschiebung $u\varphi_1$. Dann repräsentieren alle Deskriptoren die Phasendifferenz zur Phase des ersten Deskriptors und sind somit rotationsinvariant, bis auf c_1 selbst.

Die Invarianzeigenschaften folgen unmittelbar aus den Eigenschaften der Fouriertransformation, die für die diskrete Fouriertransformation in Abschnitt 2.5.2 behandelt werden.

2.4.3 N-te Einheitswurzel und Kreisteilung

Zitat aus H. Pieper: Die komplexen Zahlen, Deutscher Verlag der Wissenschaften, Berlin, 1988

„Am 1. Juni 1796 konnte man in dem damals in Jena erscheinenden *Intelligenzblatt der allgemeinen Literaturzeitung* (Nr. 66) unter der Rubrik „Neue Entdeckungen“ lesen:

„Es ist jedem Anfänger der Geometrie bekannt, daß verschiedene ordentliche Vielecke, namentlich das Dreieck, Fünfeck, Fünfzehneck, und die, welche durch wiederholte Verdoppelung der Seitenzahl eines derselben entstehen, sich geometrisch konstruieren lassen. So weit war man schon zu Euklids Zeit und es scheint, man habe sich seitdem allgemein überredet, daß das Gebiet der Elementargeometrie sich nicht weiter erstrecke: wenigstens kenne ich keinen geglückten Versuch, ihre Grenzen auf dieser Seite zu erweitern. Desto mehr, dünkt mich, verdient die Entdeckung Aufmerksamkeit, daß außer jenen ordentlichen Vielecken noch eine Menge anderer, z. B. das Siebzehneck, einer geometrischen Konstruktion fähig ist. Diese Entdeckung ist eigentlich nur ein Corollarium einer noch nicht ganz vollendeten Theorie von

größerm Umfange, und sie soll, sobald diese ihre Vollendung erhalten hat, dem Publikum vorgelegt werden. C. F. Gauß, a. Braunschweig. Stud. der Mathematik zu Göttingen.

Es verdient angemerkt zu werden, daß Hr. Gauß jetzt in seinem 18ten Jahre steht, und sich hier in Braunschweig mit eben so glücklichem Erfolge der Philosophie und der klassischen Literatur als der höheren Mathematik gewidmet hat.

Den 18. April 96. E. A. W. Zimmermann, Prof.¹⁾

Schon Euklid gab in den "Elementen" die Konstruktionsbeschreibung der regelmäßigen N -Ecke an für $3 \leq N < 300$:

$$3, 4, 5, 6, 8, 10, 12, 15, 16, 20, \dots$$

bzw. für $2^k, 3 \cdot 2^k, 5 \cdot 2^k, 15 \cdot 2^k$ mit $k \in \mathbb{N}$.

Wie aber sollte man ein regelmäßiges N -Eck für $N = 7, 9, 11$ usw. konstruieren?

"Bis zum Ende des 18. Jahrhunderts war man davon überzeugt (ohne es zu beweisen), daß man überhaupt keine weiteren N -Ecke außer die oben angegebenen regelmäßigen N -Ecke, bei denen n das Produkt von 2 oder 3 oder 5 oder 15 mit einer Potenz von 2 ist, mit Zirkel und Lineal konstruieren kann (sondern nur mechanisch mittels Anwendung eines Winkelmessers). Da erschien in Jena die zitierte Notiz des Studenten CARL FRIEDRICH GAUSS, in der er behauptete, daß "noch eine Menge anderer" regelmäßiger Vielecke, z. B. das Siebzehneck, mit Zirkel und Lineal konstruierbar wären. Zwei Jahrtausende war die Kreisteilung auf dem Stand aus der Zeit des EUKLID stehengeblieben. Nun entdeckte der junge GAUSS die erste neue Konstruktion eines regelmäßigen Vielecks seit dem griechischen Altertum: "Der Tag war der 29. März 1796; und der Zufall hatte gar keinen Anteil daran", schrieb er später. Tatsächlich war GAUSS durch vorangehende Überlegungen auf diese bahnbrechende Entdeckung vorbereitet, die nun in bestimmter Weise kommen mußte. Die Grundlagen dieser Entdeckung, die Zurückführung des Problems auf die Lösungen (Wurzeln) der Kreisteilungsgleichung und die "Zerteilung der Wurzeln der Gleichung

$$\frac{x^p - 1}{x - 1} = 0$$

in zwei Gruppen" hatte GAUSS bereits im Winter 1796, während seines ersten Semesters an der Göttinger Universität, gefunden. "Durch angestrenktes Nachdenken über den Zusammenhang aller Wurzeln untereinander nach arithmetischen Gründen glückte es mir bei meinem Ferienaufenthalt in Braunschweig, am Morgen des gedachten Tages (ehe ich aus dem Bette aufgestanden war) diesen Zusammenhang auf das klarste anzuschauen, so daß ich die spezielle Anwendung auf das 17-Eck und die numerische Bestätigung auf der Stelle machen konnte."

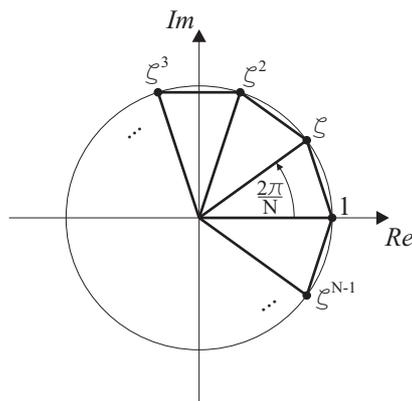
Für den Cosinus des Winkels $\frac{2\pi}{17}$ gab GAUSS übrigens den folgenden Ausdruck an:

$$\begin{aligned} \cos \frac{2\pi}{17} = & -\frac{1}{16} + \frac{1}{16}\sqrt{17} + \frac{1}{16}\sqrt{34 - 2\sqrt{17}} \\ & + \frac{1}{8}\sqrt{17 + 3\sqrt{17} - \sqrt{34 - 2\sqrt{17}} - 2\sqrt{34 + 2\sqrt{17}}}, \end{aligned}$$

woraus man die Konstruierbarkeit dieser Größe folgern kann."

Zunächst wollen wir den auf Gauß zurückgehenden algebraischen Weg zeigen, ein beliebiges N -Eck zu berechnen. Dazu bettete er den Kreis als Einheitskreis in die komplexe Ebene ein. Die komplexen Zahlen werden auf R. Bombelli (Algebra, 1572) zurückgeführt.

¹⁾Diese Ankündigung der Entdeckung der geometrischen Konstruierbarkeit des regelmäßigen 17-Ecks war die erste Veröffentlichung von GAUSS.



Die Ecken des dem Einheitskreis **einbeschriebenen** N -Eckes haben die Koordinaten

$$\begin{aligned}\zeta^n &= e^{j\frac{2\pi}{N}n} = \left(e^{j\frac{2\pi}{N}}\right)^n \quad \text{mit } 0 \leq n < N-1 \\ &= \cos \frac{2\pi}{N}n + j \sin \frac{2\pi}{N}n = \left(\cos \frac{2\pi}{N} + j \sin \frac{2\pi}{N}\right)^n \\ \zeta^N &= \left(\cos \frac{2\pi}{N} + j \sin \frac{2\pi}{N}\right)^N = \cos 2\pi + j \sin 2\pi = 1\end{aligned}$$

Für jede Potenz ζ^n gilt

$$\begin{aligned}(\zeta^n)^N &= \zeta^{nN} = \zeta^{Nn} = 1^n = 1 \quad \text{mit } 0 \leq n \leq N-1 \\ \zeta^{N+n} &= \zeta^n. \quad (\text{Periodizität}) \\ \zeta^{-n} &= \frac{1}{\zeta^n} = \frac{\zeta^N}{\zeta^n} = \zeta^{N-n} \quad (\text{entgegengesetzte Orientierung})\end{aligned}$$

Ist $z = r(\cos \varphi + j \sin \varphi) = re^{j\varphi}$ eine komplexe Zahl mit der Forderung $z^N = 1$, so gilt

$$r^N (\cos N\varphi + j \sin N\varphi) = 1,$$

also $r^N = 1$, d.h. $r = 1$, und $\cos N\varphi = 1$, $\sin N\varphi = 0$. Folglich ist $N\varphi$ ein Vielfaches von 2π , $\varphi = n \cdot \frac{2\pi}{N}$ mit $n \in \mathbb{Z}$; somit ist $z = \zeta^n$.

Damit ist folgender Satz bewiesen:

Es sei $N \geq 1$. Es gibt genau N verschiedene komplexe Zahlen z mit $z^N = 1$, nämlich

$$\begin{aligned}z_0 &= 1, \quad z_1 = \zeta = e^{j\frac{2\pi}{N}}, \quad z_2 = \zeta^2 = \dots, \\ z_n &= \zeta^n = e^{j\frac{2\pi}{N}n}, \dots, \quad z_{N-1} = \zeta^{N-1}\end{aligned}$$

Dies sind die Ecken des dem Einheitskreis einbeschriebenen regelmäßigen N -Ecks, dessen eine Ecke im Punkt $z_0 = 1$ liegt.

Die Zahlen $1, \zeta, \zeta^2, \dots, \zeta^{N-1}$ heißen **N -te Einheitswurzeln**. Sie sind Lösungen der **Kreisteilungsgleichung**

$$z^N - 1 = 0$$

Aus $z^N - 1 = (z - 1)(z^{N-1} + z^{N-2} + \dots + z + 1)$ folgt, daß entweder $z - 1 = 0$ oder $z^{N-1} + z^{N-2} + \dots + z + 1 = 0$. Die Lösungen der letzten Gleichung sind $\zeta, \zeta^2, \dots, \zeta^{N-1}$. Hieraus folgt

$$1 + \zeta + \zeta^2 + \dots + \zeta^{N-1} = 0.$$

„In seinem 1801 in Leipzig erschienenen Zahlentheoriebuch *Disquisitiones arithmeticae* (Arithmetische Untersuchungen) löste GAUSS sein in der Jenaer Zeitung gemachtes Versprechen ein und behandelte ausführlich die Problematik der Konstruktion regelmäßiger N -Ecke mit Zirkel und Lineal. Im Vorwort bemerkte er: *„Die Theorie der Kreisteilung oder die der regelmäßigen Vielecke gehört eigentlich nicht zur Arithmetica, aber ihre Prinzipien können nur der transzendenten Arithmetica entnommen werden: dieses Ergebnis wird den Geometern genau so unerwartet erscheinen wie die neuen Wahrheiten, die daraus hervorgehen, und die sie, hoffe ich, mit Vergnügen sehen werden.“*

Im siebenten Abschnitt des für die Entwicklung der Zahlentheorie grundlegenden Buches schrieb er: *„Da die Teilung des Kreises in 3 und 5 Teile zur Zeit des Euklid bekannt war, gibt es sicher Gründe sich zu wundern, daß man in einer Periode von 2000 Jahren nichts zu den Entdeckungen hinzugefügt hat, und daß alle Geometer mit Gewißheit verkündet haben, daß außer diesen Teilungen und denen welche man daraus ableitet (die Teilungen in 2^k , 15 , $3 \cdot 2^k$, $5 \cdot 2^k$, $15 \cdot 2^k$ Teile) keine anderen mit geometrischer Konstruktion ausgeführt werden können.“* Mit zahlentheoretischen und algebraischen Hilfsmitteln erkannte GAUSS, daß die Konstruierbarkeit des regelmäßigen N -Ecks mit Zirkel und Lineal allein von der zahlentheoretischen Natur der Zahl N abhängt.

Sein Ergebnis über die Kreisteilung, welches er aus seiner Methode gewonnen hatte, die Kreisteilungsgleichung aufzulösen, lautet:

Die Teilung des Kreisbogens in p gleiche Teile (p Primzahl), bzw. die Konstruktion des regelmäßigen p -Ecks, ist mit Zirkel und Lineal ausführbar, wenn p eine Primzahl der Form $p = 2^k + 1$ ist.“

Primzahlen	:	2, 3, 5, 7, 11, 13, 17, 19, 23, 31, 37, ...
Zahlen der Form $2^k + 1$:	3, 5, 9, 17, 33, 65, 129, 257, ...
Primzahlen der Form $2^k + 1$:	3, 5, 17, 257

Nur, wenn k selbst eine Zweierpotenz ist, kann p eine Primzahl sein!

Welche Zahlen $F_m = 2^{2^m} + 1$ sind Primzahlen? Diese Zahlen F_m heißen heute **Fermatsche Primzahlen**. Sie sind bis heute nur bekannt für $m = 0, 1, 2, 3, 4$:

$$F_0 = 3, F_1 = 5, F_2 = 17, F_3 = 257, F_4 = 65537.$$

Nur für Primzahlen p der Form $2^{2^m} + 1$ kann das regelmäßige p -Eck mit Zirkel und Lineal konstruiert werden. Das geometrische Problem der Kreisteilung ist damit zurückgeführt auf das (noch ungelöste) zahlentheoretische Problem, eine Übersicht über alle Fermatischen Primzahlen zu geben.

Für Signale dieser Länge lassen sich zahlentheoretische, sehr schnelle Verallgemeinerungen der Fouriertransformation angeben, die *Fermat-Transformation*.

Für eine weitere Menge von Primzahlen, die Mersenne-Zahlen, existiert auch eine *Mersenne-Transformation*.

2.5 Diskrete Fouriertransformation (DFT)

Die diskrete Fouriertransformation (DFT) geht aus der diskreten Fourierreihenentwicklung (DFR) durch Ausblenden der Grundperiode M hervor, ohne daß sie die der DFR zugrunde liegende Periodizität aufgibt. (Dies bereitet manchmal Schwierigkeiten im Umgang mit der DFT.)

Außerdem werden beim Übergang von der DFR zur DFT die Fourierkoeffizienten ebenso mit M multipliziert, wie dies beim Übergang von der Fourierreihe (FR) nach der Fouriertransformation (FT) mit der Grundperiode P geschah:

$$F_u = M\tilde{F}_u \quad \text{für } 0 \leq u \leq M-1.$$

Die diskrete Fouriertransformation ist also im wesentlichen eine ausgeblendete und skalierte diskrete Fourierreihe.

2.5.1 Herleitung der Diskreten Fouriertransformation

Andererseits ist die diskrete Fouriertransformation auch interpretierbar als eine an M äquidistanten Positionen $\omega_u = \frac{2\pi u}{M}$ abgetastete Fouriertransformation diskreter Signale (FTD). Dies wollen wir nun zeigen.

Sei $f_m, m = 0, 1, \dots, M-1$, eine Folge der Länge M . Seine FTD ist

$$F(e^{j\omega}) = \sum_{m=-\infty}^{\infty} f_m e^{-j\omega m} = \sum_{m=0}^{M-1} f_m e^{-j\omega m}.$$

Diese kontinuierliche Funktion ist periodisch mit der Periode 2π . Die inverse FTD greift auf Frequenzen $\omega \in (-\pi, \pi]$ zurück. Hiervon existieren unendlich viele. Aber nur endlich viele Frequenzen ω sind voneinander verschieden.

Für ein Signal der Länge M seien dies ebenfalls M äquidistante Frequenzen

$$\omega_u = \frac{2\pi u}{M}.$$

Zur Vereinfachung werden diese aber aus dem Intervall $[0, 2\pi)$ gewählt. Ist $F(e^{j\omega})$ mit $\omega \in [0, 2\pi)$ berechnet, so kann $F(e^{j\omega})$ durch periodische Erweiterung leicht auch für $\omega \in (-\pi, \pi]$ angegeben werden.

Sei $F_u = F(e^{j\omega_u})$ mit $\omega_u = \frac{2\pi u}{M}$ für $u = 0, 1, 2, \dots, M-1$. Dann folgt

$$F_u = F(e^{j\omega_u}) = \sum_{m=0}^{M-1} f_m e^{-j\frac{2\pi mu}{M}} \quad \text{für alle } u.$$

Dies ist die **diskrete Fouriertransformation** von f_m .

Für die Berechnung der **inversen diskreten Fouriertransformation** aus der inversen FTD berücksichtigen wir, daß $F(e^{j\omega})$ nur an den Frequenzen $\omega_u = \frac{2\pi u}{M}, u = 0, 1, \dots, M-1$, verfügbar ist. Dies führt zu der Näherung

$$\begin{aligned} f_m &= \frac{1}{2\pi} \int_0^{2\pi} F(e^{j\omega}) e^{j\omega m} d\omega \\ &\approx \frac{1}{2\pi} \left(\sum_{u=0}^{M-1} F(e^{j\omega_u}) e^{j2\pi \frac{um}{M}} \right) \left(\frac{2\pi}{M} \right), \end{aligned}$$

wobei $\omega_0 = \frac{2\pi}{M}$ das Frequenzintervall zwischen den diskreten Frequenzen ω_u ist. Diese Gleichung kann geschrieben werden

$$f_m = \frac{1}{M} \sum_{u=0}^{M-1} F_u e^{j2\pi \frac{um}{M}} \quad \text{für alle } m.$$

Dies ist die inverse DFT.

Obwohl bei dieser Herleitung lediglich eine Approximation der Folge f_m entstand, bilden f_m und F_u echte Transformationspaare. Wir zeigen dies, in dem wir zunächst die für die DFT übliche Notation einführen.

Sei $f_m, m = 0, 1, \dots, M-1$, eine Folge von M reell- oder komplexwertigen Zahlen, wobei implizit angenommen sei $f_m = 0$ für m außerhalb des Intervalls der Breite M . Wir definieren

$$W_M = e^{-j2\pi/M}$$

als System von Basisvektoren $W_M^u = e^{-j2\pi u/M}, u = 0, 1, \dots, M-1$. Die W_M werden also durch eine $M \times M$ -Matrix mit komplexen Einträgen repräsentiert. Die Basisvektoren W_M^u sind selbst auf den Stützstellen $m = 0, 1, \dots, M-1$ unitäre Basisfunktionen (bezogen auf das Intervall der Breite M), d.h.

$$\sum_{m=0}^{M-1} W_M^{mu} (W_M^{mv})^{*T} = M\delta(u-v)$$

Die Matrix W_M ist unitär,

$$W_M^{*T} = W_M^{-1}, \text{ also } \frac{1}{M} W_M W_M^{*T} = I.$$

Dann ist die DFT von f_m gegeben durch

$$F_u = \sum_{m=0}^{M-1} f_m e^{-j2\pi \frac{um}{M}} = \sum_{m=0}^{M-1} f_m W_M^{um} \quad \text{für alle } u \in \{0, 1, \dots, M-1\}.$$

Die inverse DFT lautet

$$f_m = \frac{1}{M} \sum_{u=0}^{M-1} F_u e^{j2\pi \frac{um}{M}} = \frac{1}{M} \sum_{u=0}^{M-1} F_u W_M^{-um} \quad \text{für alle } m \in \{0, 1, \dots, M-1\}.$$

Hieraus folgt durch Ersetzen der F_u

$$\begin{aligned} \mathcal{F}^{-1}\{F_u\} &= \frac{1}{M} \sum_{u=0}^{M-1} \left[\sum_{n=0}^{M-1} f_n e^{-j2\pi \frac{un}{M}} \right] e^{j2\pi \frac{um}{M}} \\ &= \frac{1}{M} \sum_{n=0}^{M-1} f_n \left[\sum_{u=0}^{M-1} e^{j2\pi(m-n)u/M} \right] = f_m \end{aligned}$$

Da $\sum_{u=0}^{M-1} e^{j2\pi(m-n)u/M} = \delta_{m-n}$, folgt durch Anwendung der Ausblendregel der Deltafunktion, daß $\mathcal{F}^{-1}\{F_u\} = f_m$.

Die DFT bildet also ein eigenständiges Fouriertransformationspaar. Ihre Basisfunktionen ähneln den kontinuierlichen komplex harmonischen Funktionen. Sie ergeben sich aus abgetasteten Kosinus- und Sinusfunktionen unterschiedlicher Wellenlänge. Der Basisvektor W_M^0 ist reell

und konstant. Die Projektion des Signals f auf diesen Basisvektor liefert die Summe über alle f_m , $m = 0, 1, \dots, M - 1$. Die W_M^u , $u \neq 0$, besitzen eine 2π -Periodizität. Da

$$(W_M^{mu})^M = e^{-j2\pi mu} = 1 \text{ für jedes } m \text{ und jedes } u$$

folgt

$$W_M^{m(u+vM)} = W_M^{mu} W_M^{mvM} = W_M^{mu}.$$

Folglich ist W_M auch periodisch in M . Hieraus folgt

$$F_u = F_{u \pm M} = F_{u \pm 2M} = \dots$$

Deshalb ist auch F_u periodisch in M . Da aber u tatsächlich einer Kreisfrequenz $\omega = u \frac{2\pi}{M}$ entspricht, ist die Periode von F_u gegeben durch $M\omega_0 = M \frac{2\pi}{M} = 2\pi$ und entspricht damit derjenigen der FTD, $F(e^{j\omega})$:

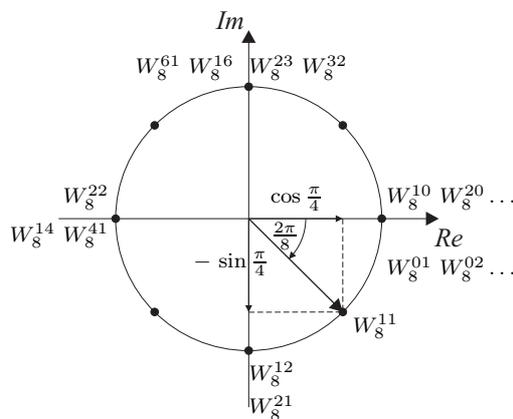
$$F_u = \mathcal{F}\{f_m\} = F(e^{j\omega})|_{\omega=u\frac{2\pi}{M}} = F_u \frac{2\pi}{M}.$$

Hieraus folgt die Interpretation der DFT als abgetastete FTD.

2.5.2 Eigenschaften der Diskreten Fouriertransformation

Abschließend wollen wir den Transformationskern der DFT etwas näher ansehen. Wir betrachten die W_M^{mu} für $M = 8$ und $u = 0, 1, 2, \dots, 7$ auf dem Einheitskreis.

$$\begin{aligned} W_8^{m0} &= e^{-j\frac{2\pi}{8}0m} = e^0 = 1 \\ W_8^{m1} &= e^{-j\frac{2\pi}{8}1m} = e^{-j\frac{\pi}{4}m} \\ W_8^{m2} &= e^{-j\frac{2\pi}{8}2m} = e^{-j\frac{\pi}{2}m} \\ W_8^{m3} &= e^{-j\frac{2\pi}{8}3m} = e^{-j\frac{3}{4}\pi m} \\ W_8^{m4} &= e^{-j\frac{2\pi}{8}4m} = e^{-j\pi m} \\ &\vdots \end{aligned}$$



$$W_8^{11} = \cos \frac{\pi}{4} - j \sin \frac{\pi}{4}$$

Für jede Position m im Ortsraum erzeugen die W_M^{mu} auf dem Einheitskreis in der komplexen Ebene beim Winkel $\varphi_M^{mu} = \frac{2\pi}{M} u m$ eine Projektion. Für aufeinander folgende Positionen m und $m + 1$ ändern sich die Orte dieser Projektionen entsprechend $\delta\varphi_M^{mu} = \frac{2\pi}{M} u$. Für $u = \frac{M}{2}$ werden

mit $\delta\varphi_M^{mu} = \pi$ die größten Sprungweiten erreicht. Für $u > \frac{M}{2}$ kommt es zu einem Phasensprung um π , was einer Umkehrung der Rotation und Verminderung der Frequenz mit wachsendem u entspricht. Die Frequenz $u_N = \frac{M}{2}$ heißt *Nyquist-Frequenz*. Ihre Bedeutung lernen wir bald kennen. Diese Frequenzabhängigkeit der DFT folgt daraus, daß wir anstelle der Periode $(-\pi, \pi]$ das Intervall $[0, 2\pi)$ in der Ableitung der DFT betrachteten.

Die zweidimensionale DFT bildet eine $M \times N$ -Matrix auf eine andere Matrix derselben Größe ab,

$$F_{uv} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_{mn} W_M^{um} W_N^{vn}$$

für alle $u \in \{0, 1, \dots, M-1\}$ and alle $v \in \{0, 1, \dots, N-1\}$.

Ist f komplexwertig, so wird für die Speicherung von f und F derselbe Platz benötigt. Im allgemeinen ist allerdings f reellwertig in der Praxis. Dann belegt natürlich F den doppelten Speicherplatz.

Die zweidimensionale inverse DFT lautet

$$f_{mn} = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F_{uv} W_M^{-um} W_N^{-vn}$$

für alle $m \in \{0, 1, \dots, M-1\}$ und alle $n \in \{0, 1, \dots, N-1\}$.

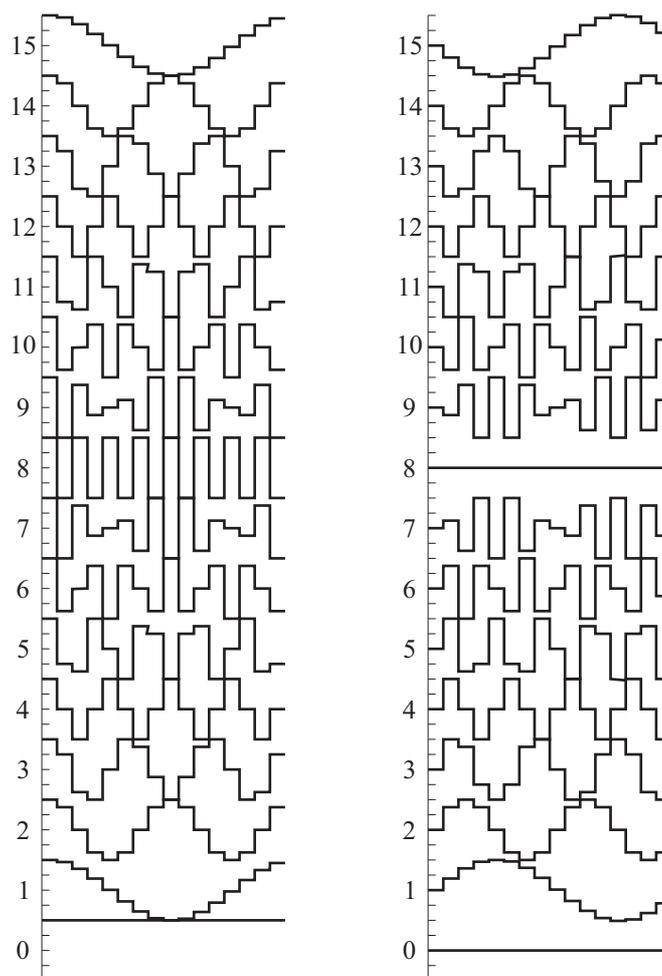
Das Basissystem der zweidimensionalen DFT sind $M \times N$ -Basismatrizen

$$W_{MN}^{uv} = W_M^u \otimes W_N^{vT} = \begin{bmatrix} 1 \\ W_M^{u1} \\ W_M^{u2} \\ \vdots \\ W_M^{u(M-1)} \end{bmatrix} \otimes [1, W_N^{v1}, W_N^{v2}, \dots, W_N^{v(N-1)}]$$

wobei \otimes das äußere Produkt der Basisvektoren der eindimensionalen DFT darstellt.

Während die 1D-DFT periodisch auf dem Einheitskreis ist, ist die 2D-DFT periodisch auf einem *Torus*. Da die 2D-DFT, wie alle anderen Repräsentationen der 2D-Fouriertransformation separabel ist, können anstelle der Basismatrizen W_{MN}^{uv} , die den $M \times N$ -dimensionalen komplexen Signalraum aufspannen, auch die entsprechenden Basisvektoren W_M^u bzw. W_N^v zur Implementierung verwendet werden.

$$F_{uv} = \sum_{m=0}^{M-1} \left(\sum_{n=0}^{N-1} f_{mn} W_N^{vn} \right) W_M^{um} = \sum_{n=0}^{N-1} \left(\sum_{m=0}^{M-1} f_{mn} W_M^{um} \right) W_N^{vn}$$



Darstellung der Basisvektoren der DFT für $M = 16$;
links Realteil (Cosinus); rechts Imaginärteil (Sinus).

Die DFT, wie alle anderen Fouriertransformationen ist eine unitäre Transformation. Allen unitären Transformationen ist gemein, daß sie eine Abbildung eines Vektorraumes (Signalraumes) auf sich selbst erzeugen, für die das Skalarprodukt invariant ist, also

$$\langle f, g \rangle = \langle \mathcal{U}f, \mathcal{U}g \rangle,$$

wenn \mathcal{U} die unitäre Transformation ist.

Abschließend noch eine Bemerkung zur Darstellung des Spektrums der Fouriertransformation (bzw. auch der anderen Varianten).

Die kanonische Darstellung als komplexe Zahl an der Position u lautet:

1. Real- und Imaginäranteil (kartesische Koordinaten)

$$F = F_R + jF_I$$

In der Praxis bedeutender ist allerdings diese Darstellung:

2. Amplitude und Phase (Polarkoordinaten)

$$F = |F| \cdot e^{j\phi}$$

Amplitude:

$$A = |F| = \sqrt{F_R^2 + F_I^2}$$

Phase:

$$\phi = \arctan \frac{F_I}{F_R}$$

Das Amplitudenspektrum A_u ist von gerader Symmetrie und das Phasenspektrum ϕ_u ist von ungerader Symmetrie. Im Gegensatz zu Real- und Imaginärteil haben Amplitude und Phase eine intuitive Interpretation. Das Amplitudenspektrum repräsentiert Luminanzeigenschaften und das Phasenspektrum Struktureigenschaften eines Signals. Tauscht man das Amplitudenspektrum eines Bildes gegen das eines anderen Bildes aus, wird das nach Anwendung der inversen Fouriertransformation rekonstruierte Bild die Strukturen des Ursprungbildes etwa beibehalten. Dies unterstreicht die Bedeutung der Phase für die Bildanalyse.

2.5.3 Theoreme der DFT bzw. FT

In diesem Abschnitt wollen wir kurz auf einige wichtige Theoreme spektraler Repräsentationen eingehen. In Abschnitt 2.7 werden weitere Zusammenhänge aufgezeigt. Die Bedeutung dieser Theoreme ist unabhängig von der gewählten Repräsentation, ihre Formulierung hingegen hängt natürlich von der Repräsentation ab.

Linearität

Die diskrete Fouriertransformation ist bezüglich der Superposition und der Multiplikation mit Konstanten linear:

$$a \cdot f + b \cdot g \quad \circ \text{---} \bullet \quad a \cdot F + b \cdot G$$

Dies läßt sich sofort aus der Definition der DFT ableiten.

Beispiel:

$$\begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & -8 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} - 9 \begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array}$$

ähnlich $\prod_{3 \times 3}(x)$ $[\delta]_{3 \times 3}(x)$
Laplaceop.

■

Ähnlichkeitssatz

gilt im Diskreten nur mit Einschränkungen

$$\mathcal{S}f_m = f_{sm} \quad \circ \text{---} \bullet \quad \frac{1}{|s|} \mathcal{S}^{-1}F_u = \frac{1}{|s|} F_{\frac{u}{s}}$$

$s > 1$: Frequenzraum wird "dichter" abgetastet

Beispiel: Resampling eines Operators (Expansion)

$$\begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|c|c|} \hline 1 & 0 & 1 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 1 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 1 & 0 & 1 \\ \hline \end{array} \approx \begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline \end{array}$$

■

Beweis:

$$\begin{aligned}
 \text{DFT}\{f_{sm}\} &= \sum_{m=0}^{M-1} f_{sm} W_M^{mu} \\
 &= \sum_{m=0}^{M-1} f_{sm} W_M^{sm \frac{u}{s}} \\
 &\stackrel{m=sl}{=} \frac{1}{|s|} \sum_{l=0}^{M-1} f_l W_M^{l \frac{u}{s}} \\
 &= \frac{1}{|s|} F_{\frac{u}{s}}
 \end{aligned}$$

Frequenzraum als "reziproker Raum" (Physik): Δ_m, Δ_u seien Distanzen im Orts- bzw. Frequenzraum. Dann gilt:

$$\Delta_m \sim \frac{1}{\Delta_u}$$

bzw.

$$\Delta_m \Delta_u = M$$

$$\Delta_m = 1 : \Delta_u = M$$

Eine Distanz von einem Bildpunkt im Ortsraum entspricht im Frequenzraum einer Distanz von M Bildpunkten (Frequenz $u = M$).

$$\Delta_m = M : \Delta_u = 1$$

Eine Distanz von $m = M$ Bildpunkten im Ortsraum entspricht im Frequenzraum einer Distanz von einem Bildpunkt (Frequenz $u = 1$).

Erinnerung : Die Basisfunktion W_M^u erfährt mit einer Periode im Ortsraum $\frac{M}{u}$ Positionen.

$$W_M^1 : \Delta_m = M, \quad W_M^2 : \Delta_m = \frac{M}{2}, \quad \dots$$

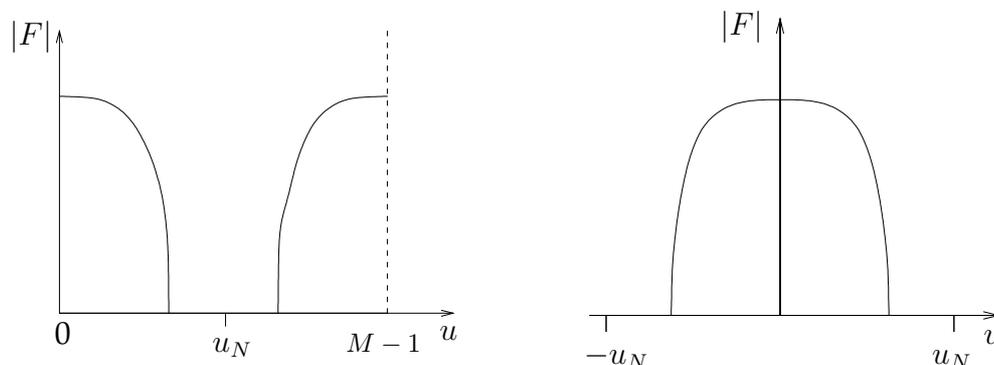
aber : Symmetrizentrum des Frequenzraumes bei $u = 0$ und $u = \frac{M}{2}$. Deshalb bezeichnet die Nyquist-Frequenz

$$u_N = \frac{M}{2}$$

die höchste darstellbare Frequenz. In der *alternativen Frequenzraumdarstellung*

$$-u_N \leq u \leq u_N$$

Das heißt, es treten auch negativen Frequenzen auf (entsprechen Phasenverschiebung von $|\Delta_u| = \pi$).



Resampling führt zu der neuen Nyquistfrequenz

$$u_N^s = \frac{1}{s} u_N$$

Für $s > 1$ schrumpft also nicht nur der Ortsraum, sondern auch der Frequenzraum. Näheres hierzu im Abschnitt 3.1 zur Abtastung.

Verschiebungssatz und Modulationssatz

$$\begin{array}{ll} \mathcal{T}^k f_m = f_{m-k} & \circ \text{---} \bullet \quad W_M^{uk} F_u \\ W_M^{-km} f_m & \circ \text{---} \bullet \quad F_{u-k} = \mathcal{T}^k F_u \end{array}$$

Eine *Verschiebung* im Ortsraum entspricht also einer frequenzproportionalen Phasenverschiebung (*Modulation*) im Frequenzraum und einer komplex-harmonische *Modulation* im Ortsraum entspricht einer *Verschiebung* des Spektrums.

Matrixform des Verschiebungssatzes (Verschiebung im Ortsraum):

$$T^k = W_M^{-1} W_M^{uk} W_M$$

mit W_M als Matrix der DFT. Matrixform des Modulationssatzes (Verschiebung im Frequenzraum):

$$W_M^{-km} = W_M^{-1} T^k W_M$$

Bedeutung/Anwendung (Verschiebungssatz):

- Bedeutung der Phase für Strukturpräsentation eines Signals

- Verschiebungsinvariante Signaldetektion :
Das Betragsspektrum eines Signals ist verschiebungsinvariant, bzw. die Fouriertransformation ist eine (bzgl. des Betrags) verschiebungsinvariante Operation.

$$\begin{array}{ccc} f_m & \text{○} \text{---} \text{●} & F_u \\ g_m = f_{m-k} & \text{○} \text{---} \text{●} & W_M^{uk} F_u = G_u \end{array}$$

Hieraus folgt $|F| = |G|$ für alle Frequenzen u .

Bedeutung/Anwendung (Modulationssatz):

- Ein (lokaler) Operator, der eine komplex-harmonische Modulation eines Signals bewirkt, erlaubt im Frequenzraum die verschiebliche Repräsentation lokaler Frequenzcharakteristika (Bandpaß-Operator). Das Gaborfilter (Abschnitt 3.10) ist ein solcher Operator. Es lassen sich Positionen im Ortsraum Frequenzeigenschaften zuordnen (lokale Amplitude, lokale Phase).
- Eine Verschiebung des Spektrums um $\frac{M}{2}$ Positionen bewirkt, daß der Koordinatenursprung des Frequenzraumes im Zentrum des Definitionsbereiches liegt.
Wegen

$$e^{\frac{j2\pi m \frac{M}{2}}{M}} = e^{j\pi m} = (-1)^m$$

folgt

$$W_M^{-m \frac{M}{2}} f_m \text{○} \text{---} \text{●} F_{u - \frac{M}{2}}$$

bzw.

$$(-1)^m f_m \text{○} \text{---} \text{●} F_{u - \frac{M}{2}}$$

Im 2D-Fall gilt

$$(-1)^{m+n} f_{mn} \text{○} \text{---} \text{●} F_{u - \frac{M}{2}, v - \frac{N}{2}}$$

Symmetrien der DFT und Zuordnungssatz

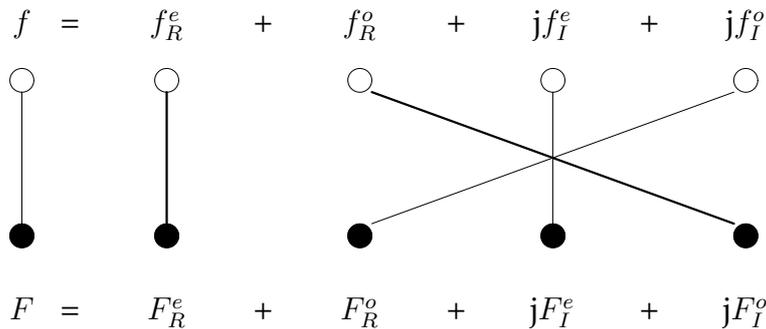
Der Spiegeloperator \mathcal{M} wirkt auf ein Signal f_m entsprechend

$$\mathcal{M}f_m = f_{-m}$$

Wegen Linearität der Fouriertransformation

$$\begin{array}{ll} f_m = f_m^e + f_m^o & F_u = F_u^e + F_u^o \\ f_m^e = \frac{1}{2}(\mathcal{I} + \mathcal{M})f_m & F_u^e = \frac{1}{2}(\mathcal{I} + \mathcal{M})F_u \\ f_m^o = \frac{1}{2}(\mathcal{I} - \mathcal{M})f_m & F_u^o = \frac{1}{2}(\mathcal{I} - \mathcal{M})F_u \\ & F_u = F_u^R + jF_u^I \end{array}$$

Zuordnungssatz:



Für ein reellwertiges Signal bildet sich der gerade Signalanteil in den Realteil ab und der ungerade Signalanteil in den Imaginärteil. Ein derartiges Spektrum heißt komplex konjugiert symmetrisch oder *hermitesch*.

Eine komplexe Funktion F heißt Hermite symmetrisch, wenn $F^* = \mathcal{M}F$, d.h.

$$F_{-u} = F_u^*$$

bzw.

$$F_u = F_{-u}^*$$

Aus dem Symmetrieeigenschaften von Real- und Imaginärteil eines reellen Signals folgt für $F_u = F_u^R + jF_u^I$

$$\mathcal{M}F_u = F_{-u} = F_u^* = F_u^R - jF_u^I,$$

also der Beweis der obigen Aussage.

Die Fouriertransformation behält also die Symmetrieeigenschaften eines Signals bei. Insbesondere gilt für reelle Signale f^e und f^o

$$\mathcal{F}\{f^e\} = F^e \equiv F_R$$

$$\mathcal{F}\{f^o\} = F^o \equiv jF_I$$

$$F_R = \mathcal{M}F_R \quad \text{Realteil ist gerade symmetrisch}$$

$$F_I = -\mathcal{M}F_I \quad \text{Imaginärteil ist ungerade symmetrisch}$$

Außerdem gilt für reelles f : $|F| = \mathcal{M}|F|$ Amplitude ist gerade symmetrisch

$$\phi = -\mathcal{M}\phi \quad \text{Phase ist ungerade symmetrisch}$$

Hieraus folgt, daß der Halbraum der positiven Frequenzen die Fouriertransformierte einer reellwertigen Funktionen vollständig charakterisiert.

Faltungssatz, Multiplikationssatz und Korrelationsatz

Faltungssatz : $g_m = (h \otimes f)_m$ $\bigcirc \text{---} \bullet G_u = (H \cdot F)_u$
Multiplikationssatz : $g_m = (h \cdot f)_m$ $\bigcirc \text{---} \bullet G_u = (H \otimes F)_u$

Anwendung: Faltung im Frequenzraum:

$$h \circledast f = \text{DFT}^{-1} \{ \text{DFT}\{h\} \text{DFT}\{f\} \}$$

2D-Faltung: M^4 Multiplikationen (bei zyklischer Faltung) \rightarrow Frequenzraumfaltung: M^2 Multiplikationen.

aber: in der Praxis (lineare Faltung: $M^2 K^2$ Multiplikationen) - Nutzen abhängig von der Ausdehnung der Impulsantwort, da Kosten zur Berechnung der Fouriertransformation. H sollte vorher berechnet sein.

Matrixform des Faltungssatzes (im Signalraum): $g_m = (h \circledast f)_m$

$$g = C_h f \quad C_h = W_M^{-1} D_h W_M$$

W_M : DFT-Matrix

D_h : Diagonalmatrix zur punktweisen Multiplikation von H mit F

$$D_h = [\delta_{kl} H_k]$$

$$= \begin{bmatrix} H_0 & 0 & \dots & & \\ 0 & H_1 & 0 & \dots & \\ 0 & 0 & H_2 & 0 & \dots \\ \vdots & & & \ddots & \\ 0 & \dots & \dots & 0 & H_{M-1} \end{bmatrix}$$

Matrixform des Multiplikationssatzes (im Signalraum): $G_u = (H \circledast F)_u$

$$G = C_H F \quad C_H = W_M D_H W_M^{-1}$$

$$D_H = [\delta_{mn} h_m]$$

Hier zeigt sich, daß \mathcal{F} und \mathcal{F}^{-1} Töplizmatrizen diagonalisieren.

Korrelationssatz Für reelle h und f gilt

$$g_m = (f \circledast h)_m \quad \circ \text{---} \bullet \quad G_u = (F H^*)_u = (F \mathcal{M}H)_u$$

da H Hermitesymmetrie besitzt ($H^* = \mathcal{M}H$)

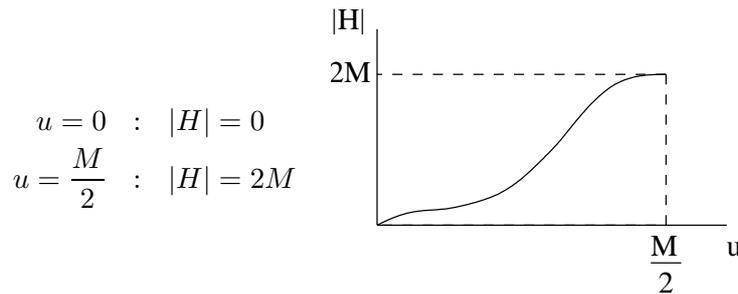
Differenz und Summe

Die Differenz und Summe sind das diskrete Analogon zu Differentiation und Integration.

Differenz : $\mathcal{D}_r f_m = f_m - \mathcal{T} f_m$

$$f_m - f_{m-1} \quad \circ \text{---} \bullet \quad F_u - W_M^{u1} F_u$$

$$\mathcal{D}_r F_u : F_u (1 - e^{-j2\pi \frac{u}{M}}) = F_u \cdot H_{D_r}$$

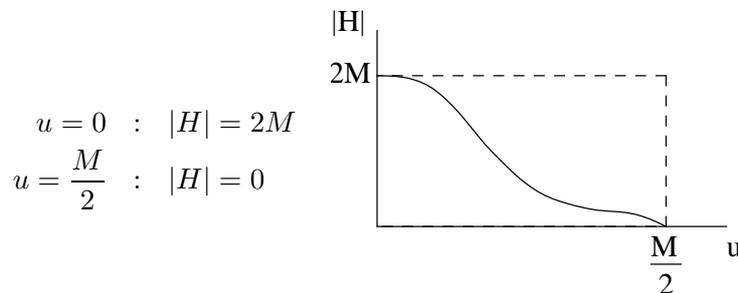


H_{D_r} ist ein Hochpaßfilter

Summe : $\Sigma_r f_m = f_m + \mathcal{T} f_m$

$$f_m + f_{m-1} \quad \circ \text{---} \bullet \quad F_u + W_M^{u1} F_u$$

$$\Sigma_r F_u : F_u(1 + e^{-j2\pi \frac{u}{M}}) = F_u \cdot H_{\Sigma_r}$$



H_{Σ_r} ist ein Tiefpaßfilter

Differentiation kontinuierlicher Signale

$$\frac{d}{dx} f(x) \quad \circ \text{---} \bullet \quad j2\pi u F(u)$$

Die Differentiation kontinuierlicher Signale führt im Frequenzraum zur Betonung hoher Frequenzen.

Beispiel: 2D-Laplace-Operator

$$\nabla^2 f(x, y) = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) f(x, y) \quad \circ \text{---} \bullet \quad -(2\pi)^2 (u^2 + v^2) F(u, v)$$



Integration kontinuierlicher Signale

$$\int_{-\infty}^x f(\xi) d\xi \quad \circ \text{---} \bullet \quad \frac{1}{j2\pi u} F(u)$$

Die Integration kontinuierlicher Signale führt im Frequenzraum zur Betonung niedriger Frequenzen.

Fourier-Korrespondenzen einiger wichtiger Funktionen

$e^{j2\pi u_0 x}$		$\delta(u - u_0)$
1		$\delta(u)$
$\cos(2\pi u_0 x)$		$\frac{1}{2} (\delta(u + u_0) + \delta(u - u_0))$
$\sin(2\pi u_0 x)$		$\frac{j}{2} (\delta(u + u_0) - \delta(u - u_0))$
$e^{-\pi x^2}$		$e^{-\pi u^2}$
$\text{III}(x)$		$\text{III}(u)$
$\text{II}(x)$		$\text{sinc}(u)$
$\Lambda(x)$		$\text{sinc}^2(u)$
$\text{sgn}(x)$		$\frac{1}{j\pi u}$

Tabelle der wichtigsten Gesetze der 1D-Fouriertransformation

Für zwei Signale $f(x)$ bzw. $g(x)$ und ihre Fouriertransformierten $F(u)$ bzw. $G(u)$ gelten folgende Sätze:

$f(x) + g(x)$		$F(u) + G(u)$	Additionssatz
$f(x)g(x)$		$F(u) * G(u)$	Multiplikationssatz
$f(x) * g(x)$		$F(u)G(u)$	Faltungssatz
$f(x) \circ g(x)$		$F^*(u)G(u), \quad f, g \text{ reell}$	Korrelationssatz
$f(ax)$		$\frac{1}{ a } F\left(\frac{u}{a}\right), \quad a \in \mathbb{R} \setminus \{0\}$	Ähnlichkeitssatz
$f(x - x_0)$		$e^{-j2\pi u x_0} F(u)$	Verschiebungssatz
$e^{j2\pi u_0 x} f(x)$		$F(u - u_0)$	Modulationssatz
$F^*(x)$		$f^*(u)$	Vertauschungssatz
$\frac{d^n}{dx^n} f(x)$		$(j2\pi u)^n F(u), \quad n \in \mathbb{N}^+$	Differentiationsatz
$\int_{-\infty}^x f(\xi) d\xi$		$\frac{F(u)}{j2\pi u} + \frac{1}{2} F(0)\delta(u)$	Integrationsatz

2.5.4 Quadratische Signalformen

[Mecklenbräuker, 1987] Signale im Ortsraum oder Frequenzraum können quadratisch miteinander verbunden werden.

Allgemein gilt: Die gemeinsamen Repräsentationen T_{mu}^f und T_{mu}^F benutzen entweder die Signale im Ortsraum oder im Frequenzraum. Es gilt

$$T_{mu}^f = \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} f_{m_1} f_{m_2}^* k_{m,u;m_1,m_2}$$

$$T_{mu}^F = \sum_{u_1=0}^{M-1} \sum_{u_2=0}^{M-1} F_{u_1} F_{u_2}^* K_{m,u;u_1,u_2}$$

Die Transformationskerne sind Fouriertransformierte zueinander:

$$k \circ \text{---} \bullet K$$

Erinnerung :

$$f \text{ reell: } f^* = f, F^* = \mathcal{M}F$$

Signalintensität I

$$k_{m,u;m_1,m_2}^I = \delta_{m-m_1} \delta_{m-m_2}$$

$$I_m^f = \sum_{m_1} \sum_{m_2} f_{m_1} f_{m_2}^* \delta_{m-m_1} \delta_{m-m_2}$$

wegen Ausblendeigenschaft der Deltafunktion

$$\sum_{m_i} f_{m_i} \delta_{m-m_i} = f_m$$

folgt

$$I_m^f = f_m f_m^* = |f_m|^2$$

Spektrale Leistungsdichte P

$$K_{m,u;u_1,u_2}^P = \delta_{u-u_1} \delta_{u-u_2}$$

damit:

$$P_u^F = |F_u|^2$$

I und P hängen nur von m bzw. u ab.

Energie des Signals

$$k_{m,u;m_1,m_2}^E = \delta_{m_1-m_2}$$

$$K_{m,u;u_1,u_2}^E = \delta_{u_1-u_2}$$

Wegen $\sum_{m_2} f_{m_2} \delta_{m_1-m_2} = f_{m_1}$ folgt

$$E^f = \sum_{m=0}^{M-1} |f_m|^2$$

$$E^F = \sum_{u=0}^{M-1} |F_u|^2$$

Die Energie hängt weder vom Ort noch von der Frequenz ab (Isometrie des Energiemaßes).

Parsevaltheorem:

Die Energie ist invariant gegenüber der Fouriertransformation:

$$E^f = E^F$$

Dies folgt aus der Unitarität der Fouriertransformation (Invarianz des Skalarproduktes gegenüber Fouriertransformation).

Autokorrelation

Der Transformationskern

$$k_{\tau; m_1, m_2}^R = \delta_{m_1 - m_2 - \tau}$$

erzeugt eine Signalrepräsentation über den Differenzkoordinaten $\tau = m_1 - m_2$.

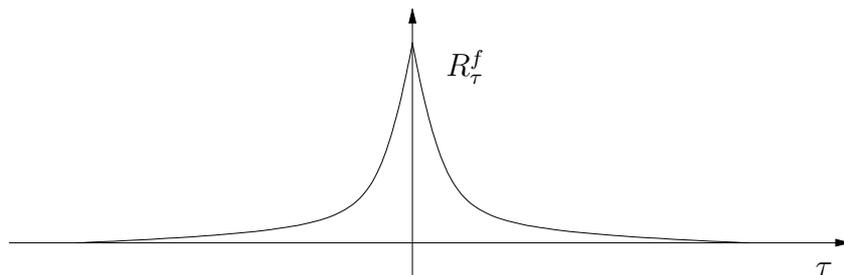
$$R_{\tau}^f = \sum_{m=0}^{M-1} f_m f_{m+\tau}$$

$$= |f_{\tau}|^2$$

R_{τ}^f ist gerade symmetrisch:

$$= (f * \mathcal{M}f)_{\tau}$$

$$= (f \circ f)_{\tau}$$



wegen Korrelationssatz $f \circ f \longleftrightarrow F F^* = |F|^2$ gilt:

$$R_{\tau}^f \longleftrightarrow P_u^F$$

Spektrale Leistungsdichte und Autokorrelationsfunktionen sind zueinander Fouriertransformierte.

2.5.5 Zyklische Faltung mittels DFT

In diesem Abschnitt soll für ein einfaches Beispiel mittels Faltungstheorem eine zyklische Faltung berechnet werden.

Erinnerung:

$$g = f \circledast h = \mathcal{F}^{-1}\{\mathcal{F}\{f\} \mathcal{F}\{h\}\}$$

Beispiel: 2D-Fall: $M = 2$

$$f = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} \quad h = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

■

Es gilt:

$$F_{uv} = \sum_{m=0}^{M-1} \sum_{n=0}^{M-1} f_{mn} W_M^{mu} W_M^{nv} \quad \text{mit} \quad W_M^{mu} = e^{-j \frac{2\pi}{M} mu}$$

und entsprechend für H_{uv} , sowie

$$g_{mn} = \frac{1}{M^2} \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} G_{uv} W_M^{-mu} W_M^{-nv}$$

mit

$$G_{uv} = F_{uv} H_{uv}$$

1. Berechnung ohne Erweiterung ($M = 2$)

$$W_2^{mu} = e^{-j\pi mu} = \begin{cases} 1 & \text{für } mu = 0 \\ -1 & \text{für } mu = 1 \end{cases}$$

$$\begin{array}{c}
 mu: \quad \begin{array}{cc} 0 & 1 \\ \downarrow & \rightarrow m \\ 0 & \begin{array}{c} 0 \\ 0 \end{array} \\ 1 & \begin{array}{c} 0 \\ 1 \end{array} \end{array}
 \end{array}$$

Für $W_2^{mu} W_2^{nv}$ erhält man folgende Tabelle zur Berechnung der X_{uv} :

		W_2^{mu}			
		1	1	1	-1
W_2^{nv}	(m, u)				
	(n, v)	(0, 0)	(1, 0)	(0, 1)	(1, 1)
1	(0, 0)	1	1	1	-1
1	(1, 0)	1	X_{00}	1	X_{10}
1	(0, 1)	1	X_{01}	1	-1
-1	(1, 1)	-1	-1	-1	1

Hieraus folgt zur Berechnung von F und H :

$$X_{00} = x_{00} + x_{10} + x_{01} + x_{11}$$

$$X_{10} = x_{00} - x_{10} + x_{01} - x_{11}$$

$$X_{01} = x_{00} + x_{10} - x_{01} - x_{11}$$

$$X_{11} = x_{00} - x_{10} - x_{01} + x_{11}$$

Also folgt

$$F = \begin{bmatrix} 4 & 2 \\ 2 & 0 \end{bmatrix} \quad H = \begin{bmatrix} 3 & 1 \\ 1 & -1 \end{bmatrix}$$

$$G = FH = \begin{bmatrix} 12 & 2 \\ 2 & 0 \end{bmatrix}$$

und schließlich

$$g = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix}$$

Dieses Ergebnis zeigt einen *Wraparound-Fehler!*

Berechnung mit Erweiterung ($M = 4$)

Erinnerung:

$$M_g \geq M_f + M_h - 1$$

$$f = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad h = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Hieraus folgt

$$g = \begin{bmatrix} 2 & 3 & 1 & 0 \\ 3 & 2 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Dieses Ergebnis hat keinen *Wraparound-Fehler*!

Wie sehen F , H und G aus?

2.5.6 Fast Fourier Transformation (FFT)

Die *Fast Fourier Transformation (FFT)* ist ein Algorithmus zur schnellen Berechnung der DFT in Zeit $O(M \cdot \log_2 M)$.

Die FFT nutzt

1. die Symmetrieeigenschaft der DFT sowie
2. die Separabilität der 2D-DFT, d.h. die Möglichkeit der Zurückführung der 2D-DFT auf 2 1D-DFTs.

1D-DFT : M^2 komplexe Multiplikationen und Additionen für beliebiges M
(1 komplexe Multiplikation = 4 reelle Multiplikationen; 1 komplexe Addition = 2 reelle Additionen.).

1D-FFT : $M \log_2 M$ komplexe Operationen, dabei aber Einschränkung für M durch $M = 2^n$, d.h. M muß Zweierpotenz sein.

Laufzeitvergleich (Anzahl komplexer Operationen):

M	DFT: M^2	FFT: $M \log_2 M$
2	4	2
8	64	24
64	4096	384
256	65536	2048

Im Falle eines 2D-Signals wird der Unterschied noch deutlicher (Ann.: $M \times M$ -Bild).

2D-DFT : M^4 komplexe Operationen
z.B. : 512×512 erfordert etwa $5 \cdot 10^{11}$ reelle Gleitpunktoperationen; bei 200 Mhz-Prozessor (50 MFLOPS) entspricht dies etwa 3h Rechenzeit.

2D-FFT : $M^2 \log_2 M$ komplexe Operationen
z.B. : 512×512 erfordert etwa $1.5 \cdot 10^7$ reelle Gleitpunktoperationen; bei 200 Mhz-Prozessor (50 MFLOPS) entspricht dies etwa 0.3 sec Rechenzeit.

Noch dramatischer ist der Unterschied im Fall der quaternionwertigen Fouriertransformation (QFT), siehe Abschnitt 2.7. Wie M. Felsberg in seiner Studienarbeit zeigte, galt auf seinem PC mit 200 MHz für die gleiche Bildgröße wie oben

2D-DQFT : etwa 12 Tage

2D-FQFT : etwa 3.5 Sekunden.

Das Prinzip der FFT ist die Methode der sukzessiven Halbierung (divide and conquer) mit $P = \frac{M}{2}$:

Dieses Verfahren wird $n = \log_2 M$ mal angewendet, bis am Ende die DFT nur noch aus 1-Punkt Transformationen besteht.

$$F_u = \sum_{m=0}^{M-1} f_m W_M^{um} = \sum_{m=0}^{2P-1} f_m W_{2P}^{um}, \quad u \in \{0, 1, \dots, 2P-1\}$$

Indem der Definitionsbereich des Signals in zwei verschiedene Intervalle mit geraden bzw. ungeraden Koordinaten zerlegt wird, zerfällt der obige Ausdruck in zwei Fouriertransformationen für Signale halber Länge.

$$F_u = \underbrace{\sum_{m=0}^{P-1} f_{2m} W_{2P}^{u2m}}_{\text{gerade Indizes}} + \underbrace{\sum_{m=0}^{P-1} f_{2m+1} W_{2P}^{u(2m+1)}}_{\text{ungerade Indizes}}$$

In diesem Ausdruck steht die Ausdehnung des Definitionsbereichs im Ortsraum im Widerspruch zu der im Frequenzraum. Da aber $W_{2P}^{2um} = W_P^{um}$, folgt

$$F_u = \sum_{m=0}^{P-1} f_{2m} W_P^{um} + \sum_{m=0}^{P-1} f_{2m+1} W_P^{um} W_{2P}^{u1}$$

mit dem Verschiebungsfaktor W_{2P}^{u1} , der die Phasenverschiebung des Anteils aus ungeraden Abtastpositionen repräsentiert. In diesem Ausdruck gilt $u \in \{0, 1, \dots, P-1\}$. Das Spektrum zerfällt in die Bereiche F_u und F_{u+P} . Wie sieht F_{u+P} aus?

Aus der Periodizität der Einheitswurzeln folgt $W_P^{u+P} = W_P^u$ und $W_{2P}^{u+P} = -W_{2P}^u$. Also gilt in den beiden Teilintervallen des Spektrums:

$$\begin{aligned} F_u &= F_u^e + F_u^o W_{2P}^{u1} \\ F_{u+P} &= F_u^e - F_u^o W_{2P}^{u1} \end{aligned}$$

wobei F_u^e das Spektrum für gerade m ist und F_u^o jenes für ungerade m . Zwei Transformationen der Länge $\frac{M}{2}$ Punkte liefern also die erste Hälfte des Spektrums. Ohne weiteren Transformationsaufwand wird die zweite Hälfte ebenfalls erhalten.

Die Berechnung auf dieser Stufe der Zerlegung wird tatsächlich nicht ausgeführt. Vielmehr erfolgt die Zerlegung solange, bis Transformationen der Länge eins erreicht werden (nach $n = \log_2 M$ Stufen). Durch Addition/Subtraktion der Ergebnisse von Stufe k wird die Transformation auf Stufe $k+1$ konstruiert, bis $k = n$ erreicht ist.

Die Anzahl der Operationen

$$\begin{aligned} m(n) &\text{ für Multiplikationen} \\ a(n) &\text{ für Additionen} \end{aligned}$$

erhält man durch Induktion

$$\begin{aligned} M = 2(n = 1) : m(1) &= 1 & a(1) &= 2 \\ M = 4(n = 2) : m(2) &= 2m(1) + 1 & a(2) &= 2a(1) + 4 \\ M = 8(n = 3) : m(3) &= 2m(2) + 4 & a(3) &= 2a(2) + 8 \end{aligned}$$

Allgemein gilt für $n \geq 1$:

$$\begin{aligned} m(n) &= 2m(n-1) + 2^{n-1} \\ a(n) &= 2a(n-1) + 2^n \end{aligned}$$

Implementierung durch Umordnen der Daten, so daß F^e und F^o sukzessive berechnet werden können:

- Umordnen:** gemäß der Bit-Umkehrregel: Beisp. für $M = 8$
Für $0 \leq m \leq M - 1$ läßt sich m darstellen durch Binärcodierung

$$m = 2^2 b_2 + 2^1 b_1 + 2^0 b_0$$

m' berechnet sich dann durch

$$m' = 2^2 b_0 + 2^1 b_1 + 2^0 b_2$$

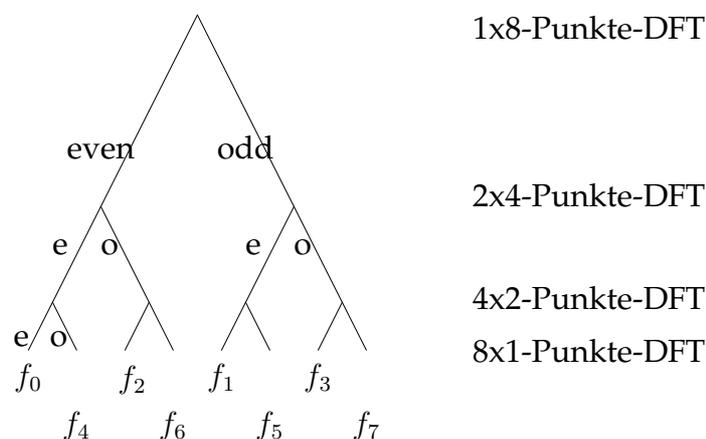
Also wird aus $m = \sum_{k=0}^{n-1} b_k 2^k$, $b_k \in \mathbb{B}$, durch Bitumkehr $m' = \sum_{k=0}^{n-1} b_{n-1-k} 2^k$.

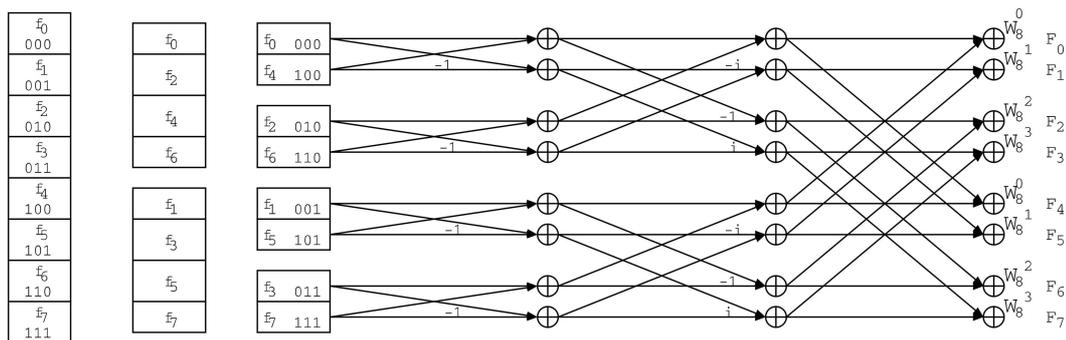
Beispiel: $M = 8$ Seien m_2 die Binärcodierung und m_{10} die dekadische Kodierung einer Zahl m .

$$\begin{aligned} m_2 = 110 \quad (m_{10} = 6) &\rightarrow m'_2 = 011 \quad (m'_{10} = 3) \\ m_2 = 001 \quad (m_{10} = 1) &\rightarrow m'_2 = 100 \quad (m'_{10} = 4) \\ m_2 = 000 \quad (m_{10} = 0) &\rightarrow m'_2 = 000 \quad (m'_{10} = 0) \\ m_2 = 111 \quad (m_{10} = 7) &\rightarrow m'_2 = 111 \quad (m'_{10} = 7) \end{aligned}$$

■

- Berechnen:** $f = [f_0, f_1, f_2, \dots, f_7]$

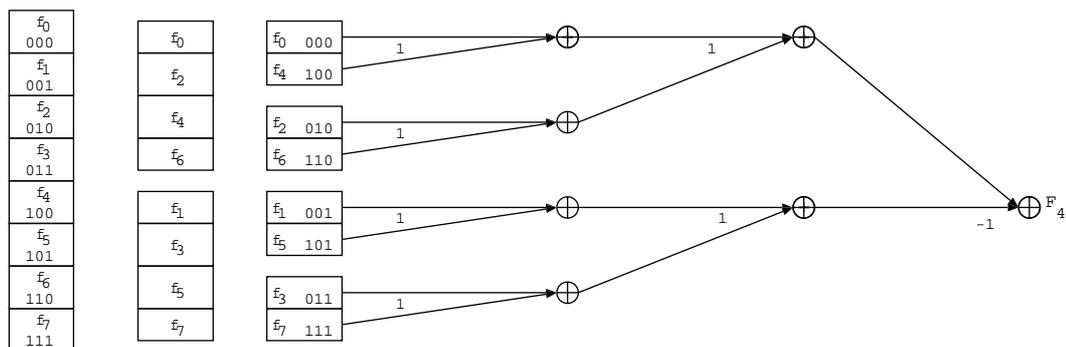
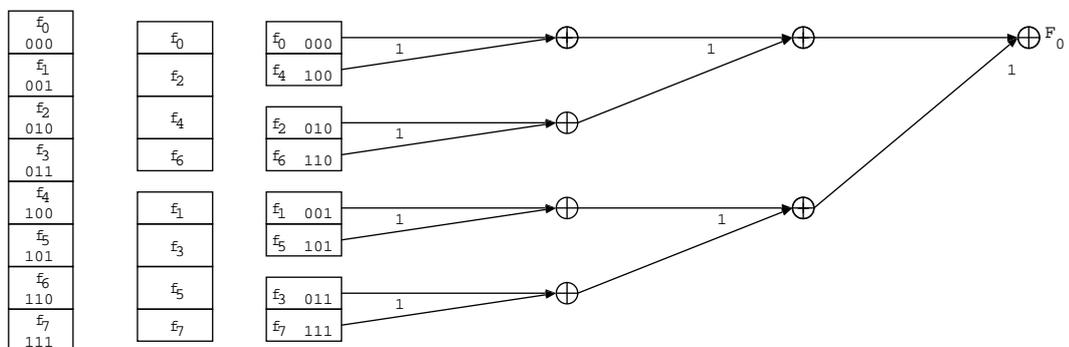




Beispiel:

$$F_0 = [(f_0 + f_4) + (f_2 + f_6)] + [(f_1 + f_5) + (f_3 + f_7)]$$

$$F_4 = [(f_0 + f_4) + (f_2 + f_6)] - [(f_1 + f_5) + (f_3 + f_7)]$$



$M = 8 : n = \log_2 8 = 3$ Stufen des FFT-Algorithmus

auf jeder Stufe : $\frac{M}{2}$ komplexe Multiplikation mit Phasenfaktoren und M komplexe Additionen

für F_0, F_4 : alle Phasenfaktoren sind Eins
sonst : auch Phasenfaktoren vom Wert j und $-j$

insgesamt : $\frac{M}{2} \log_2 M$ komplexe Multiplikationen
 $M \log_2 M$ komplexe Additionen

Die hierarchische Berechnung der FFT entspricht in dem Beispiel ($M = 8$) einer Zerlegung der unitären Transformationsmatrix der DFT in ein Produkt von drei schwach besetzten Matrizen der unitären Teiltransformationen. Die lineare Abbildung der Funktion f nach der Funktion F hat folgende Gestalt:

$$\begin{bmatrix} F_0 \\ F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \\ F_7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & W_8^1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & W_8^2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & W_8^3 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -W_8^1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -W_8^2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -W_8^3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & i & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -i & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & i \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -i \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \end{bmatrix}$$

2.6 Hartley-Transformation

[Bracewell, 1984], Hartley (1942)

siehe auch : Proc. IEEE 82(1994)3, 372-... (Spezialheft für die Hartley-Transformation)

$$\begin{aligned} \text{Fouriertransformation } \mathcal{F} : L_1(\mathbb{R}^n, \mathbb{R}) &\longrightarrow L_1(\mathbb{R}^n, \mathbb{C}) \quad (\text{bzw. } L_1(\mathbb{R}^n, \mathbb{C}) \longrightarrow L_1(\mathbb{R}^n, \mathbb{C})) \\ \text{Hartley-Transformation } \mathcal{H} : L_1(\mathbb{R}^n, \mathbb{R}) &\longrightarrow L_1(\mathbb{R}^n, \mathbb{R}) \end{aligned}$$

Basissystem der Hartley-Transformation:

$$\text{cas } 2\pi ux = \cos 2\pi ux + \sin 2\pi ux$$

$$H(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \{\cos 2\pi ux + \sin 2\pi ux\} dx$$

► inverses Basissystem = Basissystem \rightarrow gleiche Transformation

$$f(x) = \int_{-\infty}^{\infty} H(u) \{\cos 2\pi ux + \sin 2\pi ux\} du$$

► Warum ist Hartley–Transformation interessant ?

1. geringe Berechnungskomplexität:
 \mathcal{F} : komplexe Multiplikationen und Additionen
 \mathcal{H} : reelle Multiplikationen und Additionen
 1 komplexe Multiplikation = 4 reelle Multiplikationen und 2 reelle Additionen
 1 komplexe Addition = 2 reelle Additionen
2. weniger Speicherplatz
3. Fouriertransformation kann aus Hartley–Transformation konstruiert werden und umgekehrt :

$$F(u) = F_R(u) + jF_I(u) = F_e(u) + F_o(u)$$

mit

$$F_e(u) \equiv F_R(u) = \int f(x) \cos(2\pi ux) dx$$

$$F_o(u) \equiv jF_I(u) = -j \int f(x) \sin(2\pi ux) dx,$$

also

$$F(u) = H_e(u) - jH_o(u)$$

mit

$$H_e(u) = \int f(x) \cos(2\pi ux) dx \quad \text{geradsymmetr. Anteil d. Transf.}$$

$$H_o(u) = \int f(x) \sin(2\pi ux) dx \quad \text{ungeradsymmetr. Anteil d. Transf.}$$

$$H(u) = H_e(u) + H_o(u)$$

$$\rightarrow H_e(u) = \frac{1}{2}(H(u) + H(-u)) \quad \sim \cos \theta = \frac{1}{2}(\text{cas} \theta + \text{cas}(-\theta))$$

$$H_o(u) = \frac{1}{2}(H(u) - H(-u)) \quad \sim \sin \theta = \frac{1}{2}(\text{cas} \theta - \text{cas}(-\theta))$$

$$\rightarrow F_R(u) = \frac{1}{2}(H(u) + H(-u))$$

$$F_I(u) = \frac{1}{2}(-H(u) + H(-u))$$

also

$$H(u) = F_R(u) - F_I(u)$$

$$|F(u)|^2 = F(u)F^*(u) = H_e^2(u) + H_o^2(u) = \frac{1}{2}(H^2(u) + H^2(-u))$$

► diskrete Hartley–Transformation (DHT)

$$H_u = \sum_{m=0}^{M-1} f_m \text{cas} \left(\frac{2\pi mu}{M} \right)$$

$$f_m = \frac{1}{M} \sum_{u=0}^{M-1} H_u \text{cas} \left(\frac{2\pi mu}{M} \right)$$

Für die diskrete Hartley–Transformation existiert auch ein Fast Algorithmus (FHT).

► 2D-Hartley-Transformation

$$\begin{aligned}
H(u, v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \operatorname{cas}[2\pi(ux + vy)] \, dx dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \cos(2\pi ux) \cos(2\pi vy) \, dx dy \\
&\quad - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \sin(2\pi ux) \sin(2\pi vy) \, dx dy \\
&\quad + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \cos(2\pi ux) \sin(2\pi vy) \, dx dy \\
&\quad + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \sin(2\pi ux) \cos(2\pi vy) \, dx dy
\end{aligned}$$

→ der 2D-Kern ist nicht separierbar wie bei der 2D-Fouriertransformation !

→ die 2D-FHT ist gegenüber der 2D-FFT nur doppelt so schnell !

$$\left[\begin{aligned}
\text{Erinnerung : } F(u, v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} \, dx dy \\
&= \int_{-\infty}^{\infty} \underbrace{\left[\int_{-\infty}^{\infty} f(x, y) e^{-j2\pi ux} \, dx \right]}_{f_1(u, y)} e^{-j2\pi vy} \, dy \\
&= \int_{-\infty}^{\infty} f_1(u, y) e^{-j2\pi vy} \, dy
\end{aligned} \right.$$

$$H(u, v) = H_{ee}(u, v) + H_{oo}(u, v) + H_{eo}(u, v) + H_{oe}(u, v)$$

↑
gerader Anteil bzgl. x , ungerader Anteil bzgl. y

mit

$$\begin{aligned}
H_{ee}(u, v) &= \frac{1}{4} [H(u, v) + H(-u, -v) + H(-u, v) + H(u, -v)] \\
H_{oo}(u, v) &= \frac{1}{4} [H(u, v) + H(-u, -v) - H(-u, v) - H(u, -v)] \\
H_{eo}(u, v) &= \frac{1}{4} [H(u, v) - H(-u, -v) + H(-u, v) - H(u, -v)] \\
H_{oe}(u, v) &= \frac{1}{4} [H(u, v) - H(-u, -v) - H(-u, v) + H(u, -v)]
\end{aligned}$$

Konstruktion der Fourier Transformation:

$$F(u, v) = (H_{ee}(u, v) + H_{oo}(u, v)) - j(H_{eo}(u, v) + H_{oe}(u, v))$$

► Symmetrieselektivität von Fourier- und Hartley-Transformation

1D: $f(x) = f_e(x) + f_o(x)$, $f \in \mathbb{R}$ globale Separierung des Signals (in Abhängigkeit vom Symmetriezentrum) bildet sich ab in:

$$\begin{aligned}
\mathcal{F} : \quad & f_e \text{ --- } \bullet \text{ --- } F_R \equiv F_e \\
& f_o \text{ --- } \bullet \text{ --- } jF_I \equiv F_o \\
\mathcal{H} : \quad & f_e \text{ --- } \bullet \text{ --- } H_e \\
& f_o \text{ --- } \bullet \text{ --- } H_o
\end{aligned}$$

d.h., die in 1D möglichen Symmetrien werden durch die Fouriertransformation explizit gemacht.

Hingegen in der Hartley-Transformation sind die Symmetriekomponenten

$$H_e \equiv F_e \quad \text{und} \quad H_o = jF_o$$

in der Funktion $H = \mathcal{H}f$ nur vermischt enthalten. Die Hartley-Transformation kann im 1D-Fall also die globale Symmetrie von Signalen nicht darstellen.

2D: globale Symmetriezerlegung des Signals im Ortsraum (in Abhängigkeit von der Wahl des Ursprungs des Koordinatensystems):

$$f(x, y) = f_{ee}(x, y) + f_{oo}(x, y) + f_{eo}(x, y) + f_{oe}(x, y)$$

In der *2D-Hartley-Transformation* überlagern sich die korrespondierenden Symmetriekomponenten komplett:

$$H(u, v) = H_{ee}(u, v) + H_{oo}(u, v) + H_{eo}(u, v) + H_{oe}(u, v)$$

In der *2D-Fouriertransformation* entsteht eine partielle Trennung der Symmetriekomponenten:

$$F(u, v) = F_R(u, v) + jF_I(u, v)$$

mit

$$F_R(u, v) = H_{ee}(u, v) + H_{oo}(u, v)$$

$$F_I(u, v) = -(H_{eo}(u, v) + H_{oe}(u, v))$$

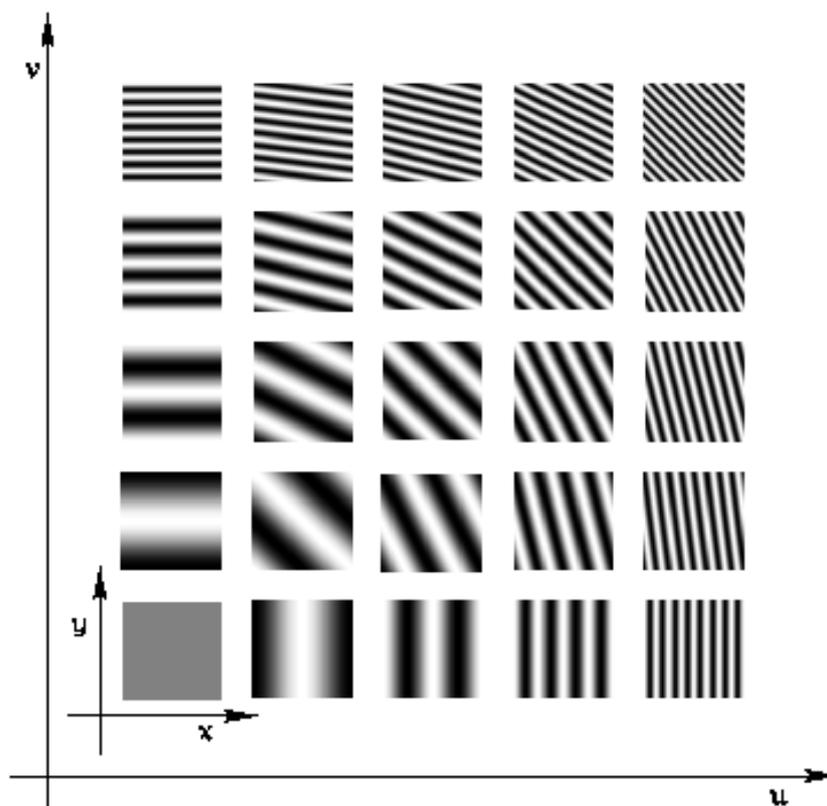
→ Zur vollständigen Separierung der Symmetriekomponenten ist eine andere Transformation erforderlich.

Ursache der Nichtadäquatheit der komplexen 2D-Fouriertransformation:

Separabilität des Basissystems der 2D-Fouriertransformation (gilt allgemein für nD -FT)

$$2D: \quad e^{-j2\pi ux} e^{-j2\pi vy} = e^{-j2\pi(ux+vy)}$$

entspricht wieder einem 1D-Konzept in der komplexen Ebene. → in \mathbb{C} gilt: nur 1D-Signale sind adäquat transformierbar, da die Algebra von \mathbb{C} nur Multiplikationen erlaubt, die in \mathbb{C} bleiben.



Die Basisfunktionen der 2D-FT

Die Symmetrien höher dimensionaler Signale sind im Komplexen nicht darstellbar. Je höher die Dimension des Signals, umso mehr Symmetriekomponenten können aber auftreten. Deren Darstellung ist in der spektralen Analyse von nD -Signalen von Bedeutung.

Ausweg: Vermeiden des Verschmelzens der Komponenten des nD -Basissystems der Fouriertransformation dadurch, daß jeder Koordinate x_i des Signals im Ortsraum eine eigene komplexe Ebene $\mathbb{C}_i = \mathbb{R}\{1, j_i\}$ zugeordnet wird. Es ergibt sich als Forderung

$$2D: \quad e^{-j_1 2\pi u x} e^{-j_2 2\pi v y} \neq e^{-j 2\pi (u x + v y)}$$

Das heißt, die Fouriertransformation von nD -Signalen muß in reichhaltigere Algebren als die der komplexen Zahlen eingebettet werden. Den Weg weist die *Clifford-Algebra* (Clifford, 1878).

2.7 Quaternionwertige Fouriertransformation

[Bülow and Sommer, 1997].

Die quaternionwertige Fouriertransformation (QFT)

$$\mathcal{F}^q : L_1(\mathbb{R}^2, \mathbb{R}) \longrightarrow L_1(\mathbb{R}^2, \mathbb{H})$$

gestattet die Separierung aller Symmetriekonzepte eines 2D-Signals:

$$F^q(u, v) = H_{ee}(u, v) - iH_{oe}(u, v) - jH_{eo}(u, v) - kH_{oo}(u, v).$$

Das Basissystem der \mathcal{F}^q lautet

$$Q_{u,v}(x, y) = e^{-i2\pi ux} e^{-j2\pi vy}$$

Die Multiplikation zweier verschiedener komplexen Zahlen

$$e^{-i2\pi ux} \in \mathbb{C}_1 = \mathbb{R}\{1, i\} \subset \mathbb{H} \quad \text{und} \quad e^{-j2\pi vy} \in \mathbb{C}_2 = \mathbb{R}\{1, j\} \subset \mathbb{H}$$

ist in der vierdimensionalen Algebra der Quaternionen (\mathbb{H}) möglich.

Die QFT macht Sinn für ein 2D-Signal. Allgemein gilt :

Ein n -dimensionales Signal erfordert die Einbettung der Fouriertransformation in eine 2^n -dimensionale Algebra. Eine derartige Fouriertransformation heißt *Clifford-Fouriertransformation* für die Dimension n .

Eine 2^n -dimensionale Algebra ist ein 2^n -dimensionaler linearer Raum, vergleichbar \mathbb{R} und \mathbb{C} für $n = 0$ bzw. $n = 1$. Wir lernen zunächst die Algebra der Quaternionen kennen.

2.7.1 Quaternionen

Hamilton, 1843 (Multiplikation zweier Vektoren im Raum)

$$\mathbb{H} = \{q = a + bi + cj + dk \mid a, b, c, d \in \mathbb{R}\}$$

Quaternionen bilden eine nicht kommutative Algebra, die man als einen vierdimensionalen Vektorraum über \mathbb{R} interpretieren kann mit der natürlichen Definition der Addition und skalaren Multiplikation. Mit den Basisvektoren

$$\begin{aligned} 1 &= (1, 0, 0, 0) & 1^2 &= 1 \\ i &= (0, 1, 0, 0) & i^2 &= -1 \\ j &= (0, 0, 1, 0) & j^2 &= -1 \\ k &= (0, 0, 0, 1) & k^2 &= -1 \end{aligned}$$

und den Multiplikationsregeln des Quaternionproduktes

$$\begin{aligned} i * j &= k = -j * i \\ j * k &= i = -k * j \\ k * i &= j = -i * k \\ i * j * k &= -1 \end{aligned}$$

erhält man die Algebra der Quaternionen als spezielle hyperkomplexe Zahlen. *Quaternion*:

$$\begin{aligned} q &= a + bi + cj + dk = a + p = s + v \\ s &\equiv a && \sim \text{skalärer Teil (s) oder Realteil (a)} \\ v &\equiv p && \sim \text{vektorieller Teil (v) oder reine Quaternion (p)} \\ v &= bi + cj + dk \end{aligned}$$

Quaternionalgebra: $(\mathbb{H}, *, +, 0, 1)$

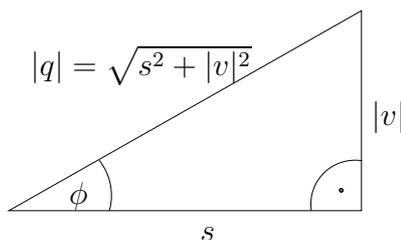
$$\begin{aligned}
 q &\stackrel{!}{=} [s, v] \\
 q_3 &= q_1 + q_2 = [s_1 + s_2, v_1 + v_2] = [s_3, v_3] \\
 q_3 &= q_1 * q_2 = [(s_1 s_2 - v_1 \cdot v_2), (s_1 v_2 + s_2 v_1 + v_1 \times v_2)] = [s_3, v_3] \\
 \rightarrow & \quad q_1 * q_2 \neq q_2 * q_1 \quad \sim \text{nicht kommutativ}
 \end{aligned}$$

Bemerkung : * Quaternionenprodukt, · Skalarprodukt, × Vektorprodukt
 Das Quaternionenprodukt erfordert zu seiner Berechnung
 16 reelle Multiplikationen und 12 reelle Additionen.

Annahme : $s_1 = s_2 = 0$
 d.h. q_1, q_2 sind Vektoren im \mathbb{R}^3 , dann folgt
 $q_3 = q_1 * q_2 = [-v_1 \cdot v_2, v_1 \times v_2] = [s_3, v_3]$

konjugierte Quaternion: $q^* = [s, -v]$

Betrag: $|q| = \sqrt{q * q^*} = \sqrt{q^* * q} \sqrt{s^2 + |v|^2} = \sqrt{a^2 + b^2 + c^2 + d^2}$



inverse Quaternion: $q^{-1} = \frac{1}{q} = \frac{1}{q} \frac{q^*}{q^*} = \frac{q^*}{|q|^2}$

Winkel zwischen zwei Quaternionen: $\alpha = \angle(q_1, q_2)$
 $\cos \alpha = \frac{s_1 q_1}{|q_1| |q_2|}$ mit $s_1 q_2 = \langle q_1, q_2 \rangle$ (inneres Produkt)

Einheitsquaternion: $|q| = 1 : q^{-1} = q^*$

polare Darstellung: $q = r(\cos \phi + I \sin \phi) = r e^{I\phi}$ wegen Eulerformel
 mit $r = |q|$, $\phi = \arctan \frac{|v|}{s}$ und

reine Einheitsquaternion: $I = \frac{v}{|v|}$, $I^2 = -1$, $|I| = \sqrt{b^2 + c^2 + d^2} = 1$
 $q_3 = [0, I] * [0, I] = -(I \cdot I) + (I \times I) = -|I|^2 = -1$
 Einheitsvektor $\hat{=}$ reine Einheitsquaternion im 3D-Raum
 (beschreibt die Orientierung einer im Raum gedrehten komplexen Ebene)

Die Algebra der Einheitsquaternionen $\mathbb{R}\{1, I\}$ ist isomorph zu der der komplexen Zahlen. Ist I eine Einheitsquaternion, so ist auch $e^{I\phi}$ eine Einheitsquaternion.

Die Exponentialfunktion für Quaternionen ist erklärt durch

$$\exp : \mathbb{H} \rightarrow \mathbb{H}$$

wegen

$$e^{(q)} = \sum_{k=0}^{\infty} \frac{q^k}{k!}, \quad q \in \mathbb{H}$$

Eine andere Darstellung der Exponentialfunktion einer Quaternion $q = [s, v]$ ist

$$e^{(q)} = e^{(s)} \left(\cos |v| + \frac{v}{|v|} \sin |v| \right).$$

Es gilt die Kommutativität

$$e^{I\phi_1} * e^{I\phi_2} = e^{I\phi_2} * e^{I\phi_1}.$$

Es gilt zwar

$$e^{I\phi_1} * e^{I\phi_2} = e^{I(\phi_1 + \phi_2)}$$

aber

$$e^{I\phi} * e^{J\phi} \neq e^{\phi(I+J)} \quad !!$$

Daher

$$e^{ix} * e^{jy} \neq e^{ix+jy} \quad x, y \in \mathbb{R}, \quad i^2 = j^2 = -1, \quad i \perp j$$

Dies ist die grundlegende Eigenschaft der Einbettung in die Algebra der Quaternionen, die für die Fouriertransformation genutzt werden kann.

2.7.2 Quaternionwertige Fouriertransformation

Die Basisfunktionen der 1D-Fouriertransformation

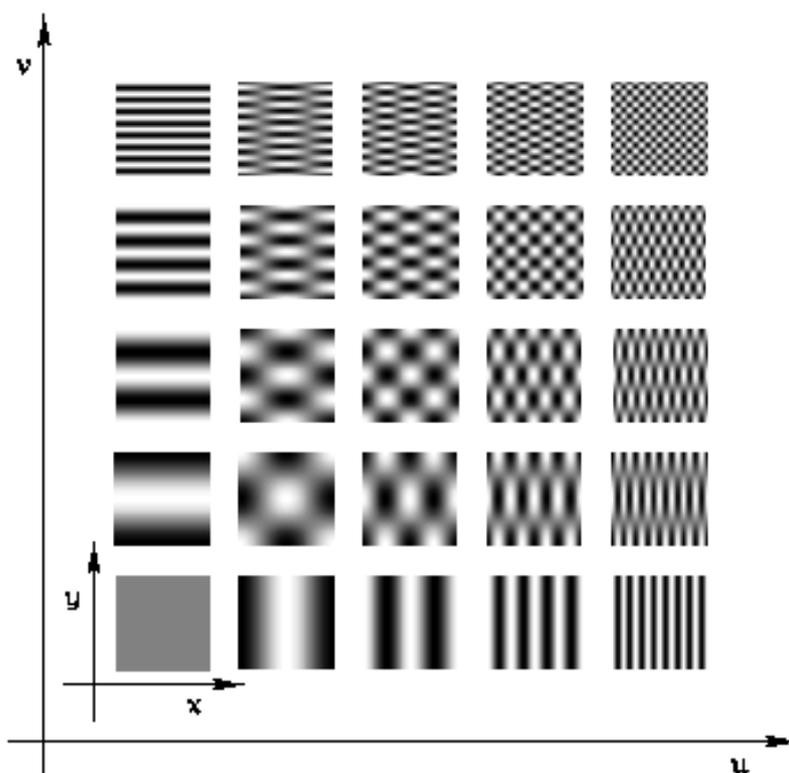
$$Q_u(x) = e^{-i2\pi ux}, \quad Q_v(y) = e^{-j2\pi vy}$$

werden nun *als Quaternionen interpretiert* und miteinander und dem Wert $f(x, y)$ multipliziert.

In der Abbildung sind die Basisfunktionen der QFT dargestellt. Während in Achsenrichtung rein eindimensionale Strukturen unterstützt werden, sind die Basisfunktionen für gemischte x - und y -Komponenten Repräsentanten zweidimensionaler Strukturen. Diese Eigenschaften wurden für die 2D-Fouriertransformation gesucht. Siehe im Gegensatz hierzu die Basisfunktionen der komplexen Fouriertransformation.

Wegen der Nicht-Kommutativität der Quaternion-Multiplikation wird bei der Multiplikation die *Ordnung der Multiplikation zu fixieren* sein, wenn $f \in \mathbb{H}$. Ist aber $f \in \mathbb{R}$, kann $f(x, y)$ an beliebiger Stelle in der Definitionsgleichung auftreten.

$$\begin{aligned} F^q(u, v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i2\pi ux} f(x, y) e^{-j2\pi vy} dx dy \\ f(x, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i2\pi ux} F^q(u, v) e^{j2\pi vy} du dv \end{aligned}$$



Die Basisfunktionen der 2D-QFT

Die Berechnung kann ähnlich wie die der 2D-Fouriertransformation als kaskadierte 1D-Transformation erfolgen:

$$\begin{aligned}
 1. \quad \tilde{f}(u, y) &= \int_{-\infty}^{\infty} e^{-i2\pi ux} f(x, y) dx \\
 2. \quad F^q(u, v) &= \int_{-\infty}^{\infty} \tilde{f}(u, y) e^{-j2\pi vy} dy
 \end{aligned}$$

Dabei gilt in \mathbb{H} :

$$F^q(u, v) = \mathcal{R}(F^q) + i\mathcal{I}(F^q) + j\mathcal{J}(F^q) + k\mathcal{K}(F^q) |_{(u,v)}$$

$\mathcal{R}, \mathcal{I}, \mathcal{J}, \mathcal{K}$ sind Operatoren, die den Realteil, \mathcal{I} -Imaginärteil u.s.w. von F_q erzeugen.

Sei $q = a + bi + cj + dk$ mit $a, b, c, d \in \mathbb{R}$. Dann gilt

$$\begin{aligned}
 \mathcal{R}(q) &\equiv a = \frac{1}{4}(q - iqi - jqj - kqk) \\
 \mathcal{I}(q) &\equiv b = \frac{1}{4}(q - iqi + jqj + kqk) \\
 \mathcal{J}(q) &\equiv c = \frac{1}{4}(q + iqi - jqj + kqk) \\
 \mathcal{K}(q) &\equiv d = \frac{1}{4}(q + iqi + jqj - kqk)
 \end{aligned}$$

Symmetrien (Zuordnungssatz):

$$1D : f = f_e + f_o \quad , f \in \mathbb{R} \text{ oder } \mathbb{C}$$



$$F = \mathcal{R}(F) + j\mathcal{J}(F) \quad F \in \mathbb{C}$$

Wenn $f \in \mathbb{C}$ und f gerade, dann transformieren sich Real- und Imaginärteil direkt.

Wenn $f \in \mathbb{C}$ und f ungerade, dann vertauschen sich Real- und Imaginärteil bei der Transformation.

	$f = f_e$	$f = f_o$
f	$\mathcal{R} \mathcal{J}$	$\mathcal{R} \mathcal{J}$
F	$\mathcal{R} \mathcal{J}$	$\mathcal{J} \mathcal{R}$

$$F(u) = H_e(u) - jH_o(u)$$

$$2D : f = f_{ee} + f_{oe} + f_{eo} + f_{oo} \quad , f \in \mathbb{R} \text{ oder } \mathbb{H}$$



$$F^q = \mathcal{R}(F^q) + i\mathcal{I}(F^q) + j\mathcal{J}(F^q) + k\mathcal{K}(F^q) \quad F^q \in \mathbb{H}$$

Wenn $f \in \mathbb{H}$ gewisse Symmetrien besitzt, kommt es zu Vertauschung der Real- bzw. Imaginärteile von F^q .

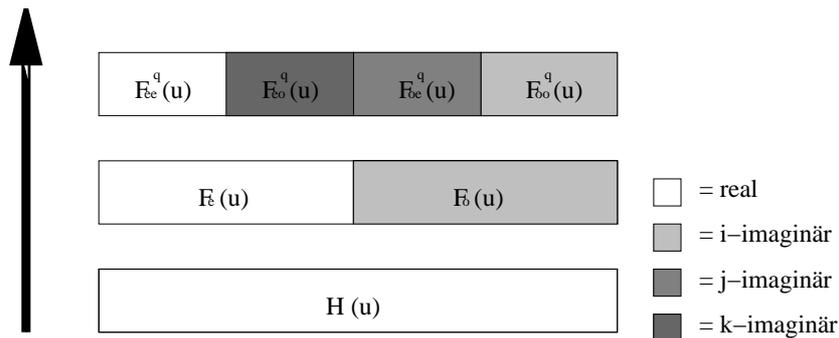
	$f = f_{ee}$	$f = f_{oe}$	$f = f_{eo}$	$f = f_{oo}$
f	$\mathcal{R} \mathcal{I} \mathcal{J} \mathcal{K}$			
F^q	$\mathcal{R} \mathcal{I} \mathcal{J} \mathcal{K}$	$\mathcal{I} \mathcal{R} \mathcal{K} \mathcal{J}$	$\mathcal{J} \mathcal{K} \mathcal{R} \mathcal{I}$	$\mathcal{K} \mathcal{J} \mathcal{I} \mathcal{R}$

Hieraus erhält man schließlich die Darstellung der quaternionischen Fouriertransformation durch die Hartley-Komponenten.

$$F^q(u, v) = H_{ee}(u, v) - iH_{oe}(u, v) - jH_{eo}(u, v) - kH_{oo}(u, v)$$

Zusammengefaßt gilt also:

1. In der quaternionwertigen 2D–Fouriertransformation werden alle Komponenten der 2D–Hartley–Transformation entkoppelt.
2. Die 2D–Fouriertransformation steht zwischen der 2D–quaternionwertigen Fouriertransformation und der 2D–Hartley–Transformation in dem Sinne, daß die 2D–QFT aus der 2D–FT in ähnlicher Weise hervorgeht wie die 2D–FT aus der 2D–HT.



2.7.3 Quaternionische Hermite-Symmetrie

Erinnerung: Eine Funktion $g : \mathbb{R}^n \rightarrow \mathbb{C}$ heißt Hermite symmetrisch, wenn $g(x) = g^*(-x)$ für alle $x \in \mathbb{R}^n$.

$$f \circ \text{---} \bullet F, \quad f \in \mathbb{R}, \quad F \in \mathbb{C}$$

$F(u) = F^*(-u)$ ist Hermite symmetrisch

Konsequenz: $f_e \circ \text{---} \bullet F_R \equiv F_e$
 $f_o \circ \text{---} \bullet jF_I \equiv F_o$

Eine Funktion $g : \mathbb{R}^2 \rightarrow \mathbb{H}$ heißt quaternionisch Hermite symmetrisch, wenn

$$g(-x, y) = \beta(g(x, y))$$

$$g(x, -y) = \alpha(g(x, y))$$

für alle $(x, y) \in \mathbb{R}^2$.

Für eine quaternionisch Hermite symmetrische Funktion g gilt auch

$$g(-x, -y) = \gamma(g(x, y)).$$

α, β, γ sind nichttriviale Algebra-Involutionen von $q \in \mathbb{H}$.

$$\alpha(q) = -iqi = a + bi - cj - dk$$

$$\beta(q) = -jqj = a - bi + cj - dk$$

$$\gamma(q) = -kqk = a - bi - cj + dk$$

$$f \circ \text{---} \bullet F^q, \quad f \in \mathbb{R}, \quad F^q \in \mathbb{H}$$

F^q ist quaternionisch Hermite symmetrisch mit obigen Eigenschaften.

Konsequenz: $f_{ee} \circ \text{---} \bullet \mathcal{R}(F^q)$

$f_{oe} \circ \text{---} \bullet \mathcal{I}(F^q)$

$f_{eo} \circ \text{---} \bullet \mathcal{J}(F^q)$

$f_{oo} \circ \text{---} \bullet \mathcal{K}(F^q)$

2.7.4 Das Phasenkonzept der 2D-QFT

Für die 2D-FT gibt es nur einen Phasenwinkel in der einzigen komplexen Ebene.

Für die 2D-QFT existieren *drei Phasenwinkel*, da drei komplexe Ebenen den 3D-Vektorraum aufspannen.

Für Rotationen in Räumen beliebiger Dimension lassen sich Rotationsmatrizen verwenden. Rotationsmatrizen sind orthogonale Matrizen. Im 3D-Fall lassen sich aber Rotationen besser mit Quaternionen beschreiben. Wir werden deshalb das Phasenkonzept der 2D-QFT aus dem *Rotationskonzept* der Quaternionen herleiten.

Eine reine Quaternion bzw. der Vektoranteil einer allgemeinen Quaternion

$$p \equiv v = bi + cj + dk$$

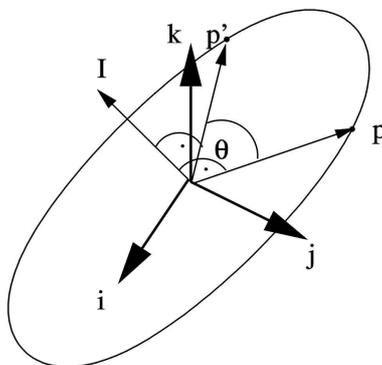
repräsentiert einen Punkt im euklidischen Raum, da $i \perp j \perp k$. Sei

$$p' \equiv v' = b'i + c'j + d'k$$

ein anderer Punkt, der aus p durch Rotation hervorgegangen ist. ($|p| = |p'|$ → reine Rotation ohne Streckung).

Orthogonale Rotation:

Sei q eine Einheitsquaternion, d.h. $|q| = 1$. Dann realisiert q eine orthogonale Rotation. Die Rotation mittels Einheitsquaternion $p' = q * p$ ist vergleichbar mit der Rotation in der komplexen Ebene durch Multiplikation eines Vektors mit einer komplexen Zahl (in polarer Darstellung). Im Falle der orthogonalen Rotation fallen die Koordinatenursprünge von $\mathbb{C} = \mathbb{R}(1, I)$ und $\mathbb{R}^3 \subset \mathbb{H}$ zusammen.



Aus der Multiplikation $p' = q * p$ lässt sich ableiten, daß

$$p' = (\cos \theta + I \sin \theta) * p$$

- ▶ I ist Rotationsachse (senkrecht auf Rotationsebene, die p und p' enthält)
- ▶ θ ist Rotationswinkel
- ▶ $q = \cos \theta + I \sin \theta$ ist eine *Einheitsquaternion* ($|q| = 1$), die eine *orthogonale Rotation* realisiert

$$q = \cos \theta + I \sin \theta = \cos \theta + iI_i \sin \theta + jI_j \sin \theta + kI_k \sin \theta$$

für $I = (I_i, I_j, I_k)$, $I_i^2 + I_j^2 + I_k^2 = -1$

- ▶ Wäre $|q| \neq 1$ → zusätzliche Skalierung (Dreh-Streckung)
- ▶ Die Forderung nach Orthogonalität der Rotation ist eine zu starke Einschränkung für die Wahl der Rotationsachse für beliebigen Punkt des Raumes. (Falls die Orthogonalität nicht eingehalten wird, dann ergibt $p' = q * p$ eine Quaternion und keinen Vektor, d.h. q bildet p nicht in einen Punkt p' ab!)

Allgemeine Rotation:

Die Rotationsachse geht nicht durch den Koordinatenursprung. Dies wird durch eine *Spinor-Multiplikation* erreicht.

Spinor-Produkt: Quaternionen sind spezielle Spinoren vom Betrag 1, $|q| = 1$.

$$p' = q * p * q^{-1} \quad |q| = 1, \quad q = \cos \frac{\theta}{2} + I \sin \frac{\theta}{2}$$

mit $I = (I_i, I_j, I_k)$.

Da $|q| = 1$, gilt $q^{-1} = q^*$.

In der Gleichung wird eine Rotation des Raumes \mathbb{R}^3 beschrieben. Man ist aber oft an der Rotation des Koordinatensystems interessiert, deshalb

- entweder Ersetzen von θ durch $-\theta$
- oder $p' = q^{-1} * p * q$

Im allgemeinen Fall muß für Spinoren nicht gelten $|q| = 1$.

Sei $q' = \lambda q$, $|q| = 1$, $\lambda \in \mathbb{R}$, $\lambda \neq 0$. Dann gilt

$$\begin{aligned} p' &= q' * p * q'^{-1} \\ &= \lambda q * p * \lambda^{-1} q^{-1} = \lambda \lambda^{-1} q * p * q^{-1} \\ &= q * p * q^{-1} \end{aligned}$$

Für $q = [s, u]$, $p = [0, v]$ erhält man

$$p' = [0, v + 2s(u \times v) + 2u \times (u \times v)]$$

Also ist p' ein Vektor im \mathbb{R}^3 .

Mehrfach-Rotation:

Quaternionen bilden eine Gruppe bezüglich der quaternionischen Multiplikation. Also lassen sich Mehrfach-Rotationen durch *Kaskadierung* von Quaternionen-Multiplikationen darstellen.

$$p'' = q_2 * p' * q_2^* = (q_2 * q_1) * p * (q_2 * q_1)^* = q_2 * (q_1 * p * q_1^*) * q_2^*$$

- Eine Mehrfach-Rotation lässt sich *kompakter* formulieren und *schneller* sowie *stabiler* mittels Quaternionen ausführen im Vergleich zur Anwendung von Rotationsmatrizen.
- Grund: Rundungsfehler in der Matrixmultiplikation erzeugen Abweichungen von der geforderten Orthogonalität. Es entstehen keine reinen Drehungen, sondern auch Skalierung und Scherung. Noch gravierender ist dies beim Schätzen von Rotationen. Zur Vermeidung der Fehlerfortpflanzung ist dabei eine fortgesetzte Normalisierung erforderlich.

Rotationsmatrix:

$$p' = q * p * q^* \stackrel{!}{=} Mp$$

$$\text{für } q = [s, v] = [a, (b, c, d)] \quad , \quad |q| = 1$$

$$M = \begin{bmatrix} 1 - 2c^2 - 2d^2 & 2bc - 2ad & 2bd + 2ac \\ 2bc + 2ad & 1 - 2b^2 - 2d^2 & 2cd - 2ab \\ 2bd - 2ac & 2cd + 2ab & 1 - 2b^2 - 2c^2 \end{bmatrix}$$

Die Rotationsmatrix des Euklidischen Raumes lässt sich in 12 verschiedenen Varianten in das Produkt dreier Rotationsmatrizen um die den Raum aufspannenden Koordinaten darstellen.

Beispiel:

$$M = M_i(\phi)M_k(\psi)M_j(\theta)$$

$\phi, \psi, \theta \sim$ Eulersche Winkel

■

Eine entsprechende Zerlegung existiert für die Quaternion der Rotation.

$$q = q_i * q_k * q_j \quad , \quad |q_i| = |q_j| = |q_k| = 1$$

mit den Rotationen um die Koordinatenachsen

$$q_i = e^{i\phi/2} = \left[\cos \frac{\phi}{2}, \left(\sin \frac{\phi}{2}, 0, 0 \right) \right]$$

$$q_j = e^{j\theta/2} = \left[\cos \frac{\theta}{2}, \left(0, \sin \frac{\theta}{2}, 0 \right) \right]$$

$$q_k = e^{k\psi/2} = \left[\cos \frac{\psi}{2}, \left(0, 0, \sin \frac{\psi}{2} \right) \right]$$

Hieraus erhält man die polare Form der quaternionwertigen Fouriertransformation:

$$F^q(u, v) = (|F^q(u, v)|, \phi(u, v), \theta(u, v), \psi(u, v)) = |F^q(u, v)|e^{i\phi(u, v)}e^{k\psi(u, v)}e^{j\theta(u, v)}$$

→ Es existieren in \mathbb{H} drei Phasenwinkel ϕ, θ, ψ .

→ Rotationswinkel: $(2\phi, 2\theta, 2\psi)$

Prozedur zum Berechnen der Phasenwinkel: für jedes (u, v)

1. Normiere F^q

2. Konstruiere die Rotationsmatrix M , die durch $F^q/|F^q|$ erzeugt wird
3. Berechne die Eulerwinkel

1D: $\phi \in [-\pi, \pi[$

2D: $(\phi, \psi, \theta) \in [-\pi, \pi[\times [-\pi/4, \pi/4] \times [-\pi/2, \pi/2[$

Kapitel 3

Anwendungen von LSI-Operatoren

- Spektrale Charakterisierung von Bildern und Operatoren
- Tiefpaßfilterung und Bildglättung
- Gauß-Filter und Unschärfeprinzip
- Hochpaßfilterung und Kantendetektion
- Hierarchien von Skala und Abstraktion
- Lokale spektrale Repräsentation

3.1 Spektrale Charakterisierung von Bildern

Digitalisierung = Diskretisierung + Quantisierung

3.1.1 Diskretisierung

Diskretisierung gestattet die Betrachtung eines ortskontinuierlichen Meßsignals f_s an *diskreten* Orten (m, n) :

$$\begin{array}{ccc} f_s(x, y) & \mapsto & \hat{f}_{m,n} \\ x, y \in \mathbb{R} & & m, n \in \mathbb{N} \end{array}$$

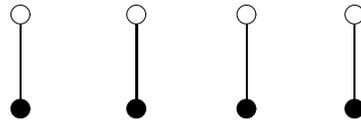
Die Diskretisierung wird realisiert durch *Abtastung* (sampling) des Signals $f_s(x, y)$ mit einem flächenhaften Sensor $a(x, y)$, der das Signal $f_s(x, y)$ über seinen Einzugsbereich lokal *integriert*. Dies entspricht der Eigenschaft lichtempfindlicher Elemente des CCD-Chips in einer Digital-kamera. Bei einem Abtastintervall $\tau \times \tau$ ergibt sich an den Positionen $(x, y) = (m\tau, n\tau)$ das diskretisierte Bildsignal $\hat{f}_{m,n}$:

$$\hat{f}_{m,n} \equiv \hat{f}_s(m\tau, n\tau) = (f_s * a)(m\tau, n\tau)$$

·	·	·	·
·	·	·	·
·	·	·	·
·	·	·	·

Die Abbildung durch integrierende Flächenelemente läßt sich als Faltung des Signals f_s mit dem Flächenelement a , gefolgt von einer *idealen Abtastung* beschreiben (Multiplikation mit der *idealen Abtastfunktion* s).

$$\hat{f} = (f_s * a) \cdot s$$



$$\hat{F} = (F_s \cdot A) * S$$

Die Funktion \hat{f} ist zunächst unendlich ausgedehnt. Durch Multiplikation mit der Aperturfunktion des Gesichtsfeldes

$$w_G = \begin{cases} 1 & \text{für } m \leq M-1, n \leq N-1 \\ 0 & \text{sonst} \end{cases}$$

entsteht folgende Korrespondenz

$$f = \hat{f} \cdot w_G \quad \circ \text{---} \bullet \quad F = \hat{F} * W_G.$$

Eine *ideale Abtastfunktion* $s(x, y)$ ist eine Folge von Einheits-Deltaimpulsen mit dem Impulsabstand τ und geht aus der Einheitsimpulsfolge (Shah-Funktion)

$$\text{III}(x) = \sum_{k \in \mathbb{Z}} \delta(x - k)$$

durch Skalierung mit $1/\tau$ hervor:

$$\begin{aligned} s(x, y) &= \frac{1}{\tau^2} \text{III}\left(\frac{x}{\tau}, \frac{y}{\tau}\right) \\ &= \frac{1}{\tau} \text{III}\left(\frac{x}{\tau}\right) \cdot \frac{1}{\tau} \text{III}\left(\frac{y}{\tau}\right) \\ &= \left(\sum_{k \in \mathbb{Z}} \delta(x - k\tau) \right) \left(\sum_{l \in \mathbb{Z}} \delta(y - l\tau) \right) \\ &= \sum_{k, l \in \mathbb{Z}} \delta(x - k\tau, y - l\tau) \end{aligned}$$

Die Fouriertransformierte der Einheitsimpulsfolge ist wieder eine Einheitsimpulsfolge:

$$\text{III}(x) \quad \circ \text{---} \bullet \quad \text{III}(u)$$

Daher folgt für die Fouriertransformierte S der Abtastfunktion aus dem Ähnlichkeitssatz der Fouriertransformation:

$$s(x, y) = \frac{1}{\tau^2} \text{III}\left(\frac{x}{\tau}, \frac{y}{\tau}\right) \quad \circ \text{---} \bullet \quad S(u, v) = \text{III}(u\tau, v\tau)$$

Eine Folge von Impulsen mit Abstand τ im Ortsraum korrespondiert mit einer Folge von Impulsen mit Abstand $1/\tau$ im Frequenzraum. Die Abtastung eines Signals f im Ortsraum (die Modulation mit der Impulsfolge s) bewirkt also die additive Überlagerung der periodischen Wiederholung des Spektrums F (die Faltung mit der Impulsfolge S). Wegen der Reziprozität von Ort und Frequenz gilt dann:

- ▶ Je größer der Abtastabstand, desto kleiner der Abstand der Wiederholungen des Originalspektrums, d.h. um so stärker deren Überlappung.

3.1.2 Abtasttheorem

Eine hinreichende Bedingung für die fehlerfreie Rekonstruktion eines kontinuierlichen Signals aus (bzw. die vollständige Repräsentation in) seiner abgetasteten Form ist die Einhaltung des *Abtasttheorems*. Danach hängt das Abtastintervall τ vom spektralen Gehalt des kontinuierlichen Signals ab. Bezeichnet man:

$$\begin{aligned} u_s &= \frac{1}{\tau} \quad \text{als Abtastfrequenz bzw. Abtastrate und} \\ u_g &\quad \text{als obere Grenzfrequenz des kontinuierlichen Signals} \end{aligned}$$

so lautet das Abtasttheorem:

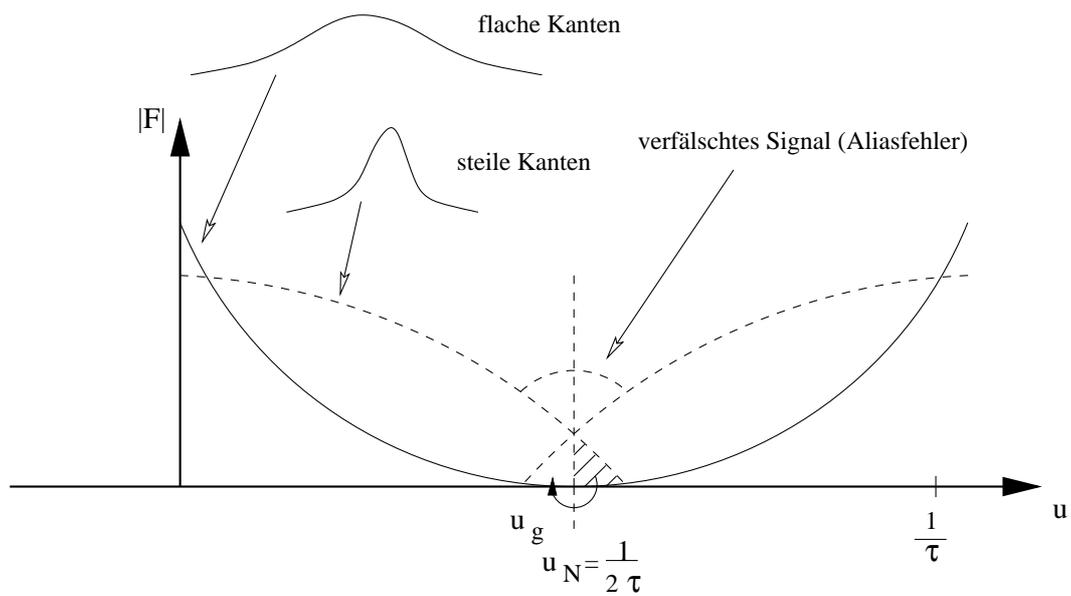
$$u_g < \frac{1}{2}u_s \quad \text{oder} \quad u_s > 2u_g$$

- ▶ Die Fouriertransformierte der Abtastrate muß mindestens doppelt so groß sein wie die obere Grenzfrequenz des Signals.
- ▶ Für die Einhaltung des Abtasttheorems muß also das kontinuierliche Signal bandbegrenzt sein. Ist diese Bedingung nicht erfüllt, muß diese Eigenschaft durch Tiefpaßfilterung vor der Abtastung hergestellt werden.
- ▶ Ein integrierendes Flächenelement bei der Abtastung besitzt die Wirkung eines Tiefpaßfilters, reicht aber oft nicht für die Bandbegrenzung aus, so daß eine zusätzliche Filterung des Signals notwendig wird.
- ▶ Bei Verletzung des Abtasttheorems kommt es zu einer Überlappung der durch die Abtastung entstandenen Wiederholungspektren, die zu einer Verfälschung des rekonstruierten Signals führt. Der von der Überlappung herrührende Fehler heißt *Aliasfehler* (*aliasing*).
- ▶ Ein Beispiel für die Verletzung des zeitlichen Abtasttheorems beim Filmen ist, daß zu schnell drehende Wagenräder so erscheinen, als würden sie sich rückwärts drehen (Wagenradedefekt).

Die *Nyquist-Frequenz*

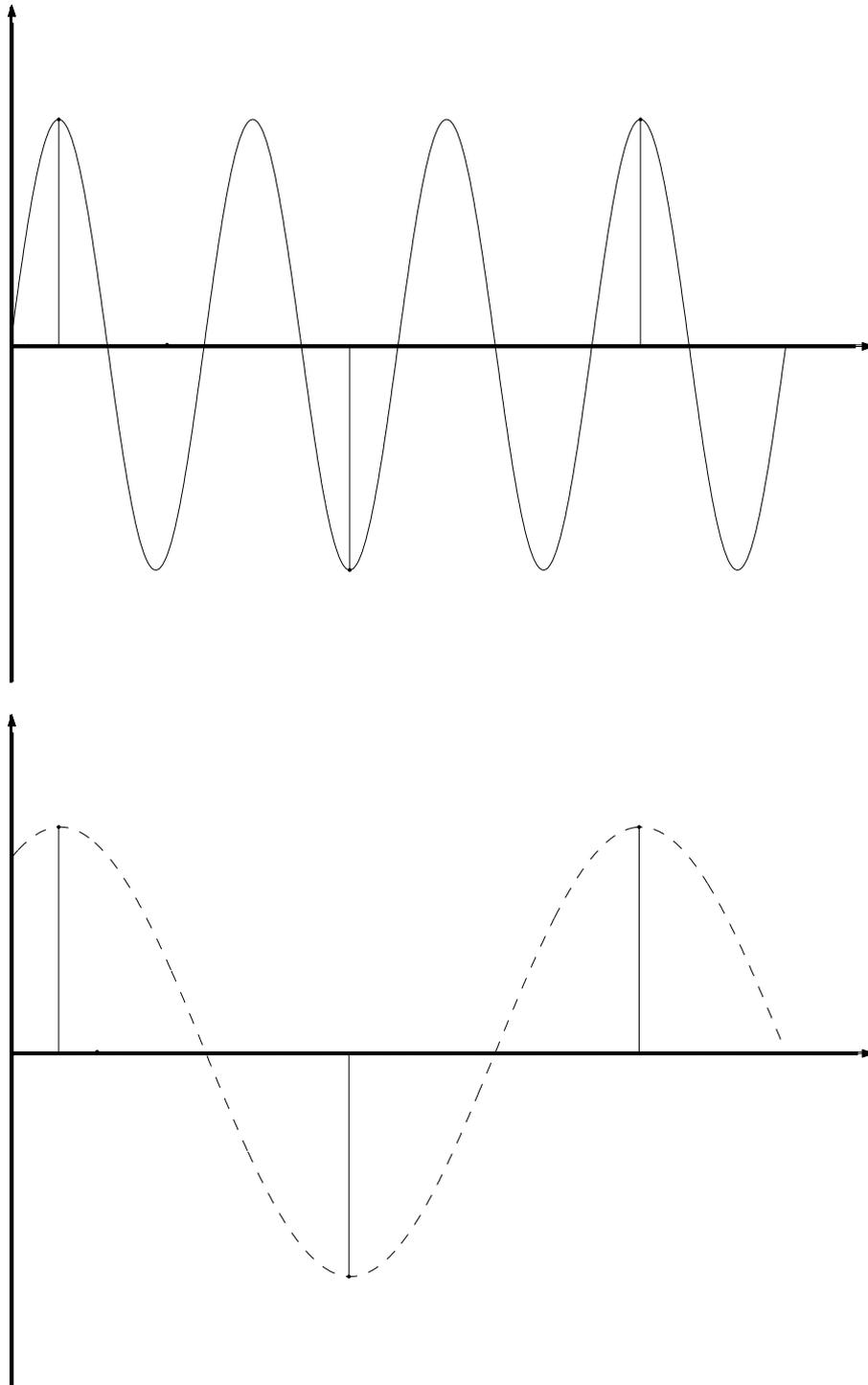
$$u_N = \frac{1}{2\tau} = \frac{u_s}{2}$$

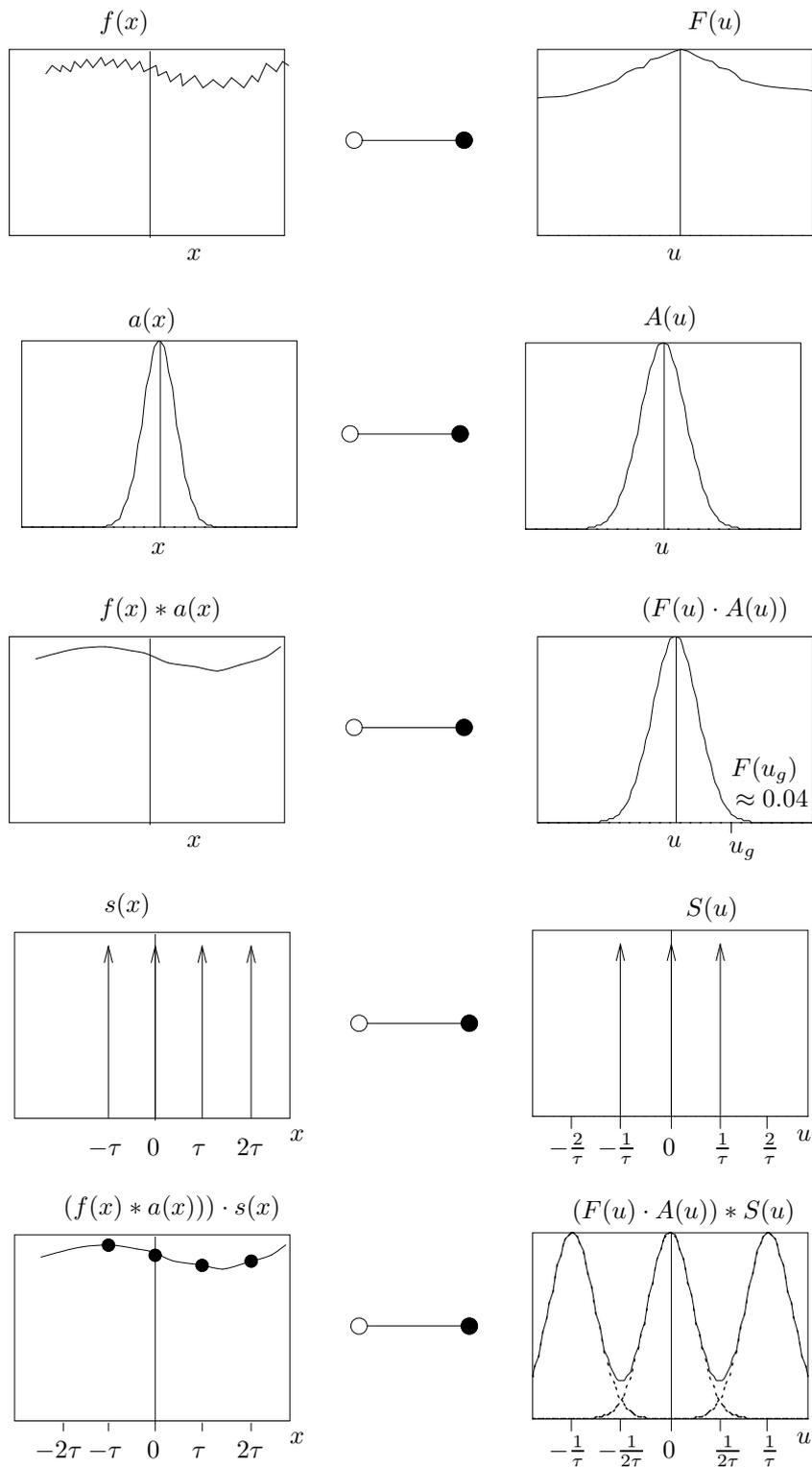
muß zur Grenzfrequenz u_g des Signals in der Relation $u_N > u_g$ stehen.



Eindimensionales Aliasing

Oben sieht man die Unterabtastung eines eindimensionalen Wellensignals und unten das rekonstruierte Signal.



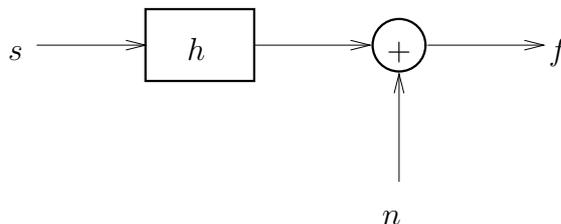


3.1.3 Bildgüteparameter

Das signaltheoretische Modell der Bilderzeugung beruht auf folgenden Annahmen:

- Homogenität von h

- additives, d.h. signalunabhängiges Rauschen n



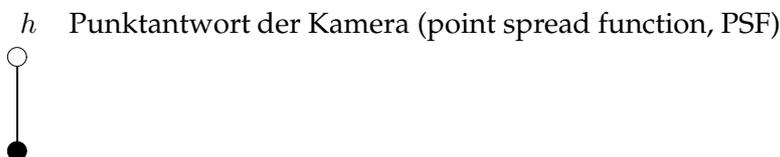
Beobachtungsgleichung: $f = s * h + n$

- h : Punktantwort der Kamera — deterministische Signalstörung, entspricht dem Flächenelement a der Abtastung
- n : Rauschen — stochastische Signalstörung

Deterministische Bildgüteparameter

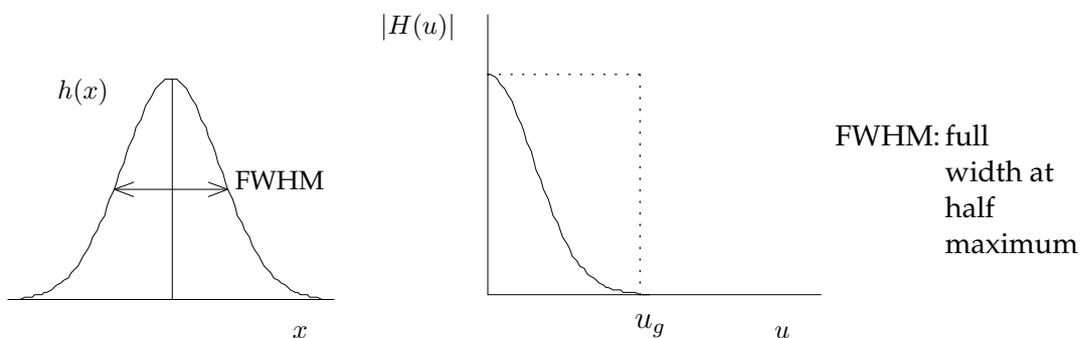
Die optische Abbildung durch die Kamera wird beschrieben durch eine systemcharakterisierende Funktion:

$$h(x, y) \quad \text{○} \text{---} \text{●} \quad H(u, v)$$



- H Übertragungsfunktion des Systems (system transfer function)
- $|H|$ Modulationsübertragungsfunktion (MTF)

Eigentlich gilt $h = g \cdot \text{PSF}$, wobei g den Gewinn (Empfindlichkeit) ausdrückt und sich demzufolge auf die Bildintensität auswirkt. Hier sei die Normierung $g = 1$ für alle Orte (x, y) angenommen.



Es gilt

$$u_g = \frac{1}{FWHM}$$

- ▶ Die räumliche Ausdehnung der Punktantwort (Punktverbreiterungsfunktion) bestimmt die Auflösung der Abbildung. Kleine Punktantworten bedeuten hohe Auflösung und umgekehrt.
- ▶ Die Modulationsübertragungsfunktion gibt die relative Kontrastdämpfung an oder die Amplitude der Abbildung einer harmonischen Schwingung mit der Frequenz u .

Stochastische Bildgüteparameter

Der *Signal–Rauschabstand* SNR (*signal to noise ratio*) oder kurz Rauschabstand ist definiert als das logarithmierte Verhältnis von Signalenergie zu Rauschenergie (Fehlerenergie). Die Einheit des Signal–Rauschabstands ist das Dezibel (dB).

$$\text{SNR} = 10 \log_{10} \frac{\sum_{k,l} f_{k,l}^2}{\sum_{k,l} n_{k,l}^2} \text{ dB}$$

- ▶ Je höher das SNR, desto geringer ist die Störung durch das Rauschen.
- ▶ Typische Werte sind:

Verfahren	Amplitudenverhältnis	Rauschabstand
Video	≈ 100	40 dB
CCD	500 ... 1000	54 ... 60 dB
Szintigraphie	1,5 ... 25	3,5 ... 28 dB

3.1.4 Darstellung von Spektren

Zentrierung

Der Algorithmus der DFT legt den Gleichanteil des Spektrums $F_{0,0}$ in eine der Bildecken. Zur korrekten Positionierung der negativen Frequenzen müssen wir die DFT entlang der Achsen um $M/2$ bzw. $N/2$ zyklisch verschieben. Gleichwertig dazu ist die Vorzeichenumkehr jedes zweiten Pixels von f :

$$f_{m,n}(-1)^{m+n} \quad \circ \text{---} \bullet \quad F_{u-\frac{M}{2}, v-\frac{N}{2}}$$

Logarithmierung

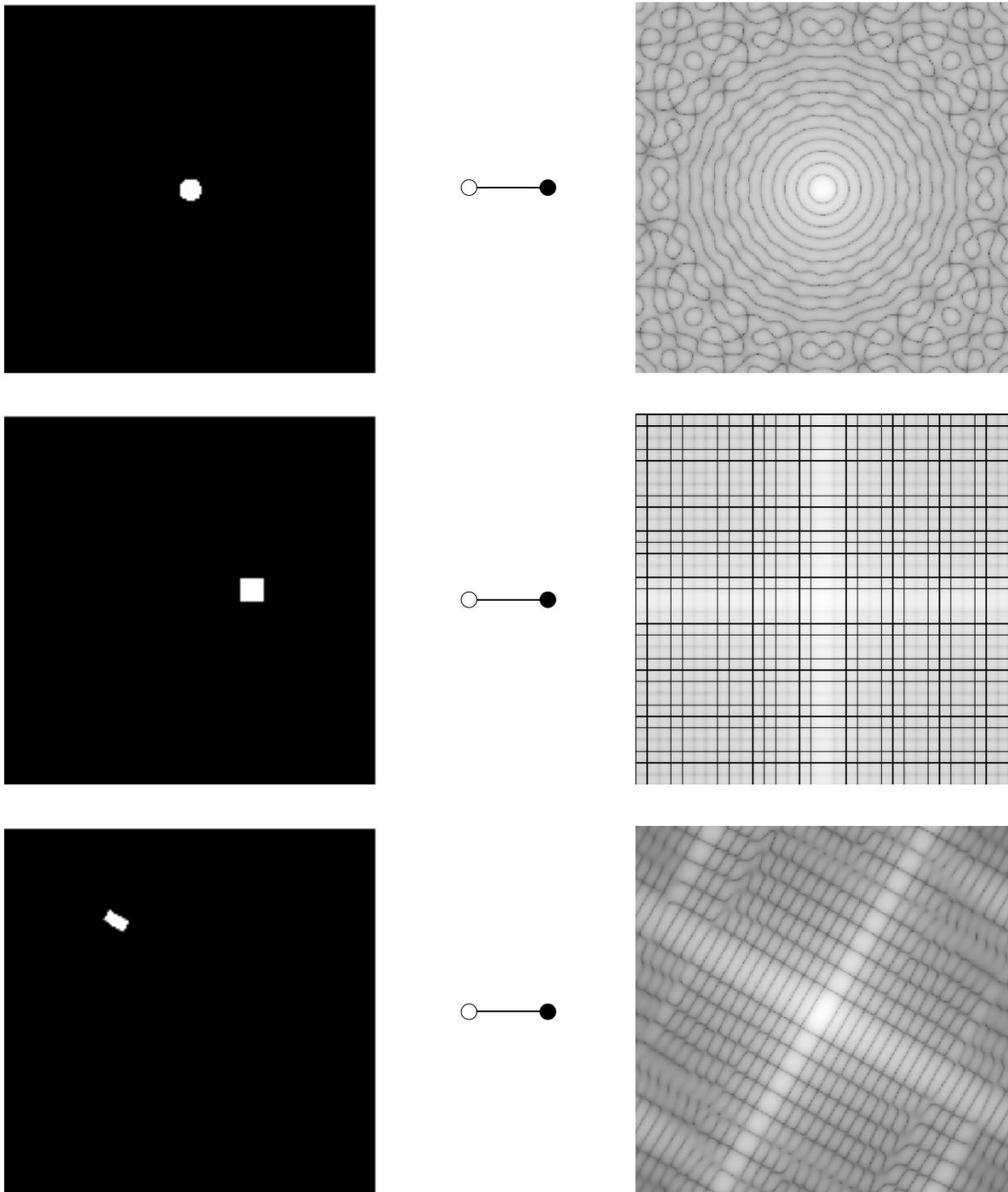
Das Spektrum natürlicher Bilder weist einen starken Abfall der Amplitude mit wachsender Frequenz auf. Typisch ist:

$$|F_{u,v}| \sim \frac{1}{\sqrt{u^2 + v^2}}$$

Erinnerung: $F_{0,0}$ repräsentiert die Summe / das Integral aller Funktionswerte des Definitionsbereiches! Der Helligkeitsumfang, den das visuelle System erfassen kann, ist auf etwa 32 Stufen begrenzt, so daß große Teile des Spektrums bei der Darstellung auf einem Bildschirm so nicht

wahrnehmbar wären. Als geeignete Hilfsmaßnahme gilt die Anhebung kleiner Amplituden durch Logarithmieren:

$$F_{u,v} \mapsto \log(1 + |F_{u,v}|)$$

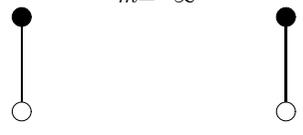


Amplitudenspektren (logarithmiert und normiert) einiger 2D-Funktionen

3.2 Spektrale Charakterisierung von LSI-Operatoren

3.2.1 Übertragungsfunktion eines LSI-Operators

Die kontinuierliche und periodische Übertragungsfunktion $H(e^{j\omega})$ eines diskreten 1D-Operators lautet:

$$H(e^{j\omega}) = \sum_{m=-\infty}^{\infty} h_m e^{-j\omega m}$$


$$h_m = \sum_{k=-\infty}^{\infty} h_k \delta(m - k)$$

- ▶ Die Übertragungsfunktion $H(e^{j\omega})$ eines *diskreten* Systems ist *periodisch* in ω mit der Periode 2π .
- ▶ Oft wird die Übertragungsfunktion eines LSI-Operators auch über den Umweg der Diskretisierung einer kontinuierlichen Impulsantwort, von der die Transformation bekannt ist, berechnet. Die Übertragungsfunktion der diskreten Impulsantwort ergibt sich dabei durch die Summation der im Spektralbereich periodisch wiederholten Übertragungsfunktion der kontinuierlichen Impulsantwort.
- ▶ Die *komplexharmonische* Funktion $f = e^{j2\pi x u_0}$ ist eine Eigenfunktion der Zirkulärmatrix C_h aller LSI-Operatoren bzw. des Faltungsoperators C_h . Die Übertragungsfunktion $H(u_0)$ ist ein Eigenwert *eines bestimmten* LSI-Operators h . Die *Eigenwertgleichung* lautet:

$$C_h f = \lambda f$$

~ Wirkung eines Operators mit Impulsantwort h auf ein spezielles Signal f

Die Eigenwertgleichung ergibt sich aus

$$g(x) = (h * f)(x) \quad \circ \text{---} \bullet \quad G(u) = (H F)(u)$$

$$g(x) = \mathcal{F}^{-1}\{(H F)(u)\}$$

$$g(x) = \int_{-\infty}^{\infty} H(u) F(u) e^{j2\pi u x} du$$

mit der Annahme $f(x) = e^{j2\pi u_0 x} \quad \circ \text{---} \bullet \quad F(u) = \delta(u - u_0)$

$$g(x) = \int_{-\infty}^{\infty} H(u) \delta(u - u_0) e^{j2\pi u x} du$$

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

$$g(x) = H(u_0) e^{j2\pi u_0 x}$$

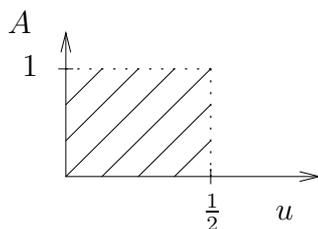
3.2.2 Amplitudenübertragung eines LSI-Operators

Hier wollen wir annehmen, daß die Impulsantwort eine kontinuierliche und unendlich ausgedehnte Funktion ist. Die *Amplitudenübertragung* (*Amplitudengang*, *Betragspektrum*) ist der Betrag der Übertragungsfunktion:

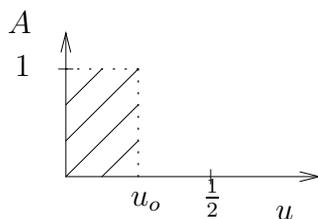
$$A(u) = |H(u)|$$

Übertragungsverhalten einiger Operatoren

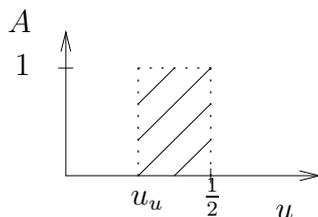
1. idealer Allpaß AP (entspricht dem Identitätsoperator im Ortsraum)



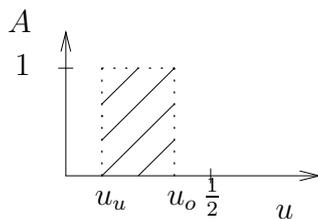
2. idealer Tiefpaß TP (eingesetzt als Glättungsfilter)



3. idealer Hochpaß HP (eingesetzt als Ableitungsfiler)



4. idealer Bandpaß BP (Serienschaltung von Hoch- und Tiefpaß)



3.2.3 Phasenübertragung eines LSI-Operators

Die Phasenübertragung (Phasengang) $\phi(u)$ ist das Argument der komplexen Übertragungsfunktion $H(u) = a(u) + j b(u)$.

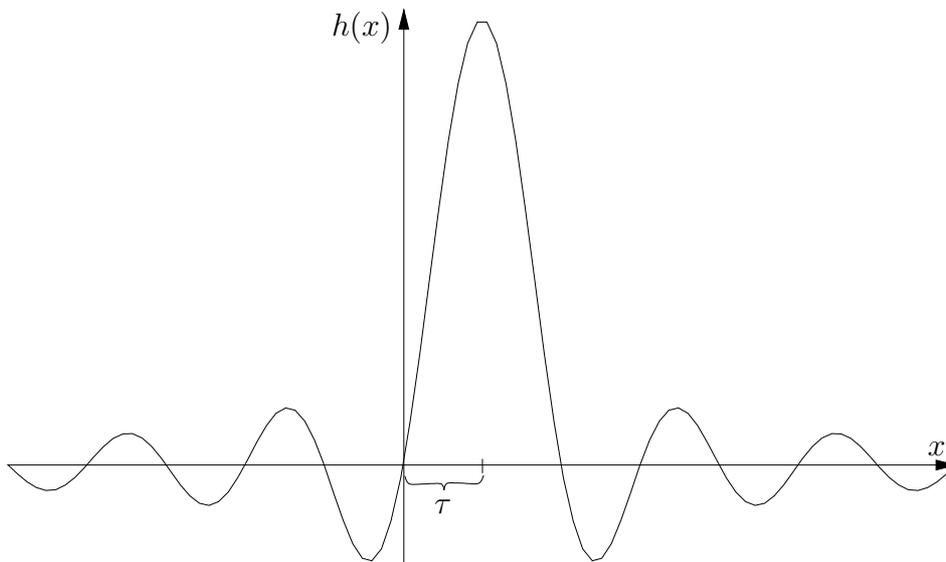
$$\begin{aligned} \phi(u) &= \arg H(u) \\ &= \begin{cases} \arctan \frac{b(u)}{a(u)} & ; a(u) \geq 0 \\ \arctan \frac{b(u)}{a(u)} + \pi & ; a(u) < 0 \wedge b(u) \geq 0 \\ \arctan \frac{b(u)}{a(u)} - \pi & ; a(u) < 0 \wedge b(u) < 0 \\ 0 & ; a(u) = 0 \wedge b(u) = 0 \end{cases} \end{aligned}$$

- ▶ Die Phase ist nur für Frequenzen definiert, die zur Amplitudenübertragung beitragen, also für $A \neq 0$. Für den Fall $A = 0$ kann die Phase willkürlich auf 0 gesetzt werden.
- ▶ Ein idealer TP hat einen streng linearen Phasengang (*Linearphasensystem*).

Beispiel eines Linearphasensystems

Ein idealer Tiefpaß $H(u)$ mit oberer Grenzfrequenz u_o entspricht im Ortsraum einer Impulsantwort in Gestalt einer (unendlich ausgedehnten) sinc-Funktion.

$$h(x) = \frac{\sin [2\pi u_o(x - \tau)]}{\pi(x - \tau)}$$



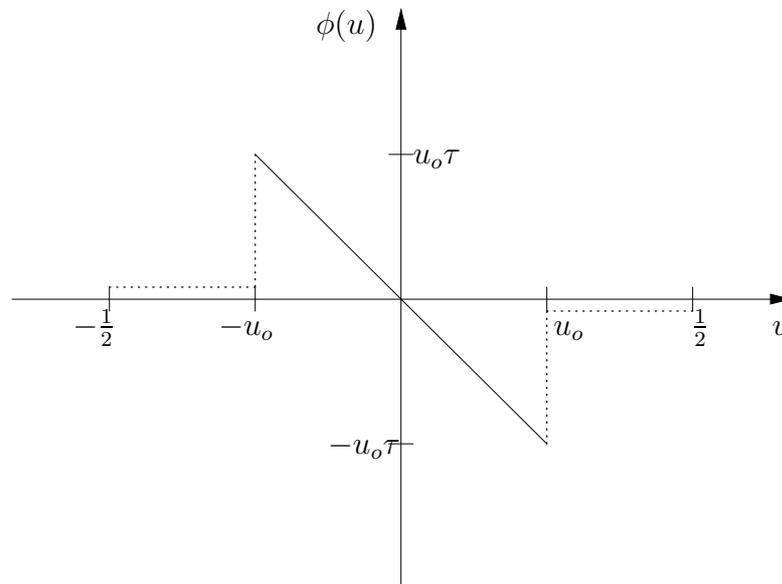
Ein idealer Tiefpaß hat einen linearen Phasengang.

$$\phi(u) = -\tau u$$

1. Die Phase ist proportional zur Frequenz

2. Die Phase ist proportional zur Verschiebung des Symmetriezentrums

► für $\tau = 0$: $\phi(u) = 0 \sim$ Nullphasensystem !



■

3.2.4 Filterung im Frequenzraum

Der Faltungssatz der FT erlaubt es, die Filterung mit einer im Ortsraum entworfenen Impulsantwort h auch im Frequenzraum durchzuführen.

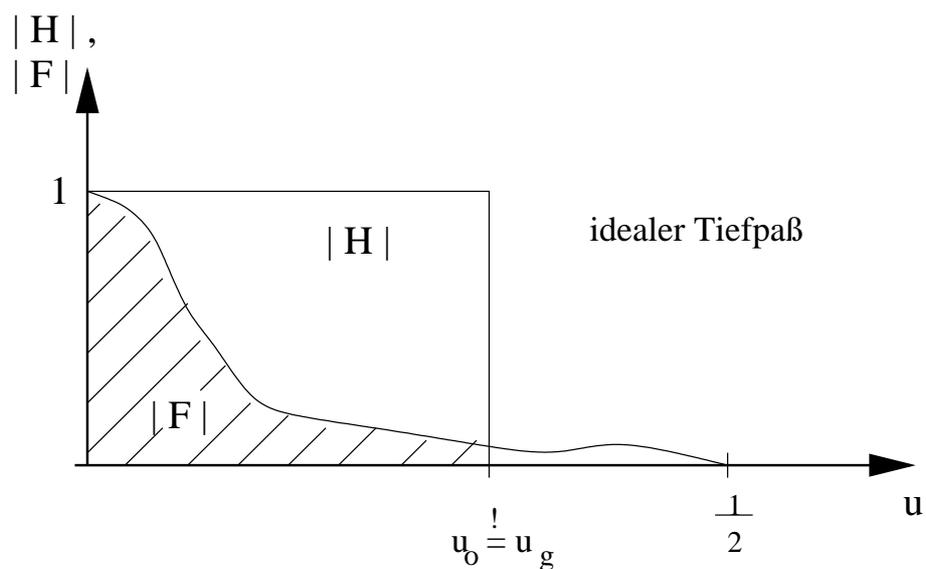
► *praktischer* Aspekt: Die Filterung im Frequenzraum besitzt ab einer gewissen Faltungsmaschengröße gegenüber der Faltung im Ortsraum eine geringere Komplexität.

$$g = \mathcal{F}^{-1} \{ \mathcal{F}\{f\} \mathcal{F}\{h\} \}$$

► Filter können auch im Frequenzraum entworfen werden. Dann gilt für ihre Anwendung:

$$g = \mathcal{F}^{-1} \{ \mathcal{F}\{f\} H \}$$

Es muß aber beachtet werden, daß im Frequenzraum entworfene Filter in manchen Fällen (z.B. "ideale" Filter mit scharfen Frequenzbegrenzungen) unbegrenzte Ortsausdehnung besitzen. Sollten derartige Filter im Ortsraum durch Faltung angewendet werden, ist ein Abschneiden der Impulsantwort erforderlich. Abschneiden im Ortsraum führt zu Approximationen der Impulsantwort mit deutlichen Konsequenzen im Frequenzraum.



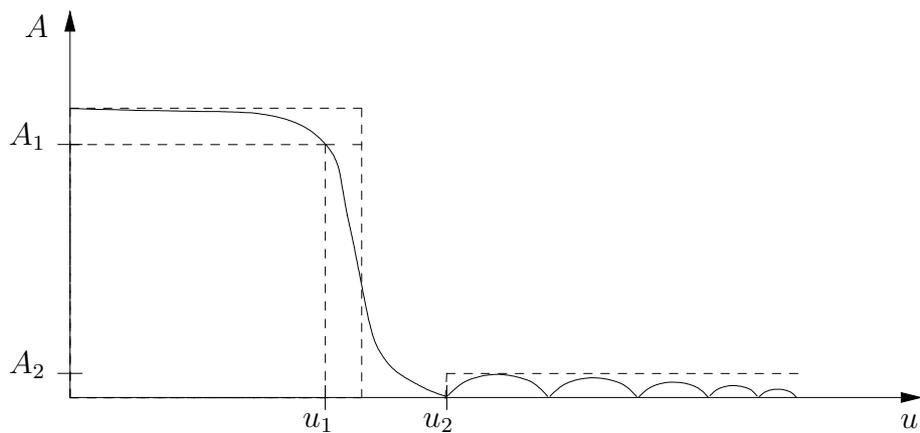
Approximation idealer Amplitudenübertragung

durch (begrenzte) Ortsraumfilter.

Die tatsächliche erreichbare Amplitudenübertragung ist gekennzeichnet durch

- unscharfe Grenzfrequenzen
- Modulation der Übertragungsfunktion

Das ist eine Folge der Faltung der Amplitudenfunktion mit der Fouriertransformierten der Abschneidefunktion im Ortsraum (sinc-Funktion).



Durchlaßbereich: $0 < u < u_1$ mit $A_1 \leq A \leq 1$

Sperrbereich: $u > u_2$ mit $0 \leq A \leq A_2$

Übergangsbereich: $u_1 \leq u \leq u_2$

3.3 Bildglättung mit Rechteckfiltern

Das Bild „Lena“ wurde durch Rauschen gestört (links) und danach mit einem Bildglättungsoperator behandelt (rechts).



- Bildglättung durch *Tiefpaßfilter* im Frequenzraum (z.B. für 3D-Bilder)
- Bildglättung durch *Mittelwertbildung* im Ortsraum
 - Anwendung von Wahrscheinlichkeitstheorie und Statistik
 - Bild ist *stochastisches Ereignisfeld (random field)*, d.h. Realisierung eines stochastischen Prozesses mit bestimmten Eigenschaften

aber: Die Regionen eines Bildes haben unterschiedliche stochastische Eigenschaften → sie entstammen unterschiedlichen stochastischen Prozessen

- ▶ Problem bei der Anwendung von LSI-Operatoren zur Mittelwertbildung, da Meßwerte nicht einer homogenen statistischen Population zugeordnet werden können (z.B. wenn Operatorfenster eine Regionengrenze überdeckt)
- ▶ Verletzung der *Ergodenhypothese*:
Populationsmittelwert \neq lokal geschätzter Mittelwert

Ausweg:

1. Folge unterschiedlich großer Supports der Operatorfenster, um den Einflußbereich der Verletzung der Ergodenhypothese zu minimieren → Auflösungspyramiden oder lokal adaptive lineare Filter
2. nichtlineare 3–Stufenfilter
 - a) Überprüfung der Ergodenhypothese
 - b) Definition der Regionengrenzen innerhalb Support (Schätzung)
 - c) innerhalb der identifizierten Regionen nun: lineare Mittelwertbildung

Beispiel: Frequenzübertragungsfunktion einer diskreten Impulsantwort, $d = 1, K = 3$:

$$h = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

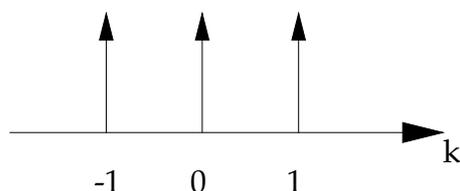
↑

Aufpunkt, d.h. Ursprung des lokalen
Koordinatensystems bei Faltung

≙ ungewichtete Mittelwertbildung in 3er Umgebung am Aufpunkt

betrachten Impulsantwort $h' = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ als Folge von Einheitsimpulsen δ_{k-m} ($m = -1, 0, 1$) mit

$$\delta_{k-m} = \begin{cases} 1 & \text{für } k = m \\ 0 & \text{sonst} \end{cases}$$



Wegen der Definition der FTD benötigen wir eine unendlich ausgedehnte diskrete Funktion h' :

$$h'_k = \begin{cases} \sum_{m=-\infty}^{\infty} \delta_{k-m} & \text{für } k = -1, 0, 1 \quad (\sim \text{drei kanonische Einheitsfunktionen}) \\ 0 & \text{sonst} \end{cases}$$

Hieraus folgt für

$$\begin{aligned} H'(e^{j\omega}) &= \sum_{k=-\infty}^{\infty} h'_k e^{-j\omega k} \\ &= \sum_{k=-1}^1 \sum_{m=-\infty}^{\infty} \delta_{k-m} e^{-j\omega k} \end{aligned}$$

da die Beiträge für alle anderen k verschwinden

$$= \sum_{k=-1}^1 \sum_{m=-1}^1 \delta_{k-m} e^{-j\omega k}$$

wegen der Definition der δ -Funktion

$$\begin{aligned} &= \sum_{k=-1}^1 (\delta_{k+1} + \delta_k + \delta_{k-1}) e^{-j\omega k} \\ &= e^{j\omega} + e^0 + e^{-j\omega} \end{aligned}$$

$$\text{da } \delta_{k+1}e^{-j\omega k} = \begin{cases} e^{+j\omega} & \text{für } k = -1 \\ 0 & \text{sonst} \end{cases} \text{ usw.}$$

$$= 1 + e^{j\omega} + e^{-j\omega} = 1 + 2 \cos \omega$$

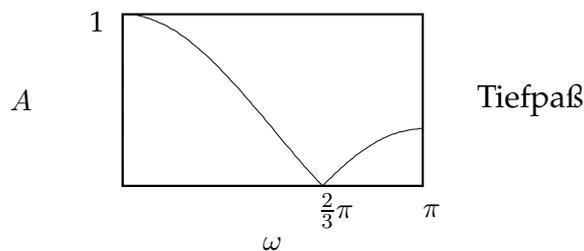
Es gilt: $H'(e^0) = 3$ für $\omega = 0$

$$\rightarrow \text{Normierung: } H(e^{j\omega}) = \frac{H'(e^{j\omega})}{H'(e^0)} = \frac{1}{3}(1 + 2 \cos \omega)$$



Diskussion der Frequenzübertragungsfunktion: $A = |H|$

ω	0	$\frac{\pi}{2}$	π	$\frac{2}{3}\pi$
A	1	$\frac{1}{3}$	$\frac{1}{3}$	0



- $K = 3$: 1 Seitenband bei $\omega_s = \pi$
- 1 Nullstelle bei $\omega_z = \frac{2}{3}\pi$

$d = 2$:

$$h = h_v * h_h = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} * \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$H(e^{j\omega_h}, e^{j\omega_v}) = \frac{1}{9}(1 + 2 \cos \omega_h, 1 + 2 \cos \omega_v)$$

~ oberhalb Frequenz $\omega_z = \frac{2}{3}\pi$ werden Strukturen nicht gelöscht, sondern um minimal $\frac{2}{3}$ gedämpft

~ unterhalb Frequenz ω_z auch nur frequenzabhängiges gedämpftes Durchlassen

Frage: Eventuell stärkere Glättung durch $K > 3$?

$$H'(e^{j\omega}) = \sum_{m=0}^{K-1} \sum_{k=0}^{K-1} \delta_{k-m} e^{-j\omega k} = \sum_{k=0}^{K-1} e^{-j\omega k}$$

$\hat{=}$ geometrische Reihe über den Gliedern der Reihe $a_k = e^{-j\omega k}$ mit Anfangsglied $a_0 = e^0 = 1$ und Quotient $q = e^{-j\omega}$.

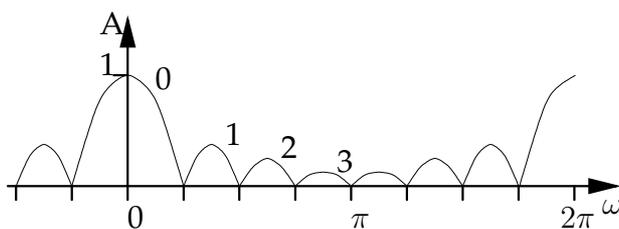
Daraus folgt für $H'(e^{j\omega})$ als Summe der geometrischen Reihe

$$\begin{aligned} H'(e^{j\omega}) &= \frac{1 - e^{-j\omega K}}{1 - e^{-j\omega}} = \frac{e^{-j\omega K/2}}{e^{-j\omega/2}} \cdot \frac{e^{j\omega K/2} - e^{-j\omega K/2}}{e^{j\omega/2} - e^{-j\omega/2}} \\ &= e^{-j\omega(K-1)/2} \cdot \frac{\sin(\omega K/2)}{\sin(\omega/2)} \\ &\quad \begin{array}{c} \uparrow \\ \text{Phasenterm} \end{array} \quad \begin{array}{c} \uparrow \\ \text{Amplitudenterm} \end{array} \end{aligned}$$

K ungerade : Phasenterm wird identisch Eins (Nullphasenfilter), da Symmetriezentrum der Impulsantwort mit Ursprung des Koordinatensystems zusammenfällt.

K gerade : Phasenterm drückt Verschiebung um 0.5 Pixel aus, da Koordinatenursprung nicht mit Symmetriezentrum zusammenfällt.

Beispiel: $K = 8$



für $0 \leq \omega \leq \pi$:

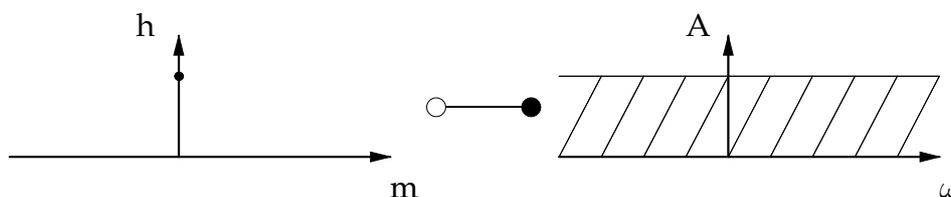
$$\begin{aligned} \text{Seitenbänder: } \omega_s &= \frac{2\pi s}{K-1} \quad \text{für } s = 1, 2, \dots, \lfloor \frac{1}{2}(K-1) \rfloor \\ \text{Nullstellen: } \omega_z &= \frac{2\pi z}{K} \quad \text{für } z = 1, 2, \dots, \lceil \frac{1}{2}(K-1) \rceil \end{aligned}$$

$\lfloor \cdot \rfloor$ ist die untere Gaußsche Klammer und $\lceil \cdot \rceil$ ist die obere Gaußsche Klammer.

► Mit wachsender Breite K schrumpft das Hauptband \sim stärkere Tiefpaßwirkung

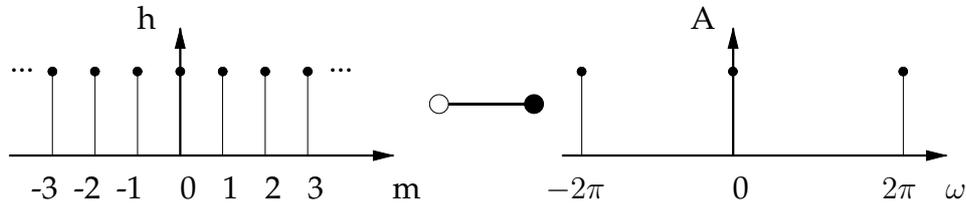
► Mit wachsender Breite K der Rechteckfolge gehen die Seitenbänder gegen Null

► einzelner Impuls: $h_m = \delta_0$  $H(e^{j\omega}) = \text{const}$ mit $1 = \int_{-\pi}^{\pi} H(e^{j\omega}) d\omega$



► unendliche Impulsfolge:

$$h_m = \sum_{k=-\infty}^{\infty} \delta_{m-k} \quad \circ \text{---} \bullet \quad H(e^{j\omega}) = \delta_0$$



3.4 Bildglättung mit Binomial- und Gaußfiltern

3.4.1 Bildung von 1D-Binomialmasken

Die 1D-Maske $h^{(n)} = 2^{-n}b^{(n)}$ heißt *Binomialfilter* n ter Ordnung. Die Binomialkoeffizienten der Ordnung n , $b^{(n)}$, folgen einer durch das Pascalsche Dreieck gegebenen Bildungsregel.

Pascalsches Dreieck

n	$b^{(n)}$	σ^2
0	1	0
1	1 1	$\frac{1}{4}$
2	1 2 1	$\frac{2}{4}$
3	1 3 3 1	$\frac{3}{4}$
4	1 4 6 4 1	$\frac{4}{4}$
5	1 5 10 10 5 1	$\frac{5}{4}$

$$1D : \quad h^{(2)} = \frac{1}{4} \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \quad \circ \text{---} \bullet \quad H(e^{j\omega}) = \frac{1}{2}(1 + \cos \omega)$$

$$2D : \quad h^{(2)} = \frac{1}{4} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix} * \frac{1}{4} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

$$\begin{matrix} \bullet \\ | \\ \circ \end{matrix} \quad H(e^{j\omega_x}, e^{j\omega_y}) = \frac{1}{4} (1 + \cos \omega_x) (1 + \cos \omega_y)$$

3.4.2 Bildung von 2D-Binomialmasken

Eine Binomialmaske der Ordnung n kann entweder direkt aus den Binomialkoeffizienten

$$h_{k,l}^{(n)} = \frac{1}{2^{2n}} \frac{n!}{(\frac{n}{2}-k)!(\frac{n}{2}+k)!} \frac{n!}{(\frac{n}{2}-l)!(\frac{n}{2}+l)!} \quad k, l = -\frac{n}{2}, \dots, \frac{n}{2}$$

$$\text{mit } K = L = n + 1$$

oder durch n -fache lineare Autofaltung des Binomialfilters 1. Ordnung (das ist ein Boxfilter)

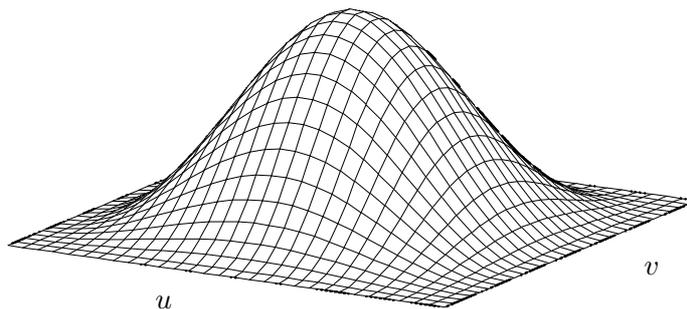
$$h^{(n)} = \frac{1}{2^{2n}} (h^{(1)})^{*n} = \frac{1}{2^{2n}} \underbrace{(h^{(1)} * h^{(1)} * \dots * h^{(1)})}_{n \text{ mal}}$$

gewonnen werden.

Beispiele:

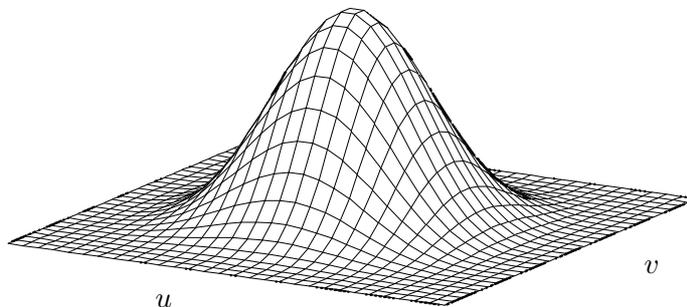
$$h^{(2)} = h_x^{(2)} * h_y^{(2)}$$

1	2	1
2	4	2
1	2	1



$$h^{(4)} = h_x^{(4)} * h_y^{(4)}$$

1	4	6	4	1
4	16	24	16	4
6	24	36	24	6
4	16	24	16	4
1	4	6	4	1



■

Verallgemeinerung:

$$h^{(1)} = 2^{-2} \Pi_2 \quad \text{mit} \quad \Pi_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\Pi_\kappa = \underbrace{\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ 1 & \cdots & & 1 \end{bmatrix}}_{\kappa \times \kappa\text{-Rechteckfilter}}$$

Die Glockenkurve Γ_K läßt sich erzeugen durch:

$$\Gamma_K = (\Pi_\kappa)^{*p} = \underbrace{\Pi_\kappa * \Pi_\kappa * \cdots * \Pi_\kappa}_{p \text{ Faktoren}} \quad \text{mit } K = p(\kappa - 1) + 1$$

und approximiert eine Gaußkurve mit der Varianz

$$\sigma_{\kappa p}^2 = \frac{p\kappa^2}{12}$$

Diese Beziehung folgt aus dem *zentralen Grenzwertsatz* der Wahrscheinlichkeitsrechnung.

Beispiel:

$$\begin{aligned} \Gamma_2 &= \Pi_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ \Gamma_3 &= \Pi_2 * \Pi_2 = \underbrace{\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}}_{\text{Dreiecksfilter}} \equiv \Lambda_3 \\ \Gamma_4 &= \Gamma_3 * \Pi_2 \\ \Gamma_5 &= \Gamma_4 * \Pi_2 \quad (\text{kubischer B-Spline}) \\ \Gamma_K &= \Gamma_{K-1} * \Pi_2 \end{aligned}$$

■

3.4.3 Eigenschaften der Gaußfunktion als LSI-Operator

Die Gaußfunktion g_σ ist die Greensche Funktion der Diffusionsgleichung. Hieraus leitet sich ihre große Bedeutung in der Physik aber auch in Computer Vision ab. Die Diffusionsgleichung einer Funktion $u(x, t)$

$$\frac{\partial}{\partial t} u = \Delta u$$

mit der Anfangsbedingung $u(x, 0) = f(x)$ hat die Lösung

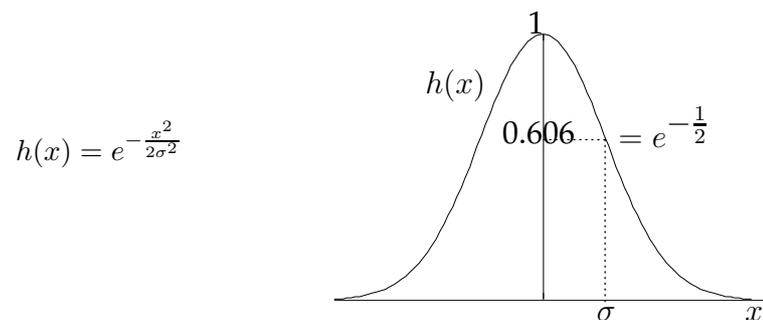
$$u(x, t) = \begin{cases} f(x) & \text{für } t = 0 \\ (g_\sigma * f)(x) & \text{für } t > 0 \end{cases}$$

mit $\sigma = \sqrt{2t}$.

Die Gaußfunktion hat wegen der folgenden Eigenschaften eine besondere Bedeutung als LSI-Operator.

1. Die Gaußfunktion ist eine Wichtungsfunktion, die durch Faltung im Ortsraum ein geglättetes Signal erzeugt (gewichtetes Glättungsfilter) bzw. durch Multiplikation in Frequenzen als ein gewichtetes Tiefpaßfilter interpretiert werden kann. Die Gaußfunktion kann in unterschiedlicher Weise repräsentiert werden.

- Auf das Maximum normierte Gaußfunktion



- Die flächennormierten Gaußfunktionen haben die Form:

$$h_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \cdot e^{-\frac{x^2}{2\sigma_x^2}}$$

$$h_y(y) = \frac{1}{\sqrt{2\pi}\sigma_y} \cdot e^{-\frac{y^2}{2\sigma_y^2}}$$

wobei $\sqrt{2\pi}\sigma$ die Fläche unter einer Gaußkurve ist.

In praktischen Anwendungen wird die Gaußfunktion wegen ihrer unendlichen Ausdehnung bei $x \geq 5\sigma \dots 8\sigma$ abgeschnitten. Hierdurch entstehen bezüglich der Fläche diese Funktion Fehler von 1% bis 0.1%.

2. Die Gaußfunktion ist die einzige reellwertige Funktion, die *separabel* und *isotrop* ist.

- *separabel* in den kartesischen Koordinaten x und y heißt:

$$h_c(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{x^2}{2\sigma_x^2}} e^{-\frac{y^2}{2\sigma_y^2}}$$

$$= h_x(x) h_y(y)$$

- *isotrop*, d.h. in Polarkoordinaten unabhängig vom Winkel (keine Richtungsabhängigkeit), heißt:

$$h_p(r, \varphi) = h_r(r) h_\varphi(\varphi) \text{ mit } h_r(r) = \frac{1}{2\pi\sigma_r^2} e^{-\frac{r^2}{2\sigma_r^2}} \text{ und } h_\varphi(\varphi) = \text{const} (\stackrel{!}{=} 1)$$

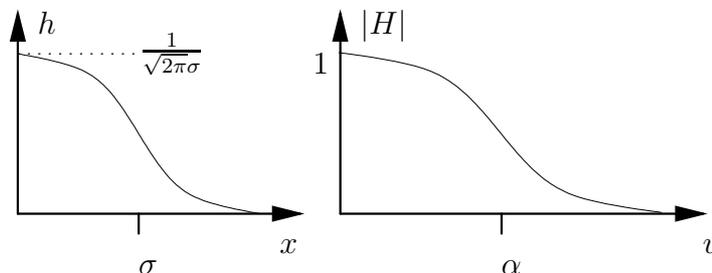
$$\begin{aligned} \text{mit } r^2 &= x^2 + y^2 \\ \sigma_r^2 &= \sigma_x^2 + \sigma_y^2 \\ \sigma = \sigma_x = \sigma_y &\implies \sigma_r^2 = 2\sigma^2 \end{aligned}$$

3. Die Gaußfunktion ist gegenüber *Faltung invariant*.

$$\left(A_1 e^{-\frac{x^2}{2\sigma_1^2}} \right) * \left(A_2 e^{-\frac{x^2}{2\sigma_2^2}} \right) = \left(A_1 A_2 \frac{\sigma_1 \sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} e^{-\frac{x^2}{2(\sigma_1^2 + \sigma_2^2)}} \right)$$

4. Die Gaußfunktion ist gegenüber *Fouriertransformation invariant*.

$$H(u) = \mathcal{F}\{h(x)\} = e^{-\frac{u^2}{2\alpha^2}} \quad \alpha = \frac{1}{2\pi\sigma}$$



5. Die Gaußfunktion besitzt *minimales Impulsbreite-Bandbreite-Produkt*.

Erinnerung: Abstände im Ortsraum verhalten sich invers zu Abständen im Frequenzraum
 → Ausdehnung von Operatoren sind im Orts- und Frequenzraum invers zueinander

Grundfrage der Mustererkennung: Was ist wo ?

wo: Lokalisierung (im Ortsraum)

was: Kategorisierung (=Lokalisation im Frequenzraum)

⇒ in beiden Räumen sollten Operatoren möglichst scharf sein !

→ Breite der Funktion in beiden Räumen berechnen: Es existieren verschiedene Maße

$$\bar{x} = \frac{\int_{-\infty}^{\infty} x |h(x)|^2 dx}{E_h} \quad \text{Ort des Mittelwertes der Intensität von } h$$

$$E_h = \int_{-\infty}^{\infty} |h(x)|^2 dx$$

$$\sigma_h^2 = \frac{\int_{-\infty}^{\infty} (x - \bar{x})^2 |h(x)|^2 dx}{E_h} \quad \text{Varianz der Intensität von } h \text{ (2. zentrales Moment)}$$

$$\bar{u} = \frac{\int_{-\infty}^{\infty} u |H(u)|^2 du}{E_H}$$

$$E_H = \int_{-\infty}^{\infty} |H(u)|^2 du$$

$$\sigma_H^2 = \frac{\int_{-\infty}^{\infty} (u - \bar{u})^2 |H(u)|^2 du}{E_H}$$

Es läßt sich (mit etwas Aufwand) zeigen, daß für eine allgemeine Funktion h bzw. für deren Fouriertransformierte H gilt

$$\sigma_h^2 \sigma_H^2 \geq \frac{1}{4\pi^2}$$

→ bzgl. Standardabweichung:

$$\sigma_h \sigma_H \geq \frac{1}{2\pi} \quad \text{Unschärferelation}$$

Standardabweichung der Intensität bzw. spektrale Leistungsdichte stehen in Beziehung zueinander → können nicht beliebig klein gemacht werden

Beispiel:

$$h(x) = e^{-a|x|}, \quad a > 0 \quad \sim \text{Exponentialfkt.}$$



$$H(u) = \frac{2a}{a^2 + u^2}$$

$$\rightarrow \sigma_h = \frac{1}{\sqrt{2}a}, \quad \sigma_H = a$$

$$\rightarrow \sigma_h \sigma_H = \frac{1}{\sqrt{2}} > \frac{1}{2} > \frac{1}{2\pi}$$

Extremfall: $h(x) = \delta(x)$ $H(u) = 1$

Beispiel:

$$h(x) = e^{-\frac{1}{2}\left(\frac{x}{a}\right)^2} \quad \text{---} \quad H(u) = \sqrt{2\pi} a e^{-2(a\pi u)^2}$$

$$\sigma_h = a, \quad \sigma_H = \frac{1}{2a\pi} \quad \rightarrow \quad \sigma_h \sigma_H = \frac{1}{2\pi}$$

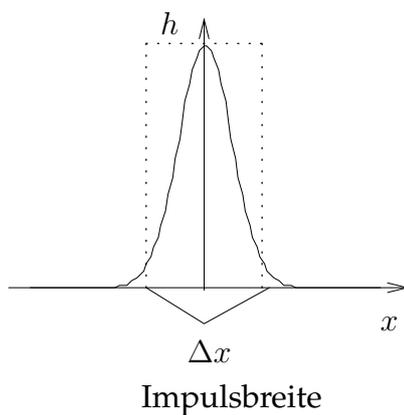
⇒ für beliebige Parameterwerte a erfüllt Gaußfunktion Unschärferelation optimal!

Bedeutung für Filterentwurf:

Es existieren verschiedene Möglichkeiten der Definition von Impulsbreite / Bandbreite

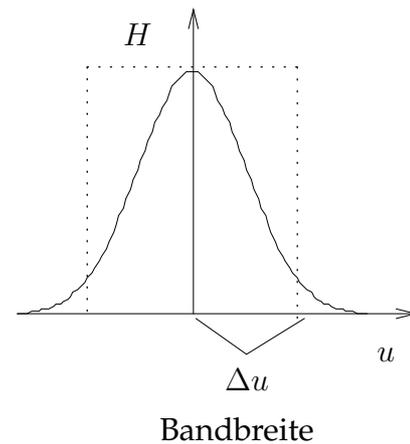
Δx : Impulsbreite

Breite eines Rechteckimpulses, dessen Höhe mit der des Impulses übereinstimmt bei $x = 0$ und der die gleiche Fläche besitzt



Δu : Bandbreite

nur positive Frequenzen haben Bedeutung



Annahme:

1. Breite des Filters im Ortsraum sei begrenzt durch Δx
 → Breite im Frequenzraum: $\Delta u \geq \frac{1}{2\pi\Delta x}$
2. Breite des Filters im Frequenzraum begrenzt auf Δu
 → Breite im Ortsraum: $\Delta x \geq \frac{1}{2\pi\Delta u}$

Δx klein ○ — ● Δu groß

hohe Ortsauflösung

große Unsicherheit (da geringe Verbesserung SNR bzw. schlechte Modellierung Signal)

Δx groß ○ — ● Δu klein

geringe Ortsauflösung

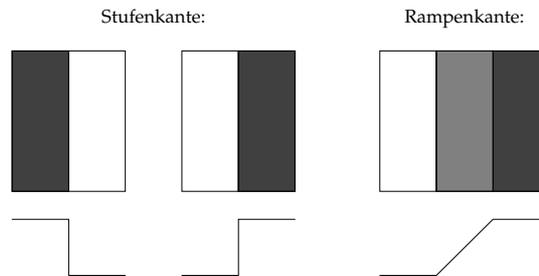
große Sicherheit (da große Verbesserung SNR bzw. gute Modellierung Signal)

Wunsch: Adaption an unterschiedliche Erfordernisse der Signalstruktur
 → Skalenpyramiden (Hierarchien)

3.5 Kantendetektion

Kanten sind *orientierte Regionen*, die in einer Richtung relativ homogene Intensitäten und in orthogonaler Richtung relativ *inhomogene Intensitäten* aufweisen mit einer ausgeprägten *Antisymmetrie*:

(mathematische) Kantenarten:



Wegen der Wirkung der PSF sind reale Kanten im Bild verschmiert (wenigstens 2 Bildpunkte breit wegen Abtasttheorem).



Linien sind orientierte Regionen der Inhomogenität in einer Richtung mit relativ ausgeprägter *Symmetrie*.



3.5.1 Kanten beleuchteter 3D-Objekte

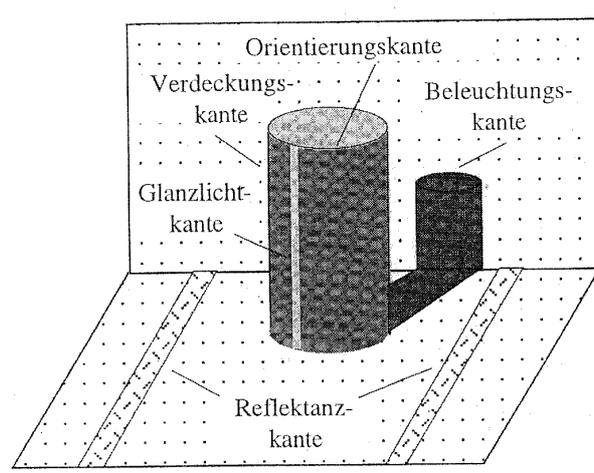
Ursprung der Kanten/Linien: Die Intensität (Luminanz) ist das Produkt aus *Beleuchtungsintensität* I_B und *Reflektanz* R . Es gilt bei diffuser Beleuchtung $f = I_B \cdot R$. Daher liegen die Ursachen der Kanten/Linien in

1. den Eigenschaften der Reflexion der Oberflächen,
2. den Beleuchtungsänderungen über dem Ort.

Diese Ursachen sind schwer zu trennen !

Begriffe:

- Reflexion: gerichtetes Rückstrahlungsvermögen eines Körpers
- Albedo: diffuses Rückstrahlungsvermögen eines Körpers
- Lambertsches Gesetz: Leuchtdichte eines diffus strahlenden Körpers ist unabhängig von der Beobachtungsrichtung.



- Verdeckungskanten
 - vom Beobachter gesehene Abgrenzung des Objektes gegen Hintergrund
 - bilden in Abhängigkeit von Beobachtungsrichtung unterschiedliche physikalische Objekte in der Szene ab
- Orientierungskanten
 - stellen Diskontinuitäten der Orientierung von Objektoberflächen dar
- Reflektanzkanten
 - stellen Diskontinuitäten der Oberflächenreflektanz dar
- Beleuchtungskanten
 - stellen Diskontinuitäten der Beleuchtung von Oberflächen dar (Schattenkanten)
- Spiegelungskanten
 - stellen gespiegelte Kanten anderen Ursprungs dar (erfordern bestimmte Eigenschaften der Reflektanz)
- Glanzlichtkanten
 - begrenzen Glanzlichter auf Oberflächen
 - in Glanzlichtern geht der lineare Zusammenhang zwischen gemessener Intensität und Beleuchtungsintensität / Reflektanz verloren

Die beiden folgenden Abbildungen zeigen das Bild "Lena" (links) und das Resultat eines Kantenoperators (rechts).



Annahme: Beleuchtungsintensität ändert sich nur langsam über dem Ort (es existieren keine Beleuchtungskanten bzw. Schattenwürfe)

→ hochfrequente Signalanteile synthetisieren Kanten, die ihre Ursache in den Objekten selbst haben:

- 3D-Gestalt → Orientierungskante, Verdeckungskante, Glanzlichtkante
- Reflektanz → Reflektanzkante

Homomorphe Filterung: Spektrum der logarithmierten Bildintensität erlaubt Trennung von Beleuchtungs- und Oberflächenanteilen an der Bildfunktion.

Ann.: $f = I_B R$, $I_B \sim$ beleuchtungsabhängiger Teil, $R \sim$ Reflektanzanteil

Hieraus folgt:

$$\log f = \log I_B + \log R$$

und

$$\mathcal{F}\{\log f\} = \mathcal{F}\{\log I_B\} + \mathcal{F}\{\log R\},$$

also für eine Schwellfrequenz u_T

$$\mathcal{F}\{\log f\} = \begin{cases} \mathcal{F}\{\log I_B\} & \text{für } u \leq u_T \\ \mathcal{F}\{\log R\} & \text{für } u > u_T. \end{cases}$$

Rekonstruktion der Objektgrenzen aus Kantenstrukturen ist ein *inverses Problem*, das im allgemeinen keine gut gestellte Aufgabe ist (*ill posed problem*).

- Unvollständige Information wegen Projektion, Verdeckung, Rauschen
Regularisierung: ill posed problem → well posed problem
durch Modelle über globale Struktur der Szene (Hinzufügen und Vernachlässigen von Strukturen)
- Außerdem wird hochfrequente Kanteninformation durch (*Sensor-*) *Rauschen* überlagert.
→ Welcher Frequenzanteil des Signals ist der Kantenstruktur und welcher dem Rauschen zuzuordnen?
Konditionierung: Reduzierung des Rauscheinflusses und damit mathematische Stabilisierung des Verfahrens

→ inverses Problem ist auch *schlecht konditioniert!*

Es bestehen zwei prinzipielle Aufgaben:

1. Klassifikation einer Kante ("Was?"):

Gegeben seien verschiedene Modelle der Interpretation der Signalstruktur. Eines davon ist das einer Kante.
Aufgabe: *Identifiziere eine lokale Signalstruktur als (Teil einer) Kante* $\hat{=}$ "Finde eine Kante."
2. Lokalisation einer Kante ("Wo?"):

Feststellen des *genauen Ortes der Kante*

Ein *Kantendetektor* ist ein lokaler Operator (im Ortsraum), der sowohl eine Bildstruktur als Kante klassifiziert als diese auch genau lokalisiert.

Die *Äquivalenzklasse "Kante"* wird durch ein Template ihrer Gestalt (als Prototyp) realisiert.

- Kantendetektor ist *Template-Matching-Operator*
 - Kreuzkorrelation zwischen Template und Bildstruktur
- differentielle Eigenschaften charakterisieren eine Kante → *Ableitungsoperator*

Forderungen an einen guten Kantendetektor:

1. *Vollständigkeit* (Sensitivität): keine falsch negativen Antworten bzgl. der Semantik der Kantenmenge
Kanten sollen nicht übersehen werden !
2. *Eindeutigkeit* (Spezifität): keine falsch positiven Antworten
 - Rauschen soll nicht als Kante interpretiert werden !
 - Es soll für eine Kante nur ein Detektionssignal erzeugt werden !
3. *Lokalisationsgüte*: Das eindeutige *Maximum der Detektorantwort* soll an der Position der Kante auftreten ! (Nullphasenfilter)
4. *Zusammenhang*: topologische Forderung \sim im Prinzip Vollständigkeit
Kanten sollen als *zusammenhängende Gebilde* detektierbar sein (insbesondere kritisch bei 2D-Kanten) → oft: Rekonstruktion des Zusammenhangs wissensbasiert im Anschluß an Kantendetektion !

Differenzierung des Signals → örtliche Differenzbildung im Diskreten (erzeugt Approximation des Kontinuierlichen)

1. Kantenorte entsprechen den Maxima der 1. Ableitung

$$x_K = \operatorname{argmax} |f'(x)|$$

2. Kantenorte entsprechen den Nullstellen der 2. Ableitung (Minima der Krümmung)

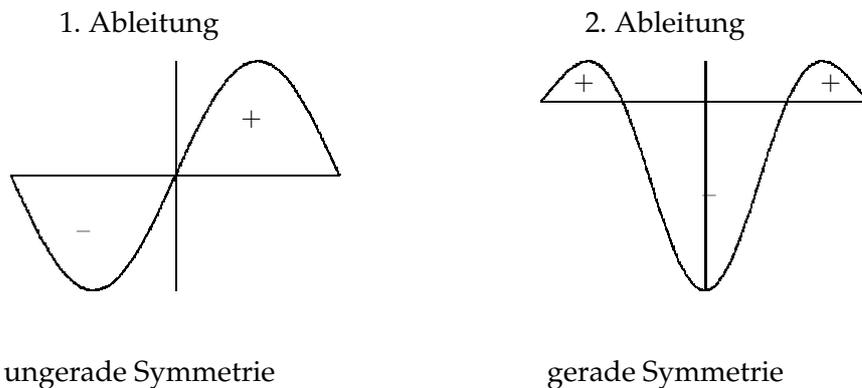
$$x_K = \operatorname{argmin}|f''(x)|$$

- kontinuierliche Ableitungen (1D):

$$\left(\frac{d}{dx}\right)^n \quad \text{○} \text{---} \text{●} \quad (j2\pi u)^n \quad \rightarrow \text{Hochpaßcharakteristik}$$

z.B. $\begin{cases} \text{Nabla-Operator} & \nabla = \frac{d}{dx} \\ \text{Laplace-Operator} & \mathcal{L} \equiv \Delta = \frac{d^2}{dx^2} \end{cases}$ $\text{○} \text{---} \text{●} \quad -4\pi^2 u^2$

- Symmetrie: wichtig für Lokalisationsgenauigkeit



$$\sum_i h_i^+ = \sum_i h_i^- \sim \sum_i h_i = 0 \quad \text{○} \text{---} \text{●} \quad |H(0)| = 0$$

→ Gleichanteil des Signals wird nicht übertragen (detektiert)
 → Alle primitiven Ableitungsoperatoren sind Hochpaßfilter !

Beispiel: Frequenzübertragungsfunktion eines 1D-Laplace-Operators

$$\mathcal{D}_x^2 = (\mathcal{D}_l \mathcal{D}_r)_x \quad h_{\mathcal{D}_x^2} = \begin{bmatrix} -1 & 2 & -1 \end{bmatrix}$$

kleinste diskrete Approximation des Laplace-Operators (für $K = 3$)
 betrachte Impulsantwort als Folge gewichteter Einheitsimpulse

$$h = \sum_{k=-\infty}^{\infty} h_k \sum_{m=-\infty}^{\infty} \delta_{k-m}$$

mit

$$h_k = \begin{cases} 2 & \text{für } k = 0 \\ -1 & \text{für } k = 1 \text{ und } k = -1 \\ 0 & \text{sonst} \end{cases}$$

und

$$\delta_{k-m} = \begin{cases} 1 & \text{für } m = k \\ 0 & \text{sonst} \end{cases}$$

bzw.

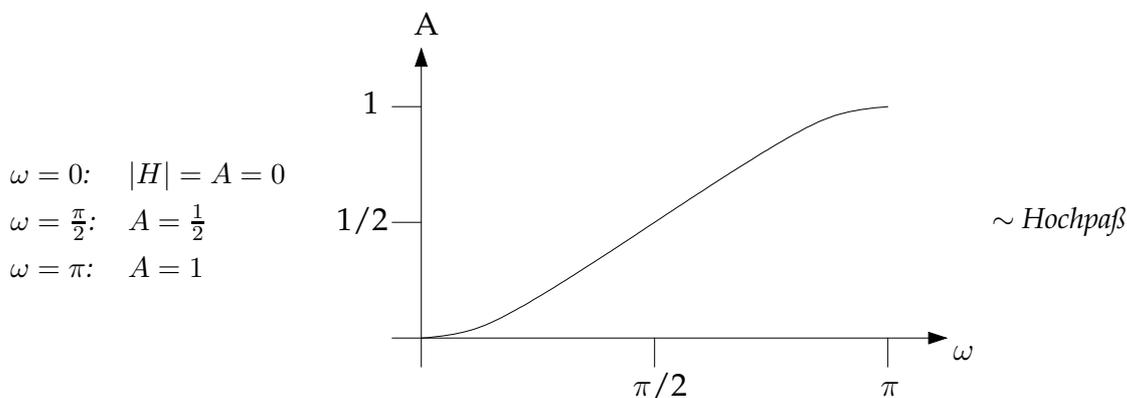
$$\sum_{m=-\infty}^{\infty} \delta_{k-m} = \begin{cases} 1 & \text{für } k = -1, 0, 1 \\ 0 & \text{sonst} \end{cases} \quad \sim \text{drei kanonische Einheits-} \\ \text{Basisfunktionen für } k = -1, 0, 1$$

Wegen

$$\begin{aligned} H'(e^{j\omega}) &= \sum_{k=-\infty}^{\infty} h_k e^{-j\omega k} && H' \text{ sei nicht normierte Frequenzübertragungsfkt.} \\ &= \sum_{k=-1}^1 h_k \sum_{m=-\infty}^{\infty} \delta_{k-m} e^{-j\omega k} && (\text{Beiträge für alle anderen } k \text{ verschwinden}) \\ &= \sum_{k=-1}^1 h_k \sum_{m=-1}^1 \delta_{k-m} e^{-j\omega k} \\ &= \sum_{k=-1}^1 h_k (\delta_{k+1} + \delta_k + \delta_{k-1}) e^{-j\omega k} \\ &= \sum_{k=-1}^1 (-\delta_{k+1} + 2\delta_k - \delta_{k-1}) e^{-j\omega k} \\ &= -e^{j\omega} + 2e^0 - e^{-j\omega} && \text{wegen} \\ &\quad -\delta_{k+1} e^{-j\omega k} = \begin{cases} -e^{j\omega} & \text{für } k = -1 \\ 0 & \text{sonst} \end{cases} \\ &\quad + 2\delta_k e^{-j\omega k} = \begin{cases} 2e^0 & \text{für } k = 0 \\ 0 & \text{sonst} \end{cases} \\ &\quad -\delta_{k-1} e^{-j\omega k} = \begin{cases} -e^{-j\omega} & \text{für } k = 1 \\ 0 & \text{sonst} \end{cases} && \text{also} \\ H'(e^{j\omega}) &= 2(1 - \cos \omega) \end{aligned}$$

Normierung auf $\omega = \pi$: $\omega = 0$: $H' = 0$, $\omega = \frac{\pi}{2}$: $H' = 2$, $\omega = \pi$: $H' = 4$, da $\cos \pi = -1$

$$H(e^{j\omega}) = \frac{H'(e^{j\omega})}{H'(e^{j\pi})} = \frac{1}{2}(1 - \cos \omega)$$



Alle AbleitungsfILTER sind Hochpaßfilter !

3.5.2 2D–Kantendetektoren

Da in der Ebene mehr Freiheitsgrade existieren, gibt es verschiedene Problemstellungen:

1. Finde alle Kanten unabhängig von ihrer Orientierung
 - Forderung: Invarianz des Operators bzgl. Rotation
 - Operator muß selbst rotationssymmetrisch sein
 - Es existiert in $\mathbb{R}^2 \times \mathbb{R}$ kein Ableitungsoperator erster Ordnung, der diese Forderungen erfüllt, aber ein Ableitungsoperator zweiter Ordnung, der Laplace-Operator.

$$\mathcal{L} \equiv \Delta \equiv \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

Der Laplace-Operator ist linear und isotrop. Die diskrete Näherung

$$\mathcal{L} = \mathcal{D}_x^2 + \mathcal{D}_y^2 \quad \text{mit} \quad h_{\mathcal{L}} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

erfüllt die Rotationsinvarianz nicht gut (nur für kleines ω).

2. Finde die Steigung in Richtung des Gradienten
 - bedeutet Filterung in zwei orthogonalen Richtungen
 - und Rekonstruktion des Gradienten in beliebiger Richtung mittels nichtlinearem Operator

Abtastung der kontinuierlichen Ableitung (in 1D)

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x-h)}{2h}$$

ergibt die diskrete Näherung ($K = 3$)

$$f'_m = \frac{1}{2}(f_{m+1} - f_{m-1})$$

also für x - bzw. y -Richtung

$$h_{\mathcal{D}_x} = \frac{1}{2} \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}, \quad h_{\mathcal{D}_y} = \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

- Kanten verlaufen senkrecht zur Operatorrichtung

$$\begin{aligned} \nabla &= \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{bmatrix} && \text{Gradientenvektor in 2D} \\ &= \begin{bmatrix} \mathcal{D}_x & \mathcal{D}_y \end{bmatrix} \end{aligned}$$

- Betrag des Gradienten

$$|\nabla| = \sqrt{\mathcal{D}_x^2 + \mathcal{D}_y^2} \quad \sim \text{nichtlinearer Operator}$$

3. Finde alle Kanten in vorgegebener Richtung
Eine *Familie* von Richtungsableitungen wird benötigt.

Beispiel: Sobel-Operator

- 4 Ableitungsrichtungen
- Mittelung in Richtung senkrecht zur Ableitungsrichtung

$$h_{\mathcal{S}_1} = \frac{1}{8} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad h_{\mathcal{S}_2} = \frac{1}{8} \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

$$h_{\mathcal{S}_3} = \frac{1}{8} \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix} \quad h_{\mathcal{S}_4} = \frac{1}{8} \begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

$h_{\mathcal{S}_1} : \quad \longleftrightarrow \text{Ableitungsrichtung} \quad \updownarrow \text{Mittelungsrichtung}$

$$\mathcal{S}_1 = \mathcal{D}_x \Pi_y^2 = \mathcal{D}_x \Lambda_y \quad \text{für } K = 3$$

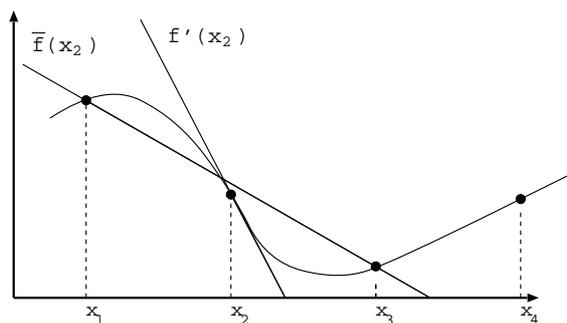
$$h_{\mathcal{D}_x} = \frac{1}{2} \begin{bmatrix} 1 & 0 & -1 \end{bmatrix} \quad ; \quad h_{\Pi_y} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad ; \quad h_{\Lambda_y} = \frac{1}{4} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

■

3.5.3 Verbesserung der lokalen Schätzung der Ableitung

Diskrete Operatoren schätzen die Eigenschaften der *Sekante* $\bar{f}(x)$ durch $(x_1, f(x_1))$ und $(x_3, f(x_3))$ und nicht die *Tangente* $f'(x_2)$.

→ Verwendung eines *Polynoms* n -ten Grades zur lokalen Repräsentation der Funktion im Aufpunkt (Interpolation)



Die Interpolation von $n+1$ paarweise verschiedenen Stützstellen mittels Polynom n -ten Grades ist eindeutig bestimmt \sim *Lagrangesches Interpolationspolynom*

Beispiel: Explizite Lösung der Bestimmungsgleichungen:

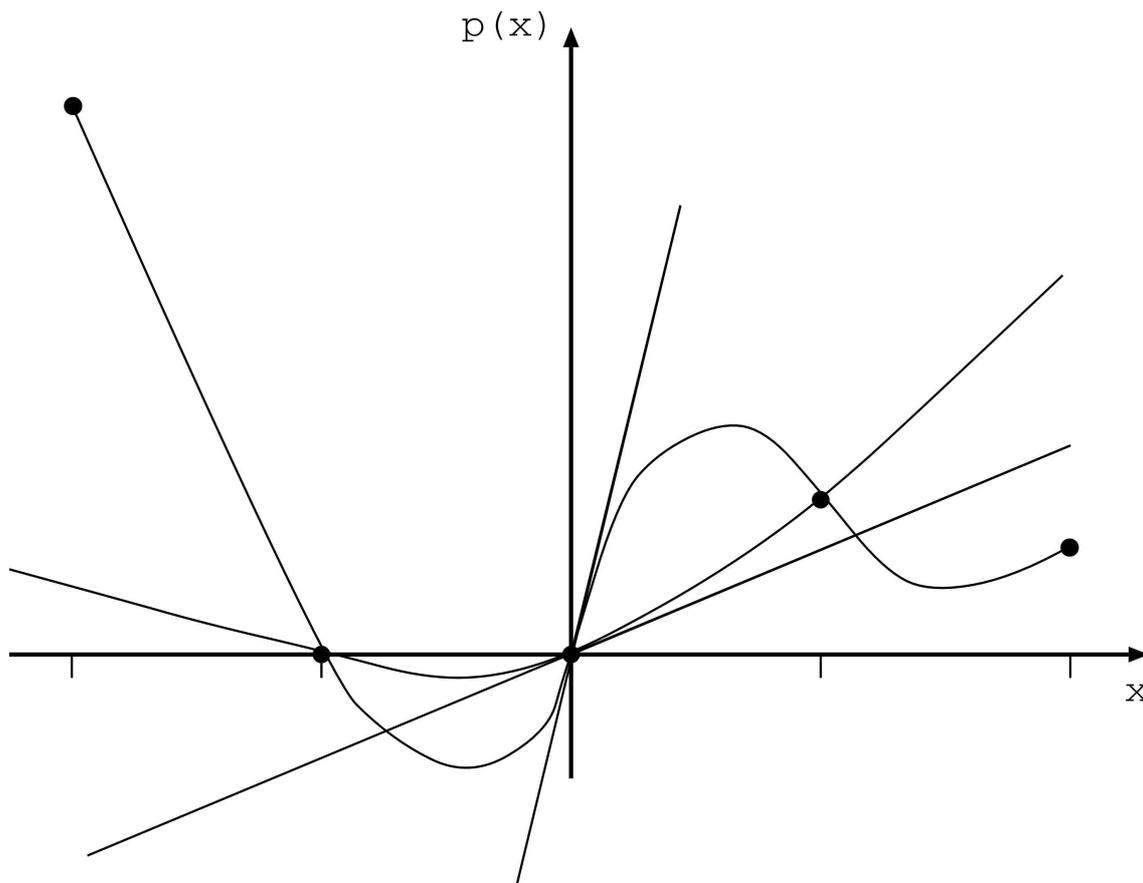
Es soll lokal eine Funktion f durch ein Polynom approximiert werden

$K = 5 : m \in \{-2, -1, 0, 1, 2\}$

Annahme: x seien lokale Koordinaten

Polynom: $p(x) = ax + bx^2 + cx^3 + dx^4$ ($\rightarrow p(0) = 0$)

Gesucht: Anstieg des Polynoms am Ort $x = 0 \rightarrow p'(0) = \frac{d}{dx}p(0) = a$



Fallunterscheidung:

1. Kurve 2. Grades (Parabel) durch drei Punkte ($c = d = 0$) \rightarrow flacher Anstieg

2. Kurve 4. Grades durch 5 Punkte \rightarrow steiler Anstieg

Bei bekanntem Polynom ist Koeffizient a die geschätzte Ableitung am Ort $x = 0$:

$$\begin{aligned} p'(x) &= a + 2bx + 3cx^2 + 4dx^3 \\ p'(0) &= a \end{aligned}$$

Der Gradient des Polynoms am Koordinatenursprung wird durch den Koeffizienten a bestimmt.

Betrachte die obigen beiden Fälle:

1. Kurve 2. Grades : $c = d = 0$

$$\left. \begin{aligned} f(-1) - f(0) &= p(-1) = -a + b \\ f(1) - f(0) &= p(1) = a + b \end{aligned} \right\} \rightarrow a = \frac{1}{2}(f(1) - f(-1))$$

$$\hat{D}_x^1 = \frac{1}{2} \begin{bmatrix} 1 & 0 & -1 \end{bmatrix} \quad \sim \text{wie gehabt}$$

2. Kurve 4. Grades

$$\left. \begin{aligned} f(-1) - f(0) &= p(-1) = -a + b - c + d \\ f(1) - f(0) &= p(1) = a + b + c + d \end{aligned} \right\} \rightarrow f(1) - f(-1) = 2a + 2c \quad (*)$$

$$\left. \begin{aligned} f(-2) - f(0) &= p(-2) = -2a + 4b - 8c + 16d \\ f(2) - f(0) &= p(2) = 2a + 4b + 8c + 16d \end{aligned} \right\} \rightarrow f(2) - f(-2) = 4a + 16c \quad (**)$$

(**) - 8(*):

$$(f(2) - f(-2)) - 8(f(1) - f(-1)) = 4a - 16a = -12a$$

$$\rightarrow a = \frac{2}{3}(f(1) - f(-1)) - \frac{1}{12}(f(2) - f(-2))$$

$$\rightarrow \hat{D}_x^2 = \frac{1}{3} \begin{bmatrix} \frac{1}{4} & -2 & 0 & 2 & -\frac{1}{4} \end{bmatrix}$$

■

Mögliche Interpolationsverfahren:

(für Subpixelgenauigkeit)

1. Abschnitte von Polynom n -ten Grades: Abschnitte verlaufen jeweils durch $\frac{n}{2}$ Punkte vor und hinter dem Aufpunkt
2. Abschnitte von Polynom 3-ten Grades: Abschnitte verlaufen jeweils durch zwei Stützstellen und schließen mit gleicher 1. und 2. Ableitung aneinander an
 \cong kubische B-Splines: minimieren die Krümmung $\int (I''(x))^2 dx$ der Interpolationsfunktion $I(x)$

3. *ausgleichende Splines* (keine eigentliche Interpolation)
minimieren ein Funktional

$$E_\alpha(I) = \alpha \sum_i (f(x_i) - I(x_i))^2 + (1 - \alpha) \int (I''(x))^2 dx$$

$\alpha \in [0, 1]$ reguliert das Verhältnis von Datentreue und Glattheit.

3.5.4 Konditionierung der Ableitungsoperation

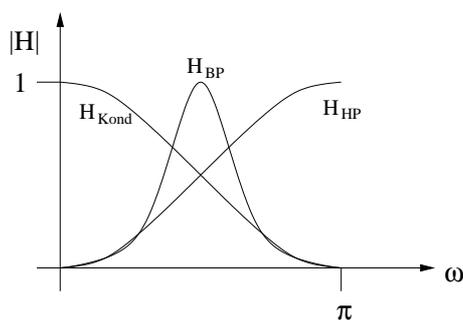
- Hochpaßcharakteristik der Ableitungsoperatoren *verstärkt relativ das Rauschen, d.h. je höher die Ableitung*, umso relativ stärker wirkt sich Rauschen aus
 - ▶ Lokalisationsgüte
 - ▶ Spezifität (Sicherheit / Eindeutigkeit)
- Rauschen kann falsche Kanten vortäuschen bzw. Kantenorte haben keinen topologischen Zusammenhang
- Rauschen kann Ort der Kantenlokalisierung verfälschen (lokaler Bias)
Regularisierung: ill posed problem \rightarrow well posed problem

1. *Modelle* über globale Strukturen der Szene: Vernachlässigung + Hinzufügen von Strukturen
2. *Konditionierung*: Reduzierung des Rausch-Einflusses

\Rightarrow *Hochpaßoperator* muß in *Bandpaßoperator* transformiert werden — Lage des Paßbandes dort, wo Kanteninformation liegt

$$h_{BP} = h_{Kondition} * h_{HP}$$

$$h_{Kondition} \sim \text{Tiefpaß (z.B. Gaußfunktion)}$$



z.B. Anpassung von $h_{Kondition}$ an PSF !

LOG-Filter (Gauß-Laplace-Operator)

Der Gaußoperator ist der optimale Konditionierungsoperator, da er das Bandbreiteprodukt minimiert.

$$h_{LOG} = \Delta * h_G$$

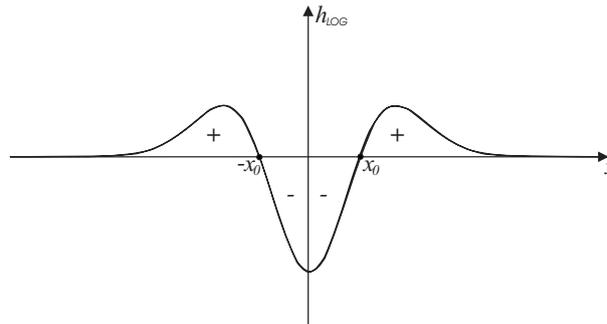
Wegen *Kommutativität* und *Assoziativität* linearer Operatoren

$$g = f * h_{\text{LOG}} = f * (\Delta * h_G) = \Delta * (f * h_G) = h_G * (f * \Delta)$$

Beispiel: 1D-LOG, kontinuierlich

$$h_{\text{LOG}}(x) = D_x^2 * h_x(x) = \frac{1}{\sigma^2} \left(\frac{x^2}{\sigma^2} - 1 \right) h_x(x)$$

$$h_x(x) = e^{-\frac{x^2}{2\sigma^2}}$$



$$\sum h_x^+ = \sum h_x^-$$

$$x_0 = \sigma$$

$$u_{\text{max}} = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma}$$

Nulldurchgang

Maximumfrequenz des Paßbandes

■

Konditionierte Filter sind Bandpaßfilter:

- ▶ σ klein \rightarrow hohe Frequenzen
- ▶ σ groß \rightarrow niedrige Frequenzen
- ▶ Stabilität gegenüber Rauschen wird mit Verlust an Ortsauflösung erkaufte !
- ▶ Detektion und Lokalisation erfordern unterschiedliche Parameter der Konditionierung \rightarrow nicht mit einem Operator realisierbar
 \rightarrow Skalenraumanalyse (σ kontinuierlich) / Auflösungspyramiden (σ diskret)

Beispiel: 2D-LOG, kontinuierlich

$$h_{\text{LOG}}(x, y) = \Delta * h_c(x, y) = -\frac{1}{2\pi\sigma^4} \left(2 - \frac{x^2 + y^2}{\sigma^2} \right) h_c(x, y)$$

nachfolgende Abtastung \rightarrow diskreter Operator

$$K = 3: \quad h_{\text{LOG}} = \begin{bmatrix} -1 & -4 & -1 \\ -4 & 20 & -4 \\ -1 & -4 & -1 \end{bmatrix}$$

■

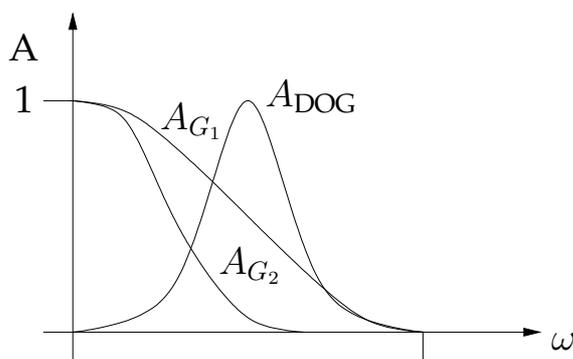
DOG–Filter (Difference of Gaussian Operator)

stellt Approximation eines LOG dar

Idee: $|H_{\text{DOG}}| = |H_{G_1}| - |H_{G_2}| \approx |H_{\text{LOG}}|$

Unschärfemaskierung: $|H_{G_1}(u)| = 1$ ● — ○ $h_{G_1}(x) = \delta(x)$

$\sigma_1 < \sigma_2$: Differenz zweier Tiefpässe ergibt einen Bandpaß



optimale Wahl bzgl. Approximation LOG:

$$\frac{\sigma_2}{\sigma_1} \approx 1.6$$

Ausführung der Kantendetektion

1. Ausführung der Faltung des Bildes mit dem Operator
2. Berechnung des Absolutbetrages der Operatorantwort
3. Suche lokaler Extrema (1. Ableitung: Maximum, 2. Ableitung: Minimum)
4. "Dünnen": Kantenregion \Rightarrow Kantenort (Zuordnung von Koordinaten)
5. Rekonstruktion zusammenhängender Kanten aus Kantensegmenten
wissensbasierter Zugang

► Nur Schritt 1 ist eine lineare Operation !

Oft: mehrere Faltungen mit unterschiedlichem σ !

3.5.5 Topologische Probleme bei 2D–Strukturen

Gekrümmte Kanten	}	"junctions"	}	intrinsische 2D–Konzepte
Schnittpunkte				
Kreuzungspunkte				
Endpunkte				

\rightarrow können mit 1D–orientierten Ableitungsoperatoren nicht erkannt werden

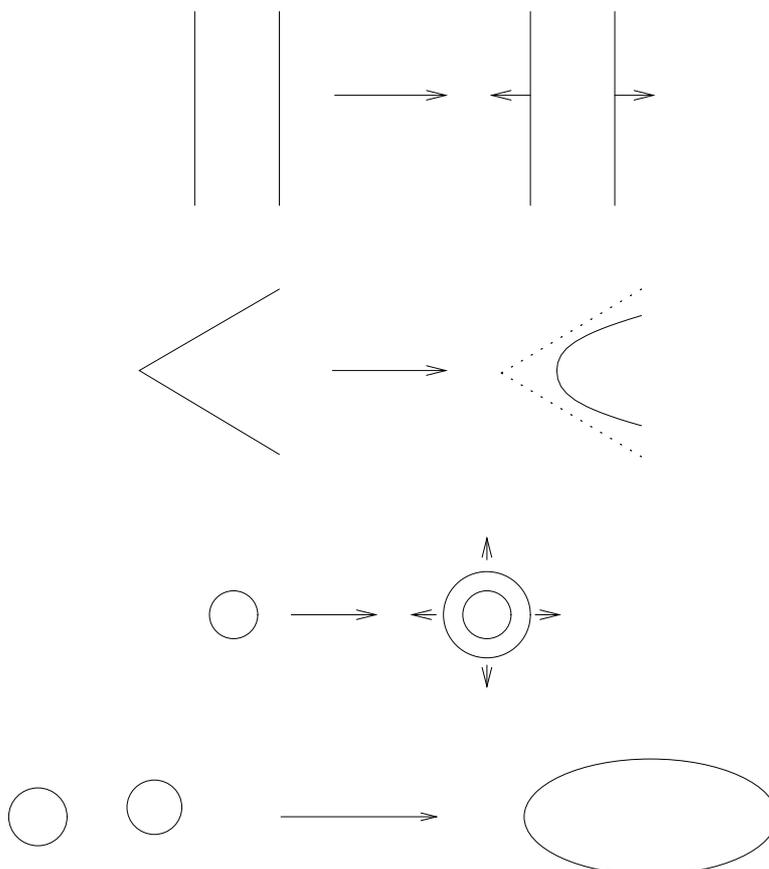
\Rightarrow *Eckendetektoren* (corner detectors)

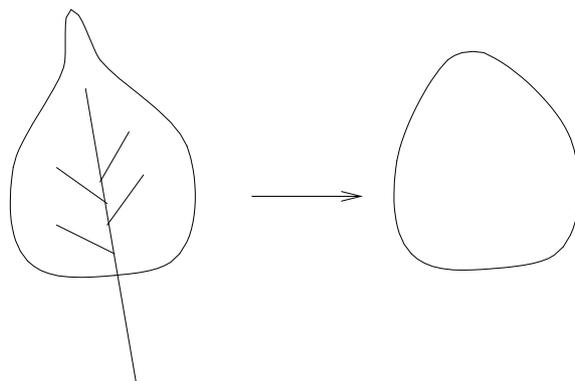
1. Innerhalb des Operatorfensters werden *Krümmungen* der Struktur erkannt
 - Nullstellen des Laplace-Operators entsprechen nicht mehr den Orten größter Steigung der Grauwertfunktion
 - ⇒ Verschieben der Nullstellen im gekrümmten Bereich in Richtung *negativer Gaußscher Krümmung* der Grauwertfunktion (hyperbolischer Bereich)
 Eckendetektor als nichtlinearer Krümmungsdetektor für Gaußsche Krümmung:

$$k_{\text{Gauß}} = \frac{\partial^2}{\partial x^2} \frac{\partial^2}{\partial y^2} - \left(\frac{\partial^2}{\partial x \partial y} \right)^2$$

2. an *Kreuzungen und Verzweigungen* (junctions) noch komplizierter
 - ⇒ "steuerbare Filter"
3. aber auch *parallele Kanten* innerhalb Operatorfenster stören sich gegenseitig ⇒ Fehllokalisierung

Je stärker konditioniert wird, umso größer ist das Operatorfenster, umso stärkere Störung der Topologie !





Ausweg: mehrere Operatoren (Pyramiden, Skalenraum)

σ groß: Feststellen einer interessanten Struktur

σ klein: Lokalisation der Struktur

3.6 Hierarchien der Skala und Abstraktion

Bildstrukturen sind auf einer einzigen Konzeptebene nicht interpretierbar.

- *Kategorisierung* und *Lokalisierung* erfordern einander widersprechende Filterparameter. Kategorisierung bedeutet *Erkennung (Interpretation)*. \sim Einbeziehung von *Kontext*. Bei kleinen Operatorfenstern (gute Lokalisierung) kann nur sehr unsicher die Interpretation der zu lokalisierenden Struktur erfolgen und umgekehrt.

Daher repräsentieren wir Bilder hierarchisch, und zwar als:

1. **Skalenhierarchie** bzw. **Auflösungshierarchie**, in der die Skala/Auflösung einen *Freiheitsgrad* darstellt.

Auflösung: im Frequenzraum

Skala: im Ortsraum

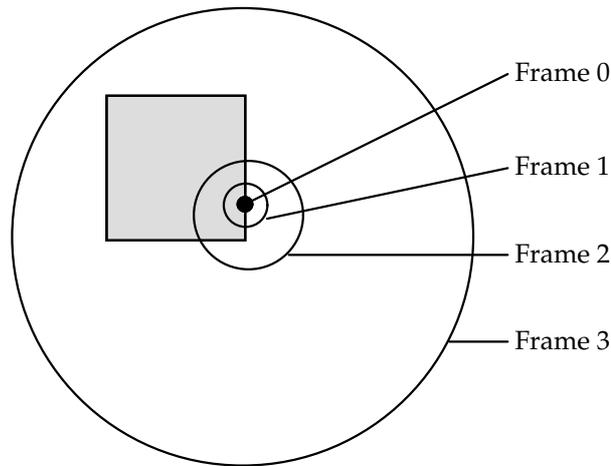
2. **Abstraktionshierarchie**, in der das Bildsignal in *Bildprimitiva* zerlegt wird.

Zu 1: Hierarchien von Skala/Auflösung entstehen durch fortgesetzte *Tiefpaßfilterung* und anschließender *Unterabtastung*. Die Tiefpaßfilterung unterdrückt höhere Ortsfrequenzen. Nach dem Abtasttheorem kann gröber abgetastet werden.

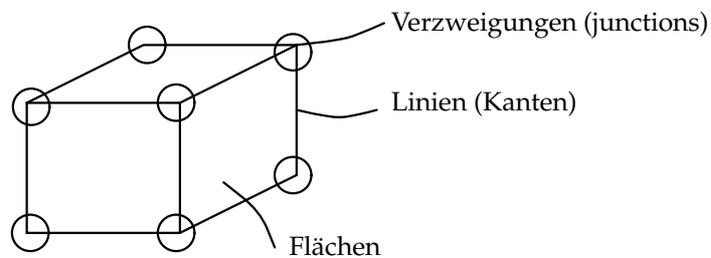
Zu 2: Eine Gruppe von Elementen auf niedriger Ebene bilden neues Element auf höherer Ebene

Ebene	Information
4	Relationen zwischen Objekten
3	Objekte
2	junctions (Krümmung, Verzweigung)
1	Linien, Kanten
Originalbild: 0	Helligkeit, Farbe

Framestruktur (für Operatoren)

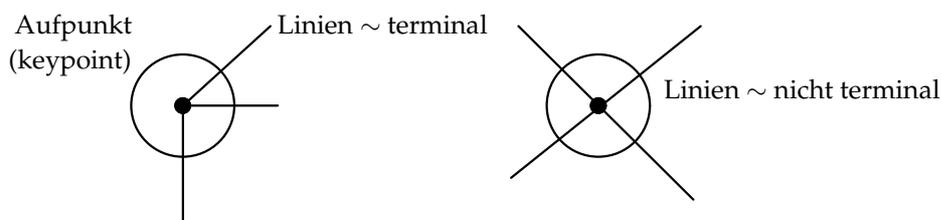


Polyederwelt



Verzweigungen: X, Y, L, T, ...

Beispiel:



verschiedene methodische Konzepte

- 2 gekrümmte Linien
- 3 Linien mit 2 Winkeln
- 1 Linienkonzept, das an 3 Orientierungswinkeln realisiert ist (steuerbare Filter)



Primitiva:

1. irreduzible Invariante (Punkte, Linien, Kanten)
2. aggregierte Invariante (Parallelen, junctions, ...)
 - bez. Projektion 3D → 2D

Beziehung zwischen beiden Hierarchiekonzepten: mit wachsender Abstraktionsebene nimmt auch Skala zu, bezüglich derer Phänomene in Bezug gesetzt werden.

Auch im Menschlichen Visuellen System sind beide Hierarchien verwirklicht, die Auflösungshierarchie in der Retina, Abstraktionshierarchie im visuellem Cortex.

Die inhomogene Struktur der Retina: Es gibt einen Ort schärfsten Sehens, d.h. kleinster Skala, die *Fovea*. Vom Zentrum der Retina nach außen nimmt die Skala zu. Alle Skalen werden parallel erzeugt. Es kommt zu *lokaler* Einbettung im *globalen Kontext*. Der inhomogene Aufbau des Retina-Sensors realisiert widerstrebende Eigenschaften wie hohe Ortsauflösung und großes Gesichtsfeld mit minimalen Aufwand. Es wird aber ein Blickkontrollsystem erforderlich. In technischen Lösungen wird Computer Vision zu Active Vision mit beweglichen Kameras. Das Blickkontrollsystem realisiert die folgenden drei Hauptaufgaben:

1. Blickverfolgung (Tracking)
2. Blickzuwendung (Attention)
3. Blickschärfung (Foveation).

In technischen Systemen ist zur Zeit nur die sequentielle Erzeugung der Auflösungshierarchien umgesetzt.

- Retina-Chip, Sandini (Genua)
- BMFT-Projekt „Elektronisches Auge“

3.6.1 Gaußpyramide

Als Tiefpaß dient ein Gaußfilter mit dem Ziel, eine *Gauß-Ableitungspyramide* zu erzeugen.

Freiheitsgrad der Auflösung ist die Skala σ der Gaußfunktion. Wird die Skala *diskret* variiert, erhalten wir einen **Stapel**, durch zusätzliche Unterabtastung eine **Pyramide**. Bei kontinuierlicher Variation der Skala σ entsteht der **Skalenraum**.

Wir unterscheiden zwei *Strategien der Filterung* im Ortsraum für die Realisierung der Gaußhierarchie:

Führen wir die **gleiche Abtastung** in allen Ebenen der Hierarchie durch, entsteht ein Stapel von Bildern *unterschiedlicher* Auflösung, aber *gleich großen* Definitionsbereichs.

Die Realisierung von reduzierter Auflösung erfolgt durch Anwendung *größerer Filtermasken*. Wir gehen von einer Realisierung des Gaußfilters durch Kaskadierung von Binomialfiltern aus. Die Faltung mit den Filtern bedeutet dann einen *hohen Rechenaufwand*. Im 1D-Fall wächst die Filtergröße (K) quadratisch mit der Grenzwellenlänge $\lambda_g \sim \frac{1}{u_g}$. Im 2D-Fall wächst der Berechnungsaufwand sogar mit der vierten Potenz. Für eine Halbierung der Grenzfrequenz des Tiefpasses ist wegen $u_g \sim \frac{1}{\sigma}$ die Verdoppelung von σ erforderlich und wir erhalten wegen $\sigma \sim \sqrt{n}$, $K \sim n \rightarrow u_g \sim \frac{1}{\sqrt{K}}$ mit n als Anzahl der Kaskaden bzw. Ordnung des Binoms

$$\frac{u_0}{u_1} = \frac{\sigma_1}{\sigma_0} = \sqrt{\frac{n_1}{n_0}} = 2 \rightarrow \frac{n_1}{n_0} = 4 \sim \frac{K_1}{K_2}$$

eine Vervierfachung der Binomordnung und damit vierfache Ausdehnung K_1 gegenüber K_0 .

Der Auflösungsstapel enthält *Redundanz*, da ein breites Filter Strukturunterschiede benachbarter Bildpunkte nicht mehr unterscheiden kann.

Zweite Möglichkeit ist die **Anpassung der Abtastung** an die Grenzfrequenz in der jeweiligen Ebene der Hierarchie gemäß dem Abtasttheorem. Reduzierte Auflösung wird dann durch *gleich große Filtermasken* und *Unterabtastung* (Resampling) realisiert. Der Rechenaufwands sinkt dabei mit abnehmender Auflösung bzw. wachsender Skala, da gleich große Filtermasken auf immer weniger Abtastpunkte angewendet werden. Wir erhalten eine *Pyramidenstruktur*.

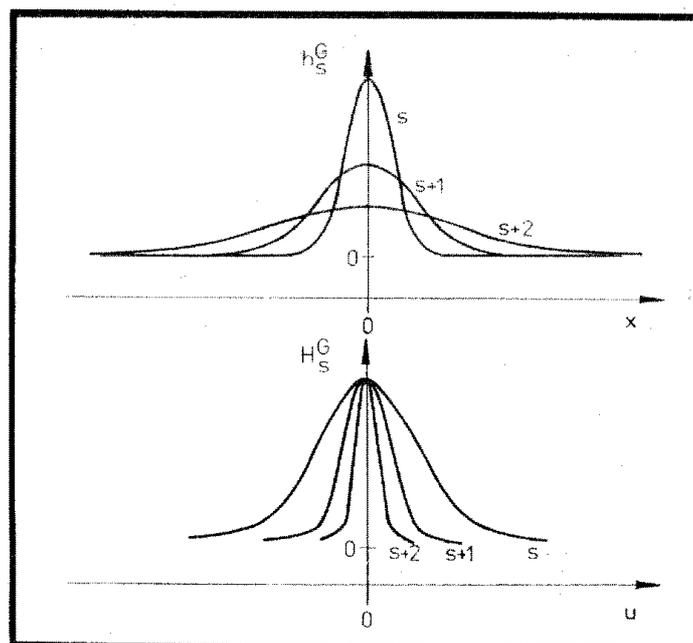
Besonders verbreitet ist die oktav Diskretisierung der Skala, bei der die Grenzfrequenz von Ebene zu Ebene halbiert wird, während sich der Abtastabstand verdoppelt. Die so gebildete Pyramide heißt **Oktavpyramide**. Auf einer *logarithmischen Frequenzskala* besitzen die Durchlaßbereiche (Paßbänder) der Oktavpyramide gleiche Abstände.

Berechnung der Gaußpyramide:

Burt (1981), Crowley (1981)

$$\begin{aligned} g^{(0)} &= f \\ g^{(s)} &= \mathcal{S}_2 \mathcal{W} g^{(s-1)} \quad s > 0 \\ &= \mathcal{S}_2^s \mathcal{W}^s g^{(0)} \\ \text{mit } \mathcal{W}^s &= \mathcal{W}^{(s)} * \mathcal{W}^{s-1} * \dots * \mathcal{W}^{(1)} \end{aligned}$$

[Burt and Aselson, 1983] wählen $\mathcal{W} = \frac{1}{16}[1 \ 4 \ 6 \ 4 \ 1]$, also eine Binomialmaske der Ordnung 4.



3.6.2 Laplacepyramide

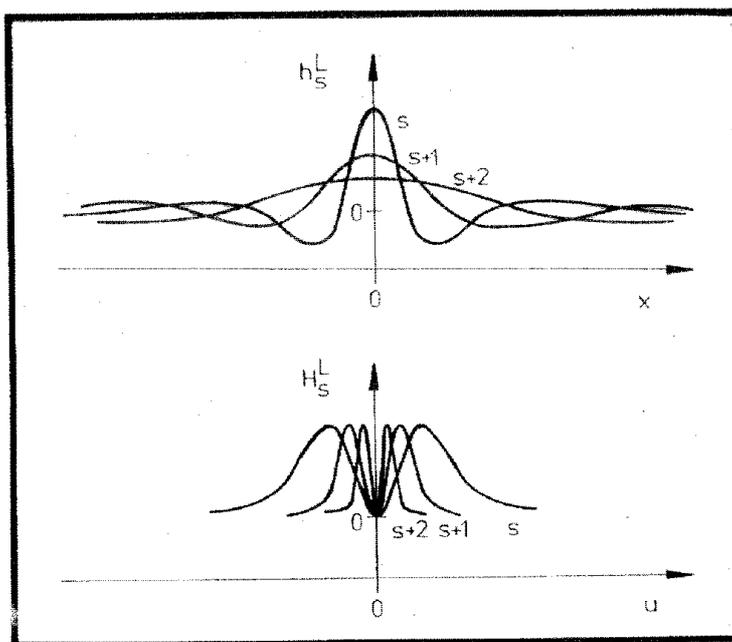
Die Berechnung der Laplace-Pyramide erfolgt entweder durch Ableitung der Gaußpyramide (LOG, Laplacian of Gaussian) oder durch Approximation mit Differenzen von Gaußebenen (DOG, difference of Gaussian). Im Stapelkonzept ergibt sich folgender Zusammenhang zwischen den Ebenen:

$$l^{(s)} = g^{(s)} - \mathcal{W}g^{(s)} = g^{(s)} - g^{(s+1)} \quad s \geq 0$$

Rekonstruktion des Originals $g^{(0)} = \sum_{s=0}^S l^{(s)} + g^{(S)}$

Die Laplacefilter bilden ein (nichtorthogonales) *Basissystem* der Frequenzdarstellung.

Die Faltung läßt sich hier als *lokale Projektion* der Bildfunktion auf Basissystem interpretieren. Der Frequenzraum wird durch Paßbänder (und nicht Dirac-Funktionen) abgetastet! Aufgrund der Nichtorthogonalität der Basisfunktionen überlappen sich die Paßbänder. Die Darstellung ist redundant.



Berechnung einer oktaven Laplace-Pyramide:

Burt, Adelson (1983) schlagen die Berechnung durch DOG vor. Das Problem ist hier, daß die Pyramidenebenen unterschiedlich große Trägerfunktionen besitzen. Die Lösung ist der Umweg über das Stapelkonzept, allerdings ohne den dort notwendigen Aufwand wegen großer Filter.

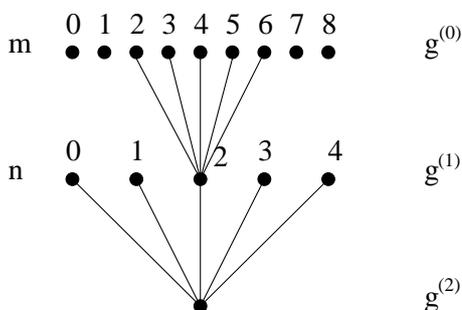
Iterative Berechnung der Ebenen in 3 Schritten:

1. Glättung und Unterabtastung $g^{(s+1)} = \mathcal{S}_2 \mathcal{W}g^{(s)} \quad s \geq 0$
2. Expansion und Interpolation $\hat{g}^{(s+1)} = \mathcal{W} \mathcal{S}_{\frac{1}{2}} g^{(s+1)} = \mathcal{W} \mathcal{S}_{\frac{1}{2}} \mathcal{S}_2 \mathcal{W}g^{(s)}$. Der Interpolationsoperator ist hier identisch mit dem Glättungsoperator. Die Forderung für Interpolationsoperator ist, daß das Integral auf Eins normiert ist.

3. Differenzbildung $l^{(s)} = g^{(s)} - \hat{g}^{(s+1)} = (\mathcal{I} - \mathcal{W}\mathcal{S}_{\frac{1}{2}}\mathcal{S}_2\mathcal{W})g^{(s)}$

Wenn das Original M Pixel groß war (1D-Fall), sind für die Pyramide $M + \frac{M}{2} + \frac{M}{4} + \dots + \frac{M}{2^S} = 2M(1 - 2^{-S-1})$ Pixel nötig. Auch hier erkennt man die Redundanz der Darstellung.

Implementierung:



Ann.: $w = \frac{1}{16}[1 \quad 4 \quad 6 \quad 4 \quad 1] \quad \rightsquigarrow K_Z = 2 \quad (K = 5)$
 $m \in \{0, 1, \dots, M - 1\}$
 $n \in \{0, 1, \dots, \frac{M-1-K_Z}{2}\}$
 d.h. $n = \frac{m-K_Z}{2}, m = 2n + K_Z$

Erinnerung:

- $g_m = \sum_{k=-K_Z}^{K_Z} f_k h_{m-k} = \sum_{k=-K_Z}^{K_Z} h_k f_{m-k}$
- Symmetrie des Operators: Faltung \rightarrow Korrelation
- Skalierung: $g_m = \mathcal{S}_{\Delta} f_m = f_{\Delta m}$

1. Transformation Originalsequenz $g^{(0)} \equiv f = [g_m^{(0)}]$

auf nächste Auflösungsstufe $g^{(1)} = [g_n^{(1)}]$

1. Bereitstellen von Träger der Ausdehnung $\frac{M-1}{2}$
2. Glättung und Unterabtastung

$$g_n^{(1)} = \sum_{k=-2}^2 w_k g_{m+k}^{(0)}$$

2. Expansion und Interpolation

1. Löschen von Träger der Ausdehnung M

2. Expansion

$$g_m^{(1)} = g_{2n+2}^{(1)} \quad \forall n$$

3. Interpolation

$$\hat{g}_m^{(1)} = \sum_{k=-2}^2 w_k g_{m+k}^{(1)} \quad \forall m$$

3. Berechnung der Differenz zwischen Träger (0) und Träger (1)

$$l_m^{(0)} = g_m^{(0)} - \hat{g}_m^{(1)} \quad \forall m$$

\Rightarrow Iteration für alle Stufen s

Pyramiden werden bei der Einbettung unterschiedlichster Probleme in eine **Grob–Fein–Strategie** bzw. **Fein–Grob–Strategie** (seltener) eingesetzt. Typische Beispiele sind

- Kantendetektion
Skalenraumfilterung nennt man ein kontinuierliches Durchstimmen von σ . Die *Kausalität* der Gaußfilter garantiert, daß bei Reduzierung der Auflösung keine neuen Strukturen (z.B. Nulldurchgänge des Laplacefilters) entstehen. Diese Eigenschaft ist eine Folge der Interpretation der Glättung als konservative Diffusion.
- Regularisiertes Matching z.B. Kontur–Modell oder Stereo–Korrespondenz

- Simulation des Attentionsmechanismus

3.7 Lokale Zerlegung der Identität

1. Die Analyse von Bildern erfordert eine **Skalenhierarchie**, die es ermöglicht, die auf verschiedenen Skalen vorhandene Bildinformation zu erfassen. Diese skalierten Strukturen sind einer Abstraktionshierarchie der Objektkonzepte zuzuordnen.

Beispiel: Das Deckblatt dieses Skripts zeigt ein Ölgemälde von Salvatore Dali, auf dem der Maler u.a. einen Rückenakt seiner Frau auf feiner Skala und das Portrait Abraham Lincolns auf grober Skala dargestellt hat. Um die Details des Aktes sehen zu können, muß man näher an das Bild herangehen. Das Portrait hingegen kann man nur aus größerer Entfernung mühelos erkennen. \leadsto Das Visuelle System kann die gesamte Information des Bildes nicht bei einem Betrachtungsabstand (= Skalenstufe) erfassen, sondern benötigt mehrere Skalenstufen (= Skalenhierarchie). ■

Wir haben gelernt, daß Kantendetektoren Modelle der gesuchten Strukturen enthalten. Sie operieren nach dem Prinzip des Template-Matching-Filters, das die Kreuzkorrelation zwischen Maske (Template) und Bildstruktur berechnet. Wollte man dieses Prinzip auf alle möglichen auftretenden Bildstrukturen, übertragen sind zur Verarbeitung dieser Information eine *unendliche Vielfalt von Operatoren* (Filtern) zu entwickeln, die der Vielfalt auftretender Objekt-Strukturen/Skalen entspricht.

Bis Ende 80er Jahre wurde dieses Ziel verfolgt, indem eine Vielzahl *heuristischer Operatortwürfe* realisiert wurde.

- ▶ Eine Systematisierung des Filterentwurfes nach allgemeinen Prinzipien ist erforderlich.
2. Beim Entwurf von Filtern besteht traditionell die Vorstellung, daß die *Antwort eines Filters* die Ausprägung des eine Struktur charakterisierenden **Merkmals** wiedergeben soll. Merkmale sollen aber bestimmte *Invarianzeigenschaften* erfüllen und auf *unterschiedlichen Abstraktionsebenen* existieren. Dies ist schwer im Filterentwurf zu realisieren (nur neuronal modellierte Operatoren erlauben dies).
 - ▶ Merkmale sind das Ergebnis der *linearen bzw. nichtlinearen Verknüpfung* von Filterantworten. Filterung dient also der Vorverarbeitung zur Gewinnung von Merkmalen.
 3. Existiert eine Basis zur Signalrepräsentation, die durch einen Satz von Filtern gebildet wird ?

Sei $\Omega = \{\mathcal{O}_l \mid l = 0, 1, \dots, L - 1\}$ eine **Operatorbasis**, d.h. die Menge der verfügbaren Operatoren, die eine Transformation $f \mapsto f^{(l)}$ induzieren, wobei

$$f^{(l)} = h^{(l)} * f \quad \forall l$$

Forderung: Ω spannt einen *Vektorraum* auf, der isomorph zu einem Teilraum des *Signalraumes* ist.

Dabei ist die Operatormenge Ω eine (vollständige) Basis des Signal-Teilraums, die *operationale Basis*, die vom Konzept der Analyse abhängt. Die Vollständigkeit von Ω garantiert, daß keine Information verloren geht, also alle Strukturen dargestellt werden können, welche im Kontext der Aufgabe interessant sind. Die mathematische Beschreibung ist im Falle orthogonaler Basisvektoren (Operatoren) einfach.

- Filterung wird als orthogonale Projektion der Signale auf ein *Basissystem* verstanden:

$$\hat{f} = \sum_{l=0}^{L-1} \langle f, h^{(l)} \rangle h^{(l)}$$

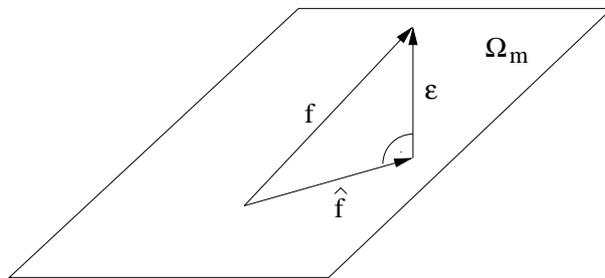
Zerlegung des Signalraums in orthogonale Teilräume

$$F = \Omega_m \oplus \Omega_m^\perp \quad (\Omega_m^\perp \text{ orthogonales Komplement zu } \Omega_m)$$

mit $f = \hat{f} + \varepsilon \quad f \in F, \hat{f} \in \Omega_m, \varepsilon \in \Omega_m^\perp$

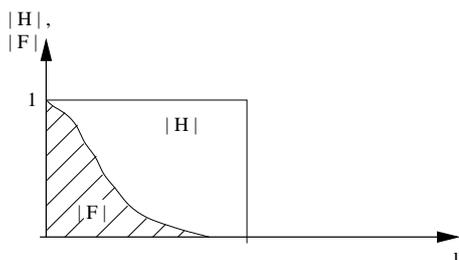
Es gilt $\varepsilon \perp \hat{f}$

und $\mathcal{P}\{f\} = \hat{f}, \mathcal{P}\{\mathcal{P}\{f\}\} = \mathcal{P}\{f\}$.
 Das heißt, \mathcal{P} ist ein Projektionsoperator.



$$\|f\|^2 = \|\hat{f}\|^2 + \|\varepsilon\|^2$$

- aber: Vollständigkeit wird auch durch nicht orthogonale (überlappende) Operatoren erreicht. Die hiermit verbundene Redundanz der Darstellung des Signals kann von Interesse sein.
 - außerdem: Nur pragmatische Vollständigkeit gefordert
 Das Basissystem hat nur die relevante Strukturinformation zu repräsentieren.
 Problem: Wie ist relevante Struktur zu modellieren ?
4. Die im visuellen Cortex gefundenen **rezeptiven Felder** bilden ein Basissystem zur Repräsentation des visuellen Reizes.
 [Hubel and Wiesel, 1959], [Hubel and Wiesel, 1962], [Hubel, 1989] sprechen von Zellen als Merkmalsdetektoren mit gewissen rezeptive Feldern.
- MVS muß sehen können, was es sehen soll (pragmatische Vollständigkeit)
 - MVS muß ökonomisch vorgehen, d.h. es kann keine unendliche Vielzahl von Operatoren hervorbringen, sondern nutzt eine systematische Hierarchiebildung.
- Basissystem entspricht *Taxonomie rezeptiver Felder*
5. Die **Vollständigkeit** der Basis zeigt sich am besten im *Frequenzraum*. Sie ist gewährleistet, wenn jeder Frequenzanteil eines Bildes durch den Satz von Filtern gleichermaßen zur Abbildung gelangt, d.h. die Summe der Übertragungsfunktionen aller Filter konstant 1 ist.
- Das Problem der Vollständigkeit verlagert sich auf die Suche nach geeigneten *Partitionierungen des Frequenzraumes*.



Beispiel: Laplace-Pyramide

$$g^{(0)} \equiv f = \sum_{s=0}^S l^{(s)} + g^{(S)}$$

Für jede Frequenz u gilt

$$|H_{\Omega}| = |H_{G_S}| + |H_{L_S}| + |H_{L_{S-1}}| + \dots + |H_{L_0}|$$

■

Beispiel: Taylorreihenentwicklung

Für jeden Bildpunkt berechne

$$\begin{aligned} f(x + \Delta x) &= f(x) + \sum_{i=1}^n \frac{(\Delta x)^i}{i!} f^{(i)}(x) + \epsilon_n \\ &= \sum_{i=0}^n \frac{(\Delta x)^i}{i!} f^{(i)}(x) + \epsilon_n \\ &\sim h_{\Omega} \approx h^{(0)} + h^{(1)} + \dots + h^{(n)} \sim \text{Operatoren für } n\text{-te Ableitung} \end{aligned}$$

■

siehe Abschnitt 3.7.2 (n -Jets) und Kapitel 4 (Facettenmodell \sim Regression von Flächen n -ter Ordnung an Signalfunktion)

6. Die **Zerlegung der Identität (Split of Identity)** entspricht einer Partitionierung des Frequenzraumes.

Wie kann der Signalwert eines Punktes spektral zerlegt werden?

Die Struktur eines *einzelnen Bildpunktes* ist in der *Korrelation mit Nachbarbildpunkten* enthalten. Ein Bildpunkt trägt nur bezüglich seiner *Nachbarn* Struktureigenschaft.

- Die Projektion der Signalwerte mittels eines *lokalen Operators* gestattet deshalb die Strukturanalyse des Aufpunktes bez. eines Konzeptes.

Nach Euklid ist ein Punkt etwas, das keine Teile hat, also verschwindend klein und nicht beobachtbar ist, da er keine Energie trägt. In der Physik ist ein Punkt eine physikalische Operation, die eine Beobachtung für ein relevantes Volumen erlaubt.

Daraus gewinnen wir eine *operationale Definition eines Punktes*: Ein Punkt ist ein *linearer isotroper Operator* nullter Ordnung, der eine Zahl einem skalaren Feld (bzw. einen Vektor einem Vektorfeld) zuordnet. Die Größe ist willkürlich und hängt vom *Konzept (frame)* ab.

- ▶ Der Punktoperator ist ein Ausblend-/Wichtungsoperator für die operationale Basis des Signalraumes.

Eine **Observable** ist eine beobachtbare physikalische Eigenschaft der Welt. Für die Beobachtung benötigen wir einen Meßapparat.

- Ein Bild ist eine physikalische Observable
- Observable haben eine intrinsische Skala (definiert durch PSF)

Die *Gaußfunktion* ist der ideale Punktoperator.

Probleme :

Die Fouriertransformation ist global. Existieren Konzepte für spektrale Charakterisierung (Amplitude, Phase, Frequenz), die lokal anwendbar sind?

- Hilberttransformation (3.8)
- Gabortransformation (3.10)

Existieren orthogonale Basissysteme für die Zerlegung?

- Hauptachsentransformation (3.7.1)
- Wavelettransformation (3.11)

Was bedeutet lokal?

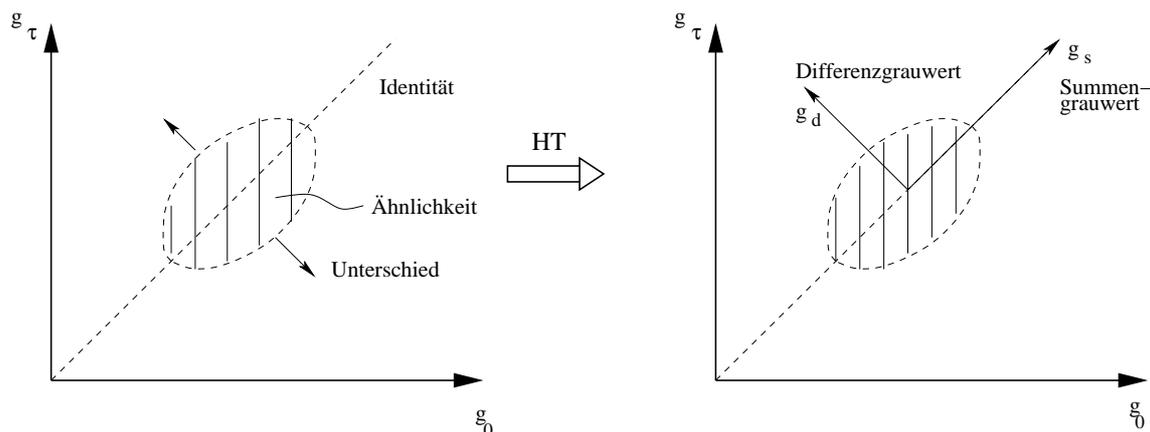
- Bezugsgröße der Lokalität ist die Ausdehnung des Operatorfensters bzw. einer Wichtungsfunktion im Operatorfenster (z.B. Parameter σ der Gaußfunktion).

3.7.1 Lokale stochastische Hauptachsentransformation

[Sommer and Meinel, 1987]. Sie ist die allgemeinste Charakterisierung eines BP bez. *Ähnlichkeit* und *Unterschied* zu seiner Nachbarschaft.

Gedankenexperiment:

Wir zerlegen den Grauwert am Aufpunkt bez. seiner Nachbarschaft in zwei *stochastisch orthogonale* Komponenten. Sei g_0 der Grauwert am Aufpunkt \mathbf{x} und g_τ der Grauwert am Testpunkt $\mathbf{x} + \tau$. Wir finden die Häufigkeit aller Paare (g_0, g_τ) in der Nachbarschaft von \mathbf{x} und tragen sie in ein 2D-Histogramm ein, das dann die Wahrscheinlichkeit $P(i, j) = P(g_0 = g_i \wedge g_\tau = g_j)$ für das Ereignis enthält, daß am Aufpunkt \mathbf{x} der Grauwert g_i und am Testpunkt $\mathbf{x} + \tau$ der Grauwert g_j auftritt.



Hauptachsentransformation (HT) heißt dann, ein in den Schwerpunkt von $P(i, j)$ verschobenes Koordinatensystem so zu drehen daß die *Varianzen* maximal werden, d.h. die Häufigkeiten bez. der Koordinaten des neuen Systems maximal entkoppelt (*entkorreliert*) sind.

Die *Autokovarianzmatrix* ist ein Maß für den Unterschied bzw. die Ähnlichkeit von g_0 und g_τ (Erwartungswerte).

$$C_{ff}(\tau) = \sigma^2(\mathbf{0}) \begin{bmatrix} 1 & \rho(\tau) \\ \rho(\tau) & 1 \end{bmatrix} \xrightarrow{HT} C_{ff}^{HT}(\tau) = \sigma^2(\mathbf{0}) \begin{bmatrix} 1 + \rho(\tau) & 0 \\ 0 & 1 - \rho(\tau) \end{bmatrix}$$

wenn $\sigma^2(\mathbf{0})$ die Varianz und $\rho(\tau)$ der Autokorrelationskoeffizient sind.

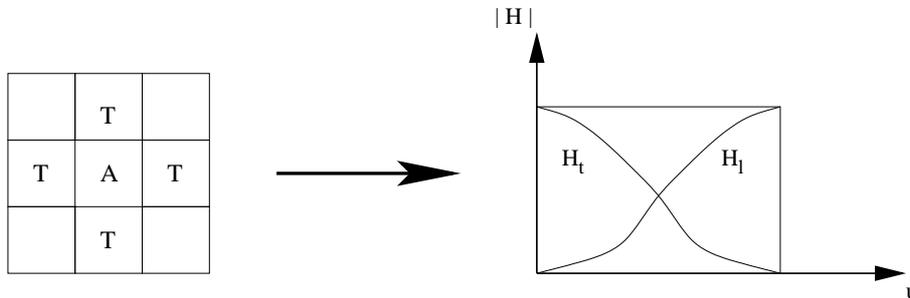
Die HT führt zu einer *Orthogonalisierung* der Zufallsgrößen g_0 und g_τ und ist daher eine *Eigenwerttransformation*. Die stochastische Eigenwerttransformation heißt auch *Karhunen-Loeve-Transformation*

$$\begin{aligned} \text{Eigenvektoren: } g_{s/d} &= \frac{1}{\sqrt{2}}(g_0 \pm g_\tau) \\ \text{Eigenwerte: } \lambda_{1/2} &= \sigma^2(\mathbf{0})(1 \pm \rho(\tau)) \end{aligned}$$

- ▶ Durch *Summen- und Differenzbildung* lassen sich Ähnlichkeit bzw. Unterschied messen.
- ▶ Die Drehung des Koordinatensystems geschieht durch *Schätzung mittels geeigneter Operatoren*

Beispiel: Realisierung in 3×3 -Nachbarschaft:

Ann.: 4-Nachbarschaft, d.h. $|\tau| = 1$ aber 4 verschiedene Richtungen überlagert



$$\begin{aligned}
 h_{HT} &= h_t + h_l \\
 \hat{f} &= \hat{f}_t + \hat{f}_l \\
 &= \underbrace{(h_t + h_l)}_{\delta\text{-Operator}} * f \approx \delta * f \equiv f
 \end{aligned}$$

$$h_t = \frac{1}{8}(\Pi_2^{*2} + \Pi_2^{T*2}) = \frac{1}{8} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{mit } \Pi_2 = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

$$h_l = \frac{1}{8}\Delta = \frac{1}{8}(\nabla_2^{*2} + \nabla_2^{T*2}) = \frac{1}{8} \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad \text{mit } \nabla_2 = \begin{bmatrix} 1 & -1 \end{bmatrix}$$

wobei Δ der Laplace- und ∇ der Nablaoperator sind.

Mit $\nabla_2 = \begin{bmatrix} 1 & -1 \end{bmatrix}$ $\circ \text{---} \bullet H_{\nabla_2}$ bzw. $\nabla_2' = \begin{bmatrix} -1 & 1 \end{bmatrix}$ $\circ \text{---} \bullet H_{\nabla_2'}$

gilt $|H_{\nabla_2}| = |H_{\nabla_2'}|$ und:

$$h_{HT} = \frac{1}{8} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix} + \frac{1}{8} \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} = \frac{1}{8} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \delta$$

■

Zur Konditionierung des Filters falten wir den Tiefpaßoperator h_t bzw. Hochpaßoperator h_l mit einem Gaußfilter

$$h_T = h_t * h_G \quad h_L = h_l * h_G$$

und erhalten einen Tiefpaßoperator h_T und einen Bandpaßoperator h_L .

Prinzip der *Unschärfemaskierung*

Wegen $\delta := h_t + h_l$ gilt

$$\begin{aligned} h_T + h_L &= (h_t + h_l) * h_G \\ &= \delta * h_G \\ &= h_G \\ \Leftrightarrow h_L &= h_G - h_T \end{aligned}$$

DOG-Realisierung des Laplace-Operators

Um das Filter an verschiedene Aufgaben anzupassen, führen wir die Parameter α und β ein und erhalten das *Unschärfemaskierungsfiler*:

$$h_U = \alpha h_T + \beta h_L = \beta h_G + (\alpha - \beta) h_T$$

Dieses Filter wirkt als *Detektor* zur Messung von *Mittelwert* und *Krümmung* der Signalfunktion.

Zunächst könnte man die Parameter α und β als unabhängig betrachten. Es zeigt sich aber, daß die Beziehung $\beta = 1 - \alpha$ sinnvoll ist.

Beispiel: *Anwendungen:*

$$1. \text{ Kantendetektion} \quad \alpha = 0 \quad \beta = 1 \quad \sum_i h_{u_i} = 0$$

$$2. \text{ Glättung} \quad \alpha = 1 \quad \beta = 0 \quad \sum_i h_{u_i} = 1$$

$$3. \text{ Schärfung} \quad \alpha \ll 1 \quad \beta \approx 1 \quad \sum_i h_{u_i} < 1$$

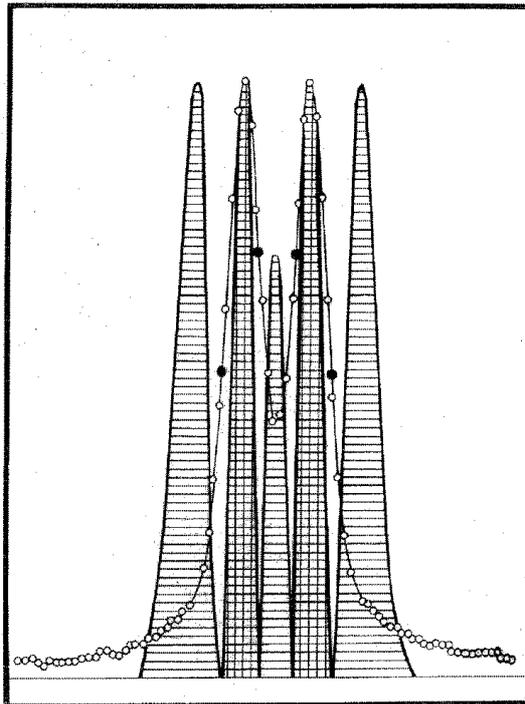
$$4. \text{ Interpolation} \quad \alpha = 1 \quad \overbrace{\beta = 1 - \alpha} \quad \beta \in \mathbb{N} \quad \sum_i h_{u_i} = 1$$

5. *adaptive Glättung (bei signalabhängigem Rauschen)*

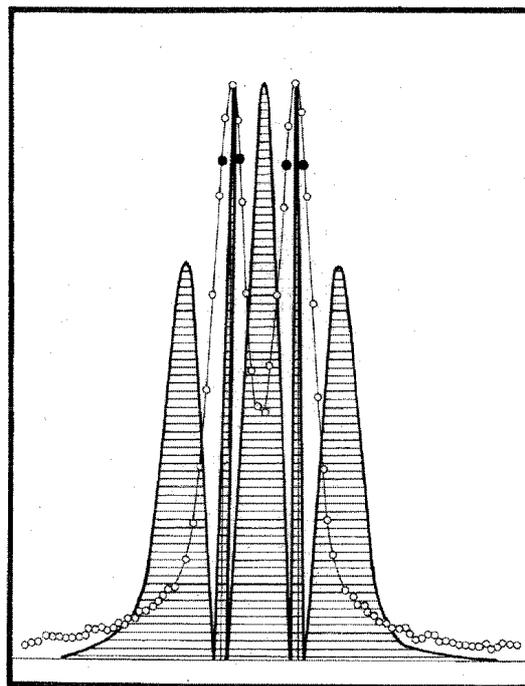
$$\beta_{mn} = 1 - \alpha_{mn} = \left(\frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2} \right)_{mn}$$

mit der Signalvarianz σ_s^2 und der Rauschvarianz σ_n^2 .

$$\begin{aligned} \beta_{mn} &\rightarrow 1, \text{ wenn } \sigma_n^2 \ll \sigma_s^2 &\rightarrow \text{ Betonung der Schärfung} \\ \beta_{mn} &\rightarrow 0, \text{ wenn } \sigma_n^2 \gg \sigma_s^2 &\rightarrow \text{ Betonung der Glättung} \end{aligned}$$



Unschärfemaskierungsfiler als Kantendetektor



Unschärfemaskierungsfiler als Restaurierungsfiler



3.7.2 Unschärfe lokale Ableitungen und Jets

[Koenderink and van Doorn, 1987].

Verschmierung und Diffusionsprozeß

Mit der Einführung des Skalenparameters s kommen wir zu einer Darstellung der Signalfunktion im *Skalenraum* $f(x; s)$. Mit zunehmender Verschmierung (wachsendem s) der Signalfunktion $f(x; s)$ werden lokale Maxima immer weiter abgebaut und lokale Minima immer mehr aufgefüllt, so daß $f(x; s)$ im Grenzwert für große s gegen die *Gleichverteilung* strebt.

Satz (Monotoniebedingung):

Nur die *Gaußfunktion* als erzeugender Operator der Verschmierung hat zur Folge, daß für Extrema 1. Ordnung ($\frac{\partial f}{\partial x} = 0$, $\frac{\partial^2 f}{\partial x^2} \neq 0$) die *Monotoniebedingung*

$$\frac{\partial f(x; s)}{\partial s} \frac{\partial^2 f(x; s)}{\partial x^2} > 0$$

gilt. ■

- ▶ Die Monotoniebedingung gilt streng nur für eindimensionale Funktionen.

Satz (Diffusionsgleichung):

Die Monotoniebedingung gilt für alle Funktionen f , die Lösungen der *Diffusionsgleichung*

$$\frac{\partial^2 f(x; s)}{\partial x^2} = \frac{\partial f(x; s)}{\partial s}$$

sind. ■

- ▶ Im Unterschied zur Darstellung der Diffusionsgleichung in Abschnitt 3.4.3 wird hier die Skala anstelle der Zeit verwendet, um explizit die mit der Zeit fortschreitende Verschmierung zu repräsentieren.

Die Lösungen sind berechenbar als Randwertproblem der partiellen DGL mit der *Greenschen Methode* und heißen **Greensche Funktionen**. Die Gaußfunktion ist die Greensche Funktion der Diffusionsgleichung. Sie bildet einen skalierbaren Punktoperator.

$$\varphi_0(x; s) = \frac{1}{\sqrt{4\pi s}} e^{-x^2/4s} \quad \text{Punktoperator}$$

Diese in der Skalentheorie übliche Darstellung entspricht der Standarddarstellung der Gaußfunktion mit $2s = \sigma^2$.

- ▶ Die Diffusionsgleichung beschreibt die Signalfunktion für alle Auflösungen *gleichzeitig*.
- ▶ Die Signalfunktion wird ersetzt durch eine *einparametrische Familie von Signalfunktionen*, die **Skalenraum-Darstellung**.

Multiresolution mit unscharfen lokalen Ableitungen

Unser Ziel ist die Beschreibung lokaler Geometrie, wobei der Begriff „lokal“ von der Auflösungsstufe abhängt. Die lokale Geometrie eines Punktes berechnen wir aus der *differentiellen Struktur* der Signalfunktion:

1. Berechnung *partieller lokaler Ableitungen* (kann auch für gesamtes Bild erfolgen)
2. Konstruktion *invarianter differentialgeometrischer Merkmale* durch Linearkombination bzw. Nichtlinearkombination partieller Ableitungen.

Der Gegenstand der *Differentialgeometrie* ist die Beschreibung der Geometrie skalarer mehrdimensionaler Funktionen bez. ihrer lokalen Eigenschaften. Bei Grauwertfunktion beziehen sich differentialgeometrische Eigenschaften sowohl auf den *Wertebereich* (Grauwert) als auch auf den *Definitionsbereich* (Ort, Zeit).

Um zu den geforderten Ableitungen der verschmierten Funktion zu kommen, falten wir die Funktion mit verschmierten Ableitungen des Punktoperators.

Diese **unscharfen Ableitungen** sind Ableitungen der Ordnung n des Punktoperators:

$$\begin{aligned}\varphi_n(x; s) &= \frac{\partial^n}{\partial x^n} \varphi_0(x; s) \quad \text{mit} \quad \varphi_0(x; s) = \frac{e^{-\frac{x^2}{4s}}}{\sqrt{4\pi s}} \\ &= \left(\frac{-1}{\sqrt{4s}} \right)^n H_n \left(\frac{x}{\sqrt{4s}} \right) \varphi_0(x; s) \sim \text{Funktion mit } n \text{ Nulldurchgängen} \\ H_n(z) &= (-1)^n e^{z^2} \frac{d^n e^{-z^2}}{dz^n} \quad \text{mit} \quad z = \frac{x}{\sqrt{4s}}\end{aligned}$$

Dabei heißen die H_n *Hermite Polynome* und entstehen in der Ableitung der Gaußfunktion durch die Anwendung der Kettenregel.

Wegen der Linearität aller beteiligten Operatoren folgt:

$$\begin{aligned}\frac{\partial^n}{\partial x^n} [f(x; 0) * \varphi_0(x; s)] &= \left(\frac{\partial^n}{\partial x^n} f(x; 0) \right) * \varphi_0(x; s) \\ &= f(x; 0) * \frac{\partial^n}{\partial x^n} \varphi_0(x; s) \\ &= f(x; 0) * \varphi_n(x; s)\end{aligned}$$

Es gilt das *Verkettungstheorem* für die Skalen s und t :

$$\varphi_n(x; s) * \varphi_m(x; t) = \varphi_{n+m}(x; s+t)$$

Außerdem stellen wir fest:

1. Bei wachsender Ordnung bleibt die Breite der Funktion erhalten, aber die Anzahl der Nulldurchgänge wächst.

$$\frac{\partial}{\partial x} \varphi_n(x; s) = \varphi_{n+1}(x; s)$$

2. Operatoren sind Lösungen der Diffusionsgleichung, da $\frac{\partial^2}{\partial x^2} = \frac{\partial}{\partial s}$

$$\frac{\partial}{\partial s} \varphi_n(x; s) = \varphi_{n+2}(x; s)$$

Eine Multiresolutionsanalyse bez. der Originalfunktion $f(x; 0)$ ist möglich mit unscharfen Ableitungsoperatoren.

Die Operatoren φ_n zeigen *asymptotisches Verhalten* für $n \rightarrow \infty$:

$$\left. \begin{array}{l} n \text{ gerade:} \quad \cos\text{-Fkt.} \\ n \text{ ungerade:} \quad \sin\text{-Fkt.} \end{array} \right\} \text{ mit Gaußfunktion als Einhüllende}$$

$$\omega = \sqrt{\frac{1}{s} \left(\frac{n+1}{2} \right)} \quad \text{Kreisfrequenz der trigonometrischen Funktionen}$$

Die Übertragungsfunktion $\phi_n(u; s) = (j2\pi u)^n e^{-4\pi^2 u^2 s}$ zeigt *Bandpaßverhalten*.

► Eine Funktion mit ähnlicher Charakteristik ist die Gaborfunktion (Abschnitt 3.10)

Die *Erweiterung auf 2D-Fall* geschieht durch:

$$\begin{aligned} \frac{\partial^n}{\partial x^n} &\longrightarrow \frac{\partial^{p+q}}{\partial x^p \partial y^q} && \text{mit } n = p + q \\ \varphi_n &\longrightarrow \varphi_{pq} = \frac{\partial^{p+q}}{\partial x^p \partial y^q} \varphi_{00} \end{aligned}$$

Da der Punktoperator *isotrop und kartesisch separierbar* ist, gilt:

$$\begin{aligned} \varphi_{pq}(x, y; s) &= \varphi_p(x; s) \varphi_q(y; s) \\ \varphi_{pq}(x, y; s) &= \left(\frac{-1}{\sqrt{4s}} \right)^{p+q} H_p \left(\frac{x}{\sqrt{4s}} \right) H_q \left(\frac{y}{\sqrt{4s}} \right) \frac{e^{-\frac{x^2+y^2}{4s}}}{4\pi s} \end{aligned}$$

Beispiel: *unscharfer Laplace-Operator* $\sim \Delta \varphi_{00}(x, y; s)$

Erinnerung: *scharfer Laplace-Op.* $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$

$$\begin{aligned} \varphi_{20}(x, y; s) &= \frac{\partial^2}{\partial x^2} \varphi_{00}(x, y; s) \\ &= \varphi_2(x; s) \varphi_0(y; s) \\ &= \frac{1}{2s} \left(\frac{x^2}{2s} - 1 \right) \frac{e^{-\frac{x^2+y^2}{4s}}}{4\pi s} \\ &= \frac{1}{2s} \left(\frac{x^2}{2s} - 1 \right) \varphi_{00}(x, y; s) \end{aligned}$$

unscharfer Laplace-Operator: $\Delta \varphi_{00}(x, y; s) = \varphi_{20}(x, y; s) + \varphi_{02}(x, y; s)$

in Polarkoordinaten $\Delta \varphi_{00}(r, \varphi; s) = \frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} \frac{e^{-\frac{r^2}{4s}}}{4\pi s} = \Delta r \frac{\varphi_0(r; s)}{\sqrt{4\pi s}}$



Lokale Jets

Definition (lokaler n -Jet):

Ein **lokaler n -Jet** ist die Äquivalenzklasse aller verschmierten Signalfunktionen, die in allen partiellen Ableitungen bis zur Ordnung n übereinstimmen (bezogen auf einen speziellen Punkt für ein gewählte Auflösung).

■

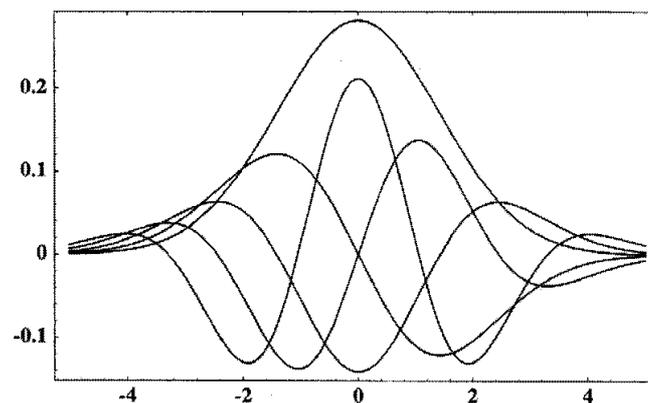
Der $2D$ n -Jet wird durch $\frac{1}{2}(n+1)(n+2)$ Zahlen charakterisiert:

$$n = 0 : \varphi_{00}$$

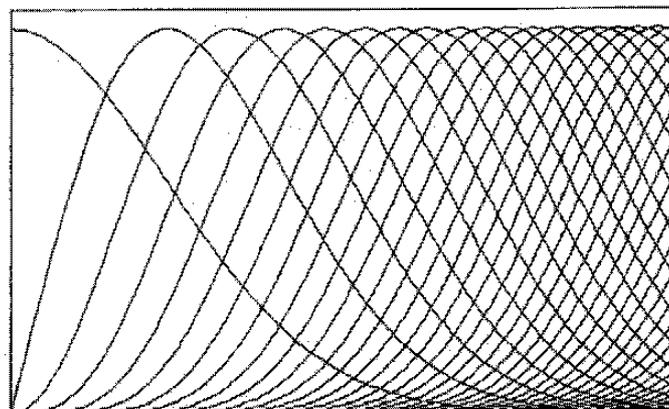
$$n = 1 : \varphi_{00}, \varphi_{10}, \varphi_{01}$$

$$n = 2 : \varphi_{00}, \varphi_{10}, \varphi_{01}, \varphi_{11}, \varphi_{20}, \varphi_{02}$$

Dies entspricht einer abgebrochenen Taylorreihenentwicklung der verschmierten Signalfunktion.



Lokaler 4-Jet : Der Punktoperator und seine Ableitungen haben gleiche Ausdehnung. Die Anzahl der Nulldurchgänge entspricht der Ordnung der Ableitung.



Übertragungsfunktion der Ableitungen des Punktoperators. Das Maximum des Paßbandes verschiebt sich mit u^n , deshalb Zunahme der Rauschempfindlichkeit mit wachsendem n .

Signalfunktionen lassen sich mit Hilfe des unscharfen Punktoperators in eine Taylorreihe entwickeln.

Im 1D-Fall:

$$\begin{aligned} f(x; s) = f(x; 0) * \varphi_0(x; s) &= f(x, 0) * \left(\sum_k \frac{\varphi_k(x; s)}{k!} x^k \right) \\ &= \sum_{k=0}^n \frac{a_k(s)}{k!} x^k \end{aligned}$$

mit

$$a_k(s) = \int_{-\infty}^{\infty} f(x; 0) \varphi_k(x; s) dx = \frac{\partial^k}{\partial x^k} f(x; s) = f(x; 0) * \varphi_k(x; s)$$

wobei die Koeffizienten a_k die Ableitungen der verschmierten Funktion sind.

Im 2D-Fall gilt:

$$\begin{aligned} x^k \rightarrow (x + y)^k &= \sum_{p=0}^k \binom{k}{p} x^p y^{k-p} \quad \text{für } k = p + q \\ f(x, y; s) &= \sum_{k=0}^n \frac{1}{k!} \sum_{p=0}^k \binom{k}{p} a_{pq}(s) x^p y^{k-p} \end{aligned}$$

mit

$$\begin{aligned} a_{pq}(s) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y; 0) \varphi_{pq}(x, y; s) dx dy \\ &= \frac{\partial^{p+q}}{\partial x^p \partial y^q} f(x, y; s) = f(x, y; 0) * \varphi_{pq}(x, y; s) \end{aligned}$$

3.7.3 Differentialgeometrische Maße

Die *Differentialgeometrie* beschreibt die Geometrie differentieller Umgebungen von Punkten auf Kurven oder Flächen auf der Grundlage von partiellen Ableitungen der Funktion.

Differentialgeometrische Maße charakterisieren die *lokale Struktur* von Funktionen.

Zwecks einer allgemeineren Beschreibung von Bildstrukturen sind *Invarianzeigenschaften* solcher Maße von Bedeutung. Verschiebungs- und Rotationsinvarianz lassen beispielsweise eine Beschreibung von Bildstrukturen unabhängig, d.h. *invariant* bezüglich Position und Drehlage in der Bildebene zu.

Aus der Differentialgeometrie bekannt sind die Invarianzen der *Gaußkrümmung* und *mittleren Krümmung*.

Da unscharfe Ableitungsoperatoren *exakte* gemischte partielle Ableitungen der *verschmierten Signalfunktion* liefern, wenn sie auf die originale (nicht verschmierte) Signalfunktion angewendet werden, sind die zu berechnenden differentialgeometrische Maße *exakt* für das *einbezogene Modell des Punktoperators*.

Für die Berechnung eines Maßes wird nicht der gesamte Jet verwendet, sondern nur die für die Maße bedeutsamen partiellen Ableitungen.

3.7.3.1 Fundamentalformen

[Besl and Jain, 1986]

Die Grauwertfunktion des Bildes sei als gekrümmte 2D-Fläche S in E^3 eingebettet. Die Fläche S ist dann in **Parameterform** beschrieben durch:

$$S = \{(x, y, z) \mid x = h(u, v), y = g(u, v), z = f(u, v); \quad (u, v) \in \mathbb{R}^2\}$$

oder, in **vektorieller Parameterform**:

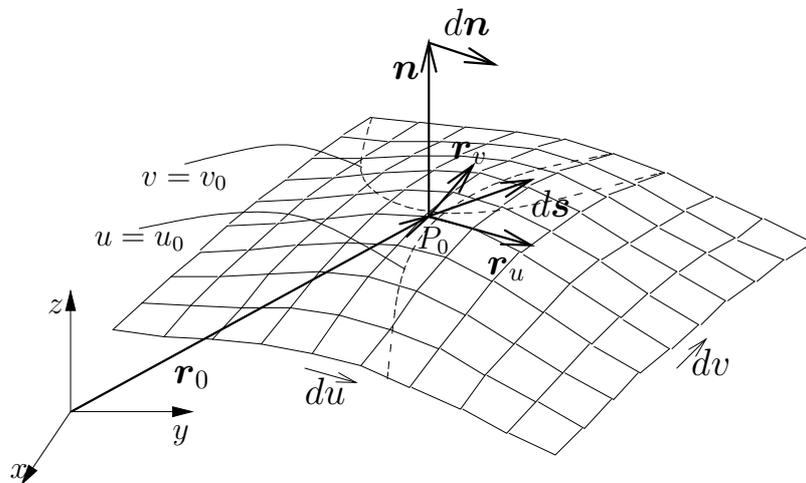
$$\begin{aligned} \mathbf{r} &= \mathbf{r}(u, v) \\ &= x(u, v)\mathbf{e}_1 + y(u, v)\mathbf{e}_2 + z(u, v)\mathbf{e}_3 \end{aligned}$$

Eine *Tangentialebene* mit *Normalenvektor* \mathbf{n} berühre die Fläche S im Punkt $P_0(u_0, v_0) \in S$, beschrieben durch den Ortsvektor $\mathbf{r}_0 = \mathbf{r}(u_0, v_0)$.

Um die Eigenschaften einer Fläche in der Umgebung des Punktes P_0 zu untersuchen, betrachtet man *Kurven* $\mathbf{r}(t)$, die auf S liegen und durch P_0 gehen.

$$\mathbf{r}(t) = \mathbf{r}(u_0 + u(t), v_0 + v(t)) \text{ mit } u(0) = v(0) = 0$$

Spezielle Kurven sind krummlinige *Koordinatenlinien* (u, v) , die durch P_0 gehen und in der Oberfläche liegen.



Der Tangentialvektor an einer durch P_0 gehenden Kurve hat im Punkt P_0 , d.h. bei $t = 0$ die Darstellung

$$\left. \frac{d\mathbf{r}}{dt} \right|_{t=0} = \frac{\partial \mathbf{r}_0}{\partial u} \frac{du(0)}{dt} + \frac{\partial \mathbf{r}_0}{\partial v} \frac{dv(0)}{dt}$$

mit

$$\begin{aligned} \mathbf{r}_u &= \frac{\partial \mathbf{r}_0}{\partial u} = \frac{\partial \mathbf{r}}{\partial u}(u_0, v_0) \\ \mathbf{r}_v &= \frac{\partial \mathbf{r}_0}{\partial v} = \frac{\partial \mathbf{r}}{\partial v}(u_0, v_0) \end{aligned}$$

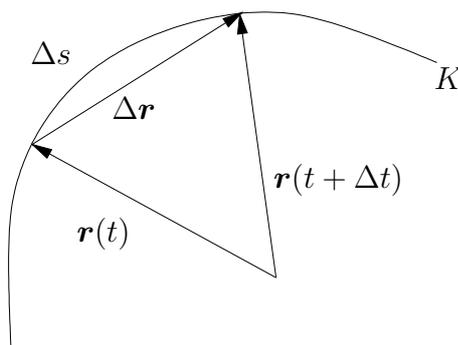
Diese partiellen Ableitungen beschreiben die Tangentialvektoren an den Koordinatenlinien u_0 und v_0 . Sind diese Tangentialvektoren an der Fläche im Punkt P_0 linear unabhängig, so heißt

der Punkt P_0 regulär, andernfalls singulär. Ist P_0 regulär, dann spannen \mathbf{r}_u und \mathbf{r}_v eine Tangentialebene an die Fläche im Punkt P_0 auf. Der *Normalenvektor* dieser Tangentialebene ist dann:

$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}$$

Die *innere Geometrie* einer Fläche wird durch das Bogenelement oder Linienelement ds^2 der Fläche S bestimmt. Für die Bogenlänge $s(t)$ einer Kurve K auf der Fläche gilt in Parameterdarstellung

$$\left(\frac{ds}{dt}\right)^2 = \frac{\partial \mathbf{r}}{\partial u} \frac{\partial \mathbf{r}}{\partial u} \left(\frac{du}{dt}\right)^2 + 2 \frac{\partial \mathbf{r}}{\partial u} \frac{\partial \mathbf{r}}{\partial v} \frac{du}{dt} \frac{dv}{dt} + \frac{\partial \mathbf{r}}{\partial v} \frac{\partial \mathbf{r}}{\partial v} \left(\frac{dv}{dt}\right)^2$$



Die Bogenlänge wird durch folgende Beziehung definiert:

$$s(t) = \int_0^t \left| \frac{\partial \mathbf{r}}{\partial \tau} \right| d\tau$$

Die **erste Fundamentalform** beschreibt das *Bogenelement* auf der Fläche.

$$\begin{aligned} \phi_1 = ds^2 &= E du^2 + 2F du dv + G dv^2 \\ &= \begin{pmatrix} du \\ dv \end{pmatrix}^T \begin{pmatrix} \mathbf{r}_u^2 & \mathbf{r}_u^T \mathbf{r}_v \\ \mathbf{r}_u^T \mathbf{r}_v & \mathbf{r}_v^2 \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix} \\ &= \mathbf{du}^T \mathbf{G} \mathbf{du} \end{aligned}$$

Die Matrix \mathbf{G} ist der *metrische Tensor* der Fläche. Er hängt nicht von der Einbettung der Fläche in E^3 ab. Ebenso bleibt \mathbf{G} bei einer *isometrischen (längenerhaltenden) Abbildung* (z.B. dünnes Papier) erhalten. Der Ausdruck für ds^2 beschreibt die *Riemannsche Metrik*.

$$\begin{aligned} g_{11} = E(u, v) &= \frac{\partial \mathbf{r}}{\partial u} \frac{\partial \mathbf{r}}{\partial u} = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\right)^2 \\ g_{12} = g_{21} = F(u, v) &= \frac{\partial \mathbf{r}}{\partial u} \frac{\partial \mathbf{r}}{\partial v} = \frac{\partial x}{\partial u} \frac{\partial x}{\partial v} + \frac{\partial y}{\partial u} \frac{\partial y}{\partial v} + \frac{\partial z}{\partial u} \frac{\partial z}{\partial v} \\ g_{22} = G(u, v) &= \frac{\partial \mathbf{r}}{\partial v} \frac{\partial \mathbf{r}}{\partial v} = \left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2 + \left(\frac{\partial z}{\partial v}\right)^2 \end{aligned}$$

Die durch den metrischen Tensor bzw. durch die erste Fundamentalform beschriebenen Eigenschaften der inneren Geometrie einer Fläche sind zum Beispiel der Flächeninhalt oder die Länge von Kurven auf der Fläche. Zum Beispiel ist

$$A = \int \int_{\Delta S} \sqrt{EG - F^2} du dv$$

der Flächeninhalt A eines Flächenstücks $\Delta S \subset S$.

Die **Monge-Darstellung** (engl. monge patch) der Fläche

$$x = u, y = v, z = f(u, v) \quad \text{oder} \quad \mathbf{r} = \begin{pmatrix} x \\ y \\ f(x, y) \end{pmatrix}$$

vereinfacht die Behandlung von Funktionen über Bildern. Die (geneigte) Ebene des Bildträgers wird Tangentialebene.

Die Monge-Darstellung liefert:

$$\begin{aligned} E &= \frac{\partial \mathbf{r}}{\partial u} \frac{\partial \mathbf{r}}{\partial u} = 1 + \left(\frac{\partial f}{\partial x} \right)^2 \\ F &= \frac{\partial \mathbf{r}}{\partial u} \frac{\partial \mathbf{r}}{\partial v} = \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \\ G &= \frac{\partial \mathbf{r}}{\partial v} \frac{\partial \mathbf{r}}{\partial v} = 1 + \left(\frac{\partial f}{\partial y} \right)^2 \end{aligned}$$

oder, mit $f_x := \frac{\partial f}{\partial x}$, $f_y := \frac{\partial f}{\partial y}$ und $f_{xy} := \frac{\partial f}{\partial x} \frac{\partial f}{\partial y}$

$$\begin{aligned} \phi_1 &= d\mathbf{x}^T \mathbf{G} d\mathbf{x} \\ &= \begin{pmatrix} dx \\ dy \end{pmatrix}^T \begin{pmatrix} 1 + f_x^2 & f_x f_y \\ f_x f_y & 1 + f_y^2 \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} \end{aligned}$$

Der metrische Tensor wird also durch nichtlineare Funktionen der partiellen Ableitungen erster Ordnung, der Richtungsgradienten, von der Bildfunktion bestimmt.

Die **zweite Fundamentalform** beschreibt die Gestalt der Fläche bez. *Krümmungscharakteristik*.

$$\begin{aligned} \phi_2 &= -d\mathbf{s}^T d\mathbf{n} = L du^2 + 2M du dv + N dv^2 \\ &= \begin{pmatrix} du \\ dv \end{pmatrix}^T \begin{pmatrix} \mathbf{r}_{uu}^T \mathbf{n} & \mathbf{r}_{uv}^T \mathbf{n} \\ \mathbf{r}_{uv}^T \mathbf{n} & \mathbf{r}_{vv}^T \mathbf{n} \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix} \\ &= d\mathbf{u}^T \mathbf{B} d\mathbf{u} \end{aligned}$$

Die zweite Fundamentalform enthält die Krümmungseigenschaften der Fläche und hängt von der Einbettung der Fläche in E^3 ab. Sie beschreibt die *äußere Geometrie* einer Fläche im Raum. Der Tensor \mathbf{B} ist der *Krümmungstensor* der Fläche. Seine Elemente enthalten die Projektionen zweiter partieller Ableitungen von S auf den Normalenvektor \mathbf{n} .

In der Monge-Darstellung erhalten wir:

$$\begin{aligned} \mathbf{n} &= \left(1 + \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right)^{-\frac{1}{2}} \left(-\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1 \right)^T \quad \text{Oberflächennormale} \\ b_{11} &\equiv L = \left(\frac{\partial^2 \mathbf{r}}{\partial u^2} \right)^T \mathbf{n} = \frac{\frac{\partial^2 f}{\partial x^2}}{\sqrt{1 + \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2}} \\ b_{12} = b_{21} &\equiv M = \left(\frac{\partial^2 \mathbf{r}}{\partial u \partial v} \right)^T \mathbf{n} = \frac{\frac{\partial^2 f}{\partial x \partial y}}{\sqrt{1 + \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2}} \\ b_{22} &\equiv N = \left(\frac{\partial^2 \mathbf{r}}{\partial v^2} \right)^T \mathbf{n} = \frac{\frac{\partial^2 f}{\partial y^2}}{\sqrt{1 + \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2}} \end{aligned}$$

und kurz:

$$\begin{aligned} \phi_2 &= d\mathbf{x}^T \mathbf{B} d\mathbf{x} \\ &= \frac{1}{\sqrt{1 + f_x^2 + f_y^2}} \begin{pmatrix} dx \\ dy \end{pmatrix}^T \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} \end{aligned}$$

Der Krümmungstensor wird also durch die partiellen Ableitungen zweiter Ordnung von der Bildfunktion bestimmt. Er gibt ein lokales Maß für die Richtungsänderung der Flächennormalen bei Bewegung in der von (u, v) aufgespannten Parameterebene an.

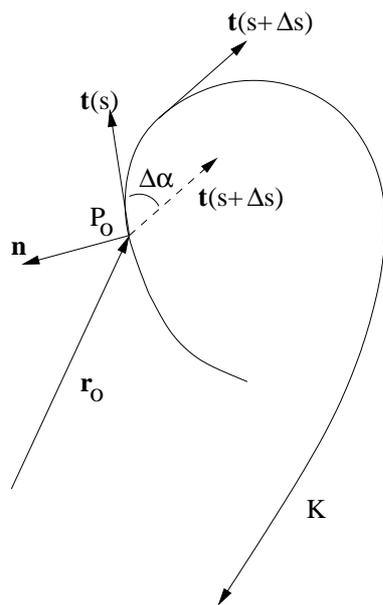
3.7.3.2 Krümmungsmaße

Krümmung, Windung und Bogenlänge einer Kurve sind *Invariante* bezüglich *Euklidischer Transformationen* (Bewegung eines starren Körpers) im Raum. Diese Eigenschaften übertragen sich auch auf Flächen im Raum. Wir haben vorn gesehen, wie spezielle in der Fläche liegende und durch Punkt P_0 gehende Kurven zur Berechnung der Fundamentalformen der Fläche führen.

Krümmung einer Kurve im Raum:

Die *Krümmung* ist ein Maß für die Abweichung einer Kurve von ihrer Tangente im betrachteten Punkt P_0 . Der Tangenteneinheitsvektor \mathbf{t} an P_0 ist definiert durch

$$\mathbf{t} = \frac{d\mathbf{r}}{ds} = \mathbf{r}', \quad |\mathbf{t}| = 1$$



Demzufolge ist die Krümmung von K in P_0 :

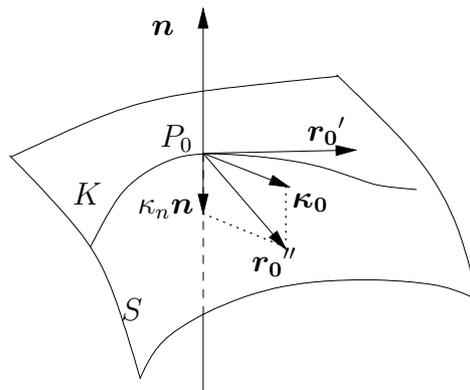
$$\kappa(s) = \frac{dt}{ds} = \lim_{\Delta s \rightarrow 0} \left| \frac{\Delta \alpha}{\Delta s} \right| = \left| \frac{d^2 \mathbf{r}}{ds^2} \right|$$

Man bezeichnet

$$\mathbf{r}'' = \kappa(s) \mathbf{n}$$

als den in Richtung der *Hauptnormalen* \mathbf{n} der Kurve am Ort P_0 weisenden *Krümmungsvektor* der Kurve.

Krümmung einer Fläche im Raum:



Um die Krümmung in der Umgebung eines Punktes P_0 auf einer Fläche S zu untersuchen, wird die Krümmung von Kurven betrachtet, die in S liegen und durch P_0 gehen.

Die Projektionen des Krümmungsvektors \mathbf{r}_0'' auf die Flächennormale \mathbf{n} führt zu folgender Zerlegung des Krümmungsvektors:

$$\mathbf{r}_0'' = \kappa_n \mathbf{n} + \boldsymbol{\kappa}_0 \quad , \text{ wobei } \mathbf{n}^\top \boldsymbol{\kappa}_0 = 0, \text{ d.h. } \boldsymbol{\kappa}_0 \text{ ist Tangentialvektor}$$

Folgende Krümmungsmaße entstehen:

$$\begin{array}{ll} \kappa_g = |\boldsymbol{\kappa}_0| & \text{geodätische Krümmung} \\ \kappa_n & \text{Normalenkrümmung} \end{array}$$

Die *geodätische Krümmung* ist ein Maß der inneren Geometrie der Fläche. Sie wird durch die *erste Fundamentalform* bestimmt. Sie ist invariant gegen Verbiegungen.

Die *Normalenkrümmung* ist von der Einbettung der Fläche im Raum abhängig, also ein Maß der äußeren Geometrie der Fläche. Sie wird durch die *zweite Fundamentalform* gegeben.

Für die folgenden Betrachtungen wird die äußere Geometrie auf die innere durch Normierung bezogen (Weingarten-Abbildung), d.h. es werden Quotienten $\frac{\mathbf{B}}{\mathbf{G}}$ (Weingarten-Matrix) für die Angabe der Krümmung verwendet.

1. Die **Normalenkrümmung** κ_n

$$\kappa_n = \frac{\phi_2}{\phi_1} = -\frac{dn}{dr}$$

ist die Projektion der Flächenkrümmung (des Krümmungsvektors *einer speziellen Kurve*) im Punkt P_0 in Richtung der Flächennormalen.

2. Die **Hauptkrümmungen** κ_1, κ_2 sind die Maximal- und Minimalwerte der Normalenkrümmungen *beliebiger Kurven* durch P_0 . Sie ergeben sich als *Eigenwerte des Krümmungstensors* \mathbf{B} bzw. von $\frac{\mathbf{B}}{\mathbf{G}}$. Die Hauptkrümmungen sind Lösungen der Gleichung

$$(EG - F^2)\kappa^2 - (EN + GL - 2FM)\kappa + (LN - M^2) = 0$$

$$\Leftrightarrow \kappa_{1/2} = H \pm \sqrt{H^2 - K} \quad \begin{array}{l} H \sim \text{mittlere Krümmung} \\ K \sim \text{Gaußsche Krümmung} \end{array}$$

Für einen regulären Punkt (bei Unabhängigkeit der Tangentialvektoren des Punktes) gilt $\kappa_1 \perp \kappa_2$.

3. Die **Gaußsche Krümmung** K

$$K = \frac{\det \mathbf{B}}{\det \mathbf{G}} = \frac{LN - M^2}{EG - F^2} = \kappa_1 \kappa_2$$

ist eine Eigenschaft der *inneren Geometrie* und bleibt unter isometrischen Abbildungen erhalten. Sie ist unabhängig von der Repräsentation der Fläche.

4. Die **mittlere Krümmung**

$$H = \frac{1}{2} \text{trace} \frac{\mathbf{B}}{\mathbf{G}} = \frac{EN + GL - 2FM}{2(EG - F^2)^{3/2}} = \frac{1}{2}(\kappa_1 + \kappa_2)$$

ist eine Eigenschaft der *äußeren Geometrie* einer Fläche.

Beide sind aber *Invariante* bezüglich Euklidischer Transformationen, d.h. gegenüber *Rotation* und *Translation*, solange die *Jacobimatrix* der Transformation nicht verschwindet. Die Jacobimatrix ist eine Matrix von ersten partiellen Ableitungen von k Funktionen über n Variablen.

$$\mathbf{J} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1} & \cdots & \frac{\partial f_k}{\partial x_n} \end{pmatrix}$$

In der Folge werden wir Regionentypen durch das Vorzeichen von mittlerer und Gaußscher Krümmung unterscheiden. Hierzu kann eine Variante dieser Krümmungsmaße verwendet werden, die auf die *Hessematrix* \mathbf{H} zurückgeführt wird.

$$\mathbf{H} = \nabla \nabla^T f = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}$$

Die Hessematrix hat die gleiche Struktur wie der Krümmungstensor in Monge-Darstellung. Sind d_1 und d_2 die Eigenwerte der Hessematrix, so erhält man als Analogon zur Gaußschen Krümmung K die *Determinante der Hessematrix*

$$D_1 = \det \mathbf{H} = d_1 d_2 = f_{xx} f_{yy} - f_{xy}^2$$

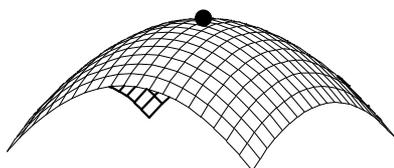
und als Analogon zur mittleren Krümmung das Ergebnis einer Anwendung des *Laplace-Operators* auf f , d.h.

$$D_2 = \Delta f = \frac{1}{2} \text{trace} \mathbf{H} = \frac{1}{2} (d_1 + d_2) = \frac{1}{2} (f_{xx} + f_{yy}).$$

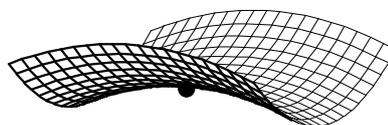
Die Werte von K und H , bzw. von D_1 und D_2 , dienen der *Klassifizierung von Flächenpunkten* $P \in S$. Die klassifizierten Muster fassen wir als **Oberflächenprimitiva** auf. In der klassischen Differentialgeometrie werden folgende vier Grundtypen von Flächen unterschieden.

Flächentyp	$H = \frac{1}{2}(\kappa_1 + \kappa_2)$	$K = \kappa_1 \kappa_2$
elliptisch		$K > 0$
hyperbolisch		$K < 0$
parabolisch	$H \neq 0$	$K = 0$
planar	$H = 0$	$K = 0$

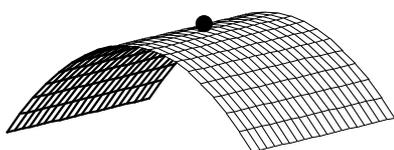
elliptisch



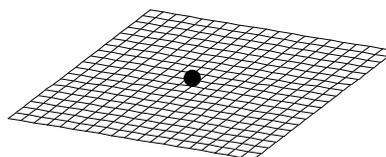
hyperbolisch



parabolisch



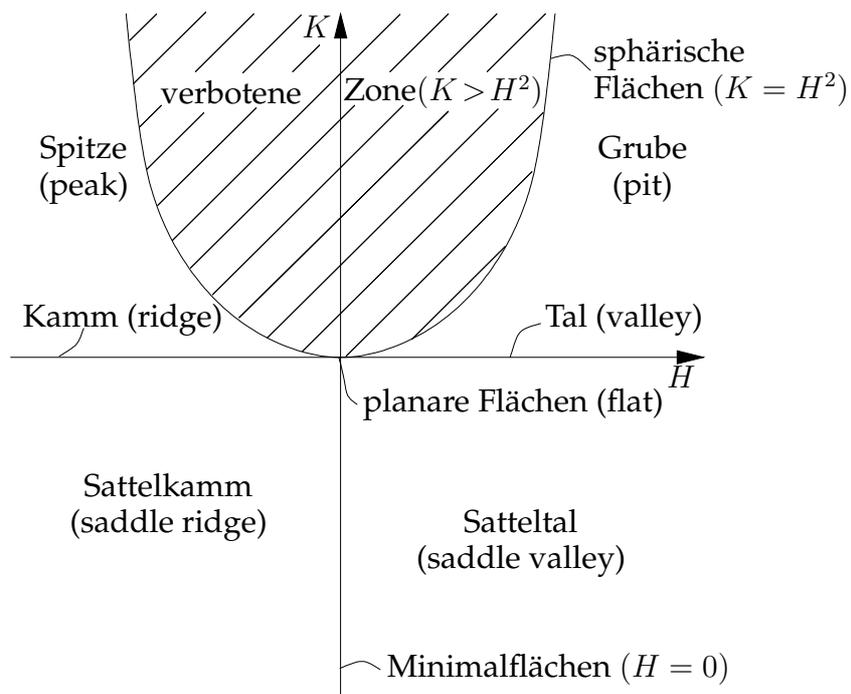
planar



Der K-H-Raum:

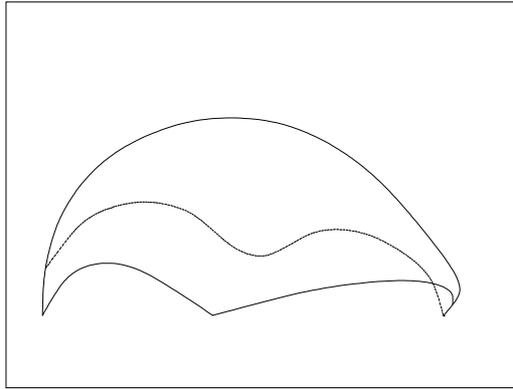
Wegen der Invarianzeigenschaft von Gaußscher und mittlerer Krümmung bezüglich der Blickrichtung auf eine 3D-Oberfläche spielt die lokale Interpretation der Signalfunktion im K-H-Raum eine interessante konzeptionelle Rolle (Terzopoulos, 1988).

Im K-H-Raum werden 8 Primitiva für Oberflächen unterschieden.

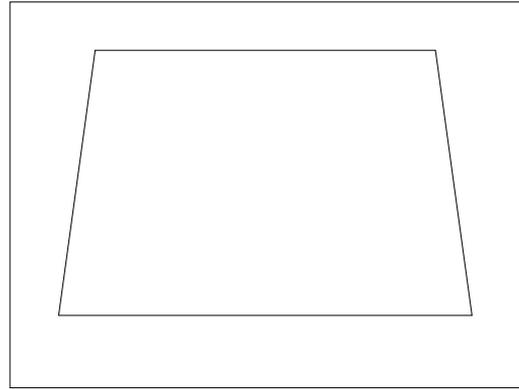


Alle regulären Punkte liegen außerhalb der verbotenen Zone.

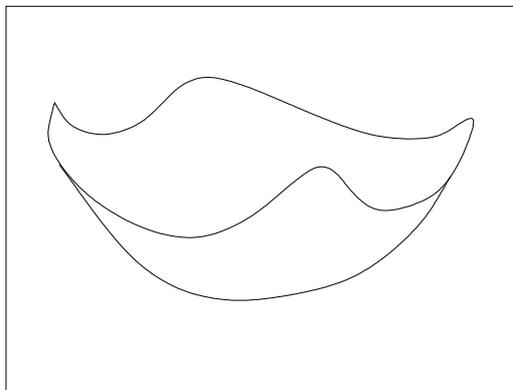
Spitzenfläche $H < 0$ $K > 0$



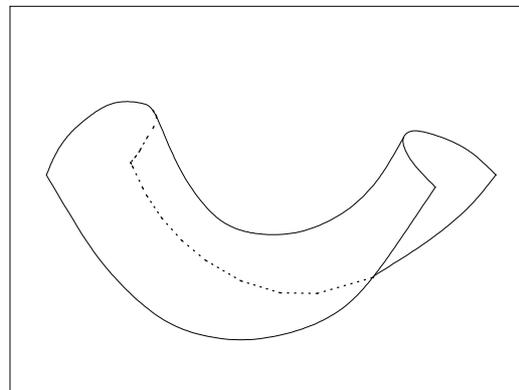
Planare Fläche $H = 0$ $K = 0$



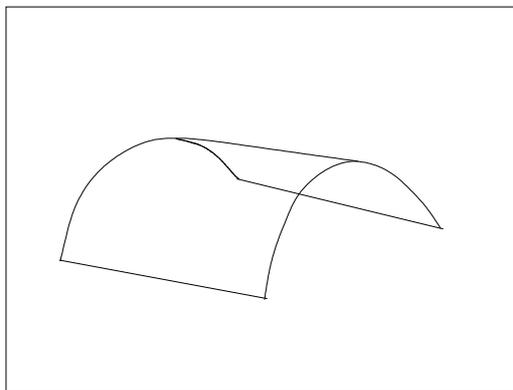
Grubenfläche $H > 0$ $K > 0$



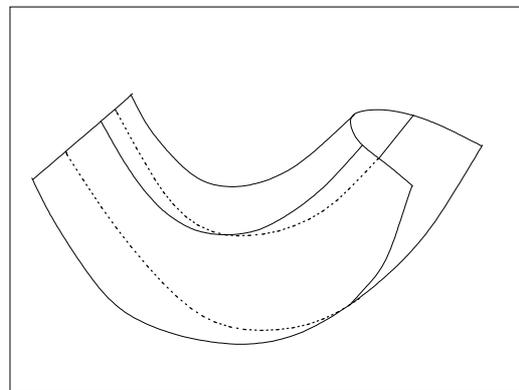
Minimalfläche $H = 0$ $K < 0$



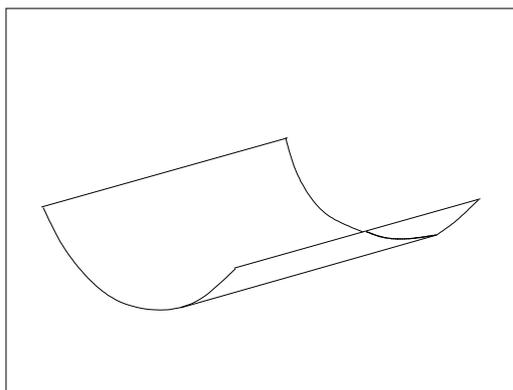
Kammfläche $H < 0$ $K = 0$



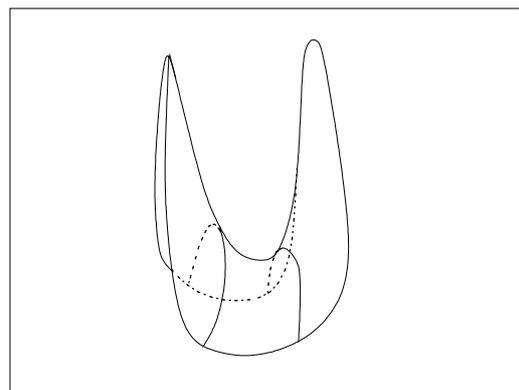
Sattelkamm-Fläche $H < 0$ $K < 0$



Talfläche $H > 0$ $K = 0$



Satteltal-Fläche $H > 0$ $K < 0$



Berechnung von Invarianten mittels partieller Ableitungen:

Koenderink, van Doorn (1982, 1986) schlagen vor, nach Geometrien zu suchen, die auf natürliche Weise die invarianten Leistungen des menschlichen visuellen Wahrnehmungssystems repräsentieren können.

Bsp.: $(f_{xx} f_{yy} - f_{xy}^2)$ sind invariant bzgl. perspektivischer Projektion.

Zur Schätzung der zur Krümmungsberechnung notwendigen partiellen Ableitungen wenden wir das Konzept der *lokalen Jets* an. Bei Monge-Darstellung erhält man mit den Koeffizienten aus 2-Jets a_{ij} :

$$\begin{aligned}
 K &= \frac{\frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2}{\left(1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\right)^2} \\
 &= \frac{a_{20}a_{02} - a_{11}^2}{(1 + a_{10}^2 + a_{01}^2)^2} \\
 H &= \frac{\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial x^2} \left(\frac{\partial f}{\partial y}\right)^2 + \frac{\partial^2 f}{\partial y^2} \left(\frac{\partial f}{\partial x}\right)^2 - 2 \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \frac{\partial^2 f}{\partial x \partial y}}{2 \left(1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\right)^{\frac{3}{2}}} \\
 &= \frac{(1 + a_{10}^2)a_{02} + (1 + a_{01}^2)a_{20} - 2a_{10}a_{01}a_{11}}{2(1 + a_{10}^2 + a_{01}^2)^{\frac{3}{2}}}
 \end{aligned}$$

Die Krümmung einer Isophote (Konturkrümmung) ist:

$$\kappa = \frac{-a_{01}^2 a_{20} + 2a_{10} a_{01} a_{11} - a_{10}^2 a_{02}}{(a_{10}^2 + a_{01}^2)^{\frac{3}{2}}}$$

Das ist die Krümmung senkrecht zum Gradienten, also entlang der Kante.

Nach [ter Haar Romeny *et al.*, 1993, Florack *et al.*, 1994] ergeben lineare oder nichtlineare *Kombinationen* einfacher Invarianten wieder Invariante, so daß die Berechnung eines *vollständigen Satzes* möglicher Invarianten für n -Jets (also unter Verwendung n -ter unscharfer Ableitungen) durch Formelmanipulation möglich ist.

Bei Ordnungen höher als zwei geht die Manipulation solcher Maße weit über die Differentialgeometrie hinaus, die Interpretation der Invarianten ist daher meist unbekannt.

Für Mannigfaltigkeiten der Dimension $n > 2$ ist eine Behandlung der Krümmung ebenfalls möglich. Derartige Betrachtungen sind beispielsweise für Bildfolgenanalyse von Interesse. Die Einbettung einer Bildfolge ($n = 3$) in einem p -dimensionalen Raum ($p \geq 4$) wird zu einer Krümmungsanalyse führen, die durch den *Riemannschen Krümmungstensor* beschrieben wird (Barth, 1995).

Da die Skala selbst eine Dimension darstellt, wird eine Multiskalen-Bildfolgenanalyse durch $n = 4$ beschrieben.

Nimmt man farbige Bilder an und ordnet den drei Farbkanälen ebenfalls je eine Dimension zu, so liegt bereits der Fall $n = 5$ vor (Sochen, 1998) und für farbige Multiskalen-Bildfolgen gilt demzufolge $n = 7$.

Wir erkennen, wie schwierig die Behandlung werden kann, wenn die Anzahl der Freiheitsgrade der Signalrepräsentation wächst.

3.8 Lokale Spektrale Repräsentation

Problem: Wie kann man erreichen, daß die spektralen Charakteristika *Amplitude* und *Phase* auch im Ortsraum für jeden Punkt $x \in D$ anwendbar sind?

Ziel: Die Größen $A(x), \phi(x)$ sollen nun *Repräsentationen der Signalfunktion* $f(x)$ sein. \sim *Zerlegung der Identität* bez. dieser beiden Konzepte

Lösung: Das reelle Signal $f(x)$ muß in ein *komplexes Signal*

$$z(x) = f(x) + j\hat{f}(x)$$

transformiert werden. Dabei bildet die reelle Funktion den Realteil und eine in geeigneter Weise transformierte Funktion $\hat{f}(x)$ den Imaginärteil des komplexen Signals. Dieses komplexe Signal heißt **analytisches Signal**.

- ▶ Das analytische Signal wurde von Gabor (1946) eingeführt. Er nannte es komplexes Signal.
- ▶ Der Begriff analytisches Signal wurde später eingeführt und ist heute üblich.
- ▶ In der Literatur findet man auch häufig die Formulierung des analytischen Signals als $z(x) = f(x) - j\hat{f}(x)$. Durch das hier gegenüber oben geänderte Vorzeichen des Imaginärteils ergeben sich in der Folge auch andere abgeleitete Beziehungen. Beide Formulierungen sind gleichwertig.

Die Funktion $\hat{f}(x)$ ist gegenüber $f(x)$ um $-\frac{\pi}{2}$ phasenverschoben.

Problem: Wie ist die komplexe Erweiterung zu realisieren, ohne daß eine Amplitudenänderung des Signals auftritt?

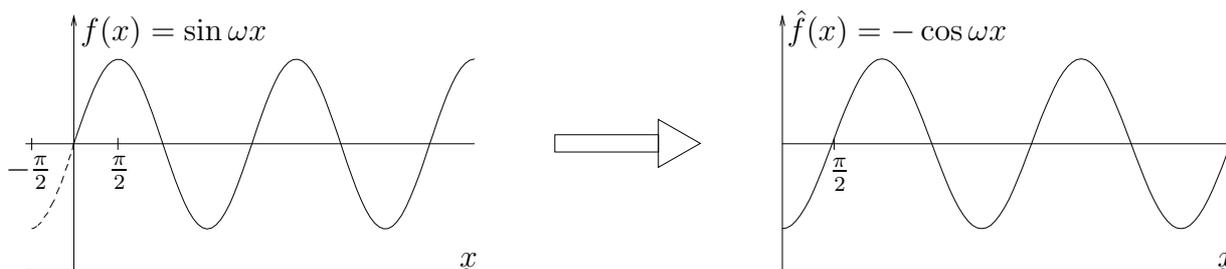
Lösung: Zerlegung der Identität nach $\cos^2 \theta + \sin^2 \theta = 1$

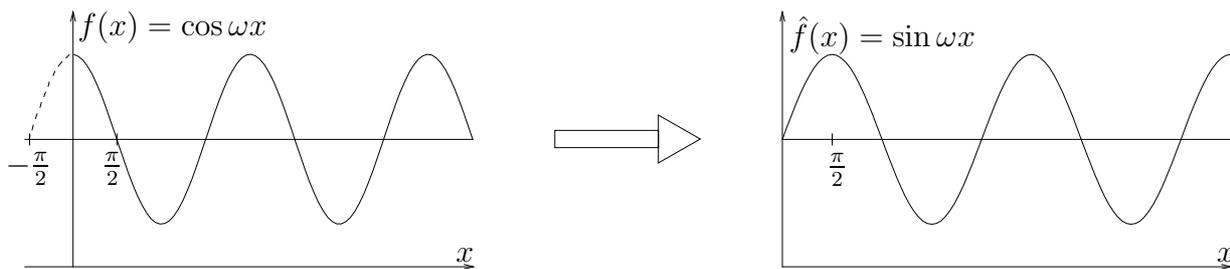
Wirkung: Die Vorstellung einer lokalen Fourierreihen-Zerlegung eines Signals impliziert, daß ein Signal durch Superposition harmonischer Funktionen unterschiedlicher Ortsfrequenz synthetisiert werden kann. Deren gerade Komponenten (cos-Funktionen) und ungeraden Komponenten (sin-Funktionen) sollen in obiger Beziehung verbunden sein.

Die geforderte Phasenverschiebung von $f(x)$ und $\hat{f}(x)$ wird durch die Berechnung der *negativen Ableitung* einer Funktion erreicht, z.B.:

$$-\frac{d}{dx} \cos(2\pi u x) = 2\pi u \sin(2\pi u x)$$

(Die Skalierung der Amplitude wird durch Normierung der Winkelfrequenz $\omega = 2\pi u$ auf deren Absolutwert vermieden.)





Daraus folgt: Wenn die komplexe Erweiterung des reellen Signals zum analytischen Signal in der Weise realisiert wird, daß die Komponente $\hat{f}(x)$ lediglich durch eine Phasenverschiebung um $|\pi/2|$ aus dem Realteil $f(x)$ gebildet wird, so ist die Summe der Quadrate beider Komponenten eine *phasenunabhängige Funktion*, die lokale Energie, welche gleich ist dem Amplitudenquadrat des Signals.

Lokale Energie und lokale Phase

$$E(x) = f^2(x) + \hat{f}^2(x) \quad \text{lokale Energie}$$

$$\phi(x) = \arg(z(x)) \quad \text{lokale Phase}$$

bilden das Konzept der **Zerlegung der Identität** für alle Ortsfrequenzen einer spektralen Analyse.

- ▶ Lokale Energie und lokale Phase sind in natürlichen Bildern stochastisch *entkorreliert*. Sie sind orthogonal im stochastischen Sinn (bzgl. der Schätzung des Erwartungswerts) und bilden eine Basis der lokalen Signalstruktur.
- ▶ Diese Zerlegung hat zur Folge, daß die *Autokorrelationsfunktionen* beider Komponenten des analytischen Signals gleich sind:

$$R_f(\tau) = R_{\hat{f}}(\tau)$$

Wegen $R_f(\tau) = \mathcal{F}^{-1}\{P_F(u)\}$ bedeutet dies eine *Verdopplung der spektralen Leistung* $P_Z(u)$ für die Fouriertransformierte des analytischen Signals:

$$Z(u) = F(u) + j\hat{F}(u)$$

Die Forderung nach der *Erhaltung der Energie* bei der komplexen Erweiterung des reellen Signals führt zu folgender Lösung für die Fouriertransformierte des analytischen Signals:

$$Z(u) = \begin{cases} 2F(u) & u > 0 \\ F(u) & u = 0 \\ 0 & u < 0 \end{cases}$$

- ▶ Sämtliche Energie des analytischen Signals liegt im positiven Frequenzbereich. Dies stellt für reelle Signale keinen Informationsverlust dar. Da die Fouriertransformierte eines reellen Signals komplex hermitesch ist ($F(-u) = F^*(u)$), können die spektralen Anteile mit negativer Frequenz rekonstruiert werden.
- ▶ Das Spektrum eines analytischen Signals ist *nicht hermitesch*, d.h. $F(-u) \neq F^*(u)$, da $Z(u) = 0$ für $u < 0$. Deshalb kann $z(x)$ nicht reell sein.

Aus der oben erwähnten Berechnung der negativen Ableitung des Spektrums folgt die *Lösung zur Berechnung des Imaginärteils des analytischen Signals*:

- Berechnung der negativen Ableitung: Multiplikation des Spektrums mit $-j2\pi u$
- Division durch $|2\pi u|$

$$\hat{F}(u) = H_I(u)F(u) = -j \frac{u}{|u|} F(u) = -j \operatorname{sgn}(u) F(u)$$

mit

$$\begin{aligned} A_{\hat{F}}(u) &= A_F(u) && \text{gleiches Amplitudenspektrum} \\ \phi_{\hat{F}}(u) &= \begin{cases} \phi_F(u) - \frac{\pi}{2} & u > 0 \\ \phi_F(u) & u = 0 \\ \phi_F(u) + \frac{\pi}{2} & u < 0 \end{cases} && \text{verschobenes Phasenspektrum} \end{aligned}$$

da $|H_I| = 1$ für alle Frequenzen und da $H_I(u)$ rein imaginär ist. Im nächsten Abschnitt bezeichnen wir den Operator \mathcal{H}_I als Hilbert-Transformation.

a) Berechnung des analytischen Signals im Frequenzraum

$$\begin{aligned} Z(u) &= \begin{cases} 2F(u) & u > 0 \\ F(u) & u = 0 \\ 0 & u < 0 \end{cases} \\ &= 2F(u) \operatorname{step}(u) && \text{mit } \operatorname{step}(u) = \begin{cases} 1 & u > 0 \\ \frac{1}{2} & u = 0 \\ 0 & u < 0 \end{cases} \\ &= F(u)(1 + \operatorname{sgn}(u)) && \text{wegen } \operatorname{step}(u) = \frac{1}{2}(1 + \operatorname{sgn}(u)) \\ &&& \text{mit } \operatorname{sgn}(u) = \begin{cases} 1 & u > 0 \\ 0 & u = 0 \\ -1 & u < 0 \end{cases} \end{aligned}$$

Hieraus folgt

$$\begin{aligned} Z(u) &= F(u) + j(-j \operatorname{sgn}(u)) F(u) = F(u) + jH_I(u)F(u) \\ &= F(u)(1 + \operatorname{sgn}(u)). \end{aligned}$$

Der Operator $H(u) = 1 + \operatorname{sgn}(u)$ erzeugt die Fouriertransformierte des analytischen Signals aus der Fouriertransformierten des reellen Signals. Es gilt

$$\begin{aligned} |H(u)| &= \begin{cases} 2 & u > 0 \\ 1 & u = 0 \\ 0 & u < 0 \end{cases} \\ |\Delta\phi(u)| &= \frac{\pi}{2} \end{aligned}$$

- b) Berechnung des analytischen Signals im Ortsraum durch inverse Fouriertransformation von $Z(u)$.

Aus

$$Z(u) = F(u) + j\hat{F}(u)$$

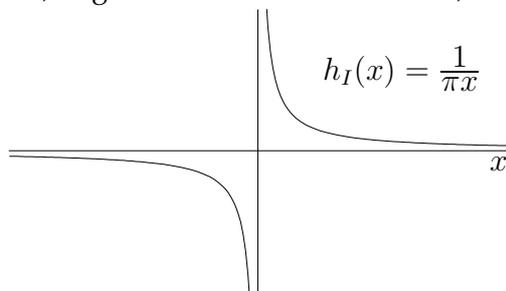
folgt

$$Z(u) = H(u) \cdot F(u) = (H_R(u) + jH_I(u)) \cdot F(u)$$

mit

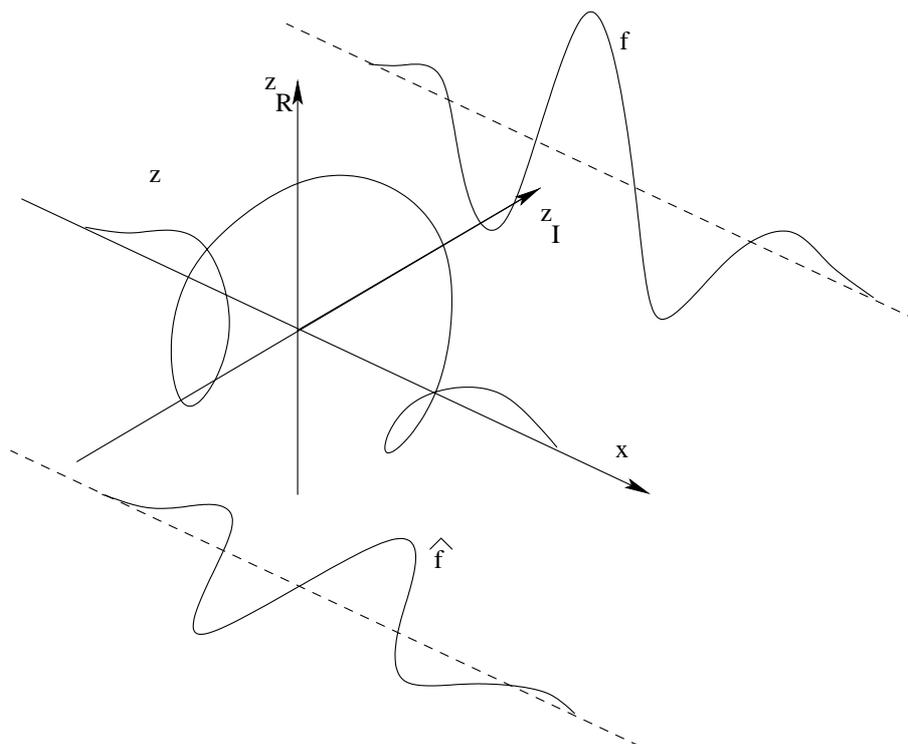
$$\begin{array}{lll} H_R(u) = 1 & \bullet \text{---} \circ & \delta(x) \equiv h_R(x) \\ H_I(u) = -j\text{sgn}(u) & \bullet \text{---} \circ & \frac{1}{\pi x} \equiv h_I(x) \end{array}$$

Erinnerung: imaginäre, ungerade Fkt. $\bullet \text{---} \circ$ reelle, ungerade Fkt.



Das analytische Signal $z(x) = f(x) + j\hat{f}(x)$ wird durch Faltung mit dem Operator $h(x) = h_R(x) + jh_I(x)$ aus dem Signal $f(x)$ erzeugt:

$$z(x) = (h * f)(x) = (f * (\delta + \frac{j}{\pi x}))(x).$$



3.8.1 Eindimensionale Hilberttransformation

Die **Hilberttransformation** \mathcal{H}_I erzeugt aus dem Realteil $f(x)$ den Imaginärteil $\hat{f}(x)$ eines analytischen Signals $z(x) = f(x) + j\hat{f}(x)$.

$$\begin{aligned} \hat{f}(x) = \mathcal{H}_I\{f(x)\} &:= (f * \frac{1}{\pi x})(x) \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(\tau)}{x - \tau} d\tau \end{aligned}$$

Dieser Ausdruck weist auf eine Singularität bei $x = \tau$ hin. Die Lösung erfolgt unter Beachtung des Cauchyschen Hauptwertes.

Die Realisierung der Hilberttransformation im Ortsraum

$$\hat{f}(x) = (f * \frac{1}{\pi x})(x)$$

ist ohne größere Fehler bzgl. der Phasenverschiebung kaum möglich. Die Hilberttransformation findet ihre Entsprechung im Frequenzraum

$$\hat{F}(u) = -j \operatorname{sgn}(u) F(u)$$

Dort bedeutet sie eine Phasendrehung der spektralen Komponenten um

$$\Delta\phi = \begin{cases} -\frac{\pi}{2} & u > 0 \\ 0 & u = 0 \\ \frac{\pi}{2} & u < 0 \end{cases} .$$

Voraussetzung ist allerdings, daß $|H_I(u)| = 1$ für alle Frequenzen ist (eine Amplitudenmodulation soll nicht stattfinden). Das heißt, H_R und H_I sind Allpaßfilter.

Einige Hilbertkorrespondenzen:

$$\begin{array}{lll} a & \bullet \longrightarrow & 0 \\ \delta(x) & \bullet \longrightarrow & \frac{1}{\pi x} \\ \sin x & \bullet \longrightarrow & -\cos x \\ \cos x & \bullet \longrightarrow & \sin x \end{array}$$

► Zweimalige Anwendung der Hilberttransformation erzeugt *Vorzeichenumkehr* für rein gerade/ungerade Funktionen.

$$\begin{array}{llll} f(x) = \sin x & \bullet \longrightarrow & \hat{f}(x) = -\cos x & \bullet \longrightarrow & \hat{\hat{f}}(x) = -\sin x \equiv -f(x) \\ f(x) = \cos x & \bullet \longrightarrow & \hat{f}(x) = \sin x & \bullet \longrightarrow & \hat{\hat{f}}(x) = -\cos x \equiv -f(x) \end{array}$$

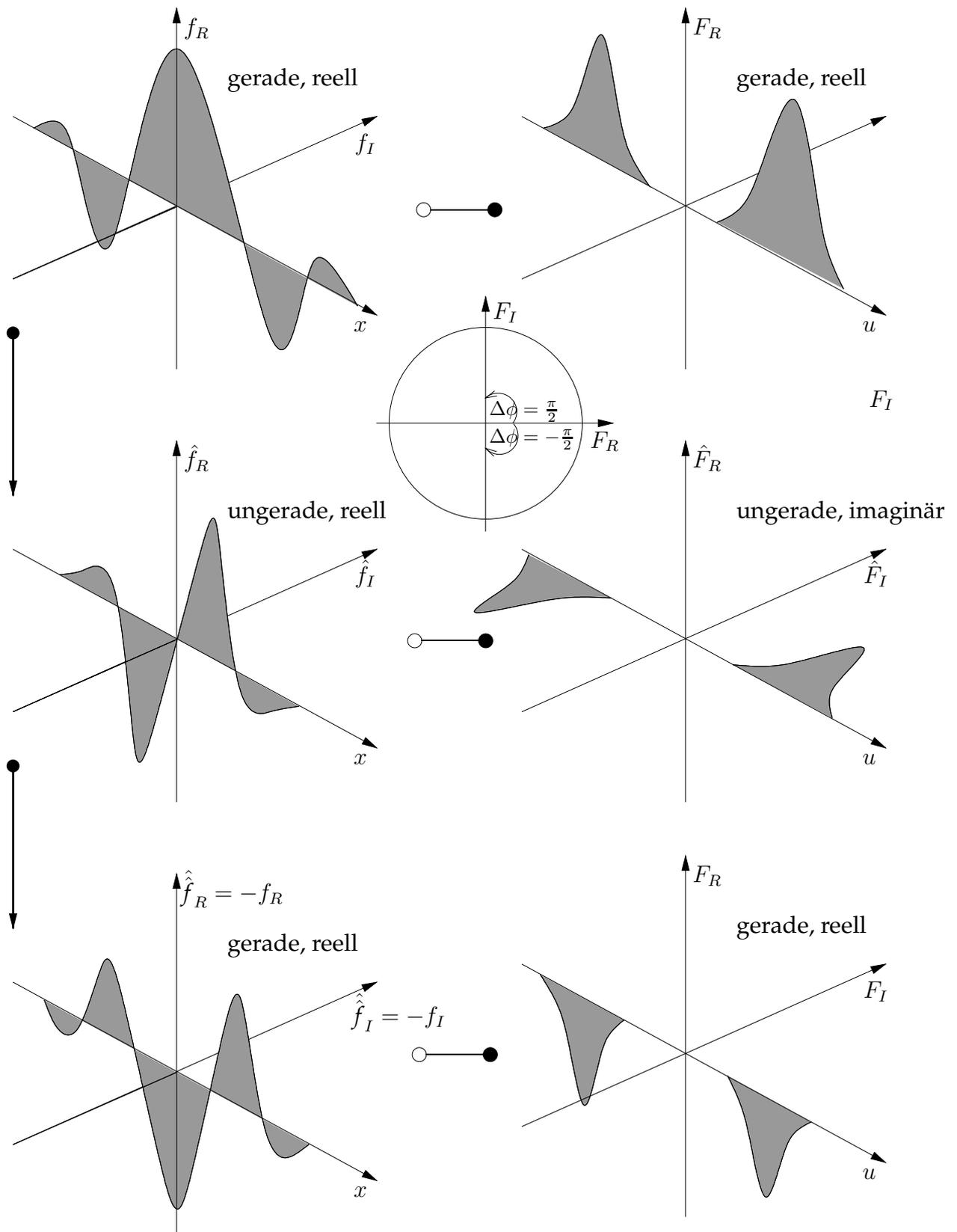
- ▶ Da $f(x)$ eine reelle Funktion ist, gilt zwischen Real- und Imaginärteil des analytischen Signals $z(x)$ stets die Hilberttransformation. Man sagt, Real- und Imaginärteil des analytischen Signals bilden ein *Hilbertpaar* oder *Quadraturpaar*

$$f(x) = z_R(x) \bullet \longrightarrow z_I(x) = \hat{f}(x)$$

Wegen ihrer Phasenverschiebung um $|\frac{\pi}{2}|$ sagt man, sie sind in *Quadratur-Phasenbeziehung*.

- ▶ Anwendung der Hilberttransformation im Frequenzraum: Zwischen Realteil $\text{Re}\{F(u)\}$ und Imaginärteil $\text{Im}\{F(u)\}$ des Spektrums $F(u)$ eines *kausalen Signals* $f(x)$, $f(x) = 0$ für $x < 0$ gilt die Hilberttransformation

$$\text{Re}\{F(u)\} \bullet \longrightarrow \text{Im}\{F(u)\}.$$



Schema der Hilbert-/Fouriertransformation eines geraden, reellen Signals

3.8.2 Anwendung des analytischen Signals

Analytische Signale werden zur lokalen Strukturanalyse auf der Grundlage *spektraler Repräsentationen* angewendet. Die lokale Blickweise der verwendeten Operatoren erfordert, daß diese an die *intrinsische Dimension* der interessierenden lokalen Signale angepaßt sind. In Bildern sind lokal folgende Dimensionen interessant:

intrinsisch 0-D: Punkt in konstanter Umgebung

intrinsisch 1-D: Kanten, Linien

intrinsisch 2-D: Ecken, Verzweigungen, gekrümmte Kanten/Linien

Definition (intrinsische Dimension):

Ein Signal $f : \mathbb{R}^N \rightarrow \mathbb{R}$ ist von instrinsischer Dimension m , wenn es konstant ist in Bezug auf $N - m$ Freiheitsgrade.

$$\text{i0D} := \{f \in L_2 : f(\mathbf{x}_i) = f(\mathbf{x}_j) \forall \mathbf{x}_i, \mathbf{x}_j \in N\}$$

$$\text{i1D} := \{f \in L_2 : f(x, y) = g(x \cos \theta + y \sin \theta) \forall (x, y) \in N\} \setminus \text{i0D}$$

$$\text{i2D} := L_2 \setminus (\text{i0D} \cup \text{i1D})$$

■

Zunächst beschränken wir uns auf 1D-Signale bzw. intrinsische 1D-Strukturen. Der Grund hierfür liegt darin, daß das *komplexwertig analytische Signal* keine intrinsischen 2D-Strukturen repräsentieren kann. Hierfür erweitern wir später das Konzept auf ein *quaternionwertiges analytisches Signal*.

Wenn wir Aussagen zur lokalen Struktur eines Signals treffen wollen, müssen wir bedenken, daß nach dem Prinzip der *Fourier-Reihenentwicklung* das Signal aus einer Vielzahl von Harmonischen mit unterschiedlichen Frequenzen synthetisiert wird.

Wenn wir also das analytische Signal nutzen wollen, um lokale Aussagen zu den spektralen Eigenschaften Amplitude und Phase zu treffen, entstehen *Probleme wegen der Frequenzbezogenheit* dieser Größen. Diese Probleme lassen sich lösen unter folgenden Annahmen:

- a) Die Struktur ist nur durch eine Frequenz synthetisierbar.

Dann müssen über den Operator, der aus dem reellen Signal das analytische Signal erzeugt, bzw. der die Hilberttransformation realisiert, keine einschränkenden Annahmen gemacht werden. Er sollte aber *Allpaßcharakter* ($|H(u)| = 1$ für alle u) besitzen, wenn die relevante strukturbildende Frequenz unbekant ist. In diesem Fall lassen sich

- *momentane Energie/Amplitude*
- *momentane Phase*
- *momentane Frequenz*

bestimmen.

- b) Die Struktur ist lokal aus einer Vielzahl von Frequenzen synthetisierbar. Um eindeutige frequenzbezogene Aussagen ableiten zu können, benötigt man Operatoren mit *Bandpaßcharakteristik*, mit denen der Frequenzraum abzutasten ist. Je enger die Paßbänder, umso besser, aber auch, umso aufwendiger. In diesem Fall spricht man von

- *lokaler Energie/Amplitude*
- *lokaler Phase*
- *lokaler Frequenz.*

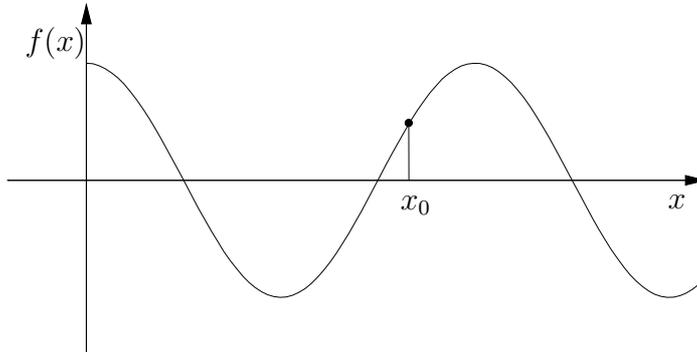
Diese lokalen spektralen Charakteristika entstehen also durch gewichtete Ausblendung aus einem Kontinuum lokaler spektraler Merkmale mittels Operatoren, die man *Quadratfilter* nennt. In den Paßbändern gelte $|H(u)| = 1$.

Gaborfilter und *Ableitungsoperatoren* sind als Bandpaßfilter hierfür geeignet.

Zunächst betrachten wir momentane spektrale Charakteristika. Wir setzen also voraus, daß ein zu analysierendes Signal am Ort x durch eine spezielle Frequenz $\omega = 2\pi u$ vollständig determiniert ist.

Motivation 1 :

Gegeben eine harmonische Funktion $f(x) = A \cos(\omega x)$, $A > 0$, können wir aus einer einzigen Messung $f(x_0)$ an der Stelle x_0 nicht beide Komponenten der Funktion, d.h. Amplitude A und Phase ϕ , getrennt erkennen. Es ist nicht erkennbar, ob die Funktion maximal, minimal, ansteigend oder abfallend ist und wir können die Amplitude A nicht ablesen.



Das analytische Signal für die Funktion $f(x)$ lautet

$$z(x) = A(\cos(\omega x) + j \sin(\omega x)) = Ae^{j\omega x}$$

Die *momentane Amplitude* $A(x)$ und die *momentane Energie* $E(x) = A^2(x)$ sind für das analytische Signal unmittelbar angebar.

$$\begin{aligned} A(x) &= |z(x)| \\ E(x) &= |z(x)|^2. \end{aligned}$$

Für unsere Funktion gilt $A(x) = A$ für alle x . Die momentane Amplitude ist also die *Einhüllende* der Funktion am Ort x .

Die *momentane Phase*

$$\phi(x) = \arg(z(x)) = \text{atan } 2(z_I(x), z_R(x))$$

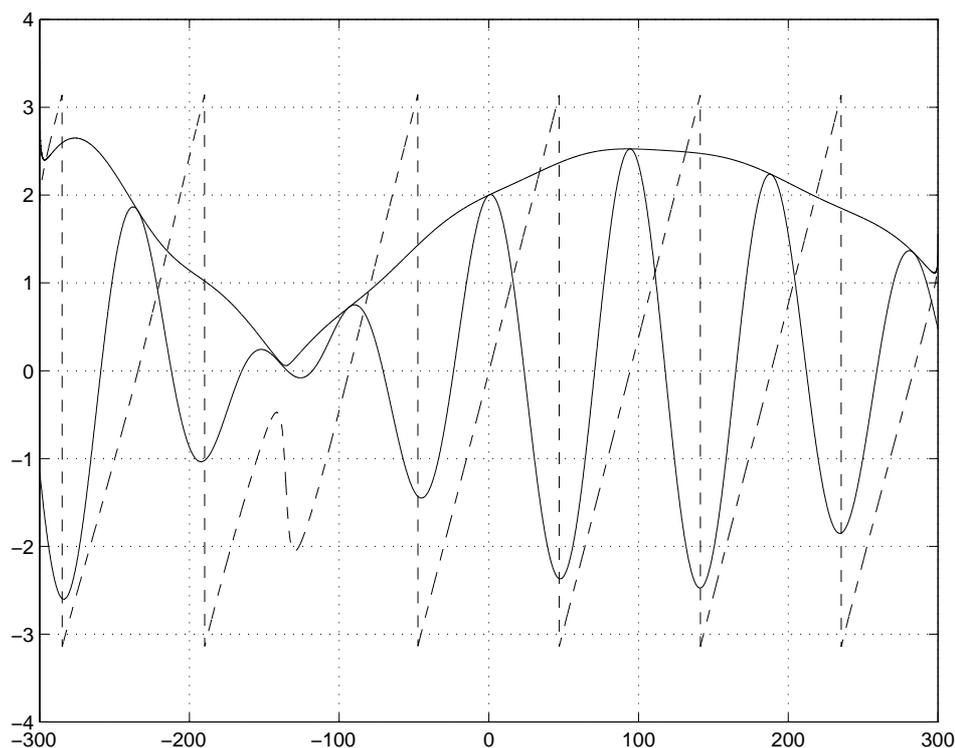
ist (nicht nur) in unserem Beispiel eine *lineare Funktion von x* (modulo 2π). Diese Eigenschaft bezeichnet man als Äquivarianz. Ihr Wert sagt etwas über den momentanen Verlauf der Funktion am Ort x aus, also über die momentane Struktur oder Qualität.

$$\phi(x_0) = \begin{cases} 2\pi k & x_0 : \text{Maximum } f(x_0) = A \\ 2\pi k + \pi & x_0 : \text{Minimum } f(x_0) = -A \\ 2\pi k + \frac{\pi}{2} & x_0 : \text{Nulldurchgang (abfallend)} \\ 2\pi k - \frac{\pi}{2} & x_0 : \text{Nulldurchgang (ansteigend)} \end{cases}$$

Hierbei ist k eine natürliche Zahl.

Beispiel: Die nächste Abbildung (Bülow, 1999) zeigt eine Funktion $f(x) = A(x) \cos(\omega x)$ mit ortsvariabler Amplitude. Die Signaleinhüllende repräsentiert die momentane Amplitude. Die momentane Phase widerspiegelt den oben beschriebenen Verlauf.

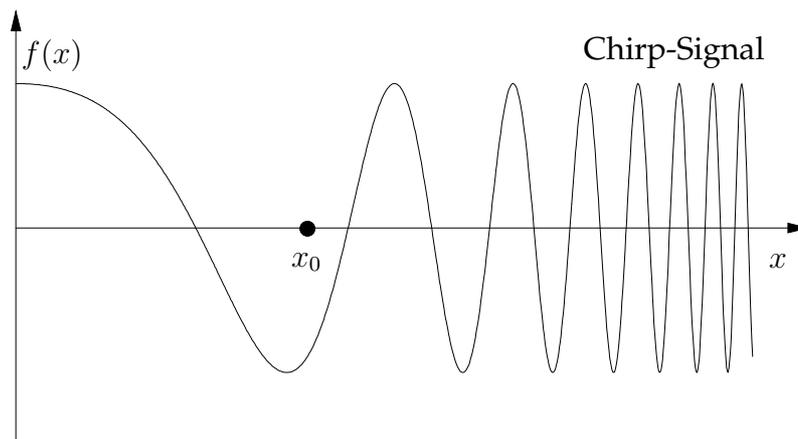
Am Ort $x \approx -150$ wird die momentane Phase falsch berechnet. Dies liegt daran, daß in Fällen, wo die Amplitude nahezu Null ist, die Phase nur sehr ungenau berechnet werden kann.



Motivation 2 :

Angenommen, die Funktion besitze eine ortsabhängige Frequenz $\omega(x)$, beispielsweise sei

$$f(x) = A \cos(\omega(x)x).$$



Die *momentane Frequenz* erhält man durch Ableitung der momentanen Phase.

$$\omega(x_0) = \left. \frac{d}{dx} \phi(x) \right|_{x=x_0}$$

Problem: Unstetigkeiten der Phase modulo 2π .

Nach Fleet (1999) ist keine algorithmische Sonderbehandlung dieser Stellen erforderlich, wenn folgende Rechnung erfolgt.

$$\omega(x_0) = \left. \frac{\left(\frac{d}{dx} z_I(x) \right) z_R(x) - z_I(x) \left(\frac{d}{dx} z_R(x) \right)}{|z(x)|} \right|_{x=x_0}$$

Bemerkungen zur Konstruktion des analytischen Signals

- Alle Faltungsterme *exakter ungerader Symmetrie* erzeugen eine Phasenverschiebung um $|\frac{\pi}{2}|$. Es ist aber schwer sicherzustellen, daß $|H(u)| = \text{const}$ für alle Frequenzen gilt. Also ist die Berechnung der momentanen spektralen Merkmale nicht einfach.
- Das Problem der konstanten Übertragungsfunktion entschärft sich, wenn die Hilbert-Transformierte von bandpaßgefilterten Signalen ermittelt werden soll.
- Werden die Operatoren zur Berechnung der Hilbert-Transformierten als Bandpaßoperatoren gewählt, so ist nur im Paßband $|H| = 1$ gefordert.

3.9 Hilbertpaar und Quadraturfilterpaar

Zwei reelle eindimensionale Funktionen f und g heißen *Hilbertpaar*, wenn die eine der Funktionen die Hilbert-Transformierte der anderen ist, d.h.

$$\hat{f} = \mathcal{H}_I\{f\} = g \quad \text{oder} \quad \hat{g} = \mathcal{H}_I\{g\} = f$$

Wenn also $g = \mathcal{H}_I\{f\}$, dann $\mathcal{H}_I\{g\} = -f$, bzw. wenn $f = \mathcal{H}_I\{g\}$, dann $\mathcal{H}_I\{f\} = -g$.

Wir haben außerdem am Beispiel der Cosinus- und Sinusfunktion gesehen, daß die Hilbert-Transformation die *Symmetrie von Funktionen vertauscht*. Sinus- und Cosinusfunktionen bilden ein Hilbertpaar. In Übereinstimmung mit der obigen Definition eines Hilbert-Paars bilden die Impulsantworten zweier Filter h_e und h_o ein Hilbert-Paar, wenn sie durch Hilbert-Transformation ineinander überführbar sind.

$$\mathcal{H}_I\{h_e\} = h_o \quad \text{und} \quad \mathcal{H}_I\{h_o\} = -h_e$$

Derartige Operatoren stehen in *Quadratur-Phasenbeziehung*, $|\Delta\phi| = \frac{\pi}{2}$. Es gilt

$$h_e^2(x) + h_o^2(x) = 1$$

und $|H_e(u)| = |H_o(u)|$. Außerdem sind sie als *orthogonal* zu betrachten,

$$h_e(x) \perp h_o(x).$$

Man nennt ein Paar von Operatoren, das ein Signal in ein Paar von Signalen zerlegt, dessen Komponenten in Quadraturphasenbeziehung stehen, auch Hilbertfilter. *Quadraturfilter* haben mit *Hilbertfiltern* die Eigenschaft gemeinsam, ein Paar zu bilden, dessen beide Komponenten in Quadratur-Phasenbeziehung stehen. Sie unterscheiden sich von Hilbertfiltern aber dadurch, daß ihre Frequenzübertragungsfunktion beliebig gewichtet sein kann. Seien $Q(u)$ die Frequenzübertragungsfunktion eines Quadraturfilters und $W(u)$ eine Gewichtsfunktion im Frequenzraum (z.B. mit Bandpaßcharakteristik), so gilt

$$Q(u) = \begin{cases} 2W(u) & u > 0 \\ 0 & \text{sonst} \end{cases}$$

Quadraturfilter transformieren ebenso ein reelles Signal in ein analytisches Signal wie der oben behandelte Operator der Hilberttransformation. Darüber hinaus bieten sie aber die Möglichkeit der *Anpassung an interessante Aufgaben*, z.B. eine *Multiskalenanalyse*, wenn sie eine skalierbare Familie von Filtern bilden. Normalerweise synthetisieren sich lokale Strukturen aus einem *Frequenzintervall*. Hier ist demzufolge der Ansatz der Quadraturfilter dem der Hilbertfilter zu bevorzugen, weil er die Verwendung von Bandpässen erlaubt.

Das Quadraturfilter $Q(u)$ ist als komplexwertig angenommen. Seine Definition weist aber darauf hin, daß es keine hermitesche Funktion ist. Also gibt es auch keinen reellen Operator mit der geforderten Eigenschaft im Ortsraum. Wohl existieren aber Lösungen für komplexwertige LSI-Operatoren.

Diese Operatoren besitzen zwangsläufig einen geraden Realteil und einen ungeraden Imaginärteil.

$$q(x) = q_e(x) + jq_o(x).$$

Das Quadraturpaar q_e und q_o steht also in Quadraturphasenbeziehung.

Beispiele für Quadraturfilter-Paare:

1. Das komplexe Filter $\varphi_a = \varphi_{2m} \mp jk\sigma^{\pm 1}\varphi_{2m\pm 1}$ mit dem Gewichtsparameter k liefert für $n_1 = 2m$ die Ableitung des Punktoperators φ_0 mit *gerader* Symmetrie und für $n_2 = 2m \pm 1$ die Ableitung des Punktoperators φ_0 mit *ungerader* Symmetrie. Diese beiden Filter bilden keine echten Quadraturfilterpaare. Daher treten Phasenstörungen (nicht linear) bei niedrigen Ordnungen auf.
2. *Gaborfilter*: Asymptoten für große Ordnungen n der Ableitungen

3.9.1 Detektion von Kanten und Linien in Bildern

Kanten und Linien sind *intrinsisch eindimensionale* Strukturen. Sie sind in der Bildebene eingebettet und haben eine *Orientierung*. Damit teilen sie die Bildebene in zwei *Halbräume*. Dies geschieht auch im Frequenzraum, aber in hierzu orthogonaler Richtung.

Die Möglichkeit, zwei Halbräume im Frequenzraum zu definieren, ist Voraussetzung für die Definition des Operators, der ein reelles Signal in ein analytisches Signal überführt. Da im 2D-Frequenzraum negative Frequenzen schwer definierbar sind, existieren im 2D-Fall unendlich viele Zerlegungen der Frequenzebene in Halbräume. Erst der natürliche Zwang durch die Orientierung einer Kante/Linie liefert eine eindeutige Lösung.

Ist $\mathbf{u} = (u, v)^T$ ein 2D-Frequenzvektor und ist $\mathbf{r} = (\cos \theta, \sin \theta)^T$ ein *Orientierungs-Einheitsvektor* einer Kante/Linie im Frequenzraum, dann weist \mathbf{r} im Frequenzraum in einen Halbraum, dem positive Frequenzen zugeordnet werden. Die entgegengesetzte Richtung wird den negativen Frequenzen zugeordnet. Also gilt

\mathbf{u} wird als positiv bezeichnet, wenn $\mathbf{r}^T \mathbf{u} > 0$

\mathbf{u} wird als negativ bezeichnet, wenn $\mathbf{r}^T \mathbf{u} < 0$.

Bezüglich der gewählten Orientierung ist die Hilberttransformation ausführbar und es existiert ein analytisches Signal.

Zu einem 2D-Signal f und seiner Fouriertransformierten F wird das *partielle oder orientierte analytische Signal* im Frequenzraum berechnet nach

$$Z(\mathbf{u}) = \begin{cases} 2F(\mathbf{u}) & , \text{ wenn } \mathbf{r}^T \mathbf{u} > 0 \\ F(\mathbf{u}) & , \text{ wenn } \mathbf{r}^T \mathbf{u} = 0 \\ 0 & , \text{ wenn } \mathbf{r}^T \mathbf{u} < 0 \end{cases}$$

$$= (1 + \text{sgn}(\mathbf{r}^T \mathbf{u}))F(\mathbf{u}).$$

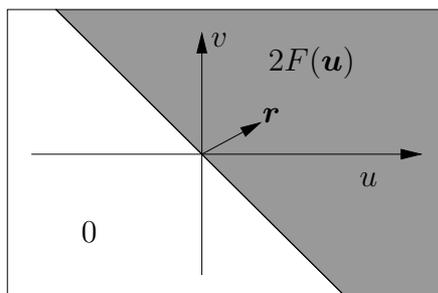
Im Ortsraum gilt hierfür

$$z(\mathbf{x}) = (f(\mathbf{x}) * (\delta(\mathbf{r}^T \mathbf{x}) + \frac{j}{\pi \mathbf{r}^T \mathbf{x}})) \delta(\mathbf{r}_\perp^T \mathbf{x}),$$

wobei \mathbf{r}_\perp ein zu \mathbf{r} orthogonaler Einheitsvektor ist.

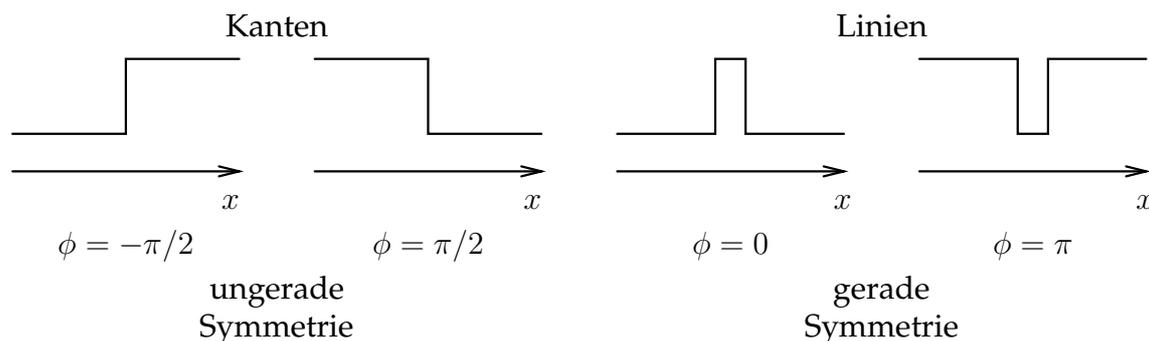
Für $\mathbf{r}^T = (1, 0)$ wird die Formulierung des analytischen Signals im 1D-Fall erreicht. Bezogen auf ein gedrehtes Koordinatensystem ist also die obige Formulierung ebenfalls *intrinsisch ein-dimensional*.

Die geeignete Orientierung wird mit *orientierten Filtern* gefunden.



Wieso sind Kanten und Linien interessante Strukturen im analytischen Signal ?

Kanten und Linien sind Signale von ungerader bzw. gerader Symmetrie.



Wegen der paarweisen Orthogonalität von Hilbert-Paaren bzw. Quadraturfilter-Paaren (h_e, h_o) prägt sich bei Faltung eines Signals mit diesen Filtern deren Symmetrie dem Output auf:

$$g_e(x) = (h_e * f)(x) \quad \text{und} \quad g_o(x) = (h_o * f)(x)$$

mit

$$\langle g_e, g_o \rangle = 0.$$

Die *Hilbert-Paare* sind als *Templates der entsprechenden Symmetrie* zu verstehen. Treffen diese Filter auf Strukturen entsprechender Symmetrie, wird deren Antwort maximal.

Um Kanten oder Linien zu detektieren, ist die Transformation des reellen in ein *analytisches Bildsignal nicht erforderlich*. Vielmehr erfolgt das Filtern des Signals f mit den symmetriespezifischen Operatoren.

3.9.2 Lokale Energie und lokale Phase

Hilbertpaare bzw. Quadraturfilterpaare sind Operatoren, die lokal eine *Zerlegung der Identität* in die Komponenten gerade und ungerade Symmetrie realisieren. Diese Symmetrie stellt ein *Basisssystem* dar (repräsentiert durch Filter-Paare), auf welches Signale projiziert werden.

Der von den Filter-Paaren *aufgespannte Vektorraum* ist geeignet, alle 1D-Strukturen zu repräsentieren. Voraussetzung ist, daß Rotation und Skala Freiheitsgrade dieser Filter sind. Während Hilbert-Paare eine Allpaßcharakteristik haben, sind die Bandpaßeigenschaften von Quadraturfiltern geeignet, der Skalenabhängigkeit von lokaler Energie und lokaler Phase zu entsprechen. Im Allgemeinen besitzt eine Struktur an beliebigem Ort Anteile unterschiedlicher Symmetrie. Folglich werden beide Komponenten eines Quadratur-Paares antworten.

Die Konzepte von momentaner Amplitude/Energie und Phase gehen über in die Konzepte von lokaler Amplitude/Energie und Phase.

Gegeben ein Paßband eines Quadraturfilter-Paares, sind *lokale Energie* und *lokale Phase* durch die Antworten dieser Filter berechenbar:

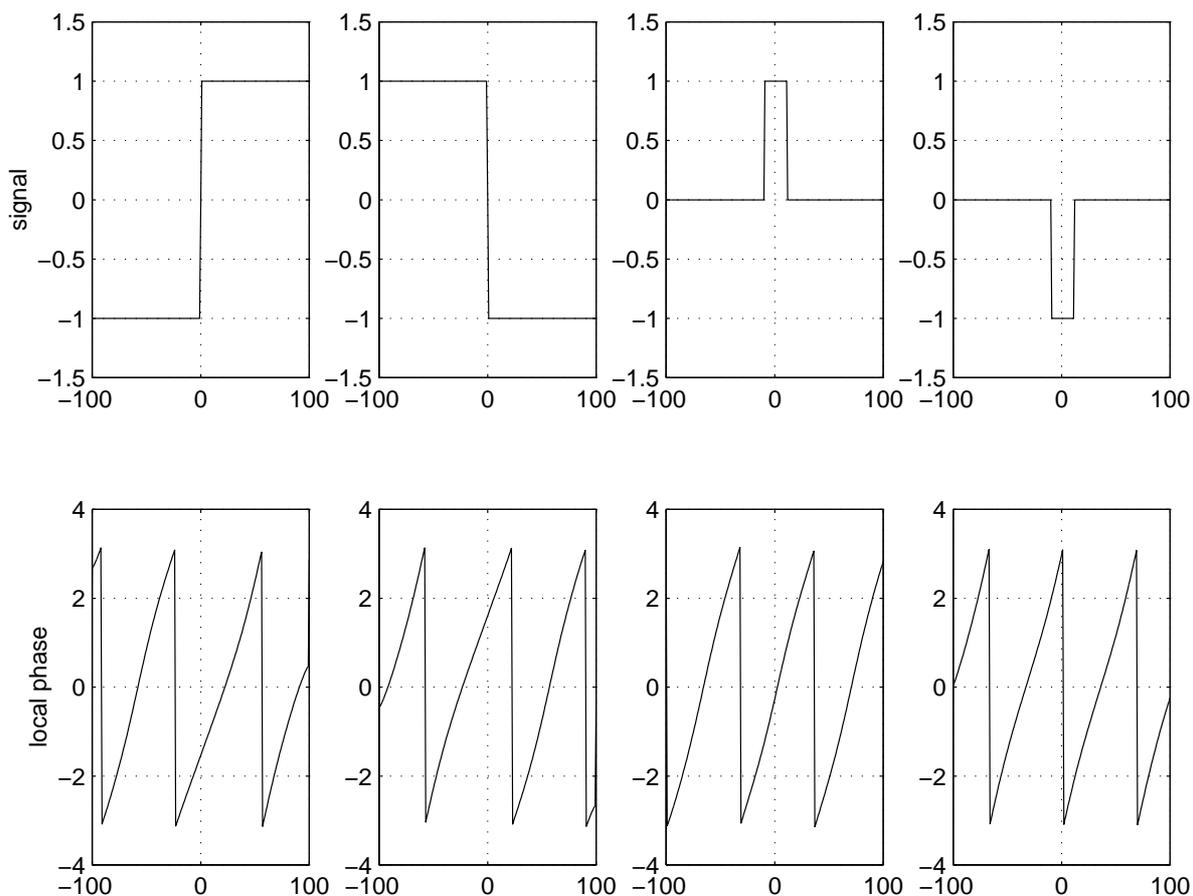
$$E(\mathbf{x}) = g_e^2(\mathbf{x}) + g_o^2(\mathbf{x})$$

$$\phi(\mathbf{x}) = \text{atan} 2(g_o(\mathbf{x}), g_e(\mathbf{x})).$$

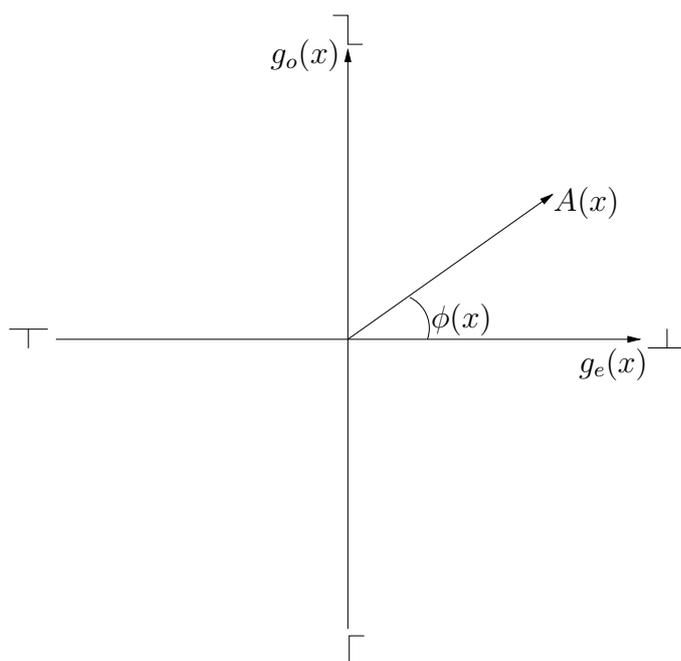
Da *Energie und Phase entkorreliert* werden, wird die Detektionsaufgabe einer 1D-Struktur in zwei Teilaufgaben zerlegt ("Was ist wo?").

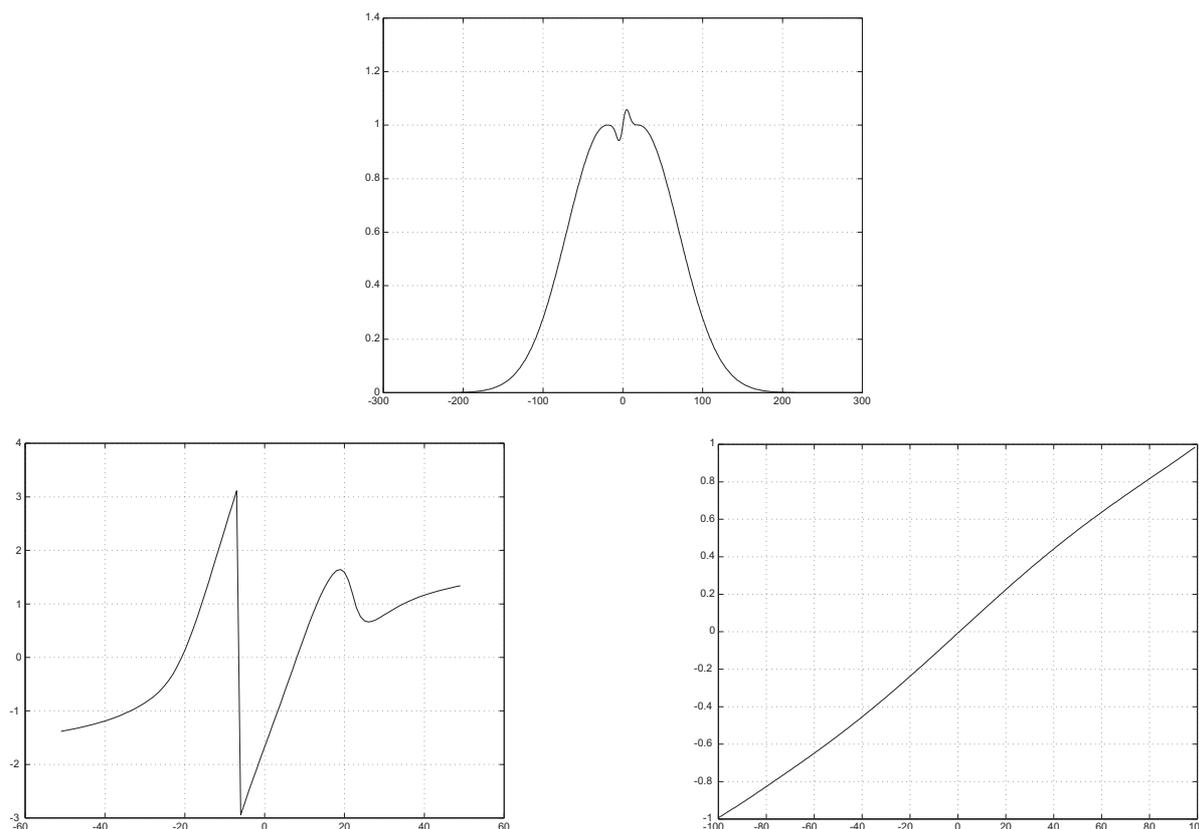
Die *lokale Energie* ist ein Maß für die *Ausprägung einer Struktur*. *Lokalisierung* bedeutet Maximierung der lokalen Energie.

Die *lokale Phase* (am Ort der Struktur) ist ein Maß für deren *Spezifik*, d.h. der *Symmetrie*. Die lokale Phase *klassifiziert* die Struktur. Umgekehrt ist h_e ein Detektor für Linien und h_o ein solcher für Kanten.



Obere Reihe: Die vier fundamentalen lokalen Strukturen (steigende und fallende Kanten, sowie positive und negative Linie) für die Phasenwinkel $-\frac{\pi}{2}$, $\frac{\pi}{2}$, 0 und π . Untere Reihe: Die Variationen der lokalen Phase bezüglich dieser Strukturen.





Oben: Ein Signal mit Strukturmerkmalen in zwei verschiedenen Skalen. Unten links: Die lokale Phase des Signals in der Umgebung des Ursprungs auf einer feinen Skala ($\sigma = 10$, $c = 2$). Unten rechts: Die lokale Phase des Signals in der Umgebung des Ursprungs auf einer groben Skala ($\sigma = 100$, $c = 2$). Am Ursprung ist die lokale Phase $-\pi/2$ (steigende Flanke) für die feine Skala und 0 (Peak) für die grobe Skala.

3.9.3 Quaternionwertiges analytisches Signal und Hilbert-Quadrupel

[Bülow(1999)]

Für *2D-Signale* (Bilder) ist die Einbettung des analytischen Signals in die komplexen Zahlen *semantisch nicht vollständig*, weil das komplexe analytische Signal nicht *intrinsisch zweidimensionale* Strukturen explizit machen kann. Dies erfordert eine *Symmetriezerlegung* nach dem in Abschnitt 2.6.0 auf Seite 160 eingeführten Schema

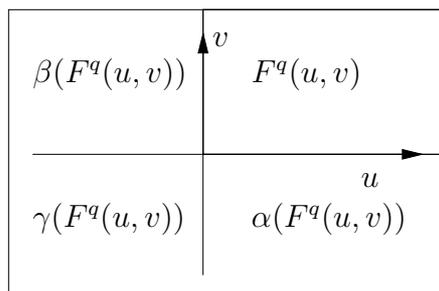
$$f(x, y) = f_{ee}(x, y) + f_{eo}(x, y) + f_{oe}(x, y) + f_{oo}(x, y).$$

In Abschnitt 2.6 wurde die *quaternionwertige Fouriertransformation* eingeführt, um eine globale explizite Repräsentation der Symmetrie *intrinsisch zweidimensionaler Strukturen* zu ermöglichen. Eine entsprechende Erweiterung der Repräsentation durch geeignete algebraische Einbettung ist für das analytische 2D-Signal erforderlich.

Wesentlich für die Konstruktion des *quaternionwertigen analytischen Signals* sind folgende Zusammenhänge:

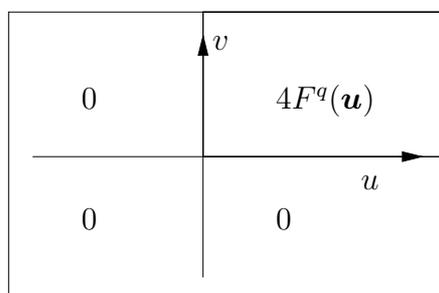
Im *1D-Fall* (komplexes Spektrum) führt die *Hermite-Symmetrie* des Spektrums eines reellen Signals zu einer Partitionierung des Frequenzraums in zwei Halbräume. Dadurch ist es möglich, ein komplexwertiges analytisches Signal zu konstruieren, das selbst nicht hermitesch ist, das aber vollständig ist bzgl. der Repräsentation von f_e und f_o .

Im *2D-Fall* (quaternionwertiges Spektrum) liegt ebenfalls *Hermite-Symmetrie* des Spektrums eines reellen Signal vor. In diesem Fall repräsentiert ein Quadrant des Frequenzraums unabhängig alle Symmetrien von f . Alle anderen Quadranten sind durch *Algebra-Involutionen* aus dem ersten Quadranten ableitbar.



Quaternionische Hermitesymmetrie des quaternionischen Spektrums eines reellen Signals

Also muß ein quaternionwertiges analytisches Signal einer Abbildung des gesamten Frequenzraums in den *ersten Quadranten* entsprechen. Dies ist die Folge der quaternionisch hermitischen Symmetrie des Spektrums F^q eines reellen Signals f . Dieses analytische Signal ist vollständig bzgl. der Symmetrien $f_{ee}, f_{eo}, f_{oe}, f_{oo}$. Es besitzt aber keine Hermite-Symmetrie.



Quaternionisches analytisches Signal im Frequenzraum

Die Hilbert-Transformation im 2D-Fall erzeugt ein *Hilbert-Quadrupel*, das in Quadratur-Phasenbeziehung steht. Sei f ein reelles 2D-Signal und F^q seine quaternionwertige Fouriertransformation. Das *quaternionwertige analytische Signal* hat folgende Definition

- a) im quaternionwertigen Frequenzraum ($\mathbf{u} = (u, v)$)

$$Z^q(\mathbf{u}) = (1 + \text{sgn}(u))(1 + \text{sgn}(v))F^q(\mathbf{u})$$

- b) im quaternionwertigen Ortsraum ($\mathbf{x} = (x, y)$)

$$z^q(\mathbf{x}) = f(\mathbf{x}) + \mathbf{n}^T \hat{\mathbf{f}}(\mathbf{x})$$

mit $\mathbf{n} = (i, j, k)^T$ und $\hat{\mathbf{f}}(\mathbf{x}) = (\hat{f}_x(\mathbf{x}), \hat{f}_y(\mathbf{x}), \hat{f}_{xy}(\mathbf{x}))^T$.

Das quaternionwertige analytische Signal besitzt also *drei imaginäre Komponenten*:

$$z^q(\mathbf{x}) = f(\mathbf{x}) + i\hat{f}_x(\mathbf{x}) + j\hat{f}_y(\mathbf{x}) + k\hat{f}_{xy}(\mathbf{x})$$

Die Komponenten $\hat{f}_x(\mathbf{x})$ und $\hat{f}_y(\mathbf{x})$ sind *partielle Hilbert-Transformation* von $f(\mathbf{x})$ in Richtung $\mathbf{r}^T = (1, 0)$, bzw. $\mathbf{r}^T = (0, 1)$.

Die Komponente $\hat{f}_{xy}(\mathbf{x})$ ist eine sogenannte *totale Hilbert-Transformierte* von $f(\mathbf{x})$.

Für ein *reelles 2D-Signal* sind die *partiellen Hilberttransformationen* in x - bzw. y -Richtung definiert durch

$$\begin{aligned}\hat{f}_x(\mathbf{x}) &= \mathcal{H}_x\{f(\mathbf{x})\} = (f * \frac{\delta(y)}{\pi x})(\mathbf{x}) \\ \hat{f}_y(\mathbf{x}) &= \mathcal{H}_y\{f(\mathbf{x})\} = (f * \frac{\delta(x)}{\pi y})(\mathbf{x}).\end{aligned}$$

Im Frequenzraum entspricht dies einer Aufteilung in Halbräume in u -, bzw. v -Richtung im Falle der Berechnung des analytischen Signals.

$$\begin{aligned}\hat{F}_u(\mathbf{u}) &= -i \operatorname{sgn}(u)F(\mathbf{u}) \\ \hat{F}_v(\mathbf{u}) &= -j \operatorname{sgn}(v)F(\mathbf{u}).\end{aligned}$$

Die imaginären Einheiten werden nur wegen der Verwendung dieser Transformation für das quaternionwertige analytische Signal verschieden gewählt.

Für ein *reelles 2D-Signal* ist die *totale Hilberttransformation* definiert durch (Hahn, 1996)

$$\hat{f}_{xy}(\mathbf{x}) = \mathcal{H}_{xy}\{f(\mathbf{x})\} = (f * \frac{1}{\pi xy})(\mathbf{x})$$

bzw.

$$\hat{F}_{uv}(\mathbf{u}) = -\operatorname{sgn}(u)\operatorname{sgn}(v)F(\mathbf{u}).$$

Ein hieraus abgeleitetes komplexwertiges analytisches Signal hat viele strukturelle Schwächen. Dies ist aber nicht das Ziel der Ausführung dieser Hilbert-Transformation. Vielmehr trägt sie dazu bei, ein quaternionwertiges analytisches Signal zu berechnen.

Im Falle von 1D-Signalen wurde festgestellt, daß zwei Funktionen, die durch Hilberttransformation auseinander hervorgehen, Hilbert-Paare bilden. Diese haben die Eigenschaften der Quadratur-Phasenbeziehung und der Invertierung der Symmetrie.

Im 2D-Fall entstehen bei quaternionwertiger Einbettung entsprechend Hilbert-Quadrupel. Sie bilden die Komponenten eines analytischen Signals. Sei ein quaternionwertiges analytisches Signal entsprechend

$$\begin{aligned}z^q(\mathbf{x}) &= z_R^q(\mathbf{x}) + iz_I^q(\mathbf{x}) + jz_J^q(\mathbf{x}) + kz_K^q(\mathbf{x}) \\ &= f_1(\mathbf{x}) + if_2(\mathbf{x}) + jf_3(\mathbf{x}) + kf_4(\mathbf{x})\end{aligned}$$

aus vier reellen zweidimensionalen Funktionen f_η , $\eta \in \{1, \dots, 4\}$, gebildet. Dann bilden diese vier Funktionen f_η ein *Hilbert-Quadrupel* mit *paarweiser Quadratur-Phasen-beziehung* zwischen den Funktionen.

Vier reelle zweidimensionale Funktionen f_η , $\eta \in \{1, \dots, 4\}$, bilden genau dann ein Hilbert-Quadrupel, wenn Komponenten der aus diesen Funktionen abgeleiteten analytischen Signale z_η^q in folgenden Beziehungen stehen

$$\begin{aligned} z_{\kappa,I}^q &= f_\lambda \\ z_{\kappa,J}^q &= f_\mu \\ z_{\kappa,K}^q &= f_\nu \end{aligned}$$

für Permutationen paarweise verschiedener $\kappa, \lambda, \mu, \nu \in \{1, \dots, 4\}$.

Wir haben im 1D-Fall gesehen, daß die Komponenten $h_e = \cos(\omega x)$ und $h_o = \sin(\omega x)$ des Kerns der (komplexen) Fouriertransformation ein Hilbert-Paar bilden.

Die entsprechenden vier Komponenten des Kerns der quaternionwertigen Fouriertransformation im 2D-Fall sind

$$\begin{aligned} h_{ee}(\mathbf{x}) &= \cos(2\pi ux) \cos(2\pi vy) \\ h_{eo}(\mathbf{x}) &= \cos(2\pi ux) \sin(2\pi vy) \\ h_{oe}(\mathbf{x}) &= \sin(2\pi ux) \cos(2\pi vy) \\ h_{oo}(\mathbf{x}) &= \sin(2\pi ux) \sin(2\pi vy) \end{aligned}$$

Das aus $h_{ee}(\mathbf{x})$ abgeleitete analytische Signal $z^q(\mathbf{x}) = e^{i2\pi ux} e^{j2\pi vy}$ stellt den QFT-Kern dar. Die vier Komponenten bilden ein Hilbert-Quadrupel.

Alles, was im 1D-Fall bezüglich der momentanen und lokalen spektralen Merkmale Energie/Amplitude und Phase, einschließlich deren Anwendung, ausgesagt wurde, kann auf den 2D-Fall bei quaternionwertige Einbettung erweitert werden.

Den imaginären Komponenten I, J und K werden die lokalen Phasenwinkel ϕ, θ und ψ zugeordnet mit

$$(\phi, \theta, \psi) \in [-\pi, \pi[\times [-\frac{\pi}{2}, \frac{\pi}{2}[\times [-\frac{\pi}{4}, \frac{\pi}{4}].$$

3.10 Gaborfunktion und lokale Fouriertransformation

[Gabor, 1946].

Entwicklung einer Signalfunktion am Aufpunkt x_0 nach einer Basis, die *optimale Lokalisierung sowohl im Ortsraum als auch im Frequenzraum* gestattet:

$$f(x) = \sum_i a_i(x) \varphi_{\sigma, u_i}(x)$$

Die **Gaborfunktion** $\varphi_{\sigma, u_i}(x)$ ist eine *komplexe Gauß-Basisfunktion*. Sie minimiert die *Unschärferelation* im Komplexen. Sie ist aber keine orthogonale Basis. Deshalb sind die durch Faltung gewonnenen Projektionskoeffizienten redundant für die Repräsentation des Signals.

3.10.1 Notation der eindimensionalen Gaborfunktion

Die Gaborfunktion stellt sich als das Produkt einer Gaußfunktion und einer komplex-harmonischen Funktion dar. Sie wird aber in verschiedener Weise notiert. Dies hat seine Ursachen in

- ▶ Normierung der Gaußfunktion
- ▶ Tradition zur Notation der Gaußfunktion
- ▶ betrachteter interessanter Aspekt der Gaborfunktion (Anwendungen).

Seien Gaußfunktion mit g und Gaborfunktion mit h bezeichnet.

- a) flächennormierte Gaußfunktion

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad \circ \text{---} \bullet \quad G(u) = e^{-2\pi^2\sigma^2 u^2}$$

- b) nicht normierte Gaußfunktion

$$g(x) = e^{-\frac{x^2}{2\sigma^2}} \quad \circ \text{---} \bullet \quad G(u) = \sqrt{2\pi}\sigma e^{-2\pi^2\sigma^2 u^2}$$

- c) "symmetrische" Gaußfunktion (in der Elektrotechnik üblich)

$$g(x) = e^{-\frac{\pi x^2}{\sigma^2}} \quad \circ \text{---} \bullet \quad G(u) = \sigma e^{-\pi u^2 \sigma^2}$$

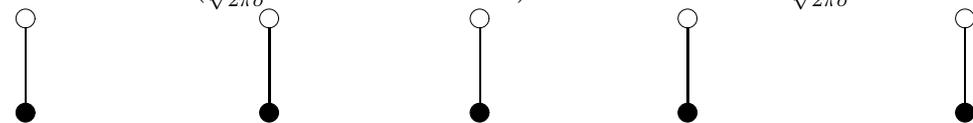
hat Fläche σ (kann auch normiert werden)

Wir verwenden Darstellung a). Also gilt folgende Fourier-Korrespondenz für die Gaborfunktion:

$$h(x; \sigma, u_0) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \cdot e^{j2\pi u_0 x}$$
$$H(u; \sigma, u_0) = e^{-2\pi^2\sigma^2 u^2} * \delta(u - u_0) = e^{-2\pi^2\sigma^2 (u - u_0)^2}$$

Hierbei ist σ der Skalierungsparameter der Gaußfunktion (fest gewählt). Die Ortskoordinate x ist lokal um den Ursprung des Koordinatensystems, also gleichzeitig globale Koordinate. Der Parameter u_0 heißt oft Mittenfrequenz des von $H(u; \sigma, u_0)$ gebildeten Paßbandes um $u = u_0$.

Aus Symmetriegründen erfolgt deshalb auch die Darstellung der Gaborfunktion im Ortsraum um einen Aufpunkt bei x_0 als weiteren Parameter. Dann sind $(x - x_0)$ und $(u - u_0)$ relative Koordinaten für die Lokalisation der Gaborfunktion im Orts- und Frequenzraum.

$$\begin{aligned}
 h(x; \sigma, x_0, u_0) &= h(x; \sigma, u_0) * \delta(x - x_0) \\
 &= \left(\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \cdot e^{j2\pi u_0 x} \right) * \delta(x - x_0) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-x_0)^2}{2\sigma^2}} \cdot e^{j2\pi u_0 (x-x_0)}
 \end{aligned}$$


$$H(u; \sigma, x_0, u_0) = (e^{-2\pi^2 \sigma^2 u^2} * \delta(u - u_0)) \cdot e^{-j2\pi u x_0} = e^{-2\pi^2 \sigma^2 (u - u_0)^2} \cdot e^{-j2\pi u x_0}$$

Die komplex-harmonische Modulation der Gaußfunktion führt zu deren Verschiebung im Frequenzraum (aus Tiefpaß wird Bandpaß). Die Verschiebung der Gaborfunktion im Ortsraum führt zu ihrer Phasenverschiebung im Frequenzraum um den Phasenwinkel $-\omega x_0$. Sie erfordert im Ortsraum eine Phasensynchronisation der Harmonischen bezüglich des neuen Aufpunktes x_0 .

In der Folge werden wir vier unterschiedliche Interpretationen aus der Gaborfunktion ableiten, die sämtlich auch auf den 2D-Fall erweiterbar sind:

1. Eine Gaborfunktion ist ein Gaußscher Bandpaßfilter
2. Eine Gaborfunktion repräsentiert ein (skalierbares) Quadraturpaar
3. Eine bezüglich u_0 erzeugte Familie von Gaborfunktionen (σ fest) stellt eine lokale Fouriertransformation dar (am Ort x_0)
4. Eine bezüglich σ erzeugte Familie von Gaborfunktionen (u_0 fest) stellt eine Gabor-Wavelettransformation dar (am Ort x_0)

3.10.2 Gaborfunktion als Bandpaßoperator

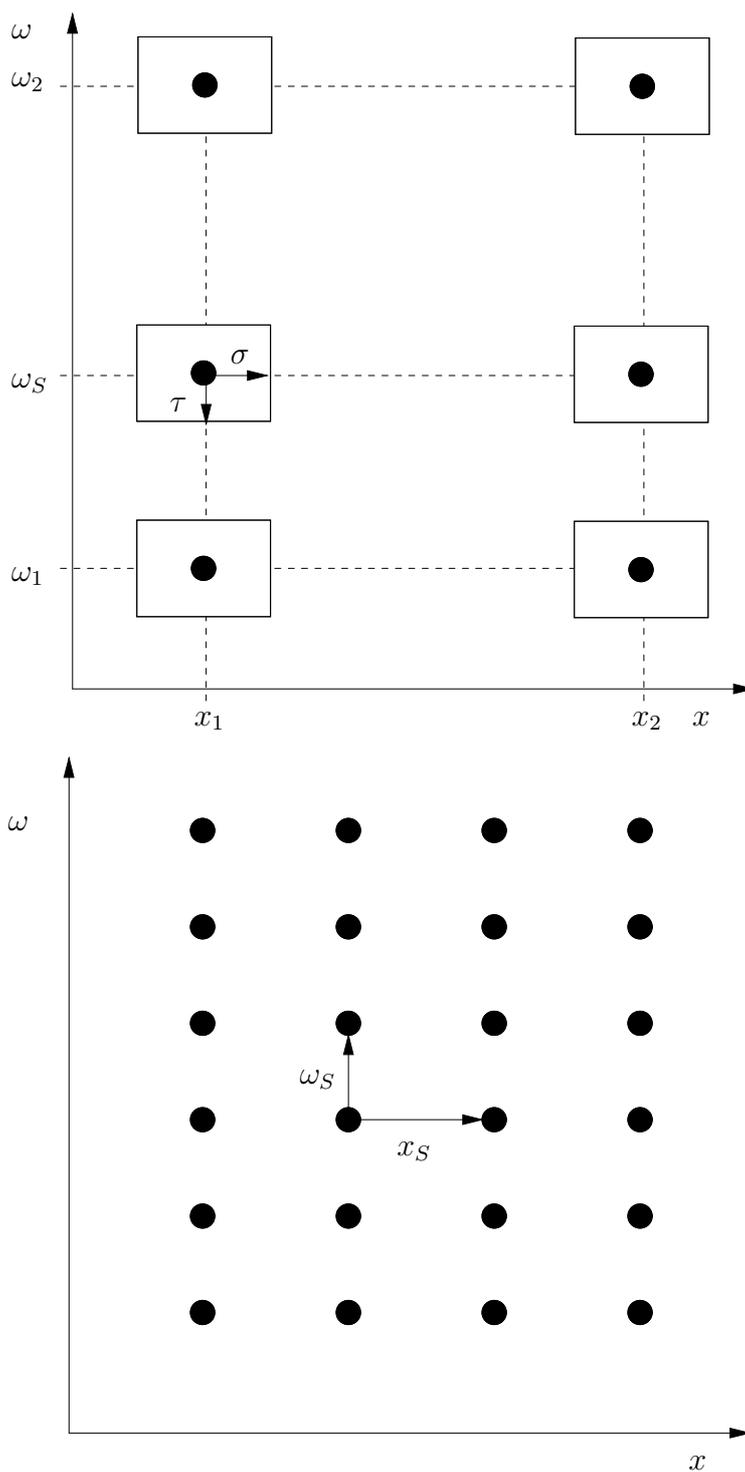
Betrachtungaspekt: Die Gaborfunktion am Ort x_0 ist eine komplex-harmonisch modulierte Gaußfunktion mit der Modulationsfrequenz u_0 .

- ▶ Sie ist ein *Bandpaßoperator*, da die harmonische Modulation der Gaußfunktion im Ortsraum nach dem Modulationssatz der Fouriertransformation zu einer Verschiebung der Gaußfunktion im Frequenzraum führt. Es entsteht ein Bandpaß mit der Mittenfrequenz u_0 .
- ▶ Die freie Wahl der Mittenfrequenz u_0 gestattet, den Frequenzraum mit Bandpässen abzutasten. Es entsteht ein Stapel von Gaborfunktionen, der bei festem σ eine konstante *Abtastbreite* in Orts- und Frequenzraum besitzt.

In der folgenden Abbildung wird der Orts-Frequenz-Phasenraum (x, ω) betrachtet, in dem im Abtastabstand (x_s, ω_s) gleich große Abtastfunktionen (Bandpässe) der Ausdehnung $2|\sigma| \times \tau$, $\tau = \sigma^{-1}$, lokalisiert sind. Also gilt für Abtastpositionen (x_i, ω_i) :

$$[x_i - \sigma, x_i + \sigma] \times [\omega_i - \tau, \omega_i + \tau].$$

Da Gaborfunktionen kein orthogonales Basissystem bilden, muß das Abtastschema redundant sein (Basisfunktionen überlappen sich).



3.10.3 Lokale Fouriertransformation und Gabortransformation

Betrachtungaspekt: Die Gaborfunktion am Ort x_0 ist eine Gauß-gewichtete komplex-harmonische Funktion mit der Frequenz u_0 .

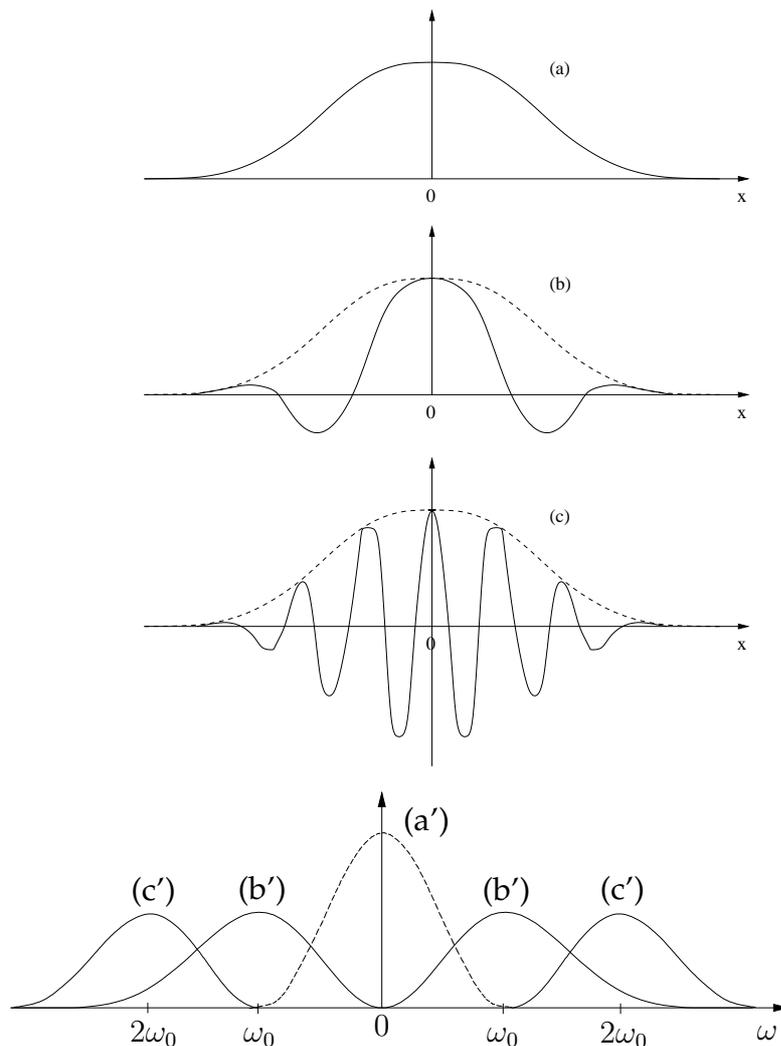
Wird die Frequenz u_0 nicht als Parameter einer Gaborfunktion betrachtet, sondern als Realisierung einer Variablen u einer Familie von Gaborfunktionen, so bilden diese das Basissystem einer speziellen lokalen Fouriertransformation, der Gabortransformation.

Die lokale Fouriertransformation (auch Kurzzeit-Fouriertransformation genannt)

$$F(u; x_0) = \mathcal{F}_l\{f(x)\} \Big|_{x=x_0} = \int_{-\infty}^{\infty} f(x)w^*(x - x_0) e^{-2\pi ux} dx$$

wird am Ort x_0 berechnet. Die Funktion $w(x - x_0)$ ist entweder eine Aperturfunktion oder eine Gewichtsfunktion. Demzufolge haben lokale Spektren unterschiedliche Gestalt in Abhängigkeit von w . Die lokale Fouriertransformation kann verschieden gelesen werden:

- ▶ Transformation über ausgeblendeten/gewichteten Signalen mit unendlich ausgedehnten Harmonischen als Basisfunktionen
- ▶ Transformation eines unbegrenzten Signals mit ausgeblendeten/gewichteten Harmonischen.



Die Kurzzeit-Fouriertransformation ist die klassische Methode der Zeit-Frequenz-Analyse eines (1D-) Signals. Dabei werden die beiden dualen Räume verbunden aufgespannt (siehe Abschnitt über quadratische Signalformen).

Das *Spektrogramm* ist das Betragsquadrat der lokalen oder Kurzzeit-Fouriertransformation im (x, u) -Phasenraum (lokales Leistungsspektrum).

$$S_f(u; x_0) = |F(u; x_0)|^2$$

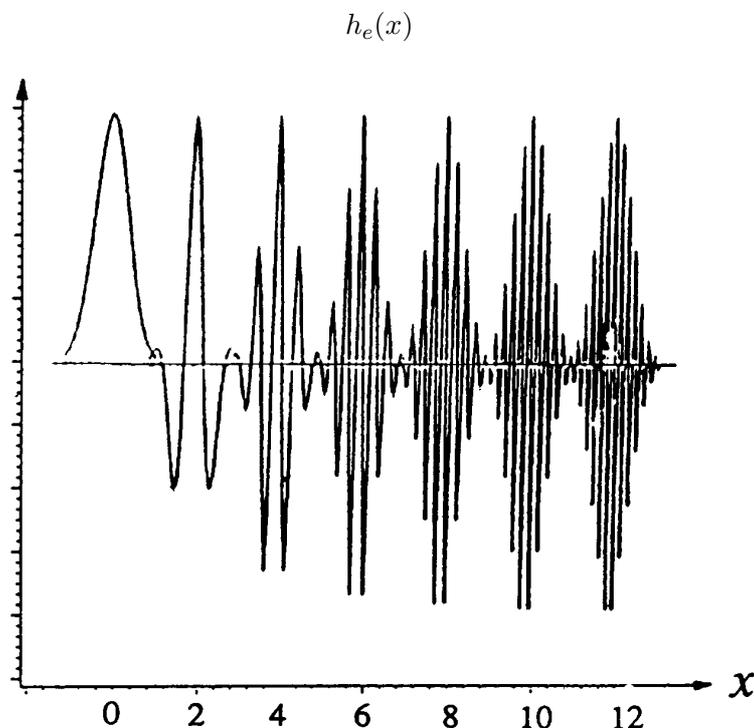
Die *Gabortransformation* ist demzufolge durch Verwendung einer Gaußschen Wichtungsfunktion $g(x; x_0)$ für den Ort x_0 zu berechnen.

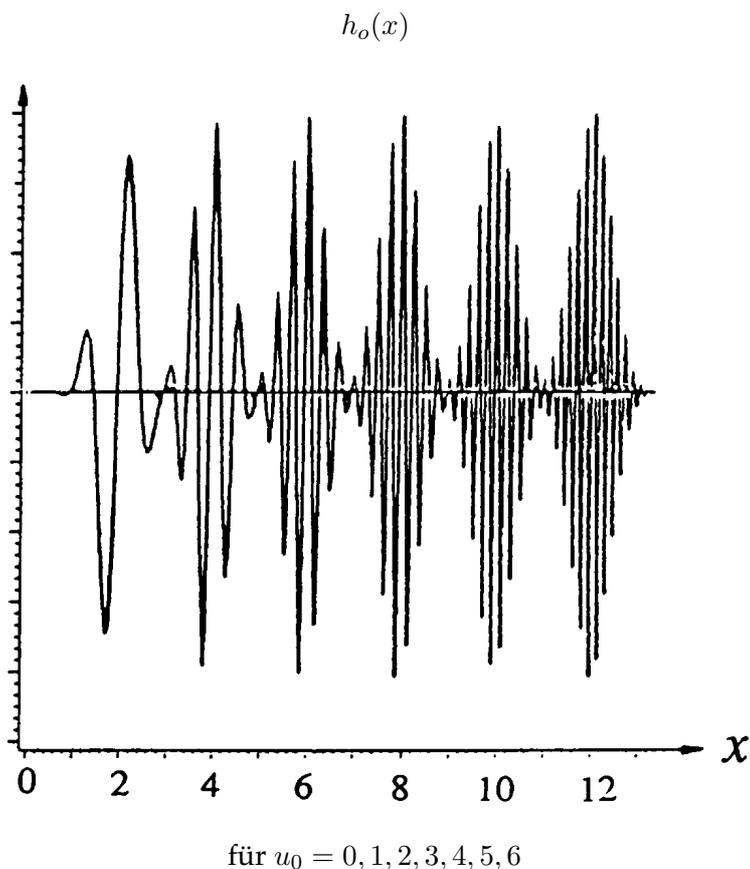
$$F_G(u; x_0) = \mathcal{G}\{f(x)\} \Big|_{x=x_0} = \int_{-\infty}^{\infty} f(x)g(x; x_0) e^{-j2\pi ux} dx$$

Die F_G sind als Projektion des Signals $f(x)$ auf die Gaborfunktionen mit den Frequenzen u am Ort x_0 zu interpretieren. Die inverse Gabortransformation

$$f_G(x_0) = \mathcal{G}^{-1}\{F_G(u; x_0)\} = \int_{-\infty}^{\infty} F_G(u; x_0)G(u; u_0) e^{j2\pi u_0 x} du_0$$

rekonstruiert den Signalwert $f_G(x_0) \approx f(x_0)$ an der Stelle x_0 . Wegen der Nichtorthogonalität der Basis ist die Rekonstruktion nicht exakt.





3.11 Gabor-Wavelettransformation

3.11.1 Wavelet-Transformation

Betrachtungaspekt: Die Gaborfunktion einer Mittenfrequenz u_0 ist eine selbstähnliche skalierbare und verschiebbliche Funktion. Die durch Skalierung σ und Verschiebung x_0 erzeugte Familie von Gaborfunktionen bildet das Basissystem einer Gabor-Wavelettransformation.

Die Waveletfunktionen wurden von Meyer und Lemarie (1986) als ein Basissystem des Hilbertraums eingeführt, von Daubechies (1988, 1990) mit der Multiskalenanalyse der Bildverarbeitung in Verbindung gebracht und von Mallat (1989) für ihre Anwendung in der Bildanalyse ausgearbeitet. Es existieren schnelle Algorithmen für die Ausführung der Wavelettransformation.

Eine Wavelettransformation hat Beziehung zur Kurzzeit-Fouriertransformation. Während letztere aber die Abtastung des Frequenzraums mit Bandpaßfiltern zur Grundlage hat, beruht die Wavelettransformation auf einer Abtastung des Freiheitsgrades Skala. Sie ist also im Ortsraum bzw. im Skalenraum erklärt.

Während das *Spektrogramm* die Signaldarstellung durch Kurzzeit-Fouriertransformation im (x, u) -Phasenraum repräsentiert, führt die Wavelettransformation zu einem *Skalogramm* im (x, σ) -Phasenraum (Skalenraum).

Sind s ein Skalierungsparameter und t ein Translationsparameter, dann ist eine kontinuierliche Waveletfunktion dargestellt durch

$$\psi(x; s, t) = |s|^{-\frac{1}{2}} \psi\left(\frac{x-t}{s}\right).$$

Die kontinuierliche Wavelettransformation eines Signals f lautet

$$W_f(t; s) = |s|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(x) \psi^*\left(\frac{x-t}{s}\right) dx.$$

Sie stellt eine Multiskalenrepräsentation (im kontinuierlichen Skalenraum) des Signals an jedem Ort t dar. Der Ausdruck für die W_f hat die Gestalt eines *Faltungsintegrals*.

Wavelets werden aus *Skalenfunktionen* (als Punktoperator) berechnet und haben bestimmte *Verträglichkeitsbedingungen* zu erfüllen. Sind Wavelets orthogonal, eignen sie sich besonders für Signalcodierung (MPEG4), für Signalanalyse ist Orthogonalität hingegen von Nachteil.

Orthogonale Wavelets bilden eine Basis. Für orthogonale Wavelettransformationen sind schnelle Algorithmen angebar. Aus der Sicht der Signalverarbeitung haben orthogonale Wavelets den Vorzug, daß die Analyse von Signalen (Hintransformation) und die Synthese von Signalen (Rücktransformation) mit denselben Filtern durchgeführt werden kann.

Biorthogonale Wavelets realisieren Analyse und Synthese mittels verschiedener Filter $H(\omega)$ und $\tilde{H}(\omega)$, die ein sogenanntes *Biorthogonalitätskriterium* (im Frequenzraum) erfüllen:

$$\overline{H(\omega)} \tilde{H}(\omega) + \overline{H(\omega + \pi)} \tilde{H}(\omega + \pi) = 1.$$

$\tilde{H}(\omega)$ heißt das zu $H(\omega)$ *duale Filter*. Im Ortsraum entsprechen diesen Filtern Wavelets ψ und *duale Wavelets* $\tilde{\psi}$.

Nichtorthogonale Wavelets bilden keine Basis, aber ein *Frame*. Wenn sie kein Frame bilden, sind diese Funktionen keine Wavelets.

Diskrete Waveletfunktionen $\psi_{k,l}$, k Skalenparameter und l Ortsparameter, erzeugen Pyramidenstrukturen in einem diskreten Skalenraum. Wavelets, egal ob kontinuierlich oder diskret, enthalten in ihrer Definition ein *Resampling*. Man erhält diskrete Wavelets aus kontinuierlichen durch folgende Parametrisierung:

Ist s_0 eine Basisskala, $s_0 > 1$, so ist $s_k = s_0^k$, $k \in \mathbb{Z}$, mit $s_0^0 = 1$ eine hieraus abgeleitete Skala. Ist x_0 die Abtastschrittweite bei der Skala s_0 , so ist $x_s = x_0 s_k = x_0 s_0^k$ die Abtastschrittweite für die Skala s_k . Sei $x_0 > 0$, dann bezeichnet $x_l = l x_s = l x_0 s_0^k$, $l \in \mathbb{Z}$, eine Abtastposition eines Wavelets mit dem Skalenparameter k .

Aus einem *Mutterwavelet* auf der Basisskala am Koordinatenursprung, $\psi(x)$, erhält man ein skaliertes diskretes Wavelet am Ort x_s durch

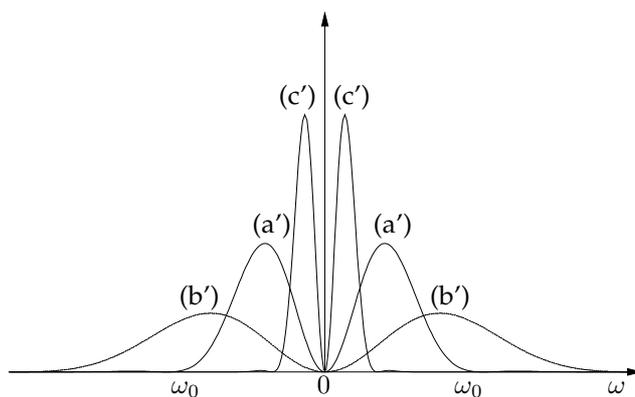
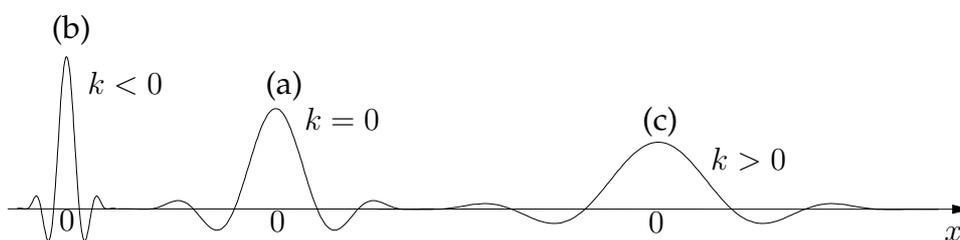
$$\psi_{k,l}(x) = s_0^{-\frac{k}{2}} \psi(s_0^{-k}(x - l x_0)).$$

Mit der Streckung/Stauchung eines Wavelets ändern sich gleichzeitig die Abtastpunkte und die Amplitude $a_s = s_0^{-\frac{k}{2}} = \frac{1}{\sqrt{s_k}}$.

Es gilt also

Fall	k	Skala	Amplitude	Abtastung
b	< 0	$s_k < 1$	$a_s > 1$	$x_s < x_0$
a	0	$s_k = 1$	$a_s = 1$	$x_s = x_0$
c	> 0	$s_k > 1$	$a_s < 1$	$x_s > x_0$

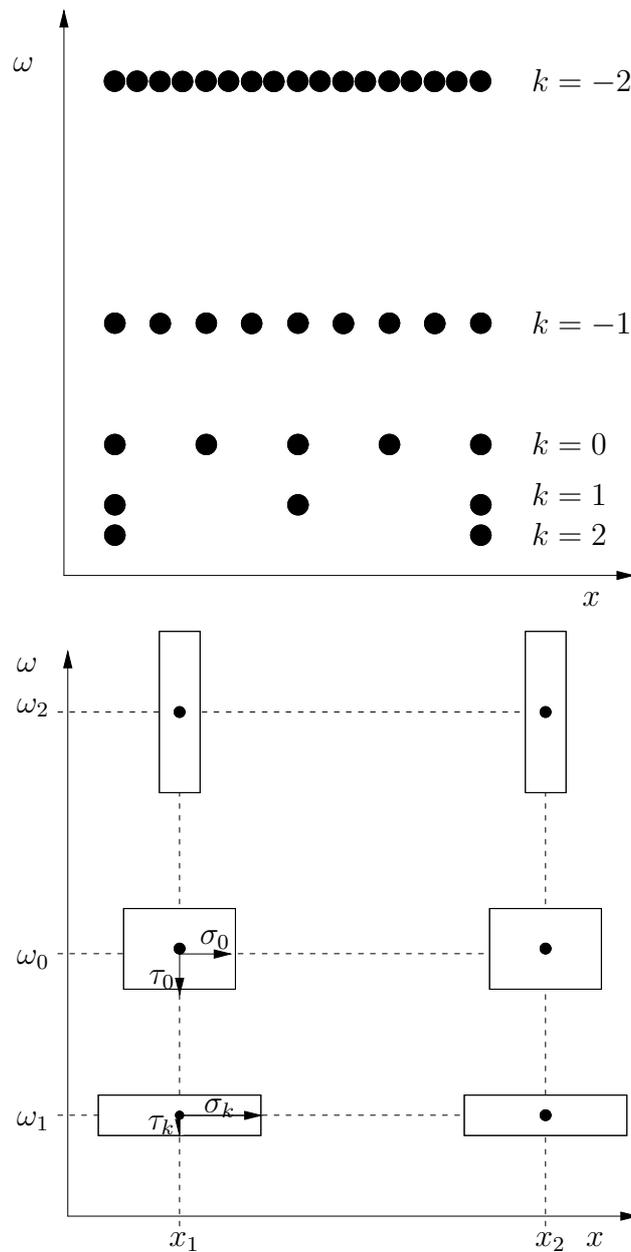
Der Skala s_k kann auch eine Frequenz $\omega_k = x_s^{-1} = \omega_0 s_k^{-1}$ zugewiesen werden, die die Abtastposition im Frequenzraum angibt. Die resultierenden Paßbänder haben eine Breite, die invers proportional zur Breite des Wavelets ist ($\Delta\omega_k \sim \frac{1}{\sigma_k}$).



Besonders interessant ist der Fall $s_0 = 2$, $x_0 = 1$. Dann entsteht eine *Oktavpyramide*. Hat das Mutterwavelet die halbe Breite σ_0 , so gilt nach Skalierung $\sigma_k = s_k \sigma_0$ und für die Breite eines Paßbandes $\tau_k = \sigma_k^{-1} = s_k^{-1} \tau_0$. Der (x, ω) -Phasenraum wird also mittels Funktionen abgetastet, die folgende Bereiche abdecken:

$$[x_l - \sigma_k, x_l + \sigma_k] \times [\omega_l - \tau_k, \omega_l + \tau_k].$$

Das hiermit verbundene Abtastschema des Phasenraums nennt man *dyadische Abtastung*.



Die Koeffizienten der diskreten Wavelettransformation sind

$$c_{k,l}(f) = \int_{-\infty}^{\infty} f(x) \psi_{k,l}(x) dx$$

Das Rekonstruktionsproblem besteht darin, s_0 , x_0 und $\psi(x)$ so zu wählen, daß gilt

$$f(x) \approx \hat{f}(x) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} c_{k,l}(f) \psi_{k,l}(x)$$

Diese Vorschrift gilt aber nur für orthogonale Wavelets. In der Signalanalyse werden nicht-orthogonale Wavelets genutzt. Hier bietet sich das Konzept der *Biorthonormalität* an, für die sogenannte *duale Wavelets* benötigt werden.

Für diesen Fall gilt :

Ob aus solcher Abtastung die *Rekonstruktion eines Signals* aus seinen Waveletkoeffizienten möglich

ist, hängt von einigen Bedingungen ab. Wird die Abtastschrittweite x_0 sehr klein gewählt, entsteht *Überabtastung*. Die Wavelets $\psi_{k,l}(x)$ sind linear abhängig und die *Waveletkoeffizienten* $c_{k,l}(f)$ enthalten viel Redundanz. Die Rekonstruktion ist prinzipiell möglich, weil es beliebig viele duale Waveletsysteme gibt, die für die Rekonstruktion genutzt werden können.

Ein *duales diskretes Wavelet*

$$\tilde{\psi}_{k,l}(x) = s_0^{-\frac{k}{2}} \tilde{\psi}(s_0^{-k}(x - lx_0)) \quad k, l \in \mathbb{Z}$$

gestattet die Signalrekonstruktion

$$\hat{f}(x) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} c_{k,l}(f) \tilde{\psi}_{k,l}(x)$$

mit

$$c_{k,l}(f) = \langle f, \psi_{k,l} \rangle.$$

Da aber alle Wavelets $\psi_{k,l}(x)$ und alle dualen Wavelets $\tilde{\psi}_{k,l}(x)$ denselben Signalraum aufspannen, gilt auch

$$\hat{f}(x) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \tilde{c}_{k,l}(f) \psi_{k,l}(x)$$

mit

$$\tilde{c}_{k,l}(f) = \langle f, \tilde{\psi}_{k,l} \rangle.$$

Der Zusammenhang zwischen Signalrekonstruktion und der Existenz eines Systems dualer Wavelets ist direkt mit dem Abtastschema verbunden. Das Kriterium für die Existenz dualer Wavelets ist die Existenz zweier Schranken A und B (*frame bounds*), $0 < A \leq B < \infty$, die folgende *Zulässigkeitsbedingung* (*admissibility condition*) einzuhalten gestatten:

$$A \|f\|^2 \leq \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} |\langle f, \psi_{k,l} \rangle|^2 \leq B \|f\|^2.$$

Gilt $A = B$ (*tight frame*), so ist $\tilde{\psi}_{k,l}(x) = \psi_{k,l}(x)$ und eine *exakte Rekonstruktion* ist möglich (auch bei Redundanz).

Wird der Abtastabstand x_0 so gewählt, daß sich die Wavelets im Phasenraum eben berühren, enthalten die Waveletkoeffizienten keine Redundanz. Man spricht von *kritischer Abtastung*. Gilt in diesem Fall auch $0 < A \leq B < \infty$, dann bilden die linear unabhängigen Wavelets $\psi_{k,l}(x)$, $k, l \in \mathbb{Z}$, eine *Riesz-Basis* des Funktionenraums $L_2(\mathbb{R})$. Gilt $f \in L_2(\mathbb{R})$, dann bilden die Wavelets also auch eine Basis für f .

Die Basis der Wavelets und die Basis der dualen Wavelets unterliegen in diesem Fall einer *Biorthogonalitätsbedingung*

$$\langle \psi_{k,l}, \tilde{\psi}_{m,n} \rangle = \delta_{km} \delta_{ln} \quad k, l, m, n \in \mathbb{Z}.$$

Derartige Wavelets heißen *biorthogonale Wavelets*.

Orthonormale Wavelets sind *selbstdual*,

$$\langle \psi_{k,l}, \psi_{m,n} \rangle = \delta_{km} \delta_{ln}.$$

Deshalb können zur Analyse (Berechnung der Projektionskoeffizienten) und zur Synthese (Rekonstruktion) eines Signals dieselben Wavelets verwendet werden. Sie besitzen gleiche Schranken (*tight frame*) und die Zulässigkeitsbedingung geht in das *Parsevaltheorem* der Fouriertheorie über.

3.11.2 Grundzüge der Multiskalenanalyse

Hier soll das von Mallat (1989) für orthonormale Wavelets vorgestellte Schema einer Mehrskalenanalyse kurz skizziert werden. Da dies auch für biorthogonale Wavelets gilt, sei dieser Fall angenommen. Auf die Darstellung der Implementierung sei verzichtet.

Unter Waveletanalyse sei die Projektion eines Signals $f \in L_2(\mathbb{R})$ auf ein biorthogonales Basissystem von Wavelets $\{\psi_{k,l}(x)\}_{k,l \in \mathbb{Z}}$ verstanden. Es sei angenommen $s_0 = 2$ und $x_0 = 1$. Also folgt für $k, l \in \mathbb{Z}$

$$\begin{aligned}\psi_{k,l}(x) &= 2^{-\frac{k}{2}} \psi(2^{-k}(x-l)) \\ \tilde{\psi}_{k,l}(x) &= 2^{-\frac{k}{2}} \tilde{\psi}(2^{-k}(x-l)).\end{aligned}$$

Sei $\text{span} \{\psi_{k,l}(x)\}_{k,l \in \mathbb{Z}} = \text{span} \{\tilde{\psi}_{k,l}(x)\}_{k,l \in \mathbb{Z}} = L_2(\mathbb{R})$.

Es gebe außerdem eine *Skalierungsfunktion* ϕ ,

$$\phi_{k,l}(x) = 2^{-\frac{k}{2}} \phi(2^{-k}(x-l)),$$

und deren duales Pendant $\tilde{\phi}$, das aus der Funktion $\phi(x)$ hervorgehe und dem *skalierbaren Punktoperator* entspreche.

Die Skalierungsfunktionen einer festen Skala s_k spannen einen *Approximations-Unterraum* V_k von $L_2(\mathbb{R})$ auf. Also gilt $V_k \subset L_2(\mathbb{R})$ mit

$$V_k = \text{span} \{\phi_{k,l}(x)\}_{l \in \mathbb{Z}}.$$

Mit wachsender Skala s_k unterscheidet sich der Approximations-Unterraum immer mehr von $L_2(\mathbb{R})$. Es gehen zunehmend Details verloren. Umgekehrt gilt

$$\lim_{k \rightarrow -\infty} V_k = L_2(\mathbb{R}).$$

Also führt die Anwendung eines *Approximationsoperators* \mathcal{A}_k , $k \in \mathbb{Z}$, zu einer Folge von Signalrepräsentationen $f_k^a(x) \in V_k$,

$$\begin{aligned}f_k^a(x) &= \mathcal{A}_k\{f(x)\} = \{\langle f(x), \phi_{k,l}(x) \rangle\}_{l \in \mathbb{Z}} \\ &= \sum_{l=-\infty}^{\infty} a_{kl}(f) \tilde{\phi}_{k,l}(x)\end{aligned}$$

Diese Folge konvergiert gegen das Signal $f(x)$ für $k \rightarrow -\infty$:

$$\lim_{k \rightarrow -\infty} f_k^a(x) = f(x)$$

Die Koeffizienten $a_{kl}(f)$ sind die *Abtastwerte eines geeignet tiefpaßgefilterten Signals*, denn jedes innere Produkt kann als Faltung mit der gespiegelten Impulsantwort geschrieben werden. Aus der Konvergenz der Folge der Approximations-Unterräume folgt die Eigenschaft

$$\dots \subset V_{k+1} \subset V_k \subset V_{k-1} \subset \dots$$

Der Raum V_k enthält mehr Struktur als der Raum V_{k+1} . Für jedes k gilt die Biorthogonalität $\langle \phi_{k,l}, \tilde{\phi}_{k,l'} \rangle = \delta_{ll'}$, demzufolge bilden die ϕ_{kl} eine Riesz-Basis von V_k . Bemerkenswert ist die Orthogonalität der Skalierungsfunktion bezüglich verschiedener Positionen auf gleicher Skala.

Die mit jedem Approximationsschritt (wachsendes k) verloren gegangenen Details werden in einen zu V_k komplementären Unterraum O_k abgespalten. Das heißt, es gilt

$$V_k = V_{k+1} \oplus O_{k+1}.$$

Der Durchschnitt aller dieser *Detail-Unterräume* O_k ist der Nullvektor. Sie ergänzen sich zum Raum $L_2(\mathbb{R})$:

$$L_2(\mathbb{R}) = \bigoplus_{k=-\infty}^{\infty} O_k.$$

Die Detail-Unterräume für alle Positionen auf gleicher Skala werden von den Wavelets aufgespannt.

$$O_k = \text{span} \{ \psi_{k,l}(x) \}_{l \in \mathbb{Z}}$$

Die Projektion des Signals $f(x)$ auf einen Detailraum O_k erfolgt durch einen *Detailoperator* \mathcal{D}_k und liefert für alle k eine Folge von *Detailsignalen* $f_k^d(x) \in O_k$,

$$\begin{aligned} f_k^d(x) &= \mathcal{D}_k \{ f(x) \} = \{ \langle f(x), \psi_{k,l}(x) \rangle \}_{l \in \mathbb{Z}} \\ &= \sum_{l=-\infty}^{\infty} d_{kl}(f) \tilde{\psi}_{k,l}(x). \end{aligned}$$

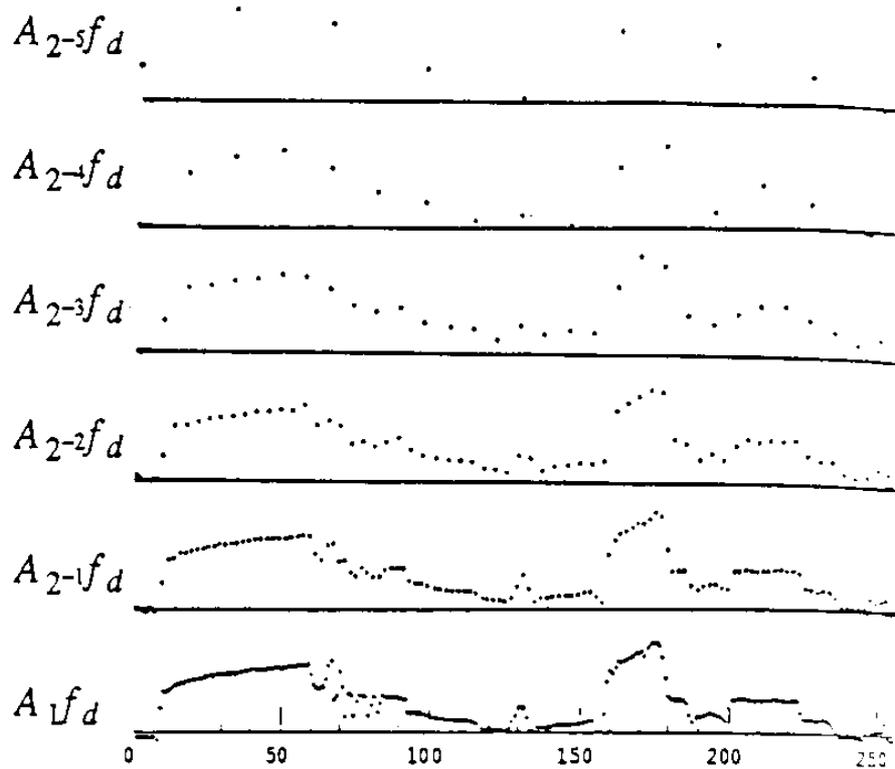
Die Koeffizienten $d_{kl}(f)$ sind die *Abtastwerte eines mittels Wavelets geeignet bandpaßgefilterten Signals*.

$$d_{kl}(f) = \langle f(x), \psi_{k,l}(x) \rangle \quad k, l \in \mathbb{Z}$$

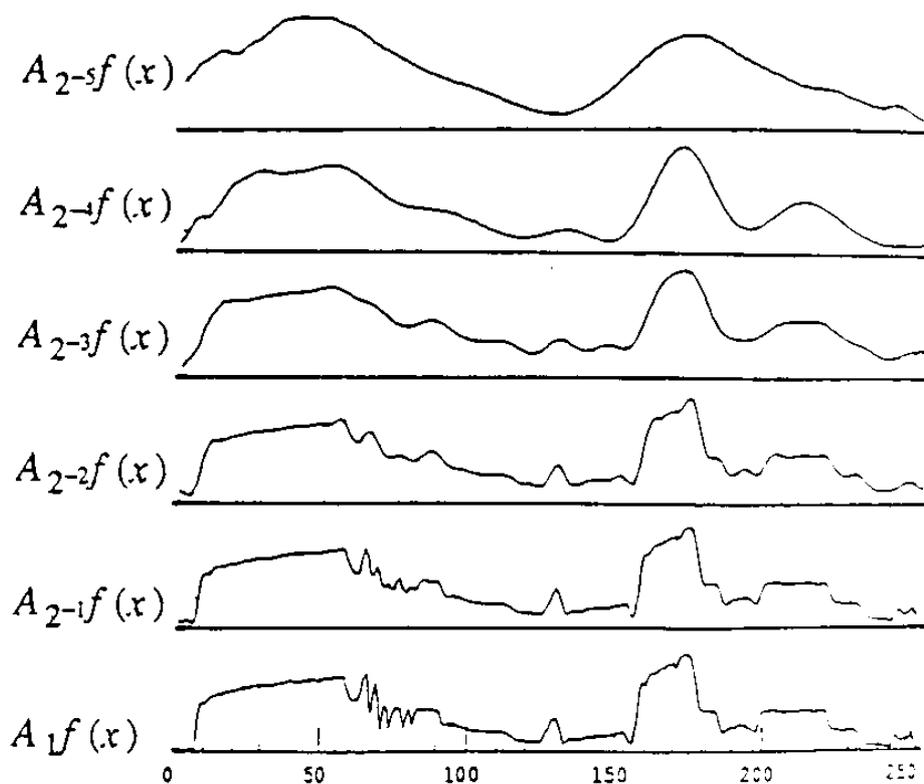
Die Detailsignale $f_k^d(x)$ sind ebenso wie die Wavelets *linear unabhängig*.

In der folgenden Abbildung aus Mallat (1989) wird das Approximationsschema der Multiskalenanalyse verdeutlicht.

Achtung! In der Abbildung ist A im Frequenzraum definiert, also kehrt sich k um: $k \rightarrow -k$.



(a)



(b)

3.11.3 Gabor-Wavelets

Gabor-Wavelets werfen einige ernsthafte theoretische Probleme auf, denn

- ▶ die Skalierungsfunktion (Gaußfunktion) ist nicht kompakt
- ▶ die Wavelets sind nicht orthogonal

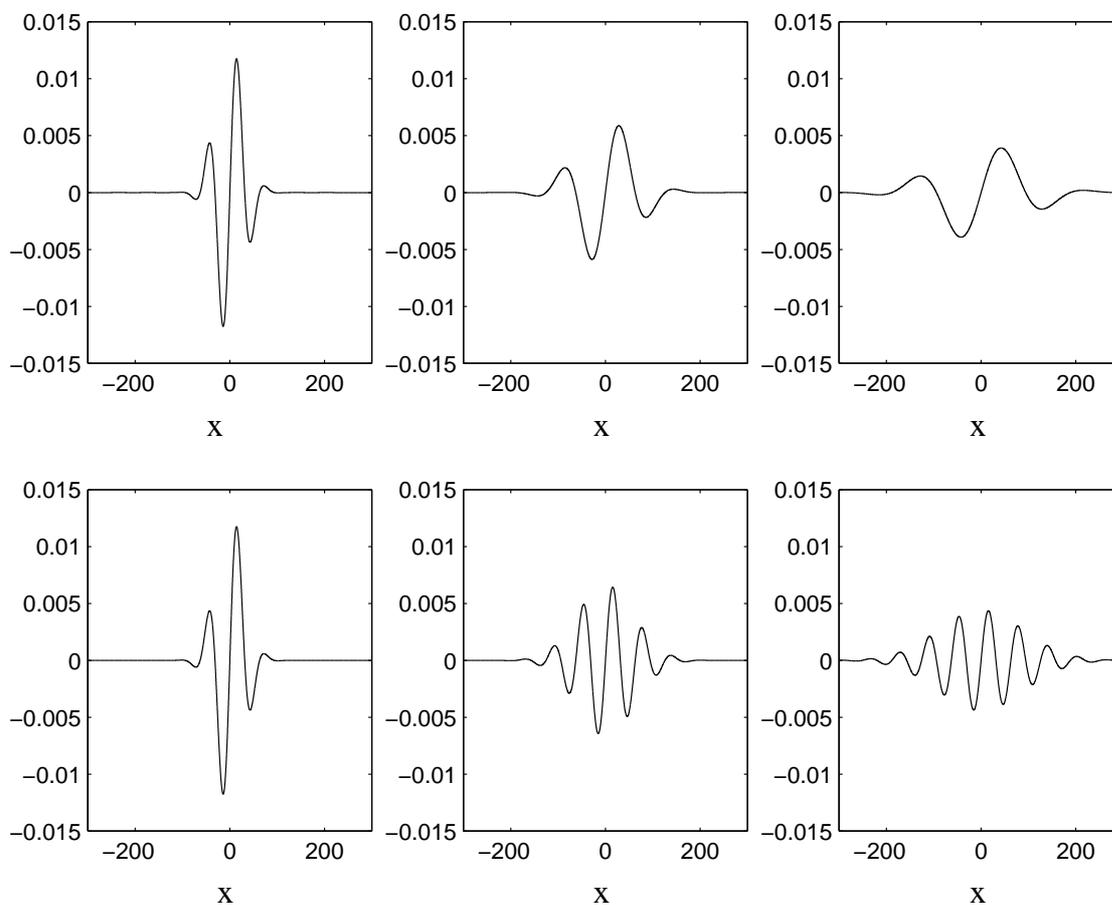
Dennoch spielen sie in der Bildverarbeitung eine hervorragende Rolle. In den letzten Jahren wurde gezeigt, daß sie zwar keine orthogonale Basis bilden, aber einen *Frame* und somit biorthogonale Eigenschaften annehmen.

Das Gabor-Wavelet heißt auch *Morlet-Wavelet*.

Werden Gaborfunktionen als Gaborwavelets verwendet, ersetzt man in der komplexen Harmonischen die Frequenz ω_0 durch den Ausdruck c/σ , wobei c eine Konstante ist und σ der Skalenparameter. Hält man die Modulationsfrequenz fest und variiert den Skalenparameter σ , so variiert auch c . Es ergibt sich das in der unteren Reihe der folgenden Abbildungen dargestellte Verhalten. Hält man dagegen $c = \text{const}$ und variiert σ , so nimmt auch die Kreisfrequenz ab, siehe obere Reihe der folgenden Abbildung. In diesem Fall wird das für Wavelets typische selbstähnliche Skalierungsverhalten erreicht. Man wählt typischerweise $c \geq 5$.

Also schreiben wir für ein Gaborwavelet

$$h(x; \sigma, c) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} e^{jc\frac{x}{\sigma}}.$$



Obere Reihe: Der imaginäre Teil von drei Gaborwavelets mit $c = 3$ und einer Varianz $\sigma = 30$

(links), $\sigma = 60$ (mitte) und $\sigma = 90$ (rechts). Hier gilt also $\omega_0 \sim \frac{1}{\sigma}$. Untere Reihe: Der imaginäre Teil von drei Funktionen der Gabor-Transformation mit $\omega = 1/10$ und Varianz $\sigma = 30$ (links), $\sigma = 60$ (mitte) und $\sigma = 90$ (rechts). Hier gilt also $c \sim \sigma$.

3.11.4 Gaborfunktion als Quadraturpaar

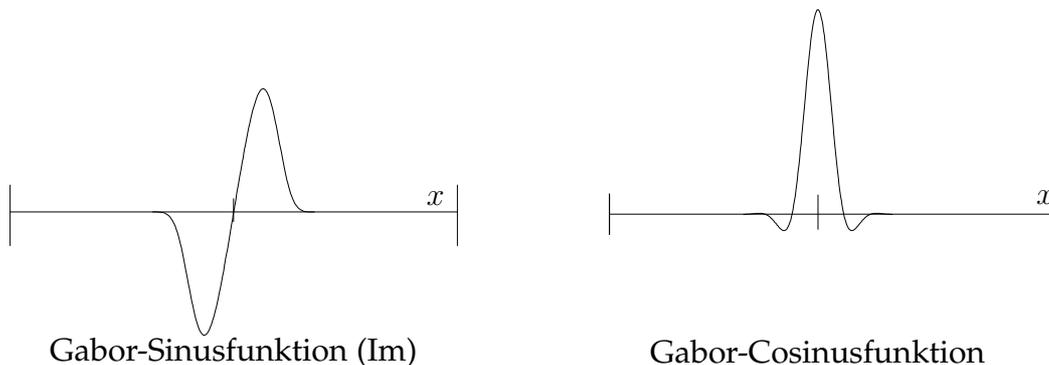
Die reelle und die imaginäre Komponente der Gaborfunktion bilden ein Quadraturpaar.

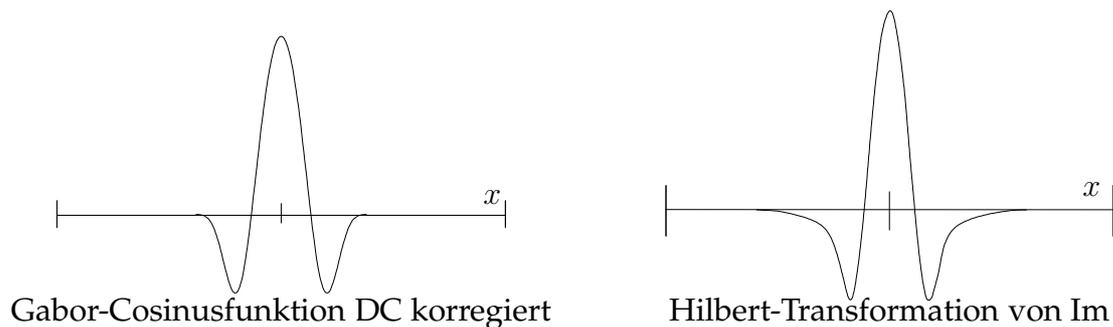
$$\begin{aligned}
 h_e(x) &= \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}} \cdot \cos(2\pi u_0 x) \\
 H_e(u) &= e^{-2\pi^2\sigma^2 u^2} * \frac{1}{2}(\delta(u + u_0) + \delta(u - u_0)) \\
 &= \frac{1}{2}(e^{-2\pi^2\sigma^2(u+u_0)^2} + e^{-2\pi^2\sigma^2(u-u_0)^2}) \\
 h_o(x) &= \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}} \cdot \sin(2\pi u_0 x) \\
 H_o(u) &= e^{-2\pi^2\sigma^2 u^2} * \frac{1}{2}j(\delta(u + u_0) - \delta(u - u_0)) \\
 &= \frac{1}{2}j(e^{-2\pi^2\sigma^2(u+u_0)^2} - e^{-2\pi^2\sigma^2(u-u_0)^2})
 \end{aligned}$$

Diese bilden aber kein echtes Hilbertpaar, da $H_e(0) \neq 0!$

Die Korrektur des Gleichanteils der Cosinus-Komponente hebt diesen Nachteil auf. (Die Folge wären nämlich nichtlineare Phasenbeziehungen). Im Falle eines Gaborwavelets (siehe Abschnitt 3.11.3), das auch als Hilbertfilter verwendet werden kann, kommt folgende Normierung zur Anwendung

$$h(x; \sigma, c) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}} (e^{jc\frac{x}{\sigma}} - e^{-\frac{1}{2}c^2}).$$





Es gilt also

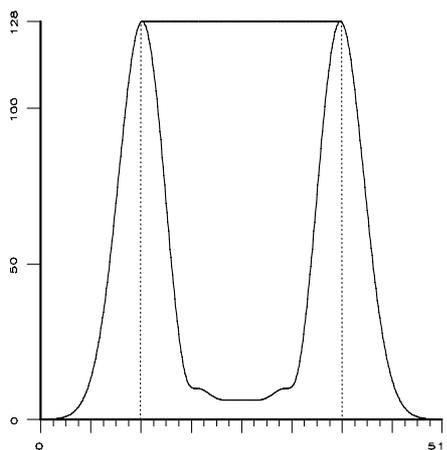
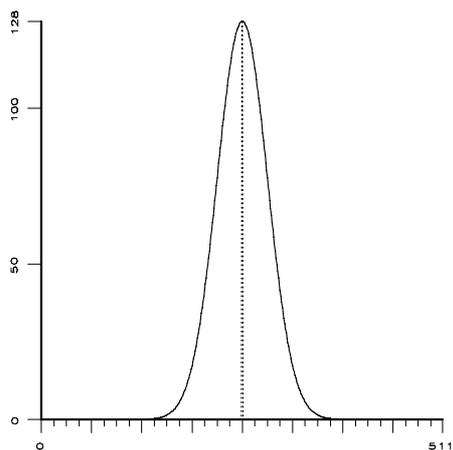
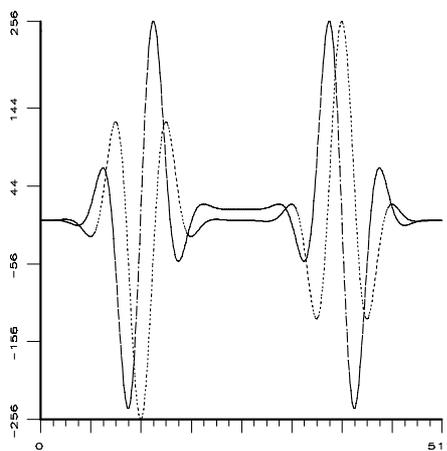
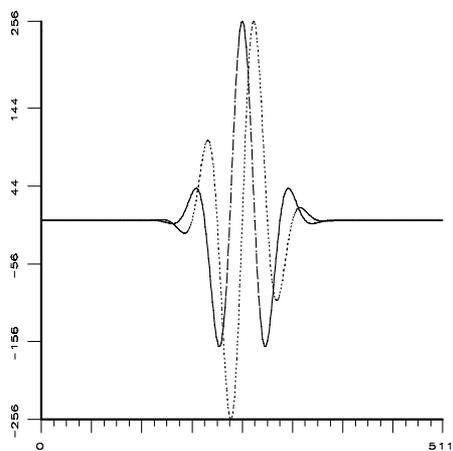
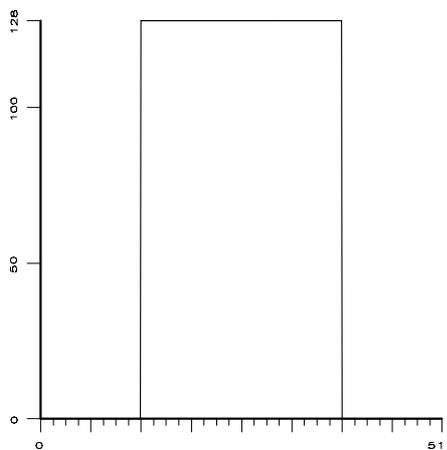
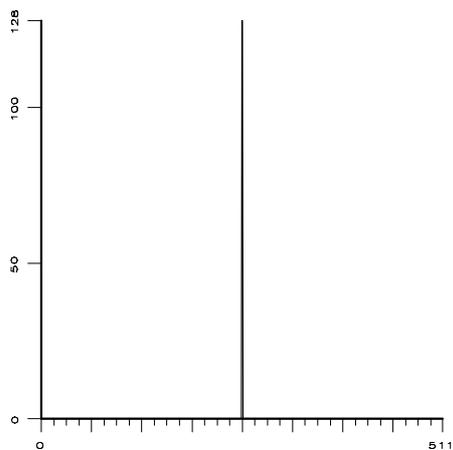
$$\begin{aligned}
 h_e &\perp h_o \\
 h_e^2 + h_o^2 &= 1 \\
 \mathcal{H}\{h_e\} &= h_o, \quad \mathcal{H}\{h_o\} = -h_e.
 \end{aligned}$$

Die Symmetrie der geraden/ungeraden Gaborfilter wird dem Output ihrer Faltung mit einem Signal f aufgeprägt. Es können die lokalen spektralen Eigenschaften Amplitude/Energie und Phase abgeleitet werden. Demzufolge sind Gabor-Quadraturpaare Kanten-/Liniendetektoren.

In den folgenden beiden Abbildungen fällt auf, daß die lokale Energie unabhängig von der Art der Kante ist. Anordnung in den Abbildungen: Oben - Signal, Mitte - Antworten der geraden/ungeraden Gaborfilter, Unten - normierte lokale Energie.

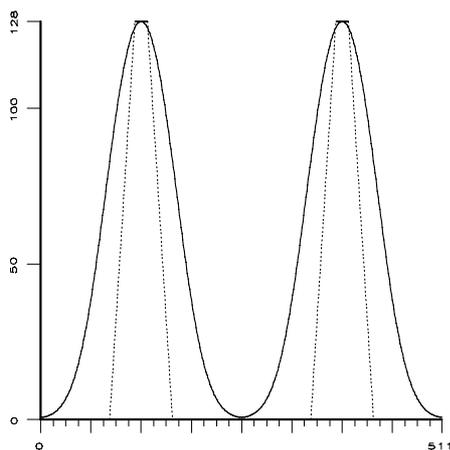
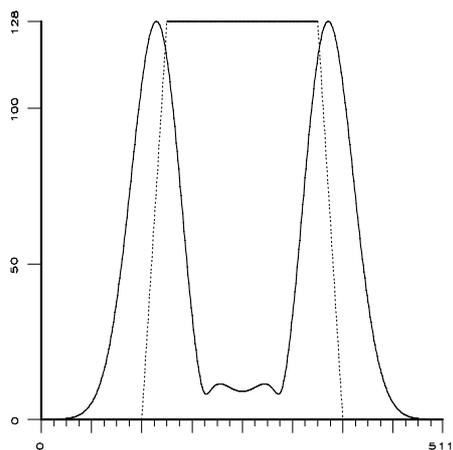
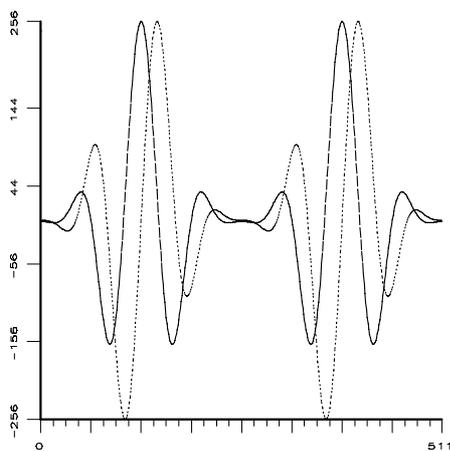
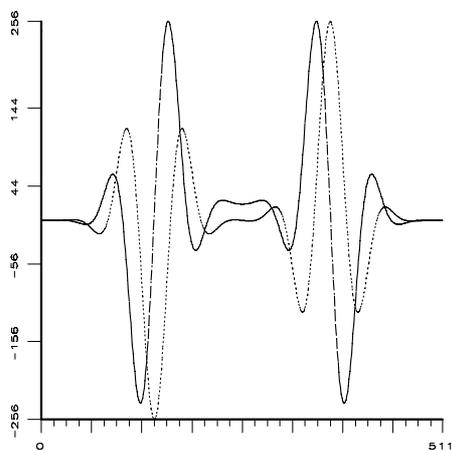
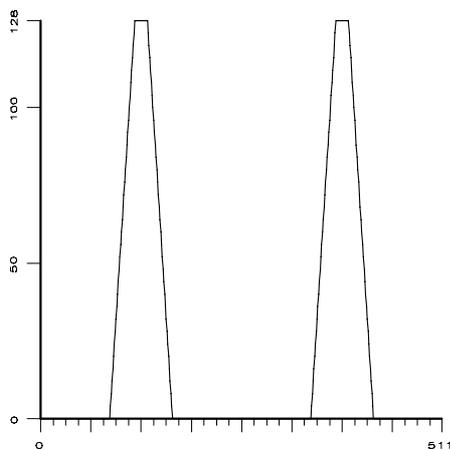
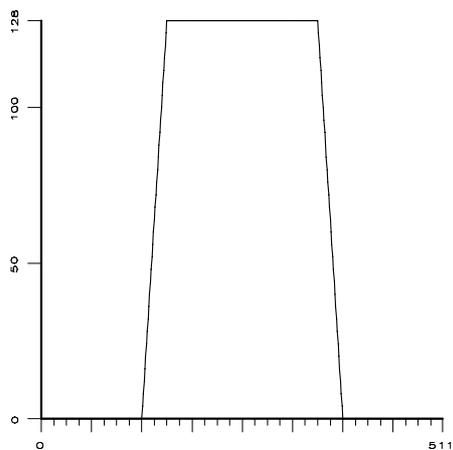
δ -Impuls

positive/negative Stufenkante

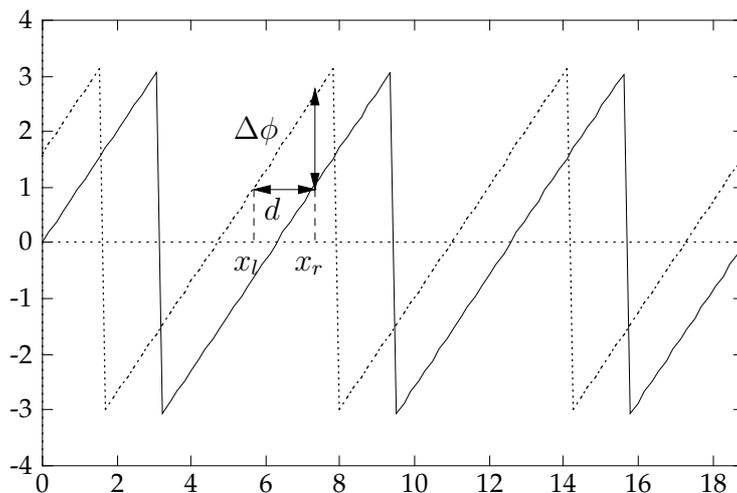


positive/negative Rampenkante

zwei Trapezimpulse

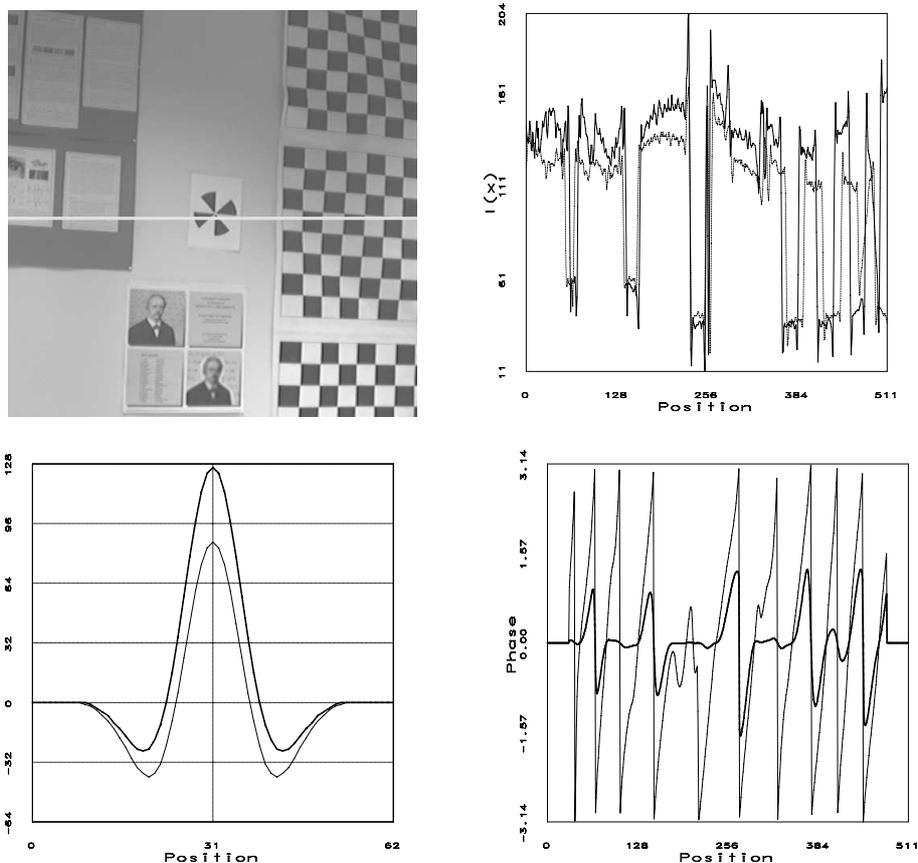


Beispiel: *Phasenbasierte Disparitätsbestimmung bei Stereosehen*



Die Idee der phasenbasierten Disparitätsschätzung basiert auf den lokalen Phasenantworten $\varphi_{l/r}(x)$. Aus der Differenz $\Delta\phi(x) = \varphi_l(x) - \varphi_r(x)$ der lokalen Phasen kann direkt die Verschiebung $d = x_l - x_r$ zwischen den korrespondierenden Phasenantworten berechnet werden.



Beispiel:

(o.l.) Eine strukturierte Wand des Labors und (o.r.) zwei reale Grauwertverläufe der markierten Bildzeile. (u.l.) Der gerade Gaborfilter für $t = 0.66$ mit und ohne Gleichanteilskorrektur (dicke Linie) mit dem die lokale Phase extrahiert wurde. (u.r.) Die dünne Linie zeigt die lokale Phase für den korrigierten Filter; die dicke Linie zeigt die Phase des unkorrigierten Filters.

■

3.12 Die 2D-Gaborfunktion

In diesem Abschnitt wird die Gaborfunktion und alle mit ihr im Zusammenhang stehenden Konzepte auf den 2D-Fall erweitert. Die Einbettung in die Ebene hat eine Reihe von Freiheitsgraden zur Folge. Wir werden außerdem zeigen, wie das Konzept auf die lokale Analyse von intrinsischen 2D-Strukturen erweitert werden kann.

3.12.1 Die Pseudo-2D-Gaborfunktion

Wir wollen in diesem Abschnitt die 1D-Gaborfunktion in die 2D-Bildebene einbetten und zeigen, wie sie als lokale Fouriertransformation oder als Gabor-Wavelettransformation zur Ab-

leitung lokaler spektraler Merkmale im Bild verwendet werden kann. Natürlich gilt die Beschränkung, daß komplexwertige Gaborfunktionen auch bei 2D-Einbettung nur intrinsisch eindimensionale Merkmale der Bildstruktur erfassen. Wir nennen deshalb die so eingebetteten Gaborfunktionen Pseudo-2D-Gaborfunktionen. Erst durch quaternionische Repräsentation der Gaborfunktion in die Bildebene erhalten wir echte 2D-Gaborfunktionen, die auch intrinsisch zweidimensionale Strukturen erfassen.

Die Pseudo-2D-Gaborfunktionen wurden von Daugman (1980) eingeführt. Sie sind 1D-Gaborfunktionen, die eine Orientierung in der Bildebene besitzen. Diese Orientierung wird durch eine eventuell vorhandene exzentrische Gestalt aber auch durch die Ausbreitungsrichtung der komplexen Harmonischen bestimmt. Pseudo-2D-Gaborfunktionen haben einen hochdimensionalen Parameterraum, der ihre Spezifik im Ortsraum und/oder im Frequenzraum bestimmt. Wir nähern uns ihrer Beschreibung schrittweise.

Die Zielstellung ist, orientierungsselektiv und skaliert eine lokale spektrale Analyse zu ermöglichen.

a) *Isotrope 2D-Gaußfunktion*

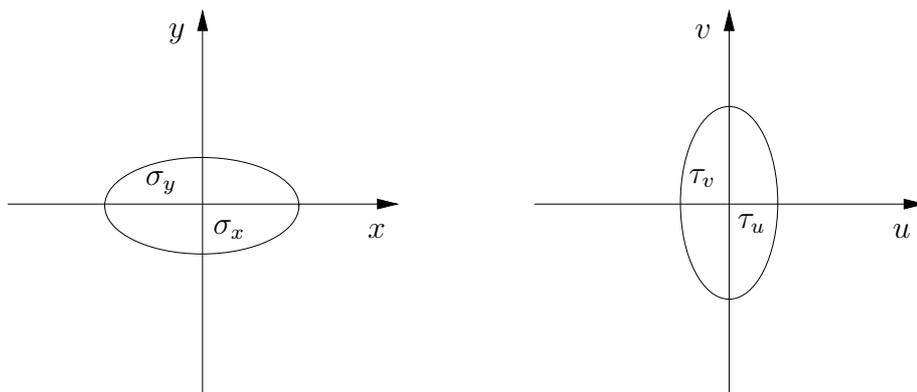
Seien $\mathbf{x} = (x, y)^T$, $\mathbf{u} = (u, v)^T$ Ortsvektoren im Ortsraum und im Frequenzraum. Es gelte $\sigma_x = \sigma_y = \sigma$.

$$\begin{aligned}
 g(\mathbf{x}; \sigma) &= g(x; \sigma) g(y; \sigma) \\
 \begin{array}{c} \circ \\ | \\ \bullet \end{array} &= \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \\
 \begin{array}{c} \bullet \\ | \\ \bullet \end{array} &= e^{-2\pi^2\sigma^2(u^2+v^2)}
 \end{aligned}$$

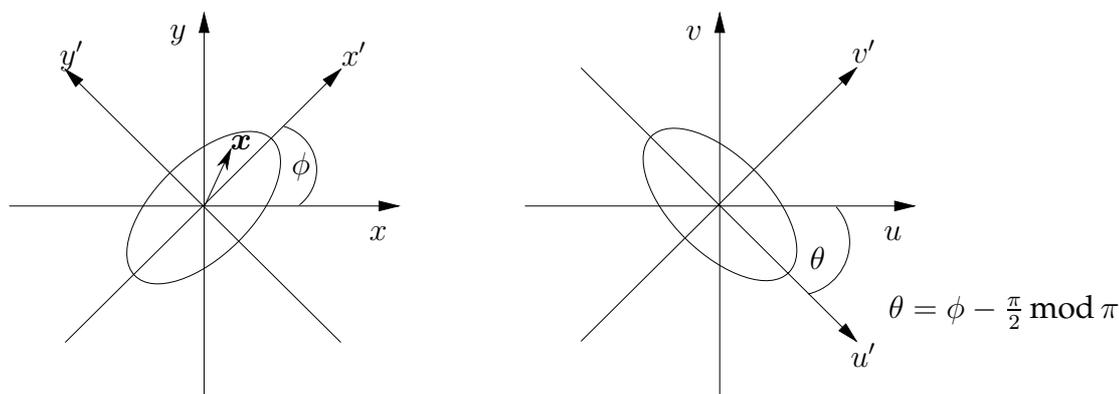
b) *Exzentrische 2D-Gaußfunktion*

Es gelte $\sigma_x = \lambda \sigma_y = \lambda \sigma$, $\lambda \geq 1$. Das *Aspektverhältnis* λ gibt das Verhältnis von großer zu kleiner Achse der einhüllenden Ellipse an. Die große Hauptachse sei zur x -Achse orientiert.

$$\begin{aligned}
 g(\mathbf{x}; \sigma, \lambda) &= g(x; \sigma, \lambda) g(y; \sigma, 1) \\
 \begin{array}{c} \circ \\ | \\ \bullet \end{array} &= \frac{1}{2\pi\lambda\sigma^2} e^{-\frac{(x/\lambda)^2+y^2}{2\sigma^2}} \\
 \begin{array}{c} \bullet \\ | \\ \bullet \end{array} &= e^{-2\pi^2\sigma^2(\lambda^2u^2+v^2)}
 \end{aligned}$$

c) *Gedrehte exzentrische 2D-Gaußfunktion*

Sei ϕ der Drehwinkel der großen Achse, der den Rotationsoperator \mathcal{R}_ϕ parametrisiert. Wir benötigen zusätzlich $\mathbf{x}' = (x', y')^T$ und $\mathbf{u}' = (u', v')^T$.



Für jeden Punkt \mathbf{x} der Gaußfunktion gilt $\mathbf{x} = \mathcal{R}_\phi\{\mathbf{x}'\}$,

$$\mathbf{x} = R(\phi)\mathbf{x}' \quad R(\phi) = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}$$

bzw.

$$\mathbf{x}' = R^{-1}(\phi)\mathbf{x} \quad R^{-1}(\phi) = R^T(\phi) = R(-\phi)$$

Die 2D-Gaußfunktion in Bildkoordinaten \mathbf{x} wird aus den an die Funktion angehefteten Koordinaten \mathbf{x}' abgeleitet.

$$g(\mathbf{x}; \sigma, \lambda, \phi) = \mathcal{R}_\phi\{g(\mathbf{x}'; \sigma, \lambda)\} = g(\mathcal{R}_\phi\{\mathbf{x}'\}; \sigma, \lambda)$$

$$\begin{array}{c} \circ \\ | \\ \bullet \end{array} = R(\phi)g(\mathbf{x}; \sigma, \lambda, 0)$$

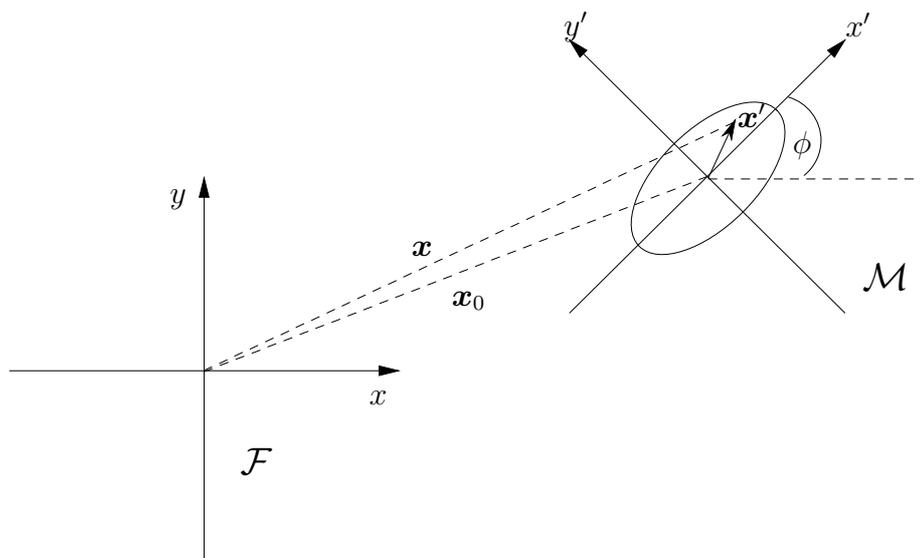
$$G(\mathbf{u}; \sigma, \lambda, \theta) = R(\phi - \frac{\pi}{2})G(\mathbf{u}; \sigma, \lambda, 0) = G(\mathcal{R}_\theta\{\mathbf{u}'\}; \sigma, \lambda)$$

D.h., $\mathbf{u} = R(\phi - \frac{\pi}{2})\mathbf{u}'$ bzw. $\mathbf{u}' = R(-\phi + \frac{\pi}{2})\mathbf{u}$.

d) *Rotierte und translatierte exzentrische 2D-Gaußfunktion*

Rotation und Translation eines starren Körpers (der Gaußfunktion) bezeichnet man als starre Verschiebung (*rigid displacement*) bzw. Euklidische Transformation. Sei $\mathcal{F} : (\mathbf{x}, \phi)$ ein festes Koordinatensystem (des Bildes) und $\mathcal{M} : (\mathbf{x}', 0)$ ein bewegliches, an die Gaußfunktion gekoppeltes Koordinatensystem.

Wollen wir die Gaußfunktion in Bildkoordinaten beschreiben, ist eine Koordinatentransformation $KT : \mathcal{F} \leftarrow \mathcal{M}$ erforderlich. Speziell gilt $KT : \mathbf{x} = R(\phi)\mathbf{x}' + \mathbf{x}_0$.



Also gilt

$$g(\mathbf{x}; \sigma, \lambda, \phi, \mathbf{x}_0) = \mathcal{D}_{\phi, \mathbf{x}_0} \{g(\mathbf{x}'; \sigma, \lambda)\} = g\{\mathcal{D}_{\phi, \mathbf{x}_0} \{\mathbf{x}'\}; \sigma, \lambda\}$$

Die Translation im Ortsraum bewirkt eine Phasenverschiebung im Frequenzraum.

$$G(\mathbf{u}; \sigma, \lambda, \theta, \mathbf{x}_0) = e^{-j2\pi \mathbf{u}^T \mathbf{x}_0} G(\mathbf{u}; \sigma, \lambda, \theta, 0)$$

Die Abtastung einer Bildfunktion $f(\mathbf{x})$ mit translatierten und rotierten Kopien einer Mutter-Gaborfunktion kann auf zwei verschiedene Wege implementiert werden:

1. Starre Verschiebung $\mathcal{D}_{\phi_j, \mathbf{x}_{0j}} = (\mathcal{R}_{\phi_j}, \mathcal{T}_{\mathbf{x}_{0j}})$ in einem Schritt
2. Inkrementierung der Verschiebungsparameter, so daß

$$\mathcal{D}_{\phi_j, \mathbf{x}_{0j}} = (\mathcal{D}_{\Delta\phi, \Delta\mathbf{x}_0} \mathcal{D}_{\phi_{j-1}, \mathbf{x}_{0,j-1}})$$

Die erste Methode wird bevorzugt. Die zweite Methode beruht auf folgendem Zusammenhang:

Ist $\mathcal{D}_1 : \mathcal{F} \rightarrow \mathcal{M}_1$ und $\mathcal{D}_2 : \mathcal{M}_1 \rightarrow \mathcal{M}_2$, dann existiert auch $\mathcal{D} = \mathcal{D}_1\mathcal{D}_2 : \mathcal{F} \rightarrow \mathcal{M}_2$. Für $\mathcal{D}_1 = (R(\phi_1), \mathbf{t}_1)$, $\mathcal{D}_2 = (R(\phi_2), \mathbf{t}_2)$ folgt

$$\mathbf{x} = [R(\phi_1)R(\phi_2)]\mathbf{x}' + [R(\phi_1)]\mathbf{t}_2 + \mathbf{t}_1$$

und folglich ist

$$\mathcal{D} = \mathcal{D}_1\mathcal{D}_2 = (R(\phi_1), \mathbf{t}_1)(R(\phi_2), \mathbf{t}_2) = (R(\phi_1)R(\phi_2), R(\phi_1)\mathbf{t}_2 + \mathbf{t}_1).$$

Rotation ist eine lineare Operation bzgl. der Wirkung auf einen Vektor $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$

$$\mathcal{R}\{\mathbf{x}\} = \mathcal{R}\{\mathbf{x}_1 + \mathbf{x}_2\} = \mathcal{R}\{\mathbf{x}_1\} + \mathcal{R}\{\mathbf{x}_2\}.$$

Translation ist (in \mathbb{R}^n) aber eine nichtlineare Operation bzgl. der Wirkung auf einen Vektor $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$.

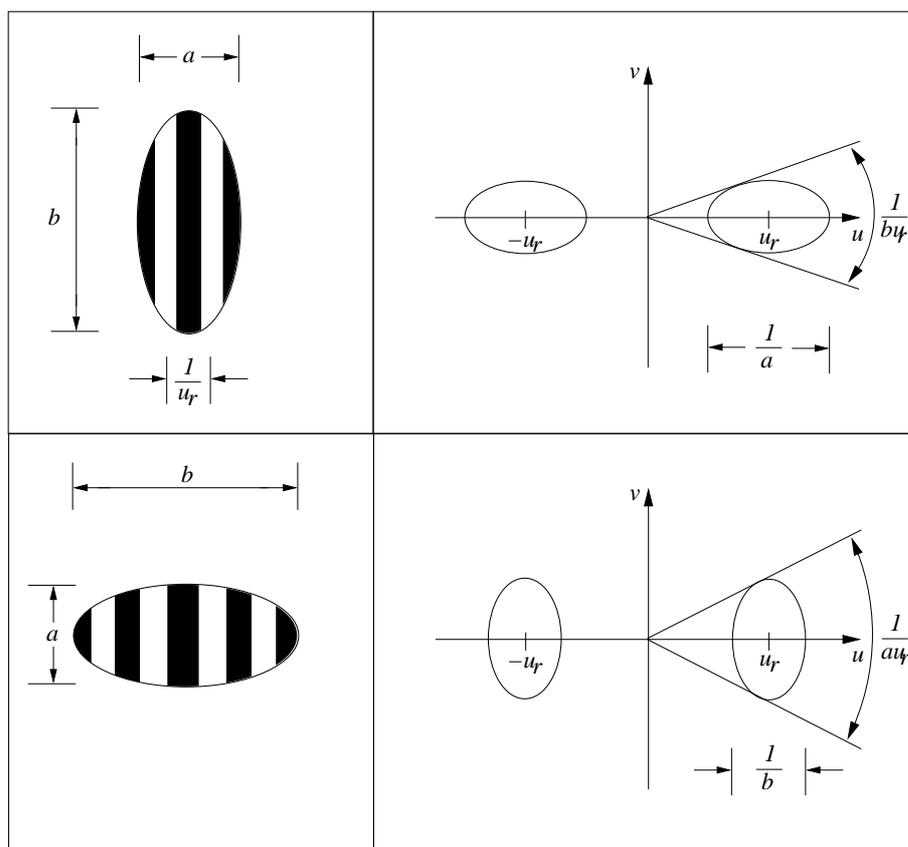
$$\mathcal{T}\{\mathbf{x}\} = \mathcal{T}\{\mathbf{x}_1 + \mathbf{x}_2\} = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{t} \neq (\mathbf{x}_1 + \mathbf{t}) + (\mathbf{x}_2 + \mathbf{t}) = \mathbf{x}_1 + \mathbf{x}_2 + 2\mathbf{t}$$

Deshalb wird Methode 1 verwendet.

e) *Pseudo-2D-Gaborfunktion*

Die Pseudo-2D-Gaborfunktion erhält man durch Modulation der (rotierten und translatierten) Gaußfunktion mit einer orientierten 1D-Harmonischen. Sei $\mathbf{u}_0 = (u_0, v_0)^T$ der Vektor der Mittenfrequenz des Paßbandes, $u_r = |\mathbf{u}_0|$.

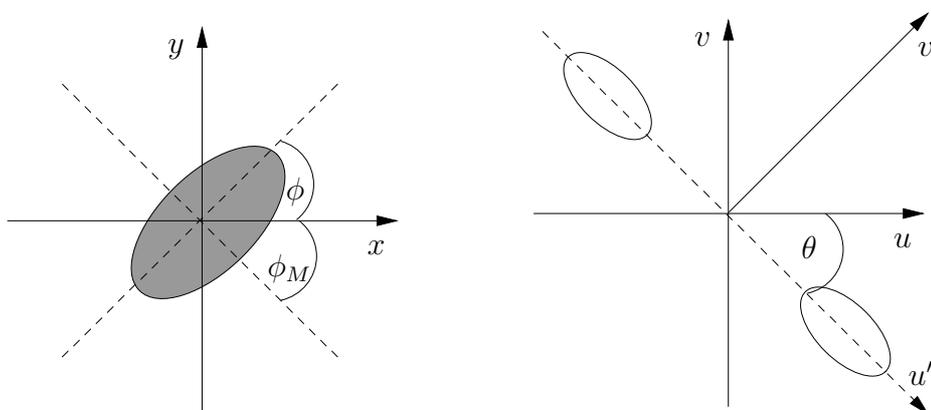
Als Erweiterung der unter d) geführten Diskussion um die Rotation der gaußschen Einhüllenden in der Bildebene muß ebenfalls die hiervon unabhängige Rotation der harmonischen Modulation betrachtet werden. Zunächst ist die Modulationsrichtung als unabhängig von der Orientierung der Gaußfunktion zu sehen. Man möchte aber eine Parkettierung des Frequenzraums im Falle einer Gabortransformation erreichen, so daß die Paßbänder unabhängig von der Orientierung gleichwertige Frequenzbereiche repräsentieren und die azimuthalen Ausdehnungen der Paßbänder klein sind (hohe Winkelauflösung). In der folgenden Abbildung sind mit a der kleine Durchmesser und mit b der große Durchmesser einer exzentrischen Einhüllenden bezeichnet. Die Abbildung zeigt die unterschiedliche Orientierung der Abtastfunktion im Frequenzraum, wenn zu der fest gewählten Modulationsrichtung (x -Achse) die Einhüllende verschieden gewählt wird.



Deshalb erfolgt eine Kopplung der Orientierung der Harmonischen an die Orientierung der Gaußfunktion in folgender Weise:

$$\phi_M = \phi - \frac{\pi}{2}.$$

Dies bewirkt eine Verschiebung der Gaußfunktion im Frequenzraum in Richtung großer Hauptachse. Damit werden die Frequenzbandbreite B und die Orientierungsbandbreite Ω unabhängig vom Orientierungswinkel ϕ .



Bei konstanter radialer Mittenfrequenz $|u_0|$ drehen sich die Paßbänder mit der Orientierung der Gaborfunktion.

Die orientierte Pseudo-2D-Gaborfunktion ist definiert durch

$$h(\mathbf{x}; \sigma, \lambda, \phi, \mathbf{x}_0, \mathbf{u}_0) = g(\mathbf{x}; \sigma, \lambda, \phi, \mathbf{x}_0) e^{j2\pi(\phi - (\pi/2))\mathbf{u}_0}$$



$$\begin{aligned} H(\mathbf{u}; \sigma, \lambda, \theta, \mathbf{x}_0, \mathbf{u}_0) &= G(\mathbf{u}; \sigma, \lambda, \theta, \mathbf{x}_0) * (\delta(u - u_0)\delta(v - v_0)) \\ &= e^{-2\pi^2\sigma^2(\lambda^2(u' - u'_0)^2 + v'^2)} \end{aligned}$$

Werden diese Funktionen für eine Gabor-Transformation verwendet, wählt man σ und λ fest und verwendet u_0 und ϕ als Freiheitsgrad:

$$h(\mathbf{x}; \sigma, \lambda, \phi, \mathbf{x}_0, \mathbf{u}_0) \longrightarrow h_{\sigma, \lambda}(\mathbf{x}; \phi, \mathbf{x}_0, \mathbf{u}_0)$$

Diese Funktionen sind Abtastfunktionen am Ort \mathbf{x}_0 mit konstanter radialer Bandbreite und Orientierungsbandbreite und mit variabler Orientierung. Der Frequenzraum wird vorzugsweise in Polarkoordinaten (u_r, θ) dargestellt.

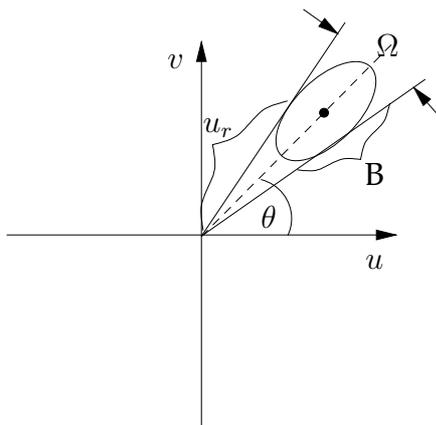
$$u_r = |\mathbf{u}_0| = \sqrt{u_0^2 + v_0^2} \quad \text{radiale Zentrumsfrequenz des Paßbandes}$$

$$\theta = \arg(u_0, v_0) \quad \text{Orientierung Zentrumsfrequenz}$$

$$B = \log_2 \left[\frac{\pi u_r \lambda \sigma + \alpha}{\pi u_r \lambda \sigma - \alpha} \right] \quad \text{Frequenzbandbreite (FWHM) [Oktaven]}$$

$$\alpha = \sqrt{\frac{\ln 2}{2}}$$

$$\Omega = 2 \arctan \left[\frac{\alpha}{\pi u_r \sigma} \right] \quad \text{Orientierungsbandbreite (FWHM) [Radian]}$$

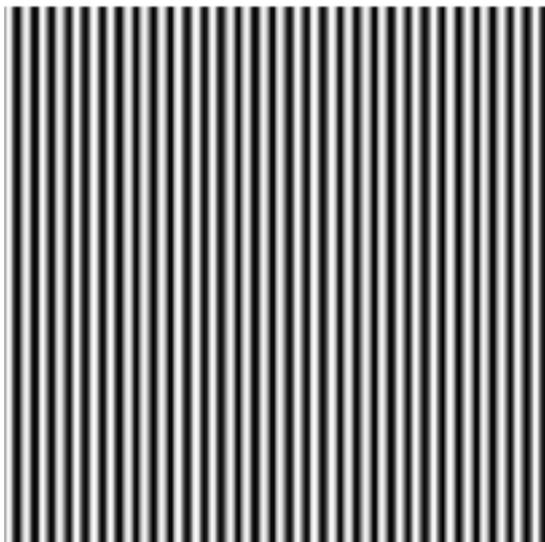


Werden orientierte Pseudo-2D-Gaborfunktionen im Sinne einer Gabor-Wavelettransformation verwendet, sind σ und ϕ Freiheitsgrade und λ, \mathbf{u}_0 sind fest gewählte Parameter. Die Kopplung von σ und \mathbf{u}_0 in einem Skalierungsparameter $c = \omega_0 \sigma$ bzw. einen Bandbreitenparameter $t = \frac{1}{c}$ skaliert die Gabor-Wavelets in selbstähnlicher Weise bei Variation von σ für $c = \text{const.}$

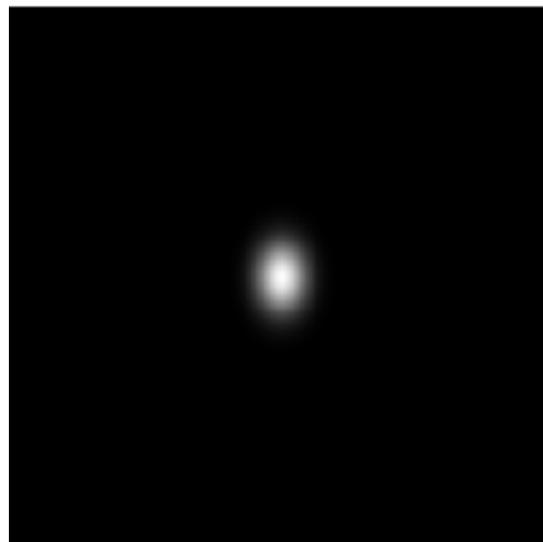
Im Frequenzraum entsteht dadurch eine Abtastung durch orientierte und mit der radialen Frequenz breiter (bezüglich B) werdende Paßbänder.

Beispiele:

Die Entstehung einer Pseudo-2D-Gaborfunktion zeigen folgende vier Bilder (invertiert, normiert).



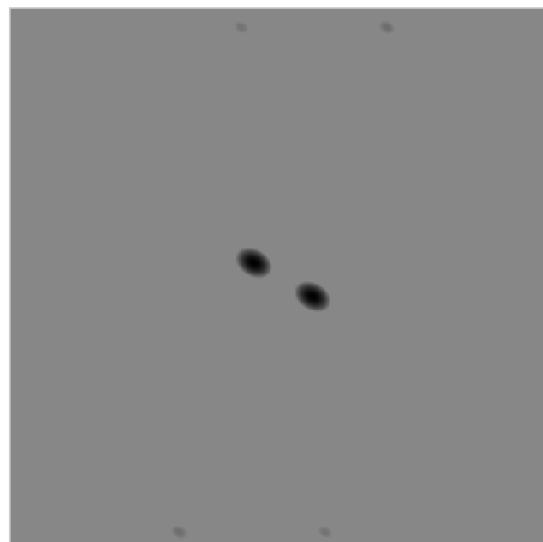
Sinusgitter



Gaußfunktion



Gaborfunktion gedreht

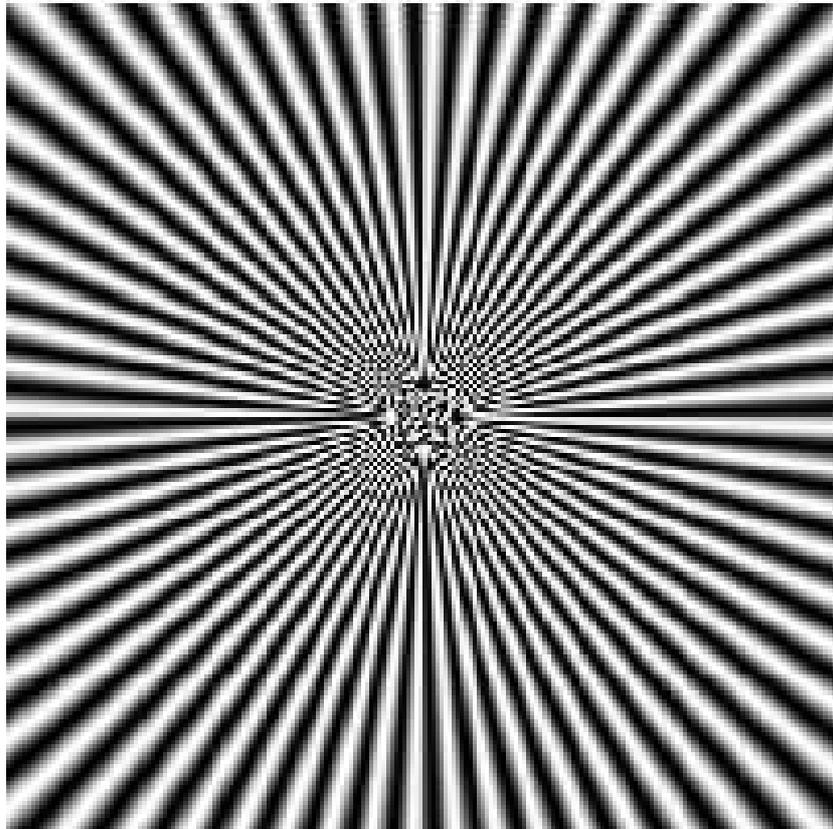


Spektrum der Gaborfunktion

Die Eignung des Gaborfilters, die lokalspektralen Größen Amplitude und Phase zu extrahieren, wollen wir am Beispiel des *Siemenssterns* als Eingangssignal zeigen. Der Siemensstern ist ein Muster, das in einem bestimmten Frequenzband alle Ortsfrequenzen in allen Orientierungen enthält. Der Betrag der Ortsfrequenz nimmt dabei von außen nach innen zu, hier, bis das Auflösungsvermögen des Druckers überschritten ist. Die Störeffekte im Zentrum des Sterns sind Aliasfehler, die von der rechtwinkligen Abtastung stammen.

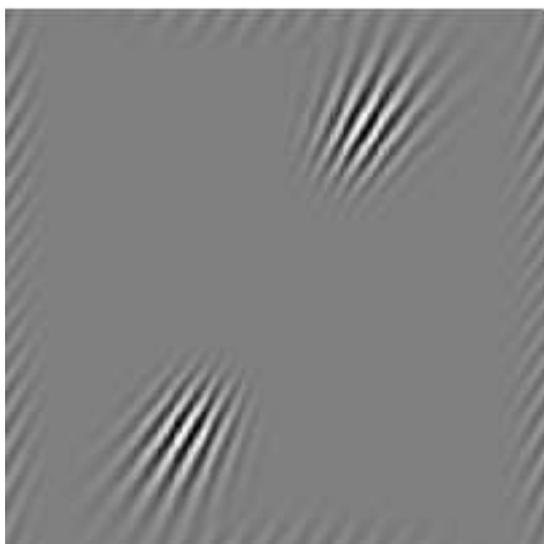


Beispiel:

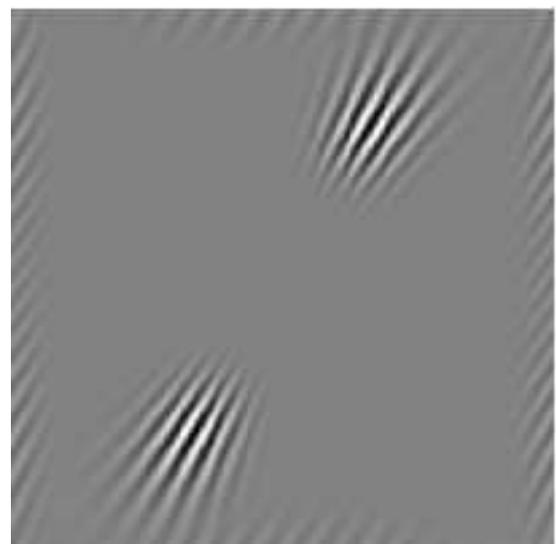


Siemensstern

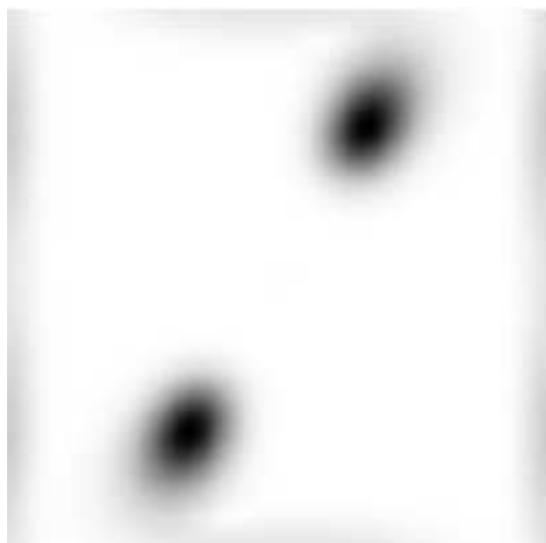
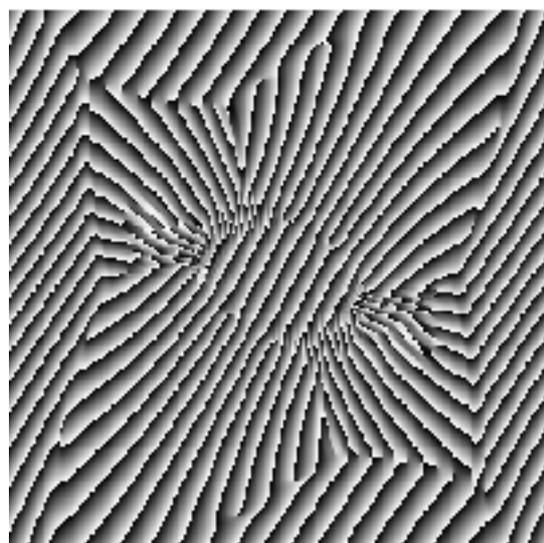
Folgendes Bildbeispiel illustriert die Erzeugung lokalspektraler Größen aus dem Ergebnis der Anwendung des oben abgebildeten Gaborfilters auf den Siemensstern.



Antwort auf gerade Gaborfunktion h_e



Antwort auf ungerade Gaborfunktion h_o

lokale Amplitude A lokale Phase ϕ 

3.12.2 Die quaternionwertige 2D-Gaborfunktion

Die komplexwertige Pseudo-2D-Gaborfunktion hat den Nachteil, lokal nur intrinsische 1D-Strukturen (gerader bzw. ungerader Symmetrie - also Linien und Kanten) erkennen zu können. Für die Detektion von intrinsischen 2D-Strukturen (Verzweigungen, Krümmungen, usw.) werden quaternionwertige 2D-Gaborfunktionen eingeführt.

Die quaternionwertige Gaborfunktion entsteht durch Erweiterung der komplexen Gaborfunktion dadurch, daß jeder Koordinatenachse des Bildträgers eine eigene komplexe Gaborfunktion zugeordnet wird. Die Beziehungen zwischen den Koordinatenachsen erhält man durch Einbettung der Gaborfunktion in die Algebra der Quaternionen. Allerdings sind auch andere gleich mächtige Einbettungen möglich. Die Wahl der Quaternionen folgt aus dem Zugang, der für die Einbettung der 2D-Fouriertransformation gewählt wurde.

Die quaternionwertige Gaborfunktion ist eine echte 2D-Gaborfunktion, weil sie alle Symmetrien von intrinsisch zweidimensionalen Signalen zu repräsentieren vermag. Sie umfaßt aber auch die 1D-Gaborfunktion, repräsentiert also als Spezialfall auch intrinsisch eindimensionale Strukturen.

Im Ortsraum ist die quaternionwertige Gaborfunktion gegeben durch

$$h^q(\mathbf{x}; \sigma, \lambda, \mathbf{u}_0) = g(\mathbf{x}; \sigma, \lambda) e^{i2\pi u_0 x} e^{j2\pi v_0 y}$$

also durch die mit der Gaußfunktion gewichteten Basisfunktion der quaternionwertigen Fouriertransformation. Im Frequenzraum ist die Übertragungsfunktion eines quaternionwertigen Gaborfilters beschrieben durch

$$H^q(\mathbf{u}; \sigma, \lambda, \mathbf{u}_0) = e^{-2\pi^2 \sigma^2 (\lambda^2 (u-u_0)^2 + (v-v_0)^2)}$$

Quaternionwertige Gaborfunktionen als Gaborwavelets sind definiert durch

$$h^q(\mathbf{x}; \sigma, \lambda, \mathbf{c}, \mathbf{u}_0) = g(\mathbf{x}; \sigma, \lambda) e^{\frac{ic_1 x}{\sigma}} e^{\frac{jc_2 y}{\sigma}}$$

mit $c_1 = 2\pi u_0\sigma$, $c_2 = 2\pi v_0\sigma$.

Explizit lautet die quaternionwertige Gaborfunktion

$$h^q(\mathbf{x}; \sigma, \lambda, \mathbf{u}_0) = (h_{ee}^q + ih_{oe}^q + jh_{eo}^q + kh_{oo}^q)(\mathbf{x}; \sigma, \lambda, \mathbf{u}_0)$$

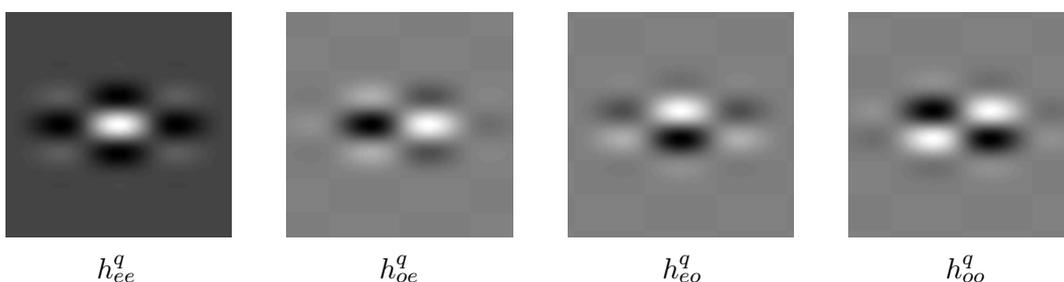
mit

$$h_{ee}^q(\mathbf{x}; \sigma, \lambda, \mathbf{u}_0) = g(\mathbf{x}; \sigma, \lambda) \cos(2\pi u_0 x) \cos(2\pi v_0 y)$$

$$h_{oe}^q(\mathbf{x}; \sigma, \lambda, \mathbf{u}_0) = g(\mathbf{x}; \sigma, \lambda) \sin(2\pi u_0 x) \cos(2\pi v_0 y)$$

$$h_{eo}^q(\mathbf{x}; \sigma, \lambda, \mathbf{u}_0) = g(\mathbf{x}; \sigma, \lambda) \cos(2\pi u_0 x) \sin(2\pi v_0 y)$$

$$h_{oo}^q(\mathbf{x}; \sigma, \lambda, \mathbf{u}_0) = g(\mathbf{x}; \sigma, \lambda) \sin(2\pi u_0 x) \sin(2\pi v_0 y)$$



Ein quaternionwertiges Gaborfilter mit den Parametern $\sigma_1 = 20$, $\sigma_2 = 10$, $c_1 = c_2 = 2$, Größe der Filtermaske: 100×100

a) lokale quaternionische Phase

Im 1D-Fall galt, daß zwischen lokaler Phase und relativer Position zum Aufpunkt der Gaborfunktion im Ortsraum ein linearer Zusammenhang besteht:

$$\phi(x) \approx x.$$

Man sagt, die Schätzungen der lokalen Phase und die räumliche Position sind äquivariant. Im 2D-Fall gilt diese Äquivarianz ebenfalls.

Sei $k^q(\mathbf{x}) = |k^q(\mathbf{x})| e^{i\phi(\mathbf{x})} e^{j\theta(\mathbf{x})} e^{k\psi(\mathbf{x})}$ die Filterantwort der Funktion $f(\mathbf{x})$ mit der Gaborfunktion $h^q(\mathbf{x})$. Dann gilt

$$\phi(\mathbf{x}) \approx x \text{ und } \theta(\mathbf{x}) \approx y \quad \text{für } (x, y) \in [-\pi, \pi] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

Die Phasenwinkel ϕ und θ zeigen also die lokale Phase in Richtung u bzw. v an, bzw. sind äquivariant bezüglich Verschiebungen in x - oder y - Richtung. Diese Phasenwinkel messen die 1D-Struktureigenschaften, bezogen auf die x - bzw. y - Achse.

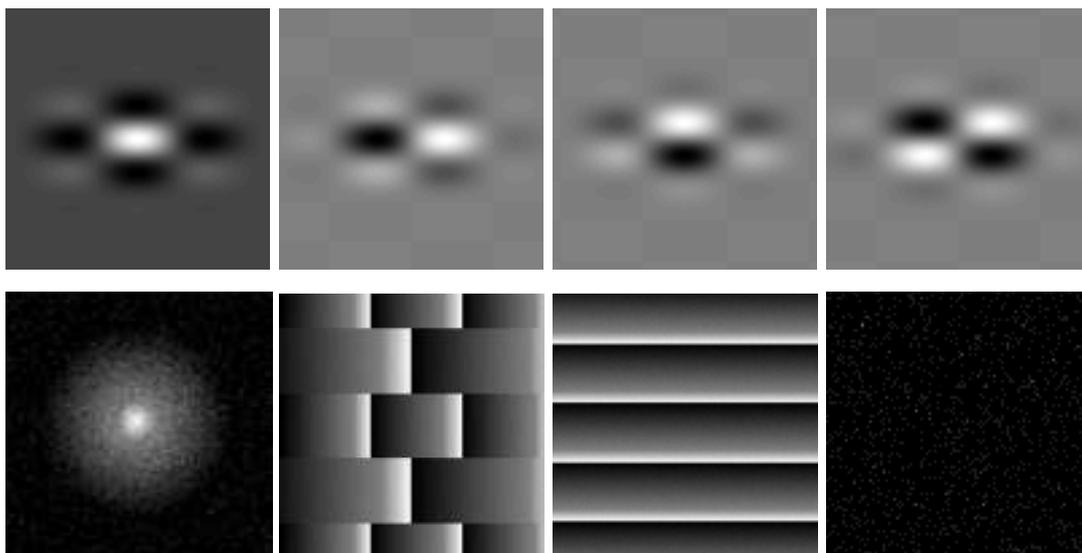
$$\phi(\mathbf{x}) = \arg(k_{oe}^q(\mathbf{x}), k_{ee}^q(\mathbf{x}))$$

$$\theta(\mathbf{x}) = \arg(k_{eo}^q(\mathbf{x}), k_{ee}^q(\mathbf{x}))$$

Der Phasenwinkel $\psi(\mathbf{x})$, $(x, y) \in [-\frac{\pi}{4}, \frac{\pi}{4}]$, bestimmt die Beziehungen dieser 1D-Konzepte zueinander:

$$\psi(\mathbf{x}) = \arg(k_{oo}^q(\mathbf{x}), k_{ee}^q(\mathbf{x})).$$

Dies geschieht in der Weise, daß eine um 45° geneigte Linie/Kante den Wert $\psi = \pm\frac{\pi}{4}$ ergibt und $\psi = 0$ gilt, wenn sich zwei orthogonal orientierte Strukturen mit gleichem Gewicht überlagern. Alle anderen Fälle führen zu $0 < |\psi| < \frac{\pi}{4}$.

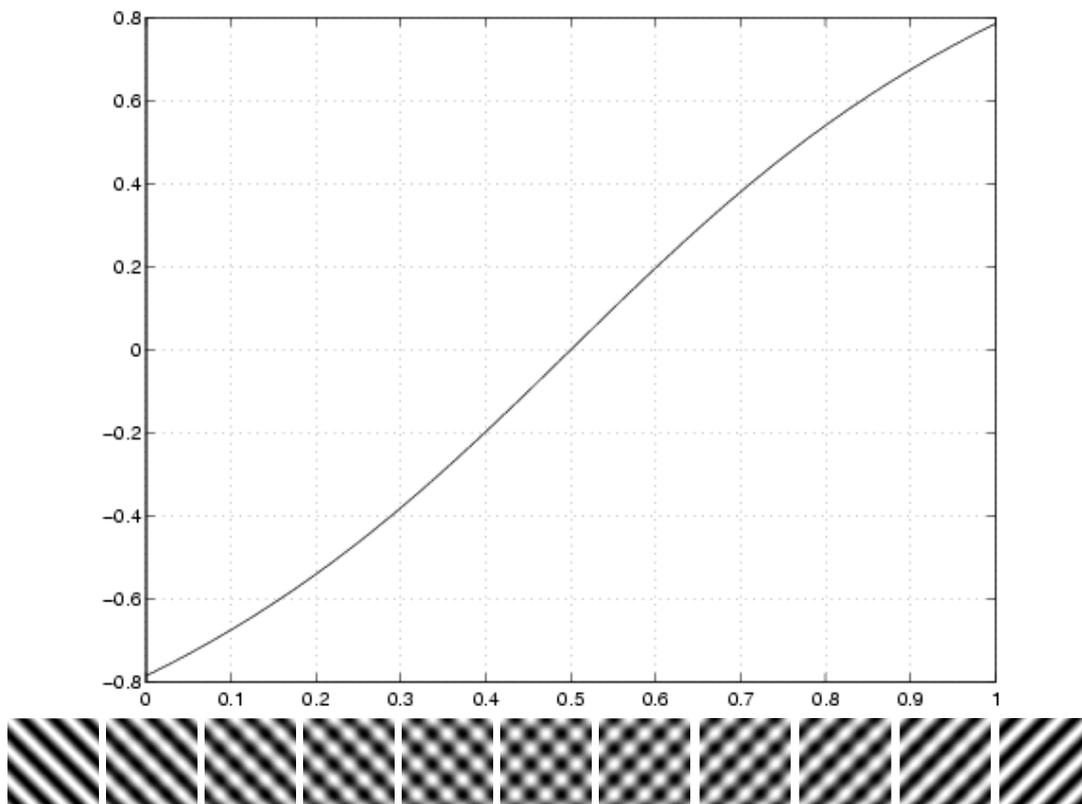


Oben: Quaternionwertiges Gaborfilter. Unten: Amplitude, Phase $\varphi(x)$, Phase $\theta(y)$, Phase $\psi(x)$.

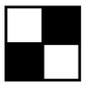
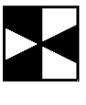
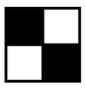
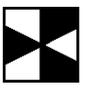
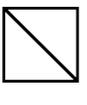
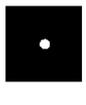
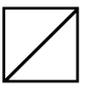
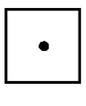
In der folgenden Abbildung wird die ψ -Phase für die Funktion

$$f(\mathbf{x}; \lambda) = (1 - \lambda) \cos(\omega_1 x + \omega_2 y) + \lambda \cos(\omega_1 x - \omega_2 y)$$

in Abhängigkeit vom Mischungsparameter λ dargestellt. Für $\lambda = 0$ und $\lambda = 1$ erhält man reine 1D-Strukturen mit $\psi = \mp \frac{\pi}{4}$. In allen anderen Fällen stellt sie Funktion intrinsische 2D-Strukturen dar, deren Phasenwinkel ψ entsprechend variiert.



In Abschnitt 3.9.2 wurde gezeigt, daß es vier Symmetrie-Grundmuster für 1D-Strukturen gibt, die von der lokalen Phase unterschieden werden können. Im 2D-Fall können die in der folgenden Abbildung dargestellten sechzehn Grundmuster durch die angegebene Phase unterschieden werden.

 $\begin{pmatrix} -\pi/2 \\ -\pi/2 \\ 0 \end{pmatrix}$	 $\begin{pmatrix} \pi/2 \\ 0 \\ 0 \end{pmatrix}$	 $\begin{pmatrix} \pi/2 \\ 0 \\ \pi/4 \end{pmatrix}$	 $\begin{pmatrix} 0 \\ 0 \\ \pi/4 \end{pmatrix}$
 $\begin{pmatrix} \pi/2 \\ -\pi/2 \\ 0 \end{pmatrix}$	 $\begin{pmatrix} -\pi/2 \\ 0 \\ 0 \end{pmatrix}$	 $\begin{pmatrix} -\pi/2 \\ 0 \\ \pi/4 \end{pmatrix}$	 $\begin{pmatrix} -\pi \\ 0 \\ \pi/4 \end{pmatrix}$
 $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$	 $\begin{pmatrix} 0 \\ \pi/2 \\ 0 \end{pmatrix}$	 $\begin{pmatrix} \pi/2 \\ 0 \\ -\pi/4 \end{pmatrix}$	 $\begin{pmatrix} -\pi \\ 0 \\ -\pi/4 \end{pmatrix}$
 $\begin{pmatrix} -\pi \\ 0 \\ 0 \end{pmatrix}$	 $\begin{pmatrix} 0 \\ -\pi/2 \\ 0 \end{pmatrix}$	 $\begin{pmatrix} -\pi/2 \\ 0 \\ -\pi/4 \end{pmatrix}$	 $\begin{pmatrix} 0 \\ 0 \\ -\pi/4 \end{pmatrix}$

Sechzehn 2D-Grundmuster und die entsprechenden Phasenwinkel $(\phi, \theta, \psi)^T$, abgeleitet am Ort des Mittelpunktes der Muster.

b) Beziehung zwischen komplexer und quaternionischer Gaborfunktion

Für $\lambda = 1$ existiert eine einfache Beziehung zwischen komplexwertigen und quaternionwertigen Gaborfunktionen in 2D.

$$\begin{aligned}
 h_e^+(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) &= g(\mathbf{x}; \sigma) \cos(\omega_{0x}x + \omega_{0y}y) \\
 &= g(\mathbf{x}; \sigma) (\cos(\omega_{0x}x) \cos(\omega_{0y}y) - \sin(\omega_{0x}x) \sin(\omega_{0y}y)) \\
 &= h_{ee}^q(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) - h_{oo}^q(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) \\
 h_o^+(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) &= g(\mathbf{x}; \sigma) \sin(\omega_{0x}x + \omega_{0y}y) \\
 &= g(\mathbf{x}; \sigma) (\cos(\omega_{0x}x) \sin(\omega_{0y}y) + \sin(\omega_{0x}x) \cos(\omega_{0y}y)) \\
 &= h_{eo}^q(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) + h_{oe}^q(\mathbf{x}; \sigma, \boldsymbol{\omega}_0)
 \end{aligned}$$

Also

$$h^+(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) = h_e^+(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) + ih_o^+(\mathbf{x}; \sigma, \boldsymbol{\omega}_0).$$

Außerdem existiert

$$h^-(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) = h_e^-(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) + ih_o^-(\mathbf{x}; \sigma, \boldsymbol{\omega}_0)$$

mit

$$\begin{aligned}
 h_e^-(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) &= g(\mathbf{x}; \sigma) \cos(\omega_{0x}x - \omega_{0y}y) \\
 &= h_{ee}^q(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) + h_{oo}^q(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) \\
 h_o^-(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) &= g(\mathbf{x}; \sigma) \sin(\omega_{0x}x - \omega_{0y}y) \\
 &= h_{oe}^q(\mathbf{x}; \sigma, \boldsymbol{\omega}_0) - h_{eo}^q(\mathbf{x}; \sigma, \boldsymbol{\omega}_0)
 \end{aligned}$$

Folglich entsprechen jedem quaternionischen Gaborfilter zwei komplexe Gaborfilter. Da sich diese Symmetrien beider Filter deren Antworten aufprägen, kann man die Antworten komplexwertiger Filter aus denen quaternionwertiger Filter ableiten.

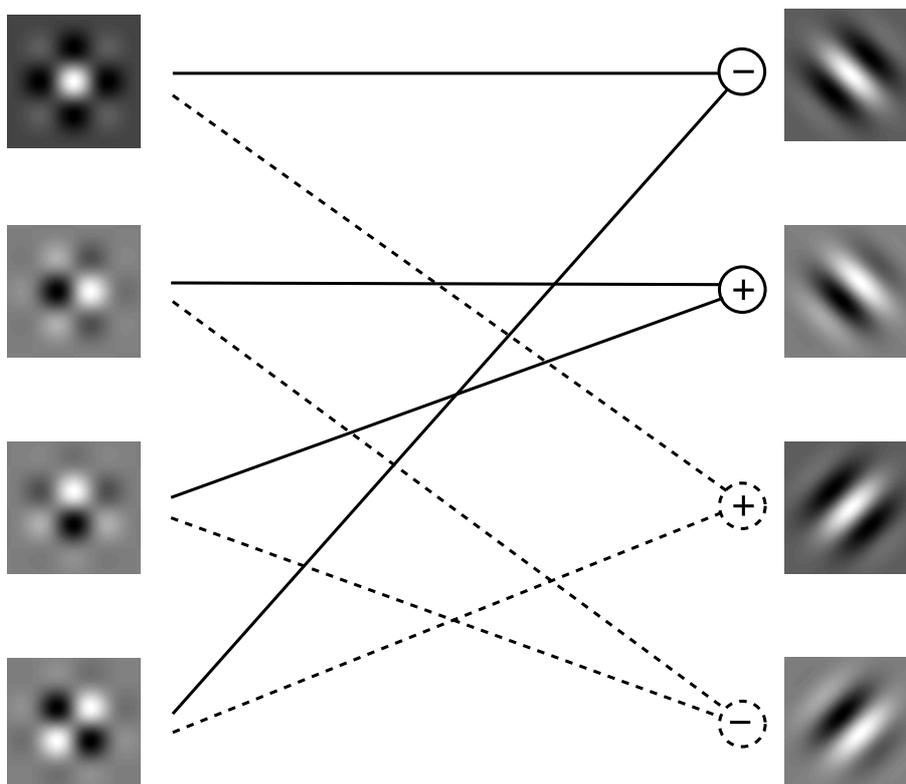
Sei

$$\begin{aligned}
 k(\mathbf{x}) &= (h * f)(\mathbf{x}) \\
 &= ((h_e + ih_o) * f)(\mathbf{x}) \\
 &= k_e(\mathbf{x}) + ik_o(\mathbf{x}) \\
 k^q(\mathbf{x}) &= (h^q * f)(\mathbf{x}) \\
 &= ((h_{ee}^q + ih_{oe}^q + jh_{eo}^q + kh_{oo}^q) * f)(\mathbf{x}) \\
 &= k_{ee}^q(\mathbf{x}) + ik_{oe}^q(\mathbf{x}) + jk_{eo}^q(\mathbf{x}) + kk_{oo}^q(\mathbf{x})
 \end{aligned}$$

Dann gilt

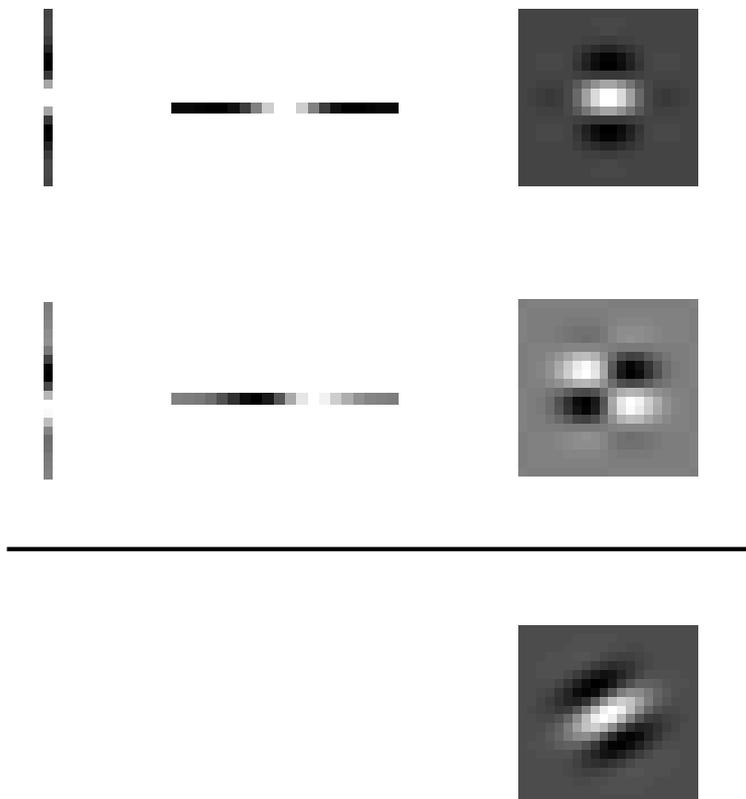
$$\begin{aligned}
 k^+(\mathbf{x}) &= (k_{ee}^q - k_{oo}^q) + i(k_{eo}^q + k_{oe}^q) \\
 k^-(\mathbf{x}) &= (k_{ee}^q + k_{oo}^q) + i(k_{oe}^q - k_{eo}^q).
 \end{aligned}$$

Die folgende Abbildung verdeutlicht dieses Bildungsschema.

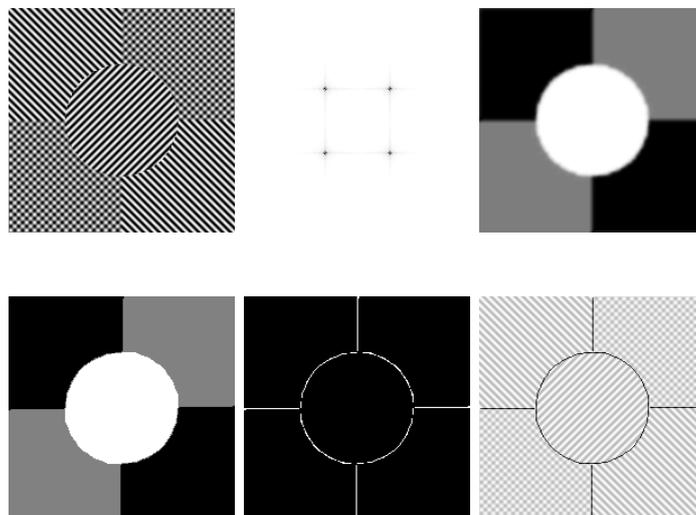


c) *Separable Implementierung*

Für $\lambda = 1$ sind quaternionwertige Gaborfilter separabel. Also lassen sich wegen b) auch komplexe Gaborfilter als Kombination separabler Komponenten quaternionischer Gaborfilter implementieren.

d) *Anwendungen quaternionwertiger Gaborfilter*

Da diese Quadratur- und Hilbertquadrupel bilden, eignen sie sich für die lokal spektrale Beschreibung intrinsisch zweidimensionaler Strukturen, z.B. für Texturanalyse.



Literaturverzeichnis

- [Ballard and Brown, 1982] D. Ballard and C. Brown: *Computer Vision*. Prentice-Hall, Englewood-Cliffs, NJ, 1982
- [Besl and Jain, 1986] P.J. Besl and R.C. Jain: Invariant surface characteristics for three-dimensional object recognition in range images. *Computer Vision, Graphics, and Image Processing*, 33:33-80, 1986
- [Bracewell, 1984] R.N. Bracewell: *The Fourier transform and its applications*. McGraw Hill, 1984
- [Bülow and Sommer, 1997] T. Bülow and G. Sommer: Multi-Dimensional Signal Processing Using an Algebraically Extended Signal Representation. In *Proceedings of the AFPAC Workshop, Kiel*, pages 148-163, 1997
- [Burt and Adelson, 1983] P.J. Burt and E.J. Adelson: The Laplacian Pyramid as a Compact Image Code. *IEEE Trans. Com.*, 32(4):532-540, 1983
- [Florack *et al.*, 1994] L.M.J. Florack, B.M. ter Haar Romeny, J.J. Koenderink, and M.A. Viergever: General intensity transformations and differential invariants. *Journal of Mathematical Imaging and Vision*, 4:171-187, 1994
- [Gabor, 1946] D. Gabor: Theory of Communication. *Journ. Inst. Electr. Eng.*, 93:429-457, 1946
- [Hubel and Wiesel, 1959] D.H. Hubel and T.N. Wiesel: Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology*, 148:574-591, 1959
- [Hubel and Wiesel, 1962] D.H. Hubel and T.N. Wiesel: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160:106-154, 1962
- [Hubel, 1989] D.H. Hubel: *Auge und Gehirn: Neurobiologie des Sehens*. Spektrum-der-Wissenschaft Verlagsgesellschaft, Heidelberg, 1989
- [Koenderink and van Doorn, 1987] J.J. Koenderink and A.J. van Doorn: Representation of Local Geometry in the Visual System. *Biol. Cybern.*, 55:367-375, 1987
- [Mecklenbräuer, 1987] W. Mecklenbräuer: A tutorial on non-parametric bilinear time-frequency signal representations. In: J.L. Lacoume, T.S. Duranne and R. Stora, eds., *Signal Processing*, Vol. II, 277-336, Elsevier Science Publ. B.V., Amsterdam, 1987
- [Oppenheim und Schafer, 1999] A.V. Oppenheim, R.W. Schafer: *Zeitdiskrete Signalverarbeitung*. Oldenburg, 1999

- [Posthoff und Woschni, 1984] C. Posthoff, E.-G. Woschni: *Funktionaltransformationen der Informationstechnik*. Akademie-Verlag, Berlin, 1984
- [Preuß, 2002] W. Preuß: *Funktionaltransformationen*. Fachbuchverlag Leipzig, 2002
- [Sommer and Meinel, 1987] G. Sommer and G. Meinel: The design of optimal Gaussian DOLP edge detectors. In L.P. Yaroslavsky, A. Posenfeld, and W. Wilhelmi, editors, *Proc. 2nd Int. Conf. CAIP'97*, volume 40, pages 82-89, Wismar, 1987. Akademie-Verlag, Berlin
- [ter Haar Romeny *et al.*, 1993] B.M. ter Haar Romeny, L.M.J. Florack, A.H. Salden, and M.A. Viergever: Higher order differential structure of images. In *Proc. 13th Int. Conf. on Information Processing in Machine*, Lecture Notes In Computer Science. Springer Verlag, Berlin *et al.*, 1993