Agenda 10.11.:

- Nachtrag letzter Termin
  - Fragen zu Polarkoordinaten & Unkorreliertheit (siehe Forum 27.10)
  1. - Frage zu 1b iii): $\mathbb{E}[X] = \sum_{j=0}^{\infty} \mathbb{P}[X > j]$ $(X \geq 0)$ $*$
- Lösungen EA 2
  2. - Monte-Carlo (A3 a))
  3. - ggf. Fragen (vgl. auch zahlreiche Forumsbeiträge)
- Beispiele, Ideen, Tipps EA 3
  4. - konvex, L-glatt, Mittelwertsatz, RKT, Testaufgaben
  5. - Hinweise zu den Aufgaben

---

1. Wenn X eine nichtnegative, diskrete ZV ist, dann gilt

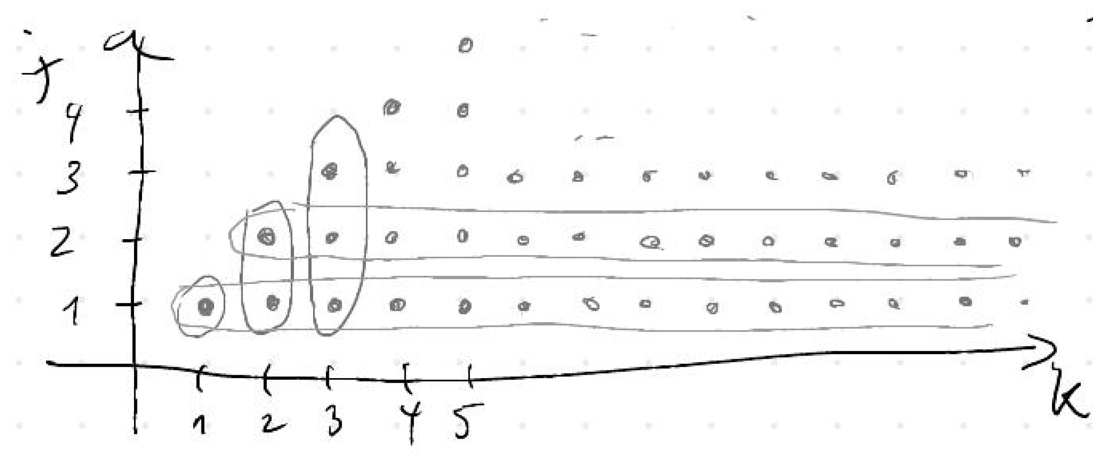
$$\mathbb{E}[X] = \sum_{k \in \mathbb{N}} k \cdot \mathbb{P}[X = k] = \sum_{j \in \mathbb{N}_0} \mathbb{P}[X > j].$$

(wichtig für 1b iii) + Testaufgabe 9 + nützlich)

Beweis:

$$\mathbb{E}[X] = \sum_{k \in \mathbb{N}} k \cdot \mathbb{P}[X = k] = \sum_{k \in \mathbb{N}_0} \Big(\underbrace{\sum_{j=1}^{k} 1}_{= k}\Big) \mathbb{P}[X = k]$$

$$= \underbrace{\sum_{k \in \mathbb{N}_0} \sum_{j=1}^{k}}_{\text{Tonelli}} \big(1 \cdot \mathbb{P}[X = k]\big) \overset{(*)}{=} \sum_{j=1}^{\infty} \underbrace{\sum_{k=j}^{\infty} \mathbb{P}[X = k]}_{= \mathbb{P}[X \geq j]} = \sum_{j=1}^{\infty} \mathbb{P}[X \geq j] = \sum_{j \in \mathbb{N}_0} \mathbb{P}[X > j]$$



$\square$

Ähnlich (siehe Forum) kann man zeigen, dass für stetige, nicht-negative ZVen gilt:

$$\mathbb{E}[X] = \int_0^{\infty} \underbrace{1 - F(x)}_{\mathbb{P}[X > x]} \, dx.$$

2.

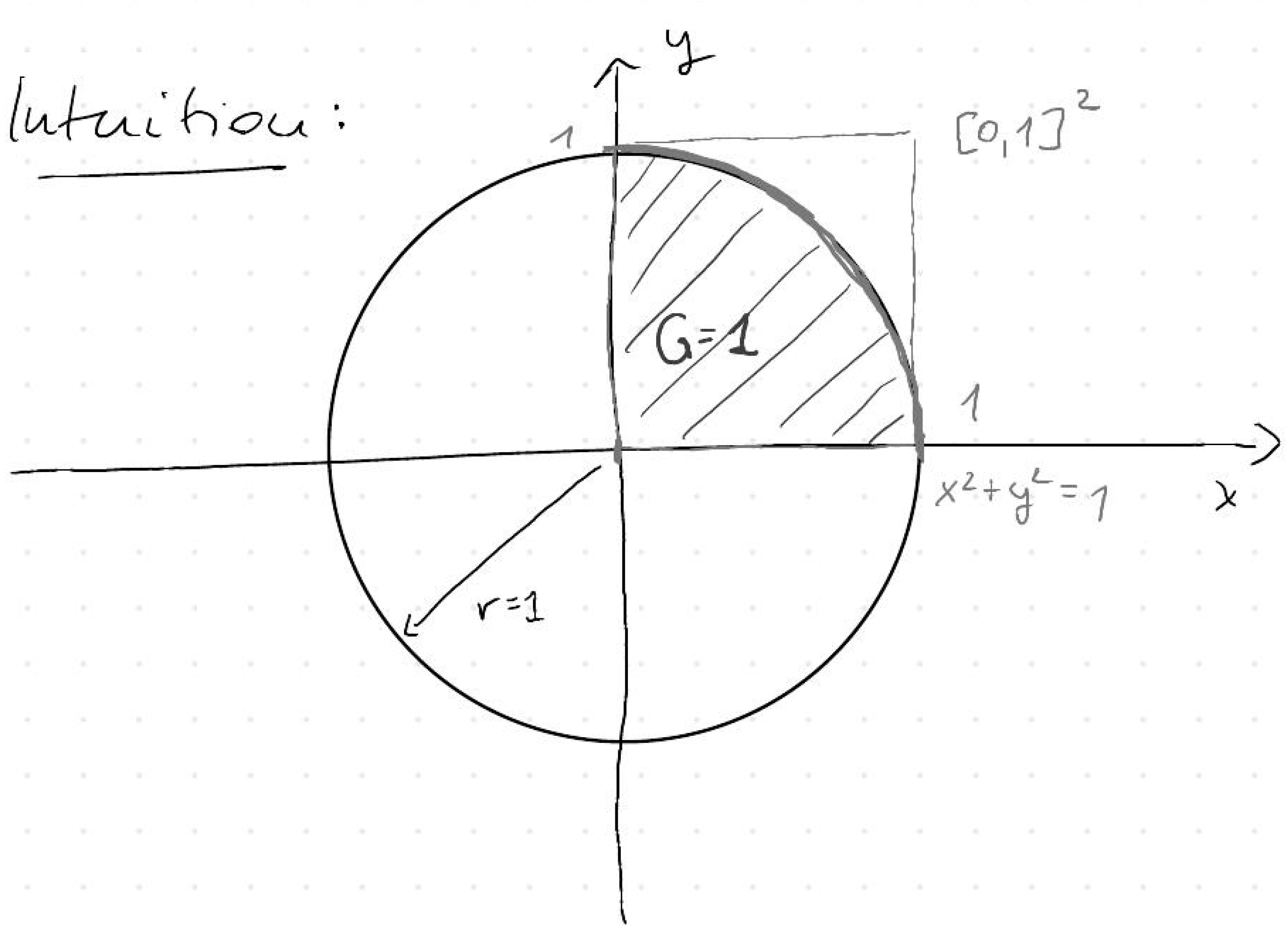
3. Monte-Carlo-Simulation (10 Punkte). Wir betrachten in dieser Aufgabe zwei Beispiele für das Verfahren.

a) (4 Punkte) Wir betrachten die Funktion $G(x, y) = \mathbb{1}_{\{x^2+y^2 \leq 1\}}$ für $x, y \in \mathbb{R}$. Finden Sie C sodass

$$\frac{\pi}{C} = \int_{[0,1]^2} G(x, y) \, dx \, dy$$

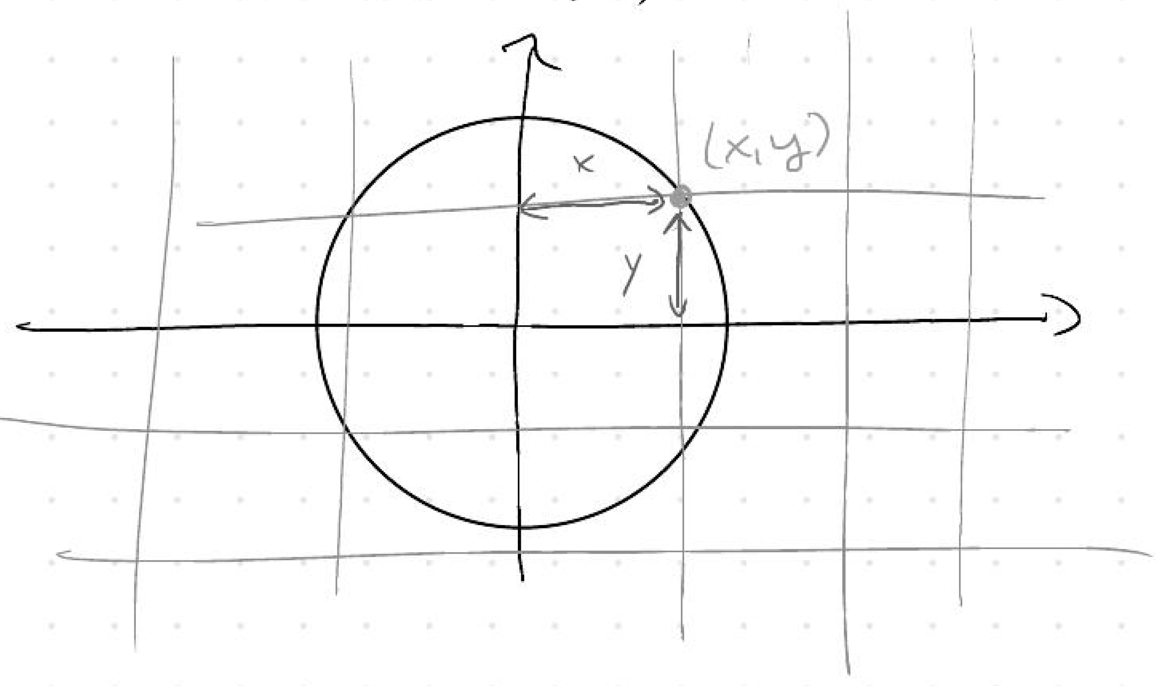
gilt und beschreiben (und begründen) Sie, wie Sie diese Gleichung nutzen können, um π numerisch zu schätzen, wenn nur stetig gleichverteilte Zufallszahlen zwischen 0 und 1 erzeugt werden können.
Hinweis: Mit der Polarkoordinatentransformation ($x = r\cos(\theta)$, $y = r\sin(\theta)$, $dx\,dy = r\,dr\,d\theta$) kann man $\int_{[0,1]^2} G(x, y)\,dx\,dy = \int_0^{\pi/2}\int_0^1 r\,dr\,d\theta$ zeigen.

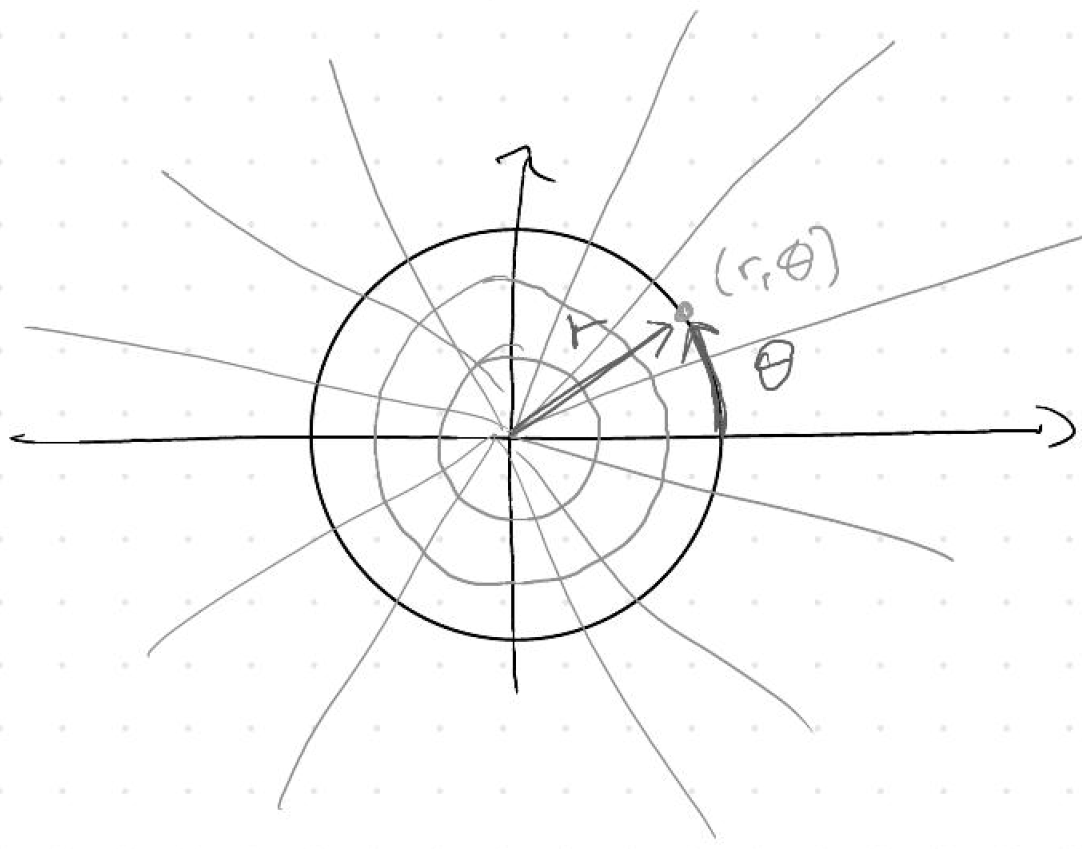
Intuition:

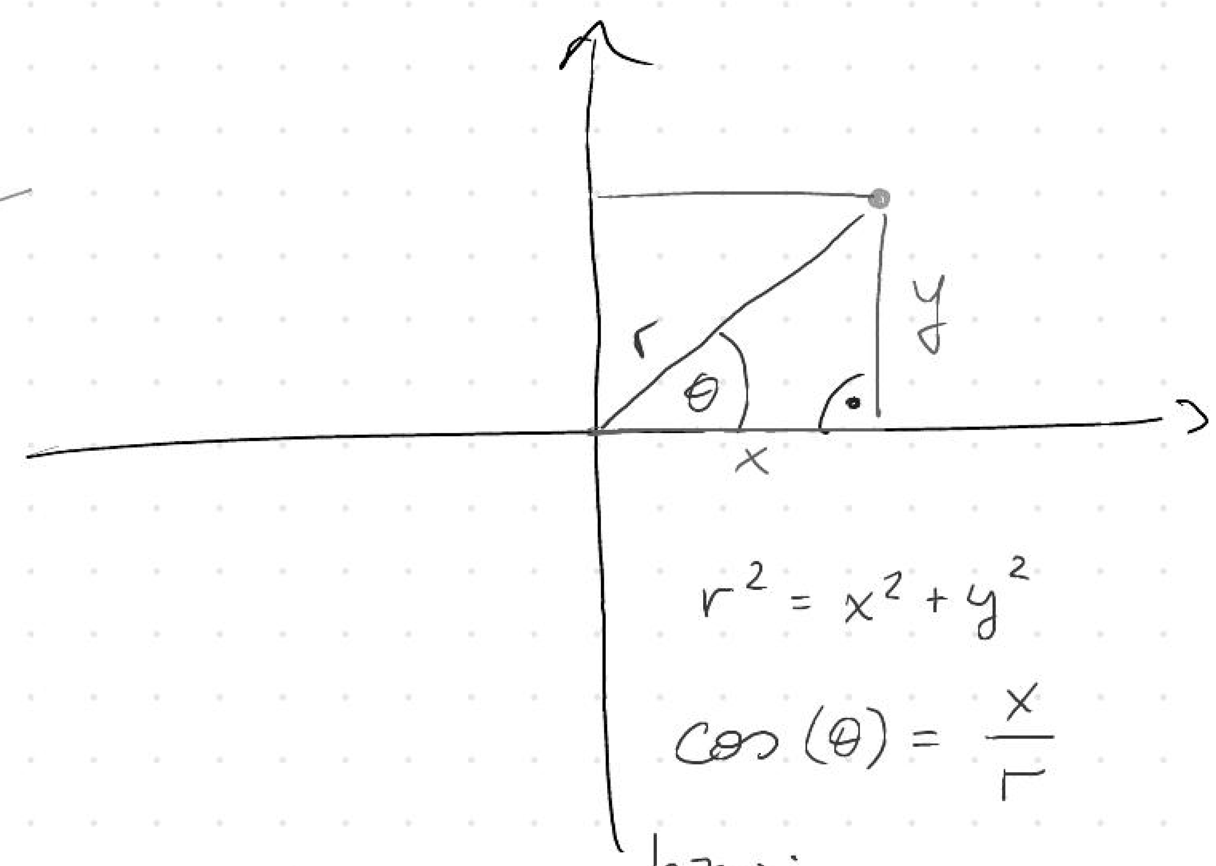


Kreisfläche: $\pi r^2 = \pi \Rightarrow C = 4$

Polarkoordinaten (durfte man hier einfach aus dem Tipp übernehmen – kommt nochmal in RE4):







$r^2 = x^2 + y^2$

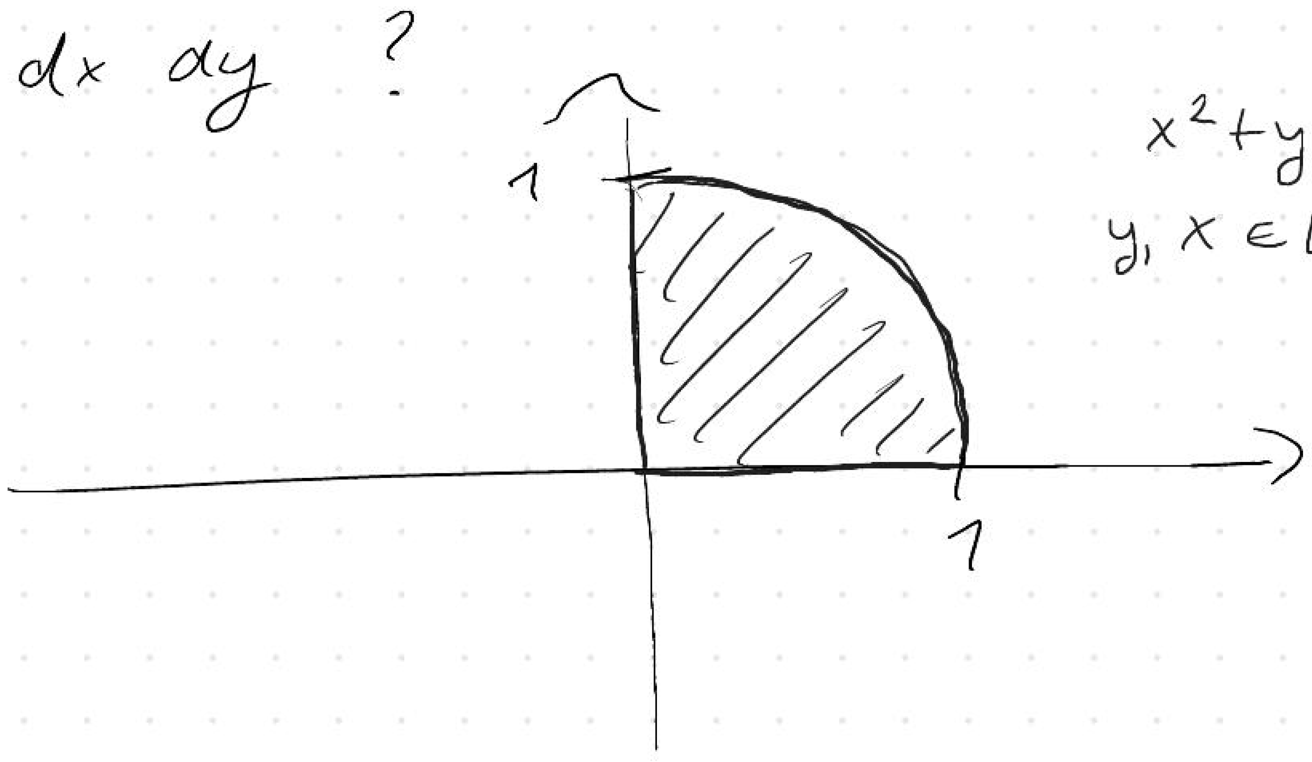
$\cos(\theta) = \frac{x}{r}$

bzw:

$x = r\cos(\theta)$

$y = r\sin(\theta)$

$\int_{[0,1]^2} G(x,y)\,dx\,dy$?

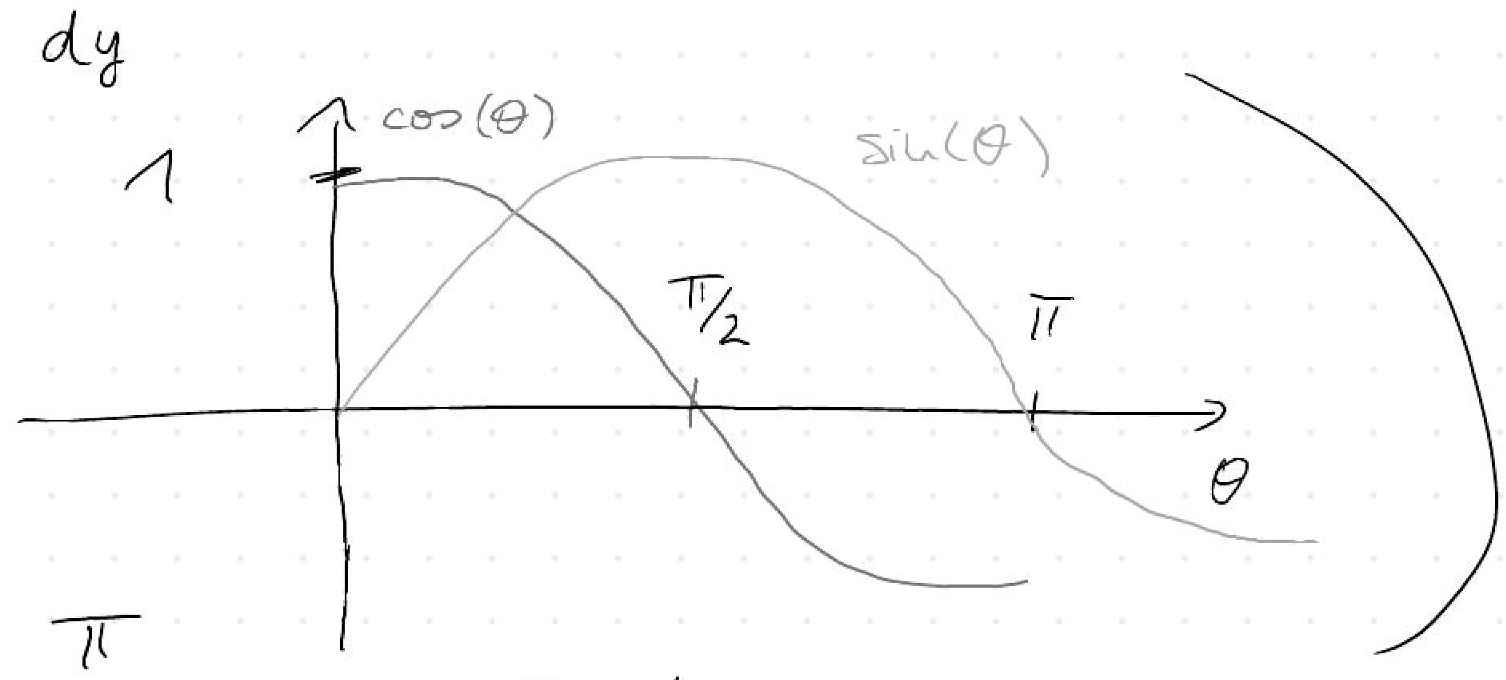


$x^2+y^2 \le 1$
$y, x \in [0,1]$ oder $r \in [0,1]$
$\theta \in [0, \underbrace{\pi/2}_{\hat{=}\,90°}]$

„mathematisch":

$$\int_{[0,1]^2} G(x,y)\,dx\,dy = \int_{[0,1]^2} \mathbb{1}_{\underbrace{x^2+y^2}_{r^2} \le 1}\,dx\,dy$$

(d.h. $x = r\cos(\theta) \in [0,1]$, $y = r\sin(\theta) \in [0,1]$ $\Big\} \Rightarrow \theta \le \pi/2$



)

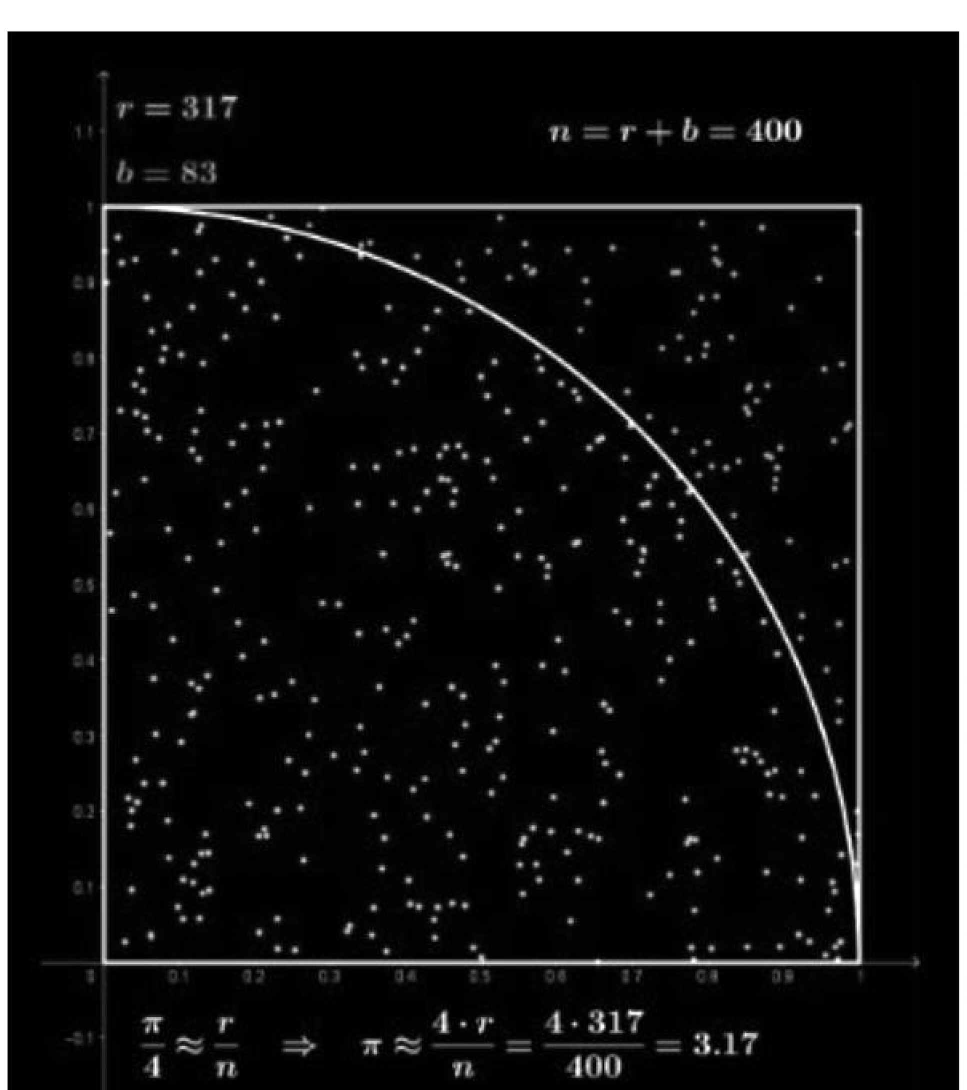
$$\overset{\text{Tipp}}{=} \int_0^{\pi/2}\int_0^1 r\,dr\,d\theta = \int_0^{\pi/2} 1/2\,d\theta = \frac{\pi}{4} \Rightarrow C = 4.$$

Monte-Carlo: Seien $(Y_k), (X_k)$ iid. $\mathcal{U}[0,1]$ verteilt. GdgZ ⇒

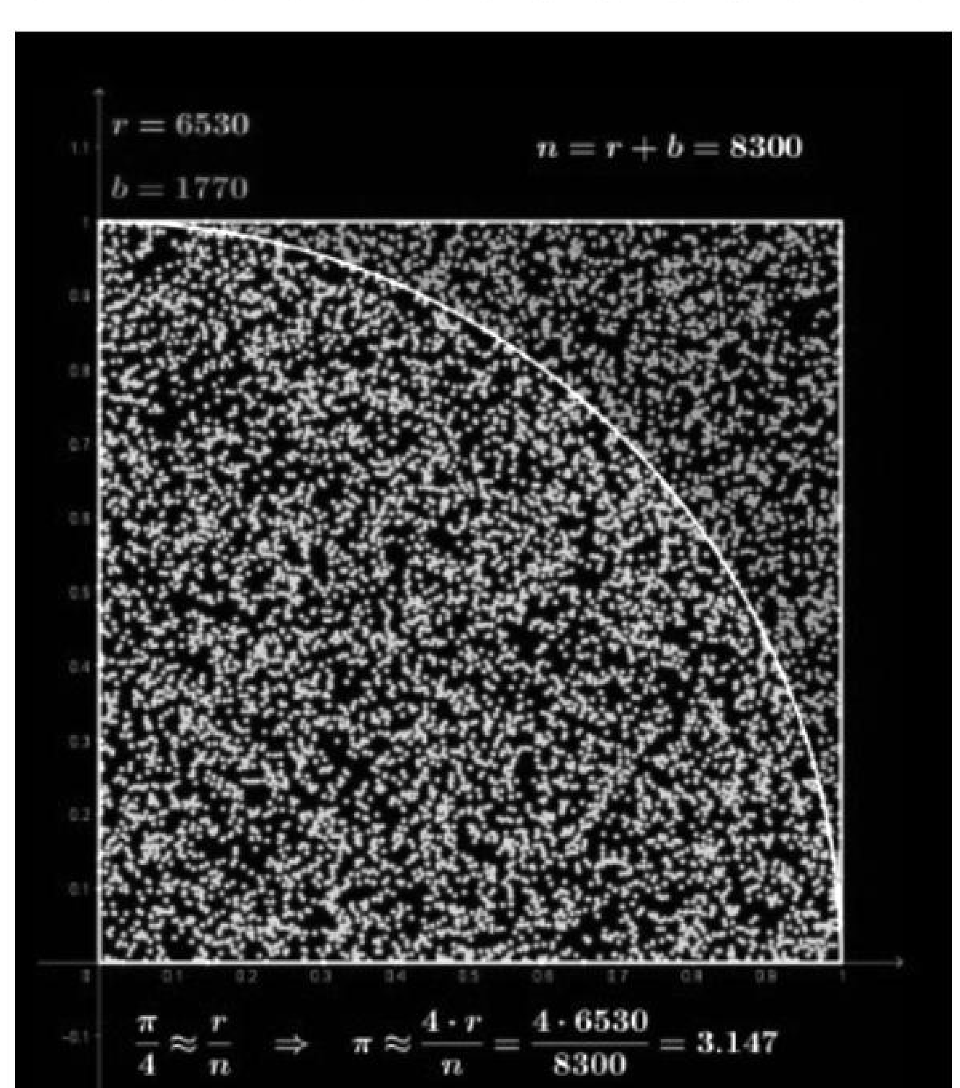
$$\frac{1}{N}\sum_{k=1}^{N} G(X_k, Y_k) \xrightarrow{N\to\infty} \mathbb{E}[G(X_1, Y_1)] \overset{\text{Dichte/Def}}{=} \int_{[0,1]^2} G(x,y)\,dx\,dy \overset{\text{s.o.}}{=} \frac{\pi}{4}$$

Setzen wir also $\hat{\pi} = \frac{4}{N}\sum_{k=1}^{N} G(X_k, Y_k)$.

Idee Monte-Carlo-Approximation: GdgZ ⇒ $\frac{1}{N}\sum_{k=1}^{N} G(X_k, Y_k) \to \mathbb{E}[G(X_k, Y_k)]$
$= \int_{[0,1]^2} G(x,y)\,dx\,dy$



Monte-Carlo-Methode zur Approximation des Werts von π.



Monte-Carlo-Methode zur Approximation des Werts von π.

Quelle: Wikipedia

4.

Definition 3.1.1. 1. Eine Menge $G \subseteq \mathbb{R}^d$ heißt *konvex*, falls für alle $x, y \in G$ und alle $t \in [0,1]$ das Element $tx + (1-t)y$ wieder in G liegt.

2. Sei G eine konvexe Menge. Eine Funktion $f\colon G \to \mathbb{R}$ heißt *konvex*, falls für alle $x, y \in G$ und alle $t \in (0,1)$ gilt, dass

$$f(tx + (1-t)y) \leq t f(x) + (1-t) f(y)$$

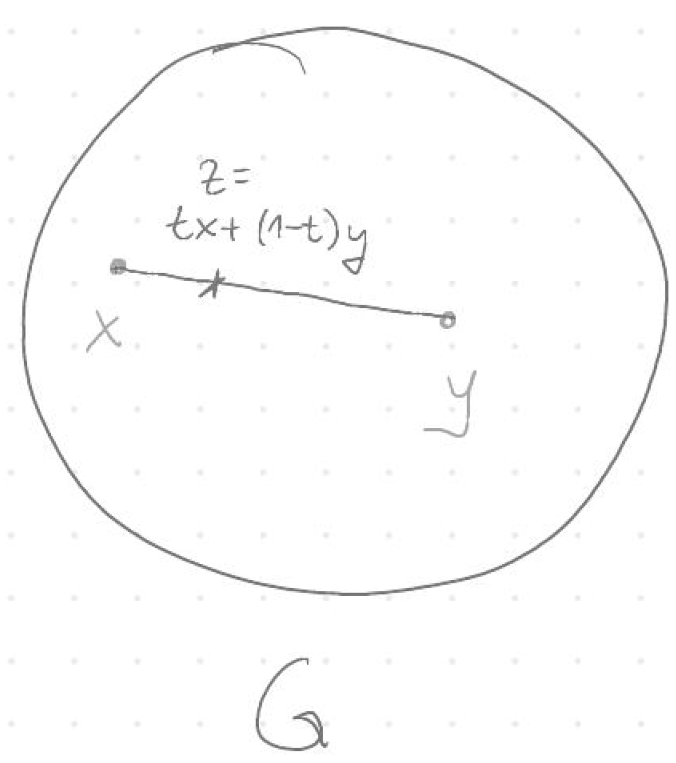
ist. $f\colon G \to \mathbb{R}$ heißt *strikt konvex*, falls für alle $x, y \in G$ und alle $t \in (0,1)$ gilt, dass

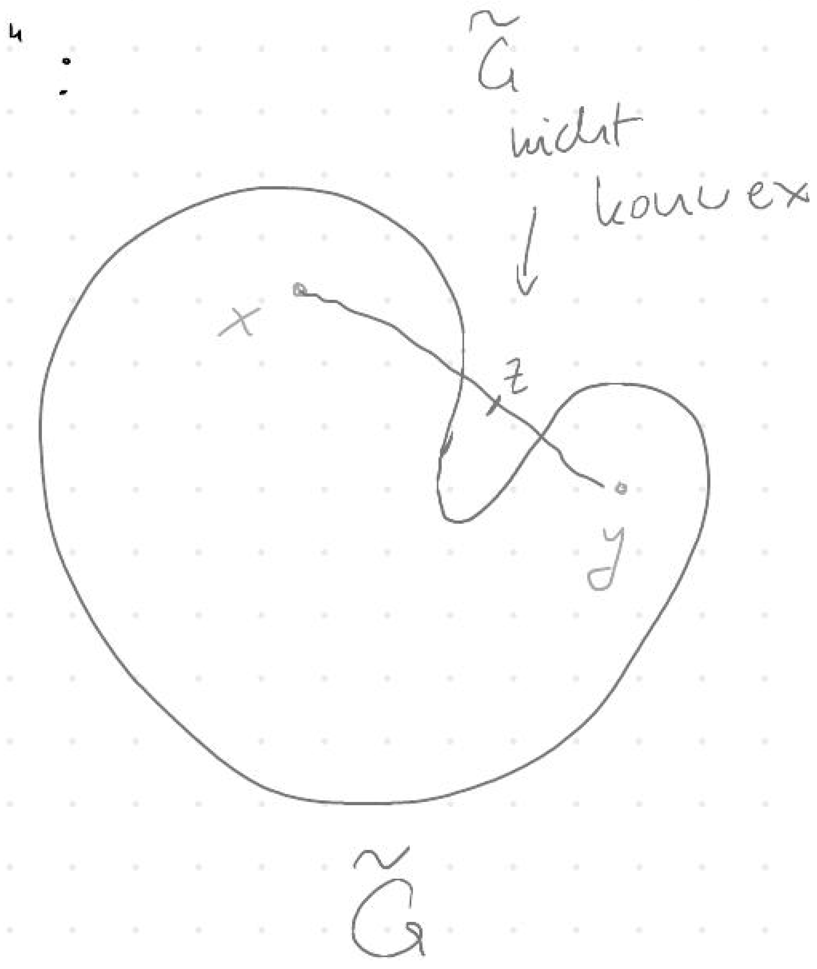
$$f(tx + (1-t)y) < t f(x) + (1-t) f(y)$$

ist.

3. Sei G eine konvexe Menge. Eine Funktion $f\colon G \to \mathbb{R}$ heißt *(strikt) konkav*, falls $-f$ (strikt) konvex ist.

Idee „konvexe Menge":





⇝ alle Punkte, die man als Konvexkombination $z = t\,x + (1-t)\,y$ aus $x, y \in G$ schreiben kann, liegen wieder in G (f.a. $x, y \in G$)

Definition 3.1.1. 1. Eine Menge $G \subseteq \mathbb{R}^d$ heißt *konvex*, falls für alle $x, y \in G$ und alle $t \in [0,1]$ das Element $tx + (1-t)y$ wieder in G liegt.

2. Sei G eine konvexe Menge. Eine Funktion $f\colon G \to \mathbb{R}$ heißt *konvex*, falls für alle $x, y \in G$ und alle $t \in (0,1)$ gilt, dass

$$f(tx + (1-t)y) \leq t f(x) + (1-t) f(y)$$

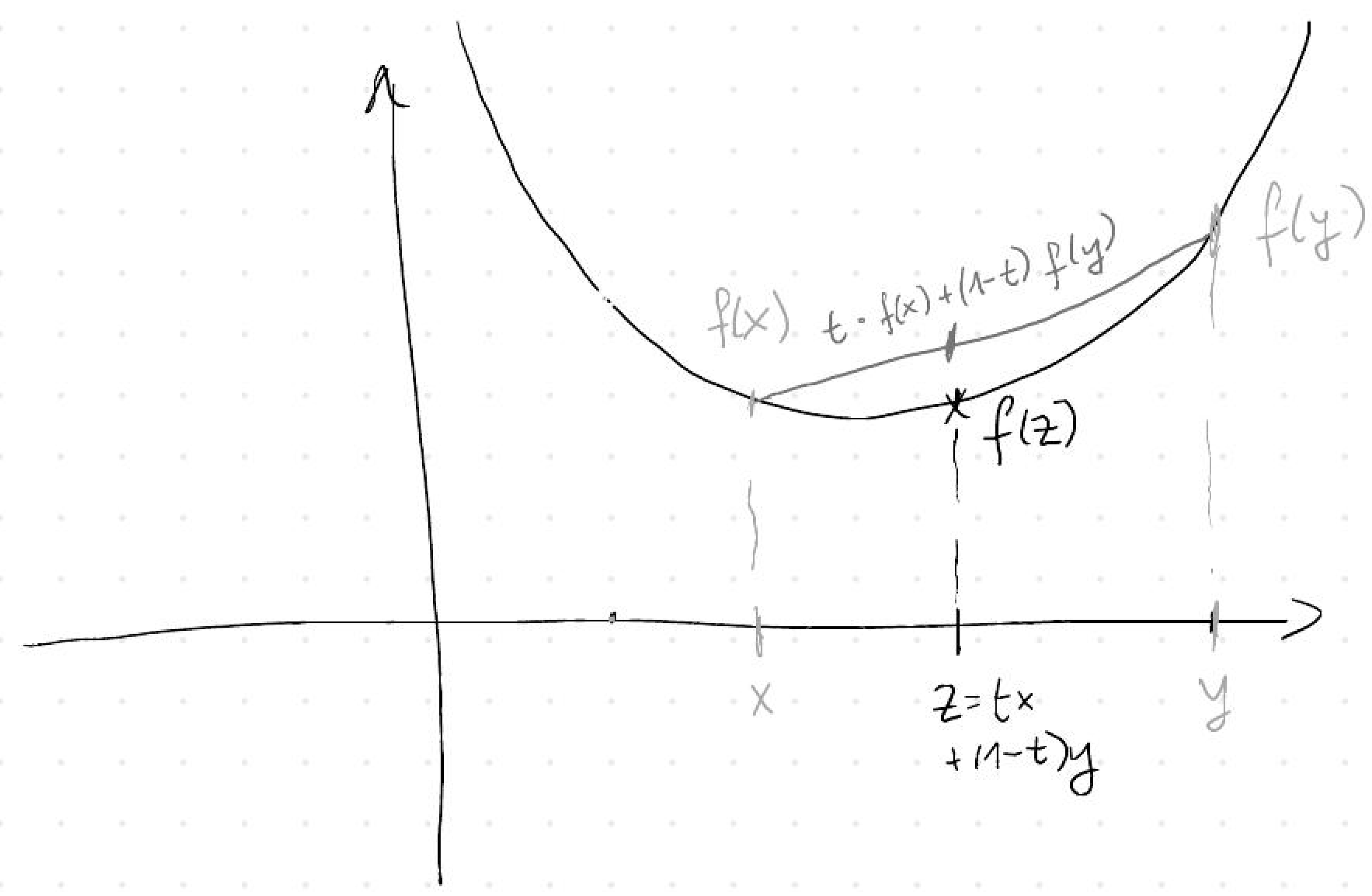
ist. $f\colon G \to \mathbb{R}$ heißt *strikt konvex*, falls für alle $x, y \in G$ und alle $t \in (0,1)$ gilt, dass

$$f(tx + (1-t)y) < t f(x) + (1-t) f(y)$$

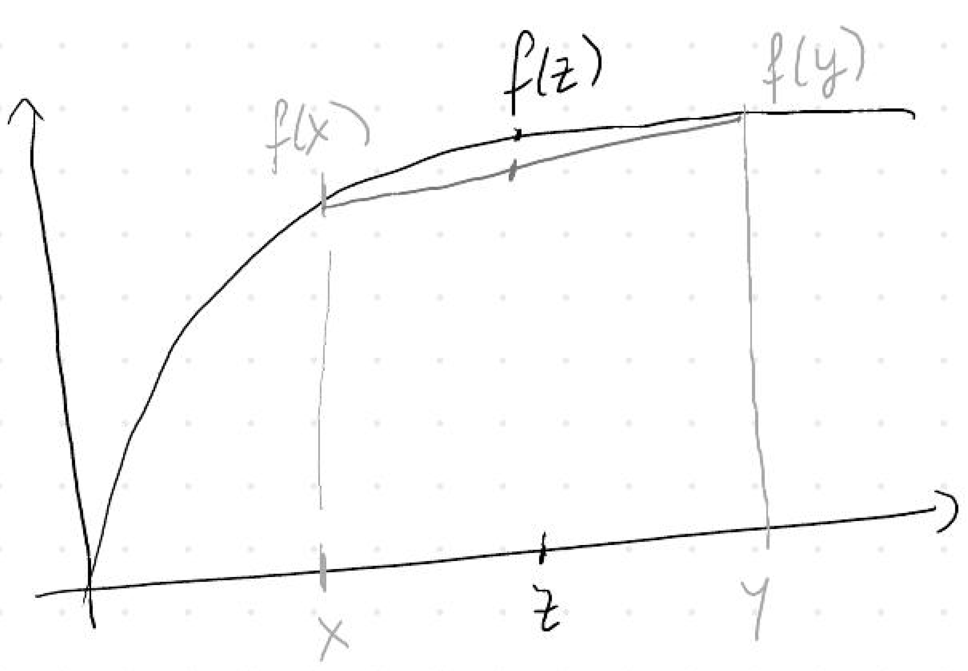
ist.

3. Sei G eine konvexe Menge. Eine Funktion $f\colon G \to \mathbb{R}$ heißt *(strikt) konkav*, falls $-f$ (strikt) konvex ist.

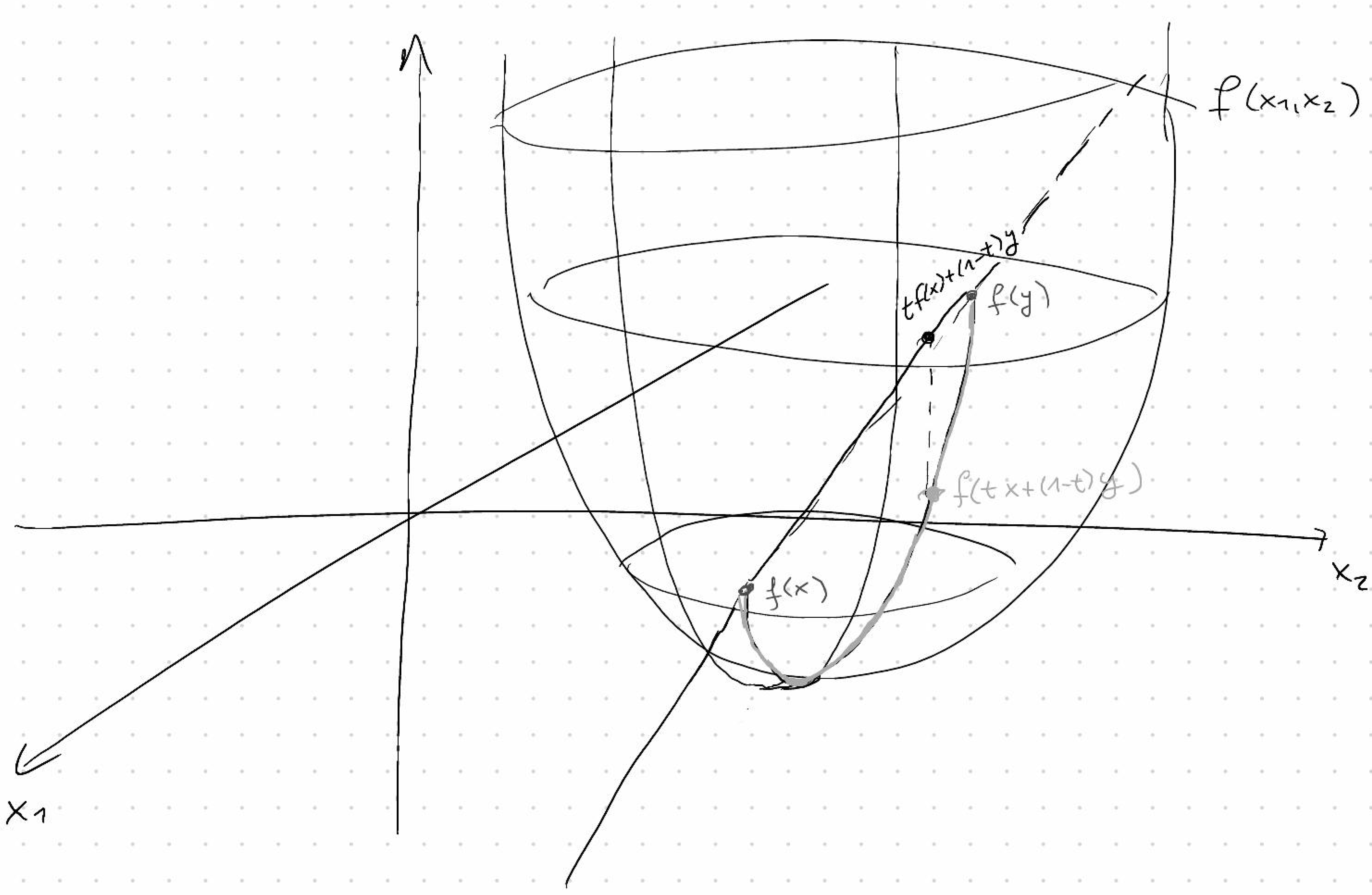
Idee „konvexe Funktion":



konvex



nicht konvex



Beispiel „konvex": Test KE3 A2

$f: (0,\infty) \times (0,\infty) \to \mathbb{R} \qquad f(x_1, x_2) = \frac{1}{x_1 \cdot x_2}$

Nützlich: Satz 3.2.13 + Def. A.3.20:

$f: G \to \mathbb{R}$ zweimal diffbar + G offen, dann

f konvex $\iff H_f(x) = \begin{pmatrix} \frac{\partial_2}{\partial x_1^2} f(x_1,x_2) & \frac{\partial_2}{\partial x_1 \partial x_2} f(x_1,x_2) \\ \frac{\partial_2}{\partial x_1 \partial x_2} f(x_1,x_2) & \frac{\partial_2}{\partial x_2^2} f(x_1,x_2) \end{pmatrix}$ pos. semidef. für alle $x \in G$.

Also: Ableitungen berechnen + Eigenwerte prüfen!

Hier $\qquad H_f(x) = \begin{pmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{pmatrix}$

$$\det(H_f(x) - \lambda I) = \left(\frac{2}{x_1^3 x_2} - \lambda\right)\left(\frac{2}{x_1 x_2^3} - \lambda\right) - \frac{1}{x_1^4 x_2^4}$$

$$= \lambda^2 - 2\left(\frac{1}{x_2^3 x_1} + \frac{1}{x_1^3 x_2}\right)\lambda + \frac{3}{x_1^4 x_2^4}$$

$$\Rightarrow \lambda_{1,2} = \underbrace{\frac{1}{x_2^3 x_1} + \frac{1}{x_1^3 x_2}}_{\geq 0} \pm \sqrt{\underbrace{\left(\frac{1}{x_2^3 x_1} + \frac{1}{x_1^3 x_2}\right)^2 - \frac{3}{x_1^4 x_2^4}}_{= \frac{(x_1^2 - x_2^2)^2}{x_2^6 x_1^6} > 0}}$$

Der Term unter der Wurzel ist zudem $\leq \left(\frac{1}{x_2^3 x_1} + \frac{1}{x_1^3 x_2}\right)^2$, also $\lambda_{1,2} \geq 0$.

$\Rightarrow$ Funktion konvex.

Beispiel "L-glatt":

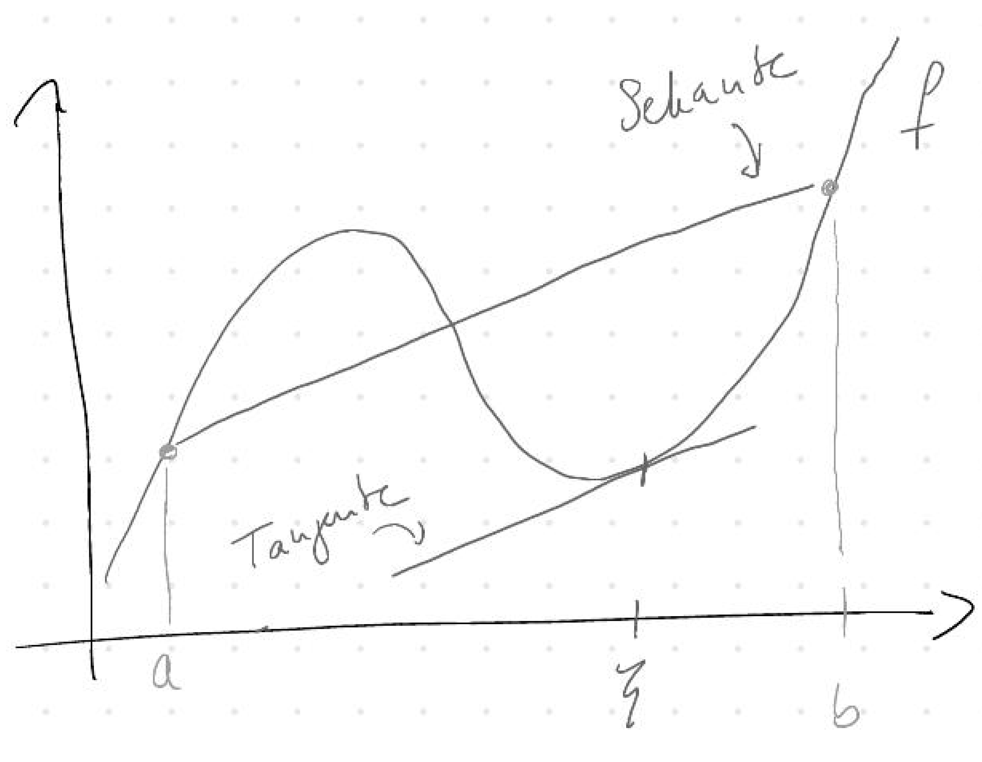
$\boxed{f: \mathbb{R}^d \to \mathbb{R} \text{ heißt } L\text{-glatt } (L>0), \text{ falls der Gradient } L\text{-Lipschitz ist: } \forall x,y: \ \|\nabla f(x) - \nabla f(y)\| \leq L\|x-y\|}$

Lipschitz-Stetigkeit kann man (oft) gut mit dem Mittelwertsatz zeigen!

Idee Mittelwertsatz (MWS):

Wenn f auf $[a,b] \subseteq \mathbb{R}$ definiert und stetig ist und auf (a,b) differenzierbar, dann gibt es mind. einen Punkt $\xi \in (a,b)$ mit

$$f(b) - f(a) = f'(\xi) \cdot (b-a).$$

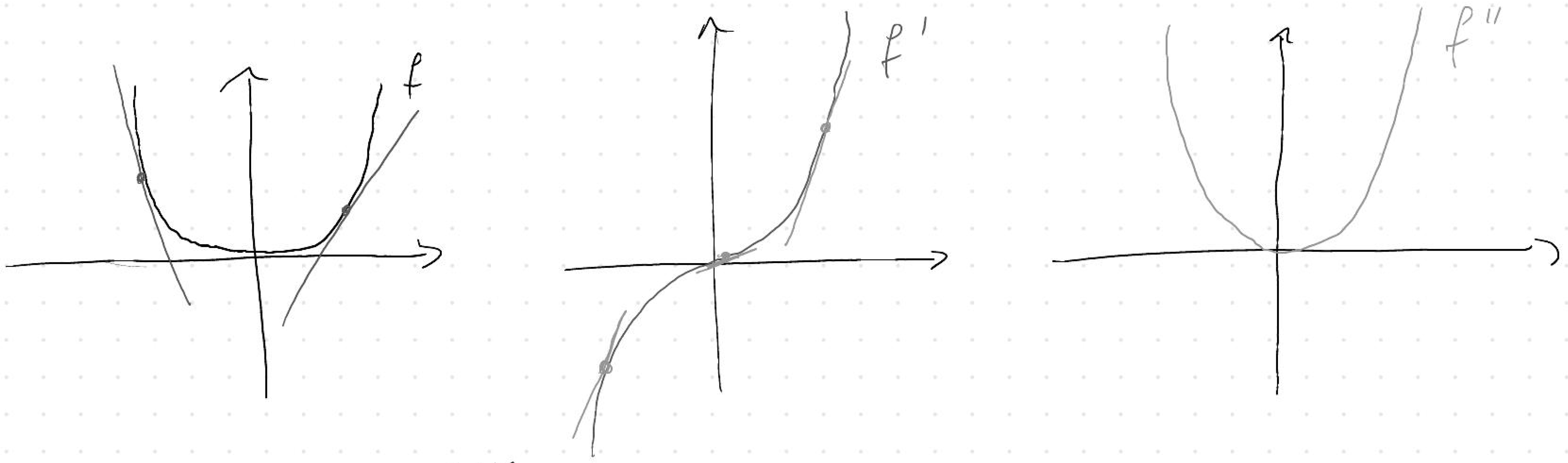


$\rightsquigarrow$ an mindestens einer Stelle ξ zwischen a und b taucht die Sekantensteigung $\frac{f(b)-f(a)}{b-a}$ als Tangentensteigung $f'(\xi)$ auf.

$$\Rightarrow |f(b) - f(a)| \leq \underbrace{\sup_{\xi \in [a,b]} |f'(\xi)|}_{L} \cdot |b-a|$$

z.B. $f : (1,2) \to \mathbb{R}$, $f(x) = x^4$ ist L-glatt mit $L = 48$.

$f'(x) = 4x^3 \qquad f''(x) = 12x^2$



$$|f'(x) - f'(y)| \overset{MWS}{=} |f''(\xi)|\,|(y-x)| \leq \sup_{\xi \in (1,2)} |f''(\xi)| \cdot |y-x| = \left(\sup_{\xi\in(1,2)} 12\xi^2\right) \cdot |y-x|$$

$$= \underset{\substack{\uparrow \\ 12 \cdot 2^2}}{48} \cdot |y-x|$$

Die gleiche Idee (aber anders aufgeschrieben) folgt aus dem Hauptsatz der Integralrechnung (Vorlesung): Setze $v = y - x$.

$$f'(x+v) - f'(x) = \int_x^{x+v} f''(y)\,dy \overset{\text{Subs.}}{=} \int_0^1 f''(x+s\cdot v)\cdot v\,ds$$

Dann

$$|f'(x+v) - f'(x)| = \left|\int_0^1 f''(x+sv)v\,ds\right| \overset{\Delta}{\leq} \int_0^1 |v| \underbrace{|f''(x+sv)|}_{\leq L} ds$$

$$= |v| \cdot \int_0^1 L\,ds = L \cdot |v| = L \cdot |y-x|.$$

Was passiert in Dimensionen > 1? ZB $G : \mathbb{R}^2 \to \mathbb{R}$, wie $G(x,y) = \sqrt{1+x^2+y^2}$.

Ähnlich wie oben haben wir (vgl. Beweis A.3.22):

$$\nabla G(x+v) - \nabla G(x) \overset{\substack{\text{komponentenweise} \\ \downarrow}}{=} \int_0^1 H_G(x+sv)\cdot v\,ds \qquad \text{wobei } x, v \in \mathbb{R}^2$$

$$\Rightarrow \|\nabla G(x+v) - \nabla G(x)\| \leq \left\|\int_0^1 H_G(x+sv)\,v\,ds\right\|$$

$$\leq \int_0^1 \|H_G(x+sv)\cdot v\|\,ds \overset{\text{Lemma A.1.9}}{\leq} \int_0^1 \|H_G(x+sv)\| \cdot \|v\|\,ds$$

Operatornorm

$$\leq \underbrace{\left(\sup_{s\in[0,1]} \|H_G(x+sv)\|\right)}_{\hat{=} L} \cdot \|x-y\|$$

Tipp für Aufgabe: versuchen, L mit Eigenwerten in Verbindung zu bringen!

Beispiel KKT:

$$p^* = \min \quad 4x^2 + y^2 - x - 2y$$
$$\text{mit} \quad 2x + y \le 1$$
$$x^2 \le 1$$

$$\left.\begin{aligned} f(x,y) &= 4x^2 + y^2 - x - 2y \\ g_1(x,y) &= 2x + y - 1 \\ g_2(x,y) &= x^2 - 1 \end{aligned}\right\} \text{Notation } 3.23$$

Theorem 3.4.20 (Karush-Kuhn-Tucker). *Wir betrachten ein allgemeines Optimierungsproblem* (3.23) *mit differenzierbaren Funktionen* $f, g_1, \dots, g_m$ *sowie* $h_1, \dots, h_p$. *Es gelte starke Dualität. Sei* $x^* \in \mathbb{R}^d$ *eine Lösung des primalen Problems und* $(\lambda^*, \nu^*) \in \mathbb{R}^{m+p}$ *eine Lösung des dualen Problems* (3.24). *Dann gelten die folgenden Aussagen:*

$$g_i(x^*) \le 0 \quad \text{für } i = 1, \dots, m, \tag{3.26}$$
$$h_j(x^*) = 0 \quad \text{für } j = 1, \dots, p, \tag{3.27}$$
$$\lambda_i^* \ge 0 \quad \text{für } i = 1, \dots, m, \tag{3.28}$$
$$\lambda_i^* g_i(x^*) = 0 \quad \text{für } i = 1, \dots, m, \tag{3.29}$$
$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^{p} \nu_j^* \nabla h_j(x^*) = 0. \tag{3.30}$$

Die Bedingungen (3.26) - (3.30) *nennt man auch* Karush-Kuhn-Tucker *oder kurz* KKT-Bedingungen.

Die Bedingungen (3.26), (3.27) und (3.27) müssen gelten, da x^* und (λ^*, ν^*) notwendigerweise zulässige Punkte des primalen bzw. dualen Problems sein müssen. Die Bedingung (3.29) haben wir schon in Satz 3.4.19 gefolgert.

Wir werden nun sehen, dass im Fall eines konvexen Problems die KKT-Bedingungen auch hinreichend sind für die Existenz von Lösungen des dualen sowie des primalen Problems.

Satz 3.4.21. *Wir betrachten ein konvexes Optimierungsproblem* (3.23). *Sei* (x^*, λ^*, ν^*) *ein Punkt, der die KKT-Bedingungen erfüllt. Dann gilt starke Dualität und* x^* *löst das primale,* (λ^*, ν^*) *löst das duale Problem.*

$$\begin{aligned} &\min_{x \in G} f(x) \\ \text{unter} \quad & g_i(x) \le 0, \quad i = 1, \dots, m, \\ & h_j(x) = 0, \quad j = 1, \dots, p. \end{aligned} \tag{3.23}$$

KKT-Bed:

(3.26) $2x + y - 1 \le 0$ A

$x^2 - 1 \le 0$ B

(3.27) —

(3.28) $\lambda_1, \lambda_2 \ge 0$ C

(3.29) $\lambda_1 (2x + y - 1) = 0$ D

$\lambda_2 (x^2 - 1) = 0$ E

(3.30) $\underbrace{8x - 1}_{f_x(x,y)} + \underbrace{\lambda_1 \cdot 2}_{\lambda_1 g_{1x}(x,y)} + \underbrace{2 \cdot \lambda_2 \cdot x}_{\lambda_2 g_{2x}(x,y)} = 0$ F

$2y - 2 + \lambda_1 \cdot 1 + 0 = 0$ G

Fallunterscheidung!

Fall 1: $\lambda_1 = \lambda_2 = 0$. Dann

$G \Rightarrow y = 1$

$F \Rightarrow x = 1/8$

↯ A kann nicht erfüllt sein

Fall 2: $\lambda_1 = 0$, $\lambda_2 > 0$. Dann

$E \Rightarrow x^2 - 1 = 0 \Rightarrow x = \pm 1$

$G \Rightarrow y = 1$

$x = 1$ ist nicht möglich wegen A. Wenn $x = -1$, dann $F \Rightarrow \lambda_2 = -9/2 < 0$

↯ C nicht erfüllt

---

usw.

Herauskommen soll $p^* = -\frac{33}{32}$.