

Merkmale von Bildregionen, Einführung in Spektraltechniken

Industrielle Bildverarbeitung, Vorlesung No. 10¹

M. O. Franz

12.12.2007

¹ falls nicht anders vermerkt, sind die Abbildungen entnommen aus Burger & Burge, 2005. 

Übersicht

- 1 Geometrische Eigenschaften binärer Bildregionen
- 2 Statistische Eigenschaften binärer Bildregionen
- 3 Fouriertransformation

Übersicht

- 1 Geometrische Eigenschaften binärer Bildregionen
- 2 Statistische Eigenschaften binärer Bildregionen
- 3 Fouriertransformation

Formmerkmale

Merkmal (Feature) einer Region: Numerische oder kategorielle Kenngröße, die aus Bildpunkten (Position und Werte) berechnet wird. Mehrere Merkmale werden in einem **Merkmalsvektor** zusammengefaßt.

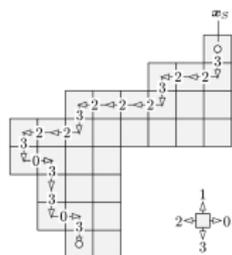
Merkmale sollten

- einfach zu berechnen,
- invariant gegenüber irrelevanten Veränderungen (z.B. Rotationen, Verschiebungen, Skalierungen) sein.

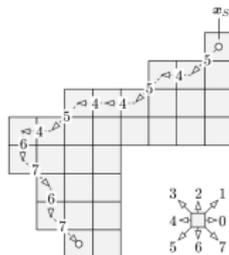
Eine Region \mathcal{R} in einem Binärbild kann als zweidimensionale Verteilung von Vordergrundpunkten $x_i = (u_i, v_i)$ beschrieben werden:

$$\mathcal{R} = \{x_1, x_2, \dots, x_N\} = \{(u_1, v_1), (u_2, v_2), \dots, (u_N, v_N)\}$$

Geometrische Eigenschaften - Umfang



4-Chain Code
3223222322303303...111
length = 28



8-Chain Code
54544546767...222
length = 18 + 5√2 ≈ 25

Umfang hängt vom Nachbarschaftstyp ab:

- bei **4er-Nachbarschaften** ist der Umfang i.A. größer als der reale.
- eine bessere Näherung wird durch die **8er-Nachbarschaft** erreicht.

$$U(\mathcal{R}) = \sum_{i=0}^{M-1} l(c_i) \quad \text{mit} \quad l(c) = \begin{cases} 1 & \text{Horizontal- und Vertikalsegmente} \\ \sqrt{2} & \text{Diagonalsegmente} \end{cases}$$

Umfang wird damit immer noch leicht überschätzt, daher Korrektur um Faktor 0.95:

$$U(\mathcal{R}) \approx 0.95 \sum_{i=0}^{M-1} l(c_i)$$

Geometrische Eigenschaften - Fläche

Die Fläche berechnet sich einfach als Anzahl N der Bildpunkte in \mathcal{R} :

$$A(\mathcal{R}) = N = |\mathcal{R}|,$$

kann auch näherungsweise bei zusammenhängenden Regionen aus der geschlossenen Kontur $\{c_0, c_1, \dots, c_{M-1}\} = \{(x_0, y_0), (x_1, y_1), \dots\}$ (**Gaußsche Flächenformel für Polygone**) berechnet werden:

$$A(\mathcal{R}) = \frac{1}{2} \left| \sum_{i=0}^{M-1} x_i y_{(i+1) \bmod M} - x_{(i+1) \bmod M} y_i \right|$$

erlaubt Berechnung der Fläche aus *Chain Codes*.

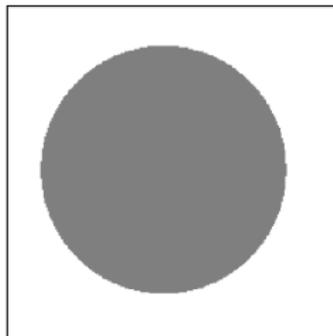
Umfang und Fläche sind invariant gegenüber Verschiebungen und Drehungen, jedoch variant gegenüber Skalierungen.

Geometrische Eigenschaften - Kompaktheit und Rundheit

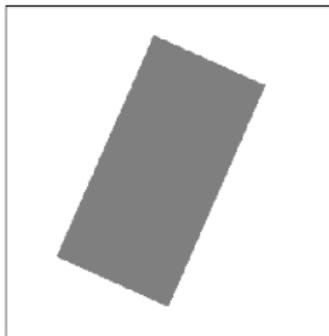
Die **Kompaktheit** einer Region beschreibt die Beziehung von Umfang und Fläche: $k = \frac{A(\mathcal{R})}{U(\mathcal{R})^2}$. Für Kreis ergibt sich $k = \frac{1}{4\pi} \Rightarrow$ Normierung auf Kreis ergibt **Rundheit**:

$$C = 4\pi \frac{A(\mathcal{R})}{U(\mathcal{R})^2}.$$

Invariant gegenüber Verschiebung, Drehung und Skalierung.



0.904
(1.001)



0.607
(0.672)



0.078
(0.086)

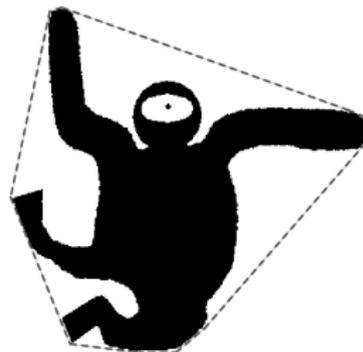
Geometrische Eigenschaften - Bounding Box und konvexe Hülle

Bounding Box bezeichnet das minimale achsenparallele Rechteck, das alle Punkte einer Region einschließt.

Konvexe Hülle: Kleinstes Polygon, das eine Region umfaßt.



(a)



(b)

Konvexität: Verhältnis der Umfänge von konvexer Hülle und Region

Dichte: Flächenverhältnis von Hülle und Region

Übersicht

- 1 Geometrische Eigenschaften binärer Bildregionen
- 2 Statistische Eigenschaften binärer Bildregionen**
- 3 Fouriertransformation

Statistische Eigenschaften

- Schwerpunkt (\Rightarrow Position der Region) :

$$\bar{x} = \frac{1}{|\mathcal{R}|} \sum_{(u_i, v_i) \in \mathcal{R}} u_i \quad \text{und} \quad \bar{y} = \frac{1}{|\mathcal{R}|} \sum_{(u_i, v_i) \in \mathcal{R}} v_i$$

- Momente:

$$m_{pq} = \sum_{(u,v) \in \mathcal{R}} I(u,v) u^p v^q \quad \text{bzw.} \quad m_{pq} = \sum_{(u,v) \in \mathcal{R}_{\text{binär}}} u^p v^q$$

z.B.

- $A = |\mathcal{R}| = \sum 1 = \sum u^0 v^0 = m_{00}$
- $\bar{x} = \frac{1}{|\mathcal{R}|} \sum u^1 v^0 = \frac{m_{10}}{m_{00}}$
- $\bar{y} = \frac{1}{|\mathcal{R}|} \sum u^0 v^1 = \frac{m_{01}}{m_{00}}$

Statistische Eigenschaften - zentrale Momente

Momente sind i.A. nicht verschiebungsinvariant \Rightarrow Verschiebung des Ursprungs an den Schwerpunkt: **zentrale Momente**

$$\mu_{pq} = \sum_{(u,v) \in \mathcal{R}} I(u,v)(u-\bar{x})^p(v-\bar{y})^q \quad \text{bzw.} \quad \mu_{pq} = \sum_{(u,v) \in \mathcal{R}_{\text{binär}}} (u-\bar{x})^p(v-\bar{y})^q$$

Die zentralen Momente hängen von der Größe der Region ab. Eine Skalierung um den Faktor s ergibt

$$\mu_{pq}(s \cdot \mathcal{R}) = s^2 \sum_{(u,v) \in \mathcal{R}} I(u,v) s^p (u-\bar{x})^p s^q (v-\bar{y})^q = s^{p+q+2} \mu_{pq}(\mathcal{R})$$

Größeninvariante Momente erhält man Normalisierung mit der Fläche

$A = m_{00} = \mu_{00}$:

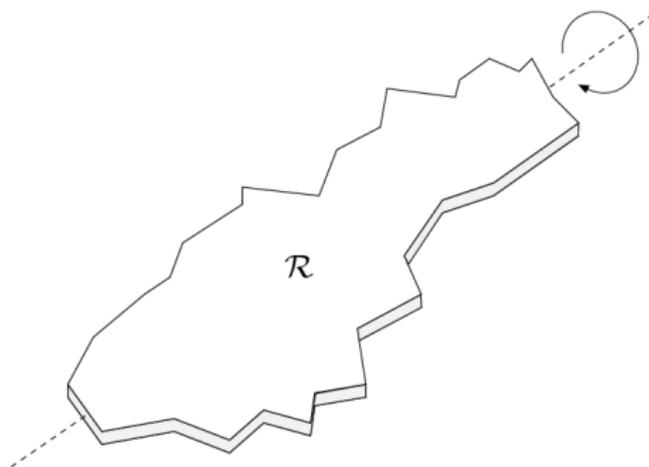
$$\bar{\mu}_{pq} = \frac{\mu_{pq}}{\mu_{00}^{(p+q+2)/2}}$$

Statistische Eigenschaften - Orientierung

Orientierung: Richtung der Hauptachse (Achse mit minimalem Drehmoment)

$$\theta = \frac{1}{2} \tan^{-1} \frac{2\mu_{11}}{\mu_{20} - \mu_{02}}$$

(Richtung des Eigenvektors zum größten Eigenwert des Trägheitstensors)

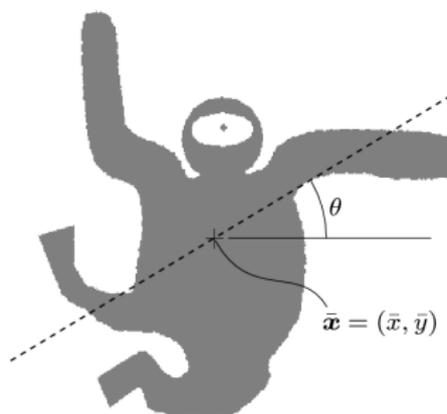


Statistische Eigenschaften - Exzentrizität

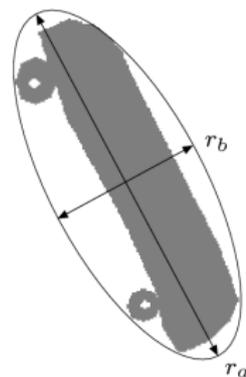
Exzentrizität: Maß für das Größenverhältnis der beiden Hauptachsen

$$E = \frac{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}{(\mu_{20} + \mu_{02})^2} \in [0, 1]$$

Bei runden Objekten ist E nahe an 0, bei länglichen Objekten nahe an 1.



(a)



(b)

Statistische Eigenschaften - Hu-Momente

Normalisierte zentrale Momente sind zwar skalierungs- und verschiebungsinvariant, aber i.A. nicht rotationsinvariant.

Hu-Momente (zusätzlich rotationsinvariant):

$$H_1 = \bar{\mu}_{20} + \bar{\mu}_{02}$$

$$H_2 = (\bar{\mu}_{20} - \bar{\mu}_{02})^2 + 4\bar{\mu}_{11}^2$$

$$H_3 = (\bar{\mu}_{30} - 3\bar{\mu}_{12})^2 + (3\bar{\mu}_{21} - \bar{\mu}_{03})^2$$

$$H_4 = (\bar{\mu}_{30} + \bar{\mu}_{12})^2 + (\bar{\mu}_{21} + \bar{\mu}_{03})^2$$

$$H_5 = (\bar{\mu}_{30} - 3\bar{\mu}_{12}) \cdot (\bar{\mu}_{30} + \bar{\mu}_{12}) \cdot [(\bar{\mu}_{30} + \bar{\mu}_{12})^2 - 3(\bar{\mu}_{21} + \bar{\mu}_{03})^2] + \\ (3\bar{\mu}_{21} - \bar{\mu}_{03}) \cdot (\bar{\mu}_{21} + \bar{\mu}_{03}) \cdot [3(\bar{\mu}_{30} + \bar{\mu}_{12})^2 - (\bar{\mu}_{21} + \bar{\mu}_{03})^2]$$

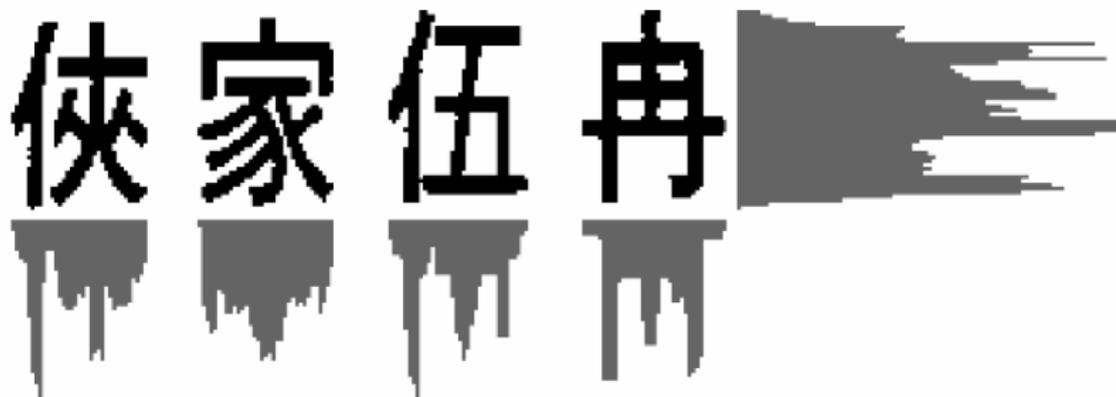
$$H_6 = (\bar{\mu}_{20} - \bar{\mu}_{02}) \cdot [(\bar{\mu}_{30} + \bar{\mu}_{12})^2 - (\bar{\mu}_{21} + \bar{\mu}_{03})^2] + \\ 4\bar{\mu}_{11} \cdot (\bar{\mu}_{30} + \bar{\mu}_{12}) \cdot (\bar{\mu}_{21} + \bar{\mu}_{03})$$

$$H_7 = (3\bar{\mu}_{21} - \bar{\mu}_{03}) \cdot (\bar{\mu}_{30} + \bar{\mu}_{12}) \cdot [(\bar{\mu}_{30} + \bar{\mu}_{12})^2 - 3(\bar{\mu}_{21} + \bar{\mu}_{03})^2] + \\ (3\bar{\mu}_{12} - \bar{\mu}_{30}) \cdot (\bar{\mu}_{21} + \bar{\mu}_{03}) \cdot [3(\bar{\mu}_{30} + \bar{\mu}_{12})^2 - (\bar{\mu}_{21} + \bar{\mu}_{03})^2]$$

meist logarithmiert wegen großen Wertebereichs.

Projektionen

$$P_{\text{hor}}(v_0) = \sum_{u=0}^{M-1} I(u, v_0) \quad \text{und} \quad P_{\text{vert}}(u_0) = \sum_{v=0}^{N-1} I(u_0, v)$$

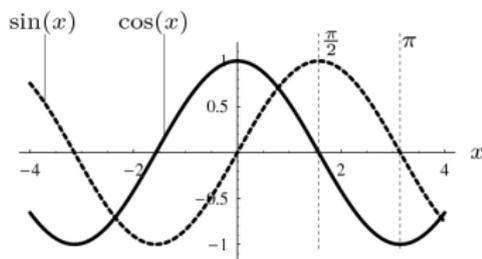


Nützlich zur Analyse von strukturierten Bildern, z.B. zur Detektion von Zeilen und einzelnen Buchstaben, oder schnellen Schätzung der Orientierung.

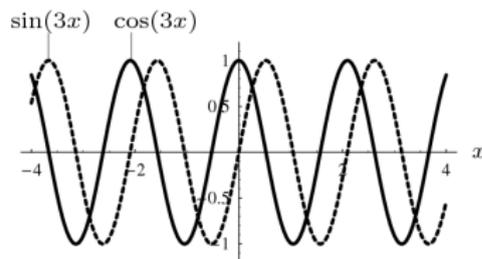
Übersicht

- 1 Geometrische Eigenschaften binärer Bildregionen
- 2 Statistische Eigenschaften binärer Bildregionen
- 3 Fouriertransformation**

Sinus- und Cosinusfunktionen



(a)



(b)

Beide Funktionen sind periodisch mit **Periodenlänge** $T = 2\pi$:

$$\cos(x) = \cos(x+2\pi) = \cos(x+k2\pi) \quad \text{und} \quad \sin(x) = \sin(x+2\pi) = \sin(x+k2\pi)$$

Kreisfrequenz ω : Anzahl der Perioden innerhalb der Strecke 2π

$$\omega = \frac{2\pi}{T}$$

Frequenz f : Anzahl der Perioden innerhalb der Strecke 1

$$f = \frac{1}{T} = \frac{\omega}{2\pi}$$

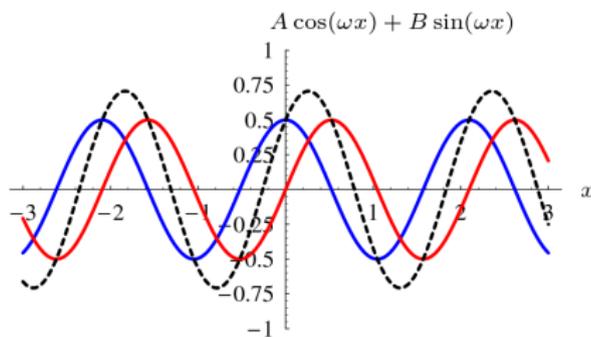
Phase

Verschiebung der Cosinusfunktion um φ entlang der x-Achse $\cos(x) \rightarrow \cos(x - \varphi)$ verändert die **Phase** der Schwingung um den **Phasenwinkel** φ , z.B. $\sin(\omega x) = \cos(\omega x - \frac{\pi}{2})$.

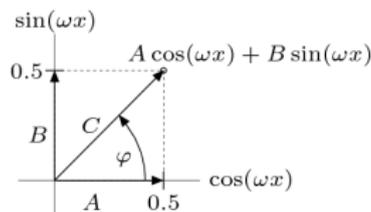
Durch Linearkombination von Sinus- und Cosinusfunktion mit **identischer** Frequenz entsteht wieder ein Sinusoid derselben Frequenz:

$$a \cos(\omega x) + b \sin(\omega x) = r \cos(\omega x - \varphi)$$

mit $r = \sqrt{a^2 + b^2}$ und $\varphi = \tan^{-1} \frac{b}{a}$.

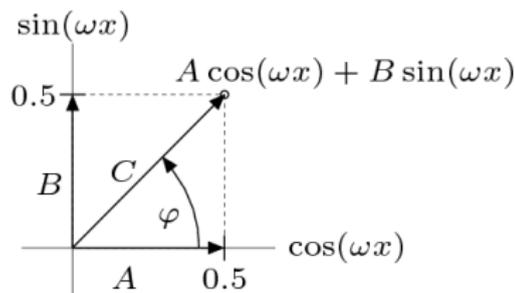


(a)



(b)

Eulersche Notation



Darstellung der Sinusoide als zweidimensionale Vektoren in der Ebene der **komplexen Zahlen**:

$$z = a + ib \in \mathbb{C}$$

mit $i^2 = -1$. Polardarstellung (**Eulersche Notation**):

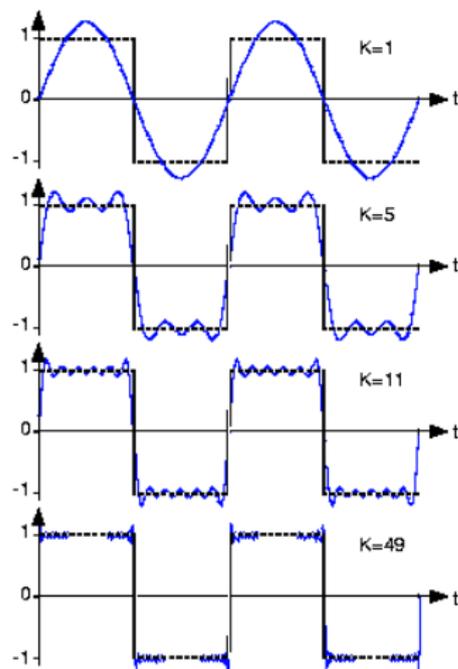
$$z = r(\cos \theta + i \sin \theta) = r e^{i\theta}$$

In der Eulerschen Notation wird die **Amplitude** des Sinusoids durch den Betrag r der komplexen Zahl repräsentiert, eine Phasenverschiebung um φ durch den Polarwinkel θ :

$$r \cos(\omega x - \varphi) = r e^{i(\omega x - \varphi)} = a \cos(\omega x) + i b \sin(\omega x)$$

Multiplikation mit einer reellen Zahl ändert nur die Amplitude,
Multiplikation mit einer komplexen Zahl mit Betrag 1 nur die Phase.

Fourierreihen



- Beliebige Sinusoide können als gewichtete Summe einer Sinus- und einer Cosinusfunktion geschrieben werden.
- Jede **periodische Funktion** $g(x)$ mit Grundfrequenz ω_0 kann als (unendliche) Summe von **harmonischen** Sinusoiden dargestellt werden:

$$g(x) = \sum_{k=0}^{\infty} a_k \cos k\omega x + b_k \sin k\omega x$$

a_k, b_k : Fourierkoeffizienten

Fourieranalyse: Berechnung der Fourierkoeffizienten aus $g(x)$.

Fourierintegral

Erweiterung auf nichtperiodische Funktionen: statt der Grundfrequenz und ihrer Harmonischen werden unendlich viele, dicht aneinander liegende Frequenzen gebraucht.

Fourierintegral:

$$g(x) = \int_0^{\infty} a(\omega) \cos(\omega x) + b(\omega) \sin(\omega x) d\omega$$

Die kontinuierlichen Koeffizientenfunktionen ergeben sich aus

$$a(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} g(x) \cos(\omega x) dx$$

$$b(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} g(x) \sin(\omega x) dx$$

beschreiben die Verteilung der im Ausgangssignal enthaltenen Frequenzkomponenten (**Spektrum**).

Fourierspektrum

Fouriertransformation behandelt im Gegensatz zum Fourierintegral sowohl Ausgangsfunktion als auch Spektrum als **komplexwertige** Funktionen.

Fourierspektrum:

$$\begin{aligned}G(\omega) &= \sqrt{\frac{\pi}{2}} (a(\omega) - i b(\omega)) \\&= \sqrt{\frac{\pi}{2}} \left[\frac{1}{\pi} \int_{-\infty}^{\infty} g(x) \cos(\omega x) dx - i \cdot \frac{1}{\pi} \int_{-\infty}^{\infty} g(x) \sin(\omega x) dx \right] \\&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(x) [\cos(\omega x) - i \sin(\omega x)] dx\end{aligned}$$

Eulersche Schreibweise:

$$G(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(x) e^{-i\omega x} dx$$

Fouriertransformation

Die Fouriertransformation liefert eine Darstellung der Funktion im **Spektralraum**:

$$G(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(x) e^{-i\omega x} dx$$

Aus dem Spektrum kann die ursprüngliche Funktion im **Ortsraum** wieder vollständig rekonstruiert werden (inverse Fouriertransformation):

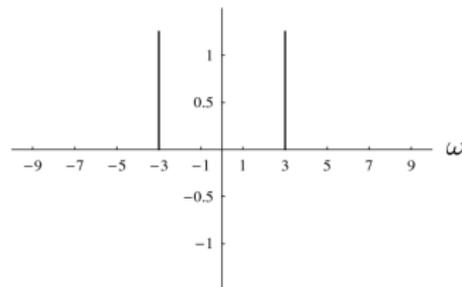
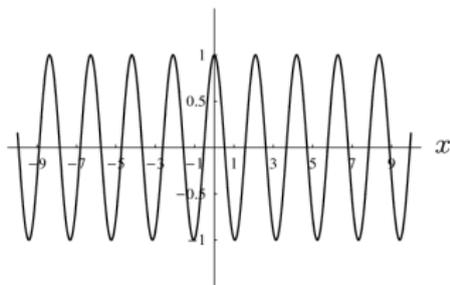
$$g(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} G(\omega) e^{i\omega x} d\omega$$

$g(x)$ und $g(\omega)$ bilden ein **Transformationspaar**:

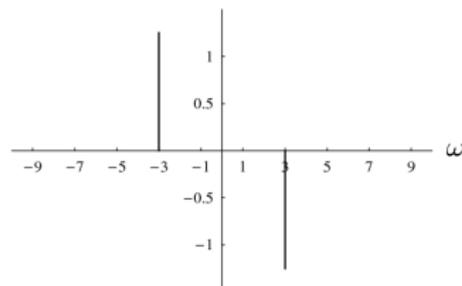
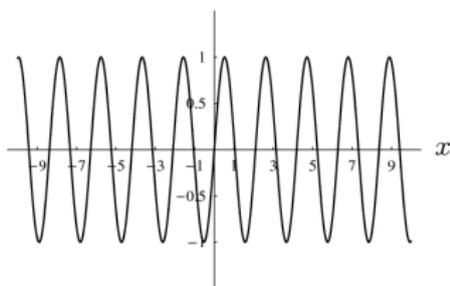
$$g(x) \circ \bullet G(\omega)$$

Achtung: Auch wenn $g(x)$ nur reellwertig ist, ist die Fouriertransformierte i.A. komplexwertig.

Fourier-Transformationspaare: Sinusoide (1)

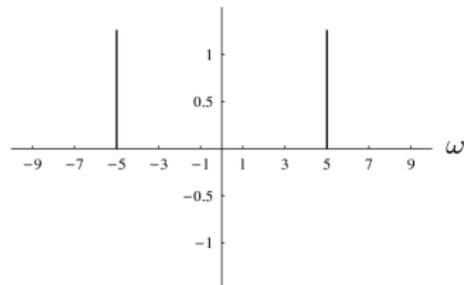
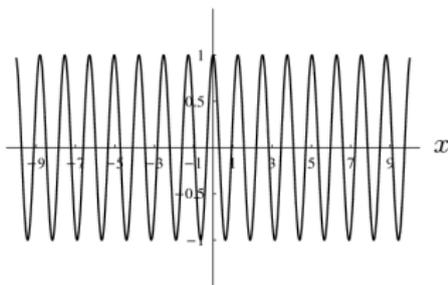


(a) Kosinus ($\omega_0=3$): $g(x) = \cos(3x)$ $\circ \bullet G(\omega) = \sqrt{\frac{\pi}{2}} \cdot (\delta(\omega-3) + \delta(\omega+3))$

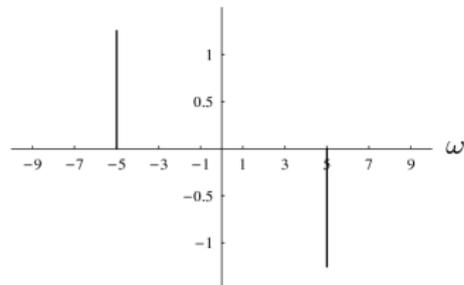
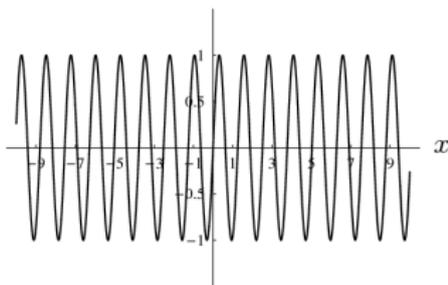


(b) Sinus ($\omega_0=3$): $g(x) = \sin(3x)$ $\circ \bullet G(\omega) = i\sqrt{\frac{\pi}{2}} \cdot (\delta(\omega-3) - \delta(\omega+3))$

Fourier-Transformationspaare: Sinusoide (2)

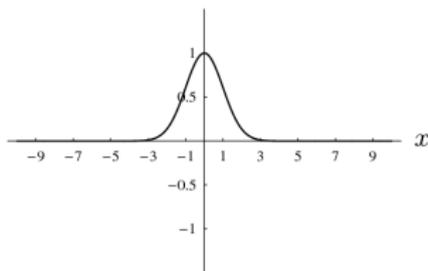


(c) Kosinus ($\omega_0 = 5$): $g(x) = \cos(5x)$ $\circ \bullet G(\omega) = \sqrt{\frac{\pi}{2}} \cdot (\delta(\omega - 5) + \delta(\omega + 5))$

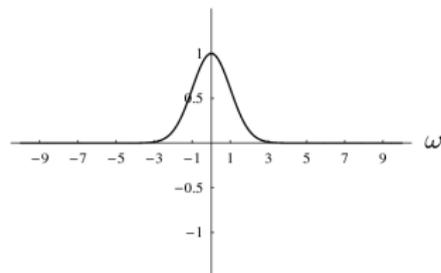


(d) Sinus ($\omega_0 = 5$): $g(x) = \sin(5x)$ $\circ \bullet G(\omega) = i\sqrt{\frac{\pi}{2}} \cdot (\delta(\omega - 5) - \delta(\omega + 5))$

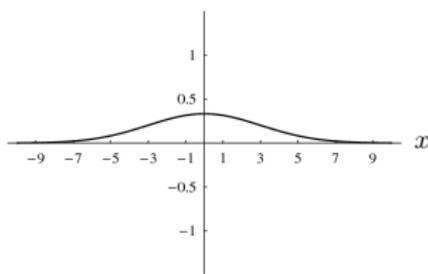
Fourier-Transformationspaare: Gaußfunktion



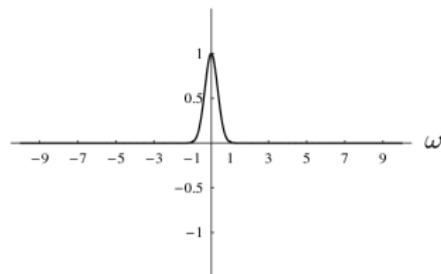
(a) Gauß ($\sigma=1$): $g(x) = e^{-\frac{x^2}{2}}$ $\circ \bullet$



$G(\omega) = e^{-\frac{\omega^2}{2}}$

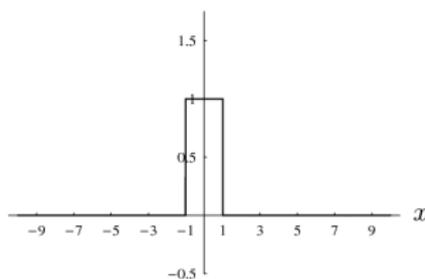


(b) Gauß ($\sigma=3$): $g(x) = \frac{1}{3} \cdot e^{-\frac{x^2}{2 \cdot 9}}$ $\circ \bullet$

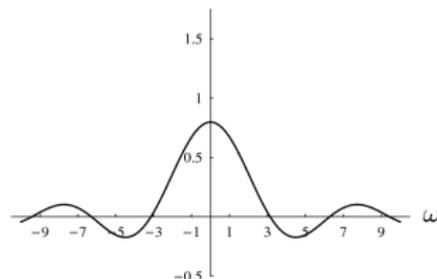


$G(\omega) = e^{-\frac{9\omega^2}{2}}$

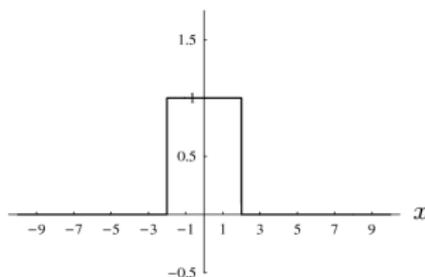
Fourier-Transformationspaare: Rechteckpuls



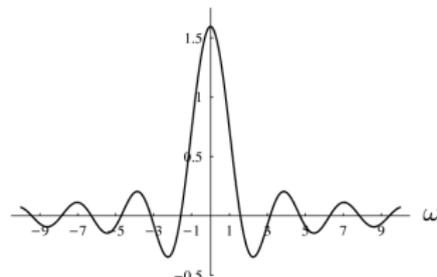
(c) Rechteckpuls ($b=1$): $g(x) = \Pi_1(x)$ $\circ \bullet$



$$G(\omega) = \frac{2 \sin(\omega)}{\sqrt{2\pi}\omega}$$



(d) Rechteckpuls ($b=2$): $g(x) = \Pi_2(x)$ $\circ \bullet$



$$G(\omega) = \frac{4 \sin(2\omega)}{\sqrt{2\pi}\omega}$$

Fourier-Transformationspaare

| <i>Funktion</i> | <i>Transformationspaar</i> $g(x) \circ \bullet G(\omega)$ | <i>Abb.</i> |
|---|--|-------------|
| Kosinusfunktion mit Frequenz ω_0 | $g(x) = \cos(\omega_0 x)$ $G(\omega) = \sqrt{\frac{\pi}{2}} \cdot (\delta(\omega - \omega_0) + \delta(\omega + \omega_0))$ | 13.3 (a,c) |
| Sinusfunktion mit Frequenz ω_0 | $g(x) = \sin(\omega_0 x)$ $G(\omega) = i\sqrt{\frac{\pi}{2}} \cdot (\delta(\omega - \omega_0) - \delta(\omega + \omega_0))$ | 13.3 (b,d) |
| Gauß-Funktion der Breite σ | $g(x) = \frac{1}{\sigma} \cdot e^{-\frac{x^2}{2\sigma^2}}$ $G(\omega) = e^{-\frac{\sigma^2 \omega^2}{2}}$ | 13.4 (a,b) |
| Rechteckpuls der Breite $2b$ | $g(x) = \Pi_b(x) = \begin{cases} 1 & x \leq b \\ 0 & \text{else} \end{cases}$ $G(\omega) = \frac{2b \sin(b\omega)}{\sqrt{2\pi\omega}}$ | 13.4 (c,d) |

Eigenschaften der Fouriertransformation (1)

Symmetrie

- Fouriertransformierte einer reellen Funktion ist symmetrisch zum Ursprung (**hermitesche Funktion**, $(a + ib)^* = a - ib$):

$$G(\omega) = G^*(-\omega)$$

- Eine reelle achsensymmetrische Funktion hat eine reelle Fouriertransformation.
- Eine reelle punktsymmetrische Funktion hat eine rein imaginäre Fouriertransformation.

Linearität



$$a \cdot g(x) \circlearrowright a \cdot G(\omega)$$



$$g_1(x) + g_2(x) \circlearrowright G_1(\omega) + G_2(\omega)$$

Eigenschaften der Fouriertransformation (2)

Ähnlichkeit

- Stauchung um Faktor s : $g(x) \rightarrow g(sx)$

$$g(sx) \circlearrowright \frac{1}{|s|} G(\omega)$$

Verschiebungseigenschaft

- Verschiebung um Distanz d entlang: $g(x) \rightarrow g(x - d)$

$$g(x - d) \circlearrowright e^{-i\omega d} G(\omega) \quad \text{Betrag bleibt unverändert.}$$

Faltungseigenschaft

-

$$g(x) * h(x) \circlearrowright G(\omega) \cdot H(\omega)$$

-

$$g(x) \cdot h(x) \circlearrowright G(\omega) * H(\omega)$$