# Computer Vision

Lingjie Zhang

2024 SS

**This course is based on the lecture EI70110 of Technical University of Munich**

## Contents

# 1 Wissenswertes über Bilder

## 1.1 Darstellung von Bildern

### Von Farbbild zum Intensitätsbild

- Farbbilder bestehen aus mehreren Kanälen
- In diesem Kurs ausschließlich Graustufenbilder

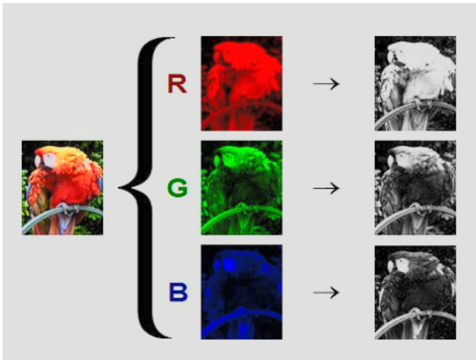



Figure 1.1: RBG image

### Kontinuierliche und diskrete Darstellung

- **Kontinuierliche** Darstellung als Funktion zweier Veränderlicher (zum Herleiten von Algorithmen) $I : \mathbb{R}^2 \supset \Omega \to \mathbb{R}, (x, y) \mapsto I(x, y)$
- Häufige Annahmen:
  - $I$ differenzierbar
  - $\Omega$ einfach zusammenhängend und beschränkt
- Diskrete Darstellung als Matrix $I \in \mathbb{R}^{m \times n}$

  Eintrag $I_{k,l}$ entspricht dem Intensitätswert
- Skalierung typischerweise zwischen [0, 255] oder [0, 1]

| |
|---|
| VGA: 480× 640 Pixel (ca. 0.3 Megapixel) |
| HD: 720× 1280 Pixel (ca. 1.0 Megapixel) |
| FHD: 1080× 1920 Pixel (ca. 2.1 Megapixel) |

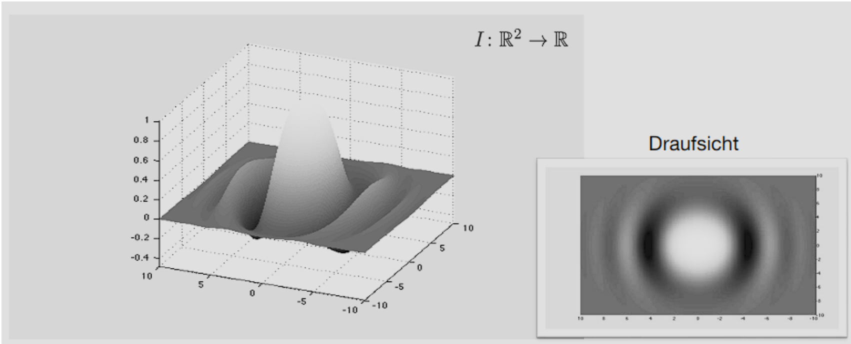



Figure 1.2: Graph einer Funktion

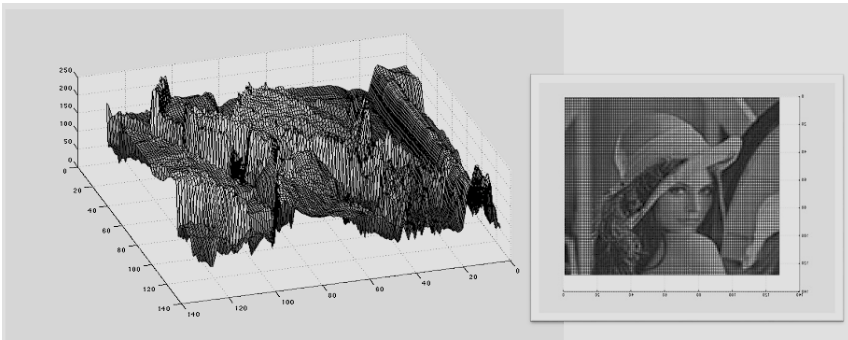


Figure 1.3: Graph eines Fotos

### Diskretes Abtasten

- Abtasten eines eindimensionalen Signals

$$S\{f(x)\} = (..., f(x-1), f(x), f(x+1), ...)$$

- Abtasten eines Bildes

$$S\{I(x,y)\} = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \\ \cdots & I(x-1,y-1) & I(x-1,y) & I(x-1,y+1) & \cdots \\ \cdots & I(x,y-1) & I(x,y) & I(x,y+1) & \cdots \\ \cdots & I(x+1,y-1) & I(x+1,y) & I(x+1,y+1) & \cdots \\ & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

### Diskrete Darstellung/Matrixdarstellung

- Annahme: Ursprung links oben
- Matrixeintrag ist $I_{k,l} = S\{I(0,0)\}_{kl}$

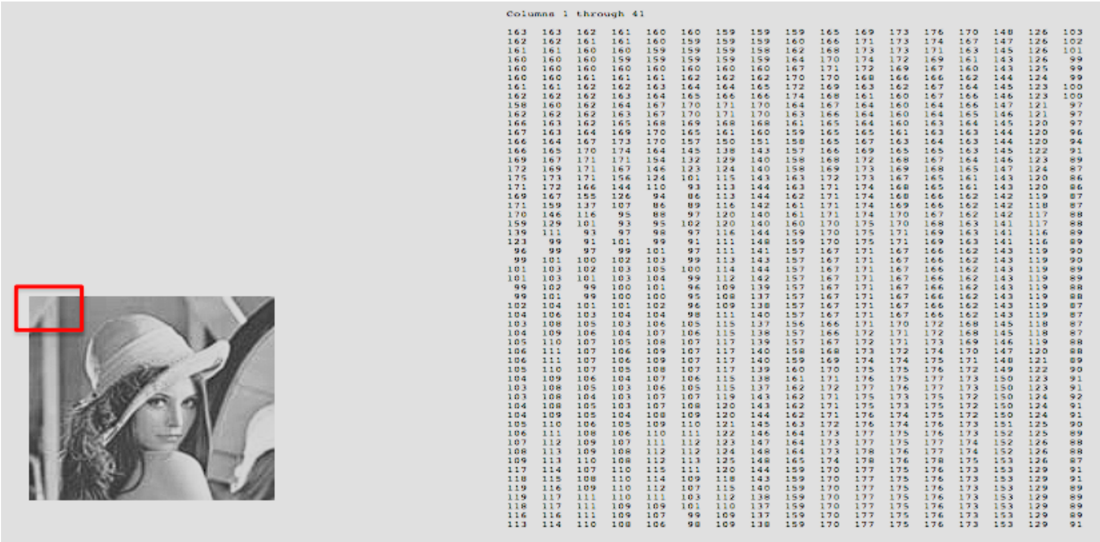


Figure 1.4: Diskrete Darstellung/Matrixdarstellung

**Zusammenfassung**

- Bilder in Grautönen
- Bilder als Matrizen
- Bilder als glatte Funktionen

## 1.2 Bildgradient

### 1.2.1 Der Gradient eines Bildes

Figure 1.5: Der Gradient eines Bildes

Kanten sind starke lokale Änderungen der Intensität. Lokale Änderungen werden durch den Gradienten beschrieben.

$$\nabla I(x,y) = \begin{bmatrix} \frac{\partial}{\partial x} I(x,y) \\ \frac{\partial}{\partial y} I(x,y) \end{bmatrix}$$

**Wie schätzt man den Gradienten?** Gegeben ist das Bild in diskreter Form $I \in \mathbb{R}^{m \times n}$.

$$\frac{\partial}{\partial x} I(x,y) \approx I(x+1,y) - I(x,y)$$
$$\frac{\partial}{\partial y} I(x,y) \approx I(x,y+1) - I(x,y)$$

### 1.2.2 Diskretes und kontinuierliches Signal

**Interpolation** Vom diskreten Signal $f[x] = S\{f(x)\}$ zum kontinuierlichen Signal $f(x)$. Interpoliertes Signal ist Faltung der Abtastwerte mit dem Interpolationsfilter.

$$f(x) \approx \sum_{k=-\infty}^{\infty} f[k]h(x-k) =: f[x] * h(x)$$

**Interpolationsfilter** Diskretes Signal: $f[x] = S\{f(x)\}$. Kontinuierliches Signal: $f(x) \approx f[x] * h(x)$.

- Gaußfilter: $h(x) = g(x)$
- Ideales Interpolationsfilter: $h(x) = \mathrm{sinc}(x)$
- Damit gilt: $f[x] * h(x) = f(x)$

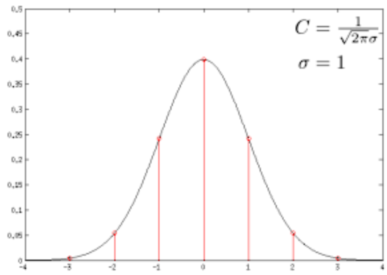



Figure 1.6: $g(x) = Ce^{\frac{-x^2}{2\sigma^2}}$

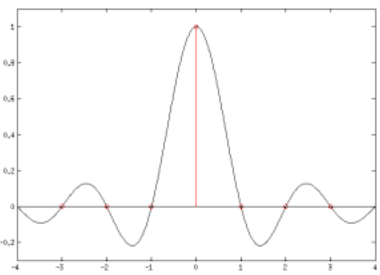


Figure 1.7: $sinc(x) = \frac{sin(\pi x)}{\pi x}, sinc(0) := 1$

### 1.2.3 Die diskrete Ableitung

**Mit Hilfe des rekonstruierten Signals**

- Algorithmisch
  1. Rekonstruktion des kontinuierlichen Signals
  2. Ableitung des kontinuierlichen Signals
  3. Abtastung der Ableitung
- Herleitung

- $f'(x) \approx \frac{d}{dx}(f[x] * h(x))$
  $f[x] * h'(x)$
- $f'[x] = f[x] * h'[x]$
  $= \sum_{k=-\infty}^{\infty} f[x-k]h'[k]$

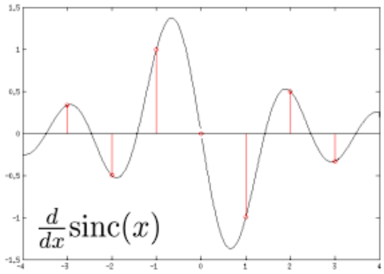



Figure 1.8: Sinc-Funktion(Langsames Abklingen)

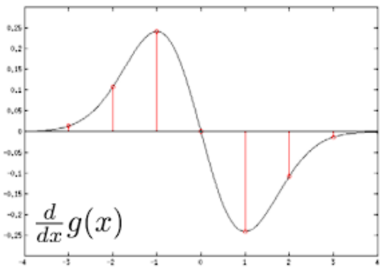



Figure 1.9: Gaußfilter(Schnelles Abklingen)

### 1.2.4 Zweidimensionale Rekonstruktion

**Separables 2D-Gaußfilter** 2D-Rekonstruktion:
$I(x,y) \approx I[x,y] * h(x,y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} I[k,l]g(x-k)g(y-l)$

### 1.2.5 Zweidimensionale Ableitung

**Ausnutzen der Separabilität** Ableitung in x-Richtung

$$\frac{d}{dx}I(x,y) \approx I[x,y] * (\frac{d}{dx}h(x,y))$$
$$= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} I[k,l]g'[x-k]g(y-l)$$

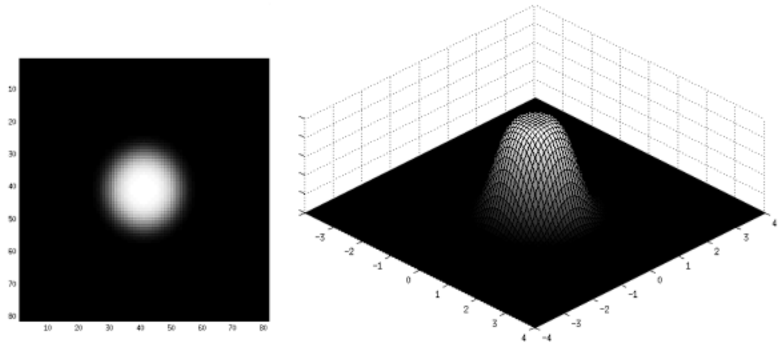


Figure 1.10: $h(x, y) := g(x)g(y)$

$$
\begin{aligned}
S\{\frac{d}{dx} I(x,y)\} &= I[x,y] * g'[x] * g[y] \\
&= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} I[x-k, y-l] g'[k] g[l]
\end{aligned}
$$

### 1.2.6 Endliche Approximation des Gaußfilters

#### Normierung des endlichen Filters

- In der Praxis wird die unendliche Summe durch wenige Summanden approximiert
- Wie wählt man eine geeignete Gewichtung $C$ des Gaußfilters $g(x) = Ce^{\frac{-x^2}{2\sigma^2}}$?
- Interpoliertes Signal: $f(x) \approx \sum_{k=-\infty}^{\infty} f[k]g(x-k)$
- Abgetastetes interpoliertes Signal: $f[x] \approx \sum_{k=-\infty}^{\infty} f[x-k]g[k]$
- Approximation durch endliche Summe: $f[x] \approx \sum_{k=-n}^{n} f[x-k]g[k]$
- Die endliche Approximation von $f[x]$ ist eine gewichtete Summe der Werte $f[x-n], ..., f[x+n]$ mit den Gewichten $g[n], ..., g[-n]$
- Normierungskonstante $C$ so gewählt, dass sich alle Gewichte zu 1 addieren
- Wähle $C = \frac{1}{\sum\limits_{-n \leq k \leq n} e^{\frac{-k^2}{2\sigma^2}}}$

### 1.2.7 Sobel-Filter

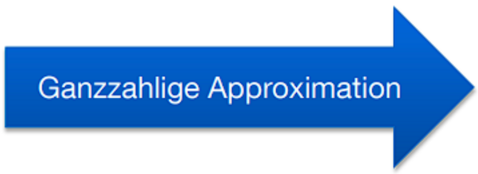
#### Herleitung

- Approximation von $S\{\frac{d}{dx}I(x,y)\} = I[x,y] * g'[x] * g[y] = \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} I[x-k,y-l]g'[k]g[l]$ durch endliche Summe $\sum_{k=-1,0,1}\sum_{l=-1,0,1} I[x-k,y-l]g'[k]g[l]$
- Daraus folgt der Normierungsfaktor $C = \frac{1}{1+2e^{-\frac{1}{2\sigma^2}}}$
- Für die Wahl $\sigma = \sqrt{\frac{1}{2\log 2}}$ ergeben sich somit die Werte

$$g[-1] = \frac{1}{4}; g[0] = \frac{1}{2}; g[1] = \frac{1}{4}$$
$$g'[-1] = \frac{1}{2}\log 2; g'[0] = 0; g'[1] = -\frac{1}{2}\log 2 \quad (\frac{1}{2}\log 2 \approx 0.35)$$

- Aus praktischen Gründen sind ganzzahlige Filterkoeffizienten erwünscht
- Für das Detektieren von Intensitätsunterschieden ist ein Vielfaches des Gradienten ausreichend

$$\frac{1}{8}\log 2 \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$




| | | |
|---|---|---|
| 1 | 0 | -1 |
| 2 | 0 | -2 |
| 1 | 0 | -1 |

Horizontales Sobel-Filter

#### Beispiel

### 1.2.8 Zusammenfassung

- Der Bildgradient ist ein wichtiges Werkzeug für die Bestimmung von lokalen Intensitätsänderungen
- Diskrete Ableitung wird durch Differenzieren des interpolierten Signals berechnet
- Sobel-Filter sind ganzzahlige Approximation eines Vielfachen des Gradienten

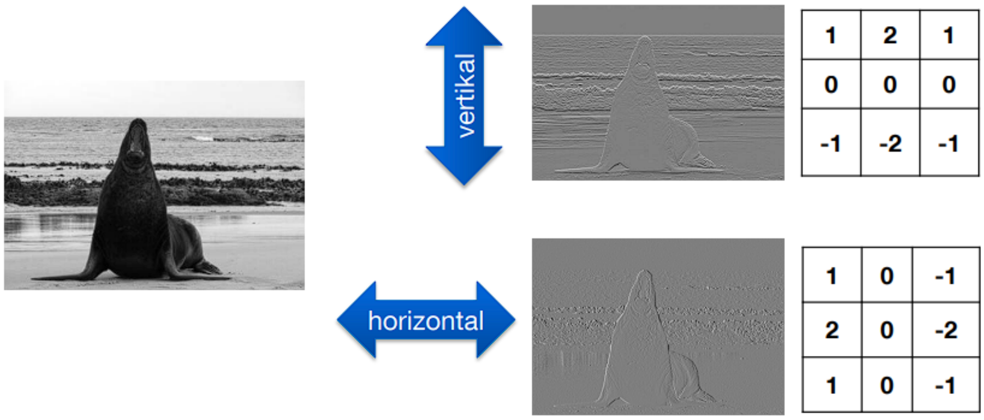



Figure 1.11: Sobel-Filterung

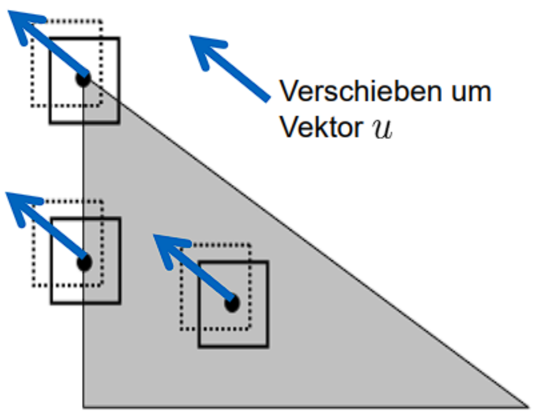
## 1.3 Merkmalspunkte-Ecken und Kanten

### 1.3.1 Ecken und Kanten

**...liefern markante Bildmerkmale**

- Bestimmung von Konturen
- Berechnungen von Bewegungen in Bildsequenzen
- Schätzen von Kamerabewegung
- Registrierung von Bildern
- 3D-Rekonstruktion

### 1.3.2 Harris Ecken- und Kantendetektor




**Änderung des Bildsegments in Abhängigkeit der Verschiebung**

- Ecke: Verschiebung in jede Richtung bewirkt Änderung
- Kante: Verschiebung in jede bis auf genau eine Richtung bewirkt Änderung
- Homogene Fläche: Keine Änderung, egal in welche Richtung

**Formelle Beschreibung der Änderung**

- Position im Bild: $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad I(x) = I(x_1, x_2)$
- Verschiebungsrichtung: $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$
- Änderung des Bildsegments:
$$S(u) = \int_W (I(x+u) - I(x))^2 dx$$
- Differenzierbarkeit von $I$:
$$\lim_{u \to 0} \frac{I(x+u) - I(x) - \nabla I(x)^\top u}{||u||} = 0$$

**Approximation der Änderung**

- Folgerung aus Differenzierbarkeit: $I(x+u) - I(x) = \nabla I(x)^\top u + o(||u||)$
- Restterm $o(||u||)$ mit der Eigenschaft $\lim_{u \to 0} o(||u||)/||u|| = 0$
- Approximation für kleine Verschiebungen: $I(x+u) - I(x) \approx \nabla I(x)^\top u$
- Approximation der Änderung im Bildsegment:
$$S(u) = \int_W (I(x+u) - I(x))^2 dx \approx \int_W (\nabla I(x)^\top u)^2 dx$$

**Die Harris-Matrix**

- Ausmultiplizieren des Integrals:
$$S(u) = \int_W (\nabla I(x)^\top u)^2 dx = u^\top (\int_W \nabla I(x) \nabla I(x)^\top dx) u$$
- Harris-Matrix: $G(x) = \int_W \nabla I(x) \nabla I(x)^\top ds$
$$\nabla I(x) \nabla I(x)^\top = \begin{bmatrix} (\frac{\partial}{\partial x_1} I(x))^2 & \frac{\partial}{\partial x_1} I(x) \frac{\partial}{\partial x_2} I(x) \\ \frac{\partial}{\partial x_2} I(x) \frac{\partial}{\partial x_1} I(x) & (\frac{\partial}{\partial x_2} I(x))^2 \end{bmatrix}$$
- Approximative Änderung des Bildsegments:
$$S(u) \approx u^\top G(x) u$$

**Eigenwertzerlegung**

- Eigenwertzerlegung der Harris-Matrix:

$$G(x) = \int_W \nabla I(x) \nabla I(x)^\top dx = V \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} V^\top$$

  mit $VV^\top = I_2$ und den Eigenwerten $\lambda_1 \leq \lambda_2 \leq 0$

- Änderung in Abhängigkeit der Eigenvektoren: $V = [v_1, v_2]$

$$S(u) \approx u^\top G(x) u = \lambda_1 (u^\top v_1)^2 + \lambda_2 (u^\top v_2)^2$$

**Art des Merkmals in Abhängigkeit der Eigenwerte**

- Beide Eigenwerte positive
  - $S(u) > 0$ für alle $u$ (Änderung in jede Richtung)
  - Untersuchtes Bildsegment enthält eine Ecke
- Ein Eigenwert positiv, ein Eigenwert gleich null
  - $S(u) \begin{cases} = 0, & \text{falls} \quad ; u = rv_2 \qquad \text{(Keine Änderung nur in Richtung} \\ & \qquad\qquad\qquad\qquad \text{des Eigenvektors zum Eigenwert 0)} \\ > 0, & \text{sonst} \end{cases}$
  - Untersuchtes Bildsegment enthält eine Kante
- Beide Eigenwerte gleich null
  - $S(u) = 0$ für alle $u$ (Keine Änderung, egal in welche Richtung)
  - Untersuchtes Bildsegment ist eine homogene Fläche

### 1.3.3 Praktische Realisierung des Harris-Detektors

**Berechnung der Harris-Matrix**

- Approximiere $G(x)$ durch endliche Summe

$$G(x) = \int_W \nabla I(x) \nabla I(x)^\top dx \approx \sum_{\tilde{x} \in W(x)} \nabla I(\tilde{x}) \nabla I(\tilde{x})^\top$$

- Gewichtete Summe in Abhängigkeit der Position von $\tilde{x}$

$$G(x) \approx \sum_{\tilde{x} \in W(x)} w(\tilde{x}) \nabla I(\tilde{x}) \nabla I(\tilde{x})^\top$$

- Gewichte $w(\tilde{x}) > 0$ betonen Einfluss der zentralen Pixel

**Eigenwerte**

- In der Realität nehmen Eigenwerte nie genau den Wert Null an, z.B. auf Grund von Rauschen, diskreter Abtastung und numerischen Ungenauigkeiten
- Charakteristik in der Praxis
  - Ecke: zwei große Eigenwerte
  - Kante: ein großer Eigenwert, ein kleiner Eigenwert
  - Homogene Fläche: zwei kleine Eigenwerte
- Entscheidung mittels empirischer Schwellwerte

**Ein einfaches Kriterium für Ecken und Kanten**

- Betrachte die Größe $H := \det(G) - k(\mathrm{tr}(G))^2 = (1-2k)\lambda_1\lambda_2 - k(\lambda_1^2 + \lambda_2^2)$
- Ecke (beide Eigenwerte groß)
  - $H$ größer als ein positiver Schwellwert
- Kante (ein Eigenwert groß, ein Eigenwert klein)
  - $H$ kleiner als ein negativer Schwellwert
- Homogene Fläche (beide Eigenwerte klein)
  - $H$ betragsmäßig klein

### 1.3.4 Zusammenfassung

**Harris-Detektor zur Bestimmung von Merkmalspunkten**

- Auswertung der (approximierten) Harris-Matrix

$$G(x) \approx \sum_{\tilde{x} \in W(x)} w(\tilde{x}) \nabla I(\tilde{x}) \nabla I(\tilde{x})^\top$$

- Eigenwertzerlegung von $G(x)$ liefert auch Info über Richtung etwaiger Kanten
- Effiziente Implementierung mit Hilfe des Ausdrucks

$$H := \det(G) - k(\mathrm{tr}(G))^2$$

- Entscheidung mittels Schwellwerten
  - Ecke: $0 < \tau_+ < H$
  - Kante: $H < \tau_- < 0$
  - Homogene Fläche: $\tau_- < H < \tau_+$