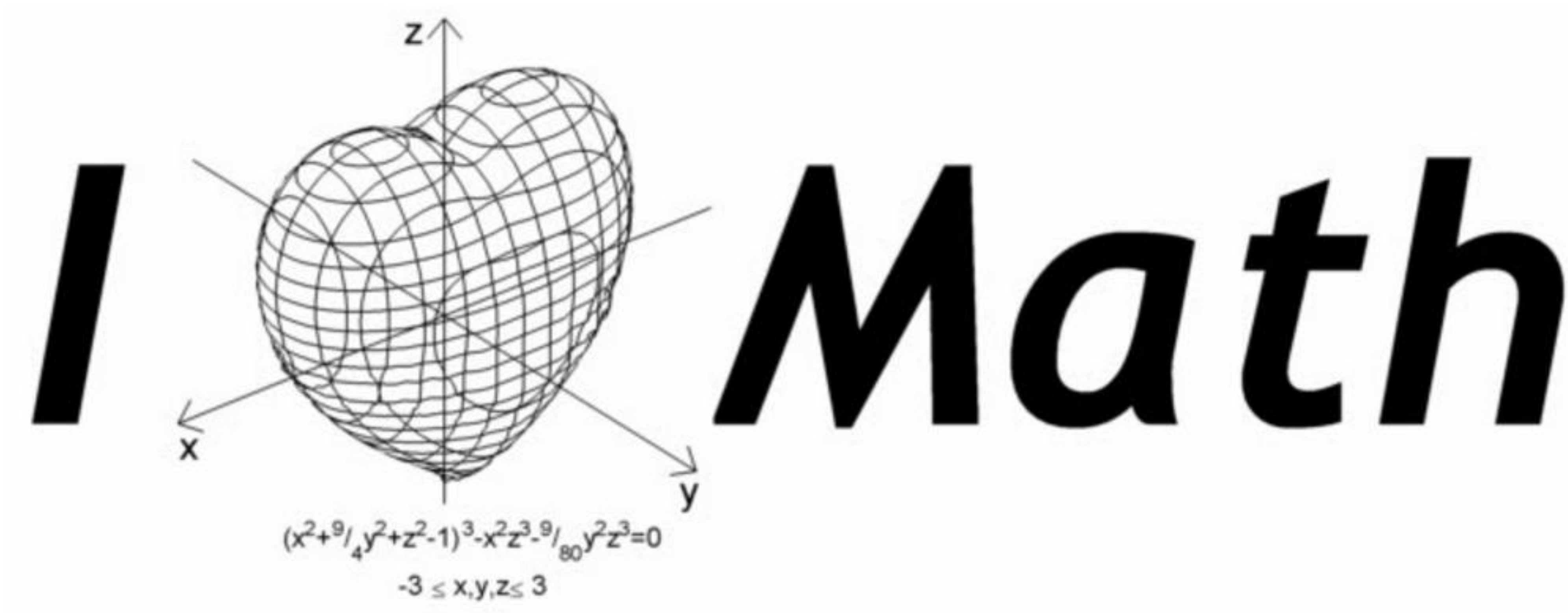
Hochschule **RheinMain**
University of Applied Sciences
Wiesbaden Rüsselsheim

Kapitel 5

Differenzialrechnung mehrdimensionaler Funktionen

Inhalt

5.1 Mehrdimensionale Funktionen



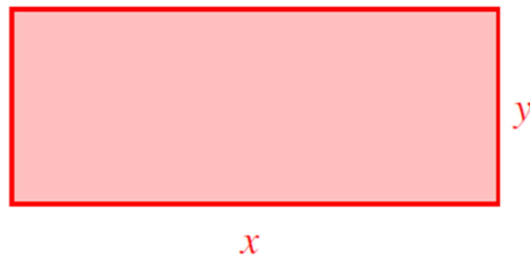
Warum Funktionen von mehreren Variablen?

Viele Funktionen in Naturwissenschaft, Technik und Informatik hängen von mehreren Variablen ab.

Einfache Beispiele:

(a) Der Flächeninhalt eines Rechtecks hängt von Länge und Breite ab:

Rechtecksfläche = $f(x, y) = x \cdot y$



(b) Die Temperatur in diesem Raum hängt von allen drei räumlichen Koordinaten ab: $T(x, y, z)$

Praxisrelevante Beispiele

Computergrafik: Darstellung von Flächen im Raum

Information Retrieval: Dokumente bestehen aus Tausenden von Termen

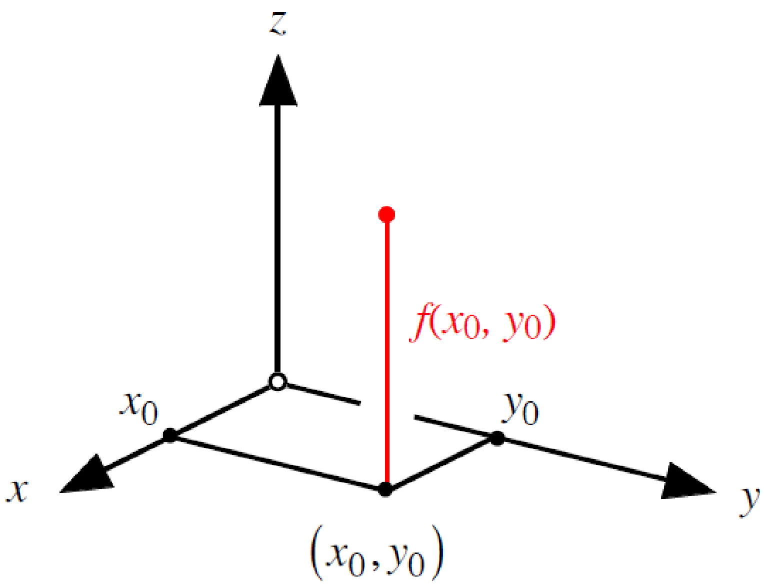
Bildverarbeitung: Bilder bestehen aus Millionen von Pixeln

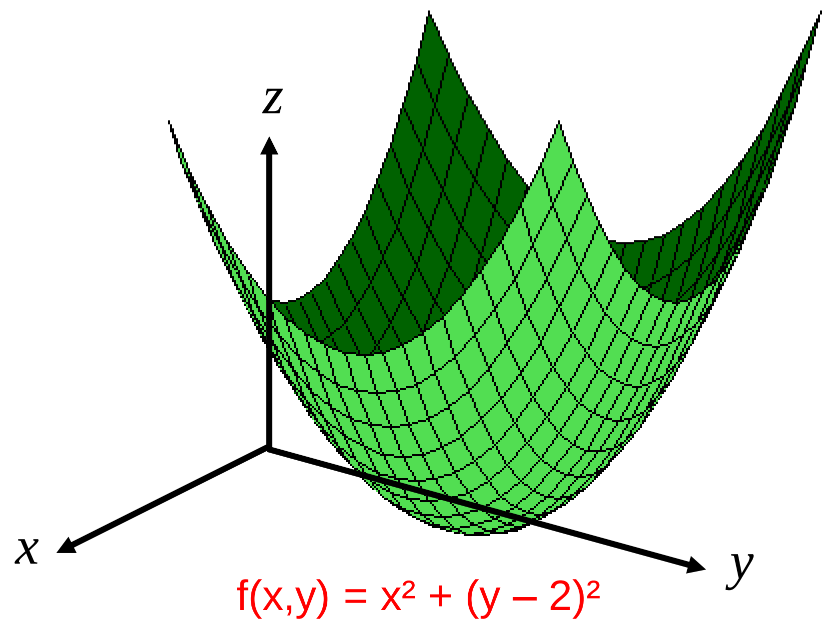
BWL: Betriebswirtschaftliche Größen wie Gewinn und Verbrauch hängen von mehreren Variablen ab

Physik: Phänomene wie Bewegung, Kräfte oder Felder sind mehrdimensional

Darstellung von f(x, y)

Darstellung von $z = f(x,y)$ im räumlichen Koordinatensystem:

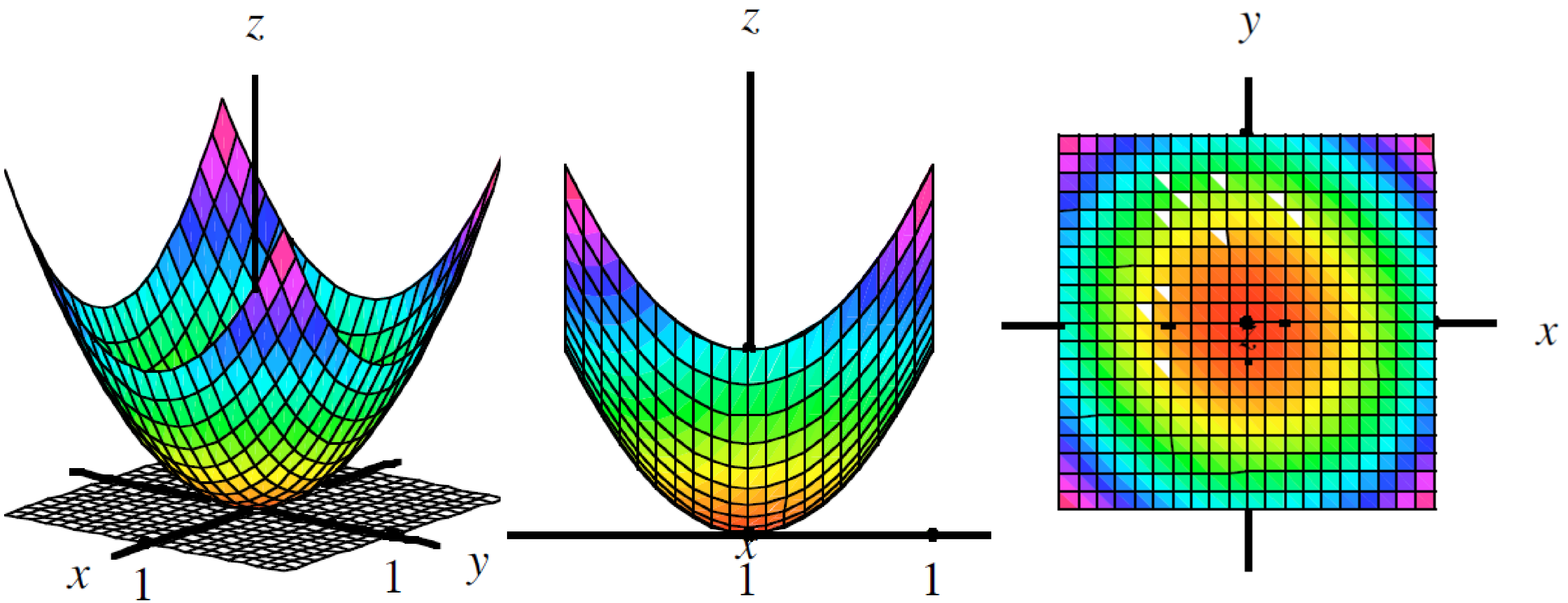




Der gesamte Graph wird zu einer Fläche über der x-y-Ebene.

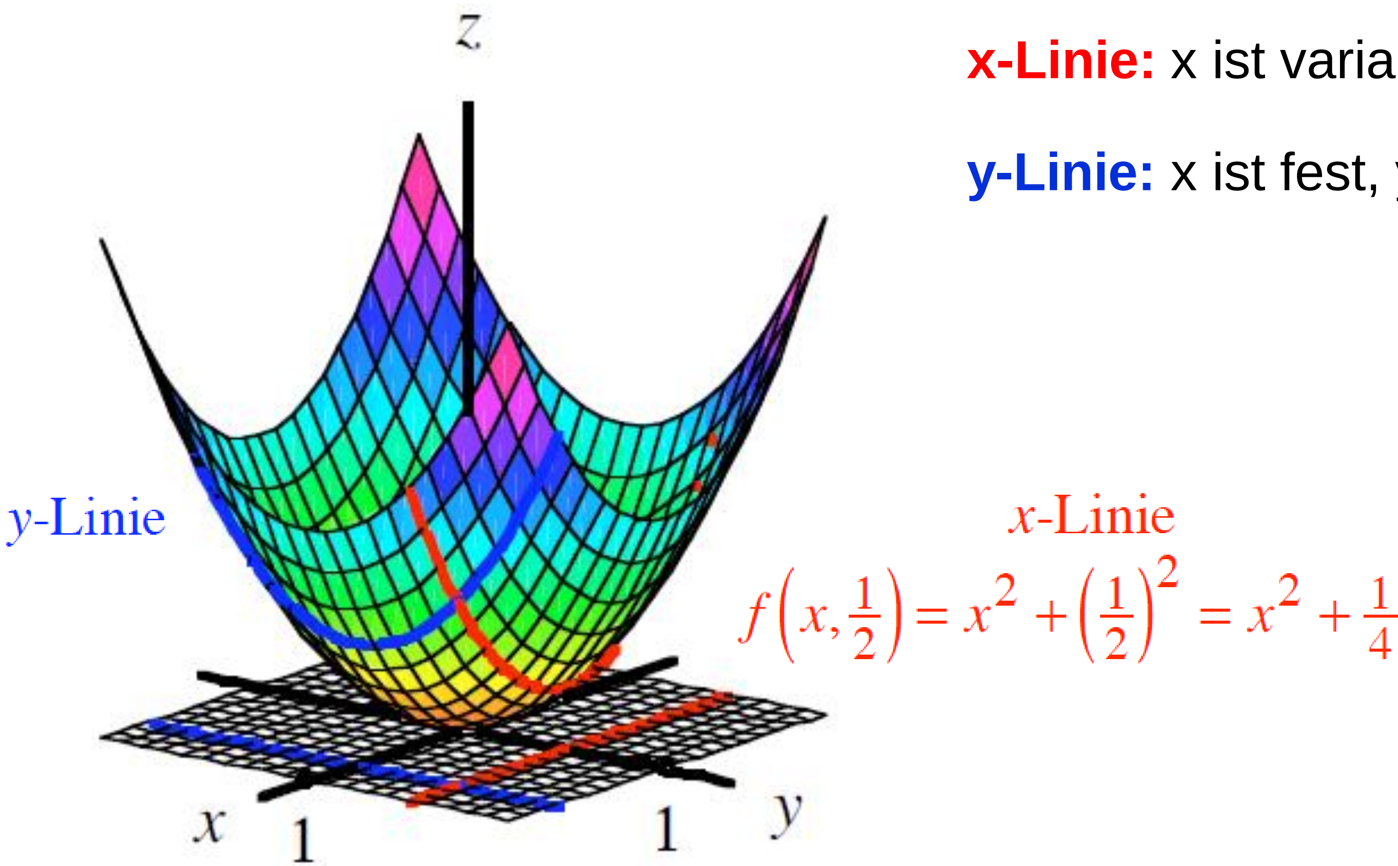
1. Beispiel: Paraboloid

Beispiel: Die Funktion **$f(x, y) = x^2 + y^2$** in verschiedenen Darstellungen:



Schrägbild und Sicht von vorne und von oben

Netzlinien



x-Linie: x ist variabel, y ist fest

y-Linie: x ist fest, y ist variabel

Niveaulinien

Niveaulinien (Höhenlinien) sind horizontale Schnitte auf bestimmten Höhen. Sie werden in der x-y-Ebene dargestellt.

Beispiel: $f(x) = x^2 + y^2$

Niveau 0: $f(x, y) = x^2 + y^2 = 0$ Niveaulinie: Nullpunkt

Niveau 1: $f(x, y) = x^2 + y^2 = 1$ Niveaulinie: Einheitskreis

Niveau 2: $f(x, y) = x^2 + y^2 = 2$ Niveaulinie: Kreis mit Radius $\sqrt{2}$

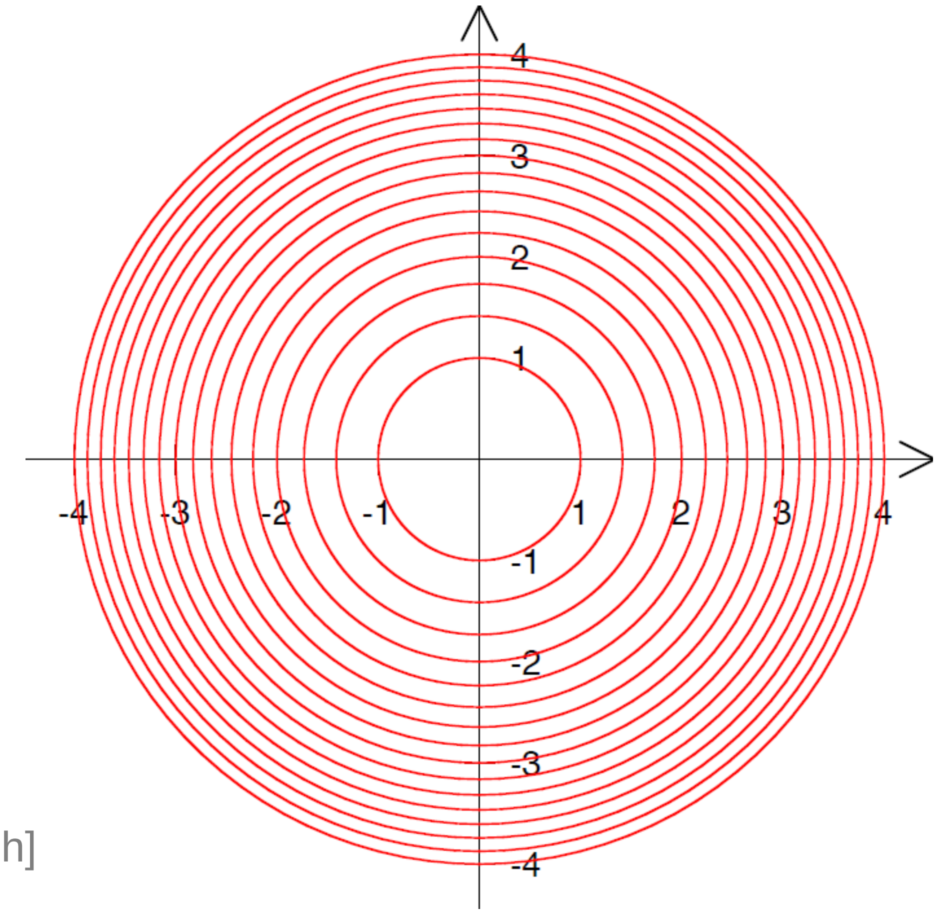
Allgemein:

Niveau c: $f(x, y) = x^2 + y^2 = c$ Niveaulinie: Kreis mit Radius $\sqrt{c}$

Niveaulinien für die Niveaus 1, 2, …, 16

Je dichter die Niveaulinien,
umso steiler ist die Fläche.

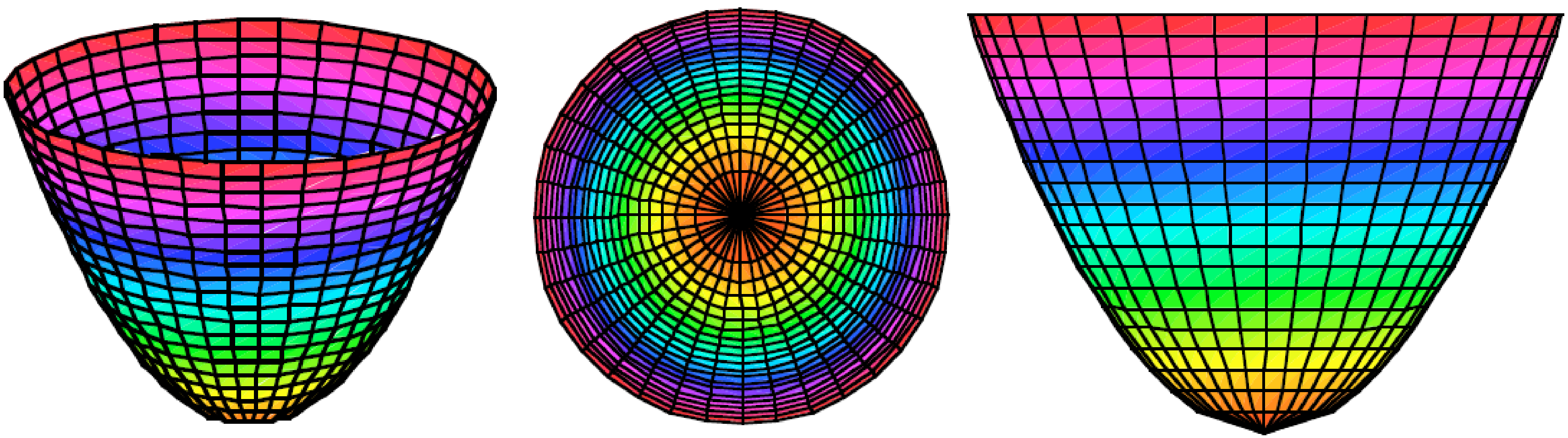
Wenn die Niveaulinien Kreise sind, ist die Fläche rotationssymmetrisch zur z-Achse.



GeoGebra: ImpliziteKurve[f(x,y) - h]

Niveaulinien im Schrägbild, Grund- und Aufriss

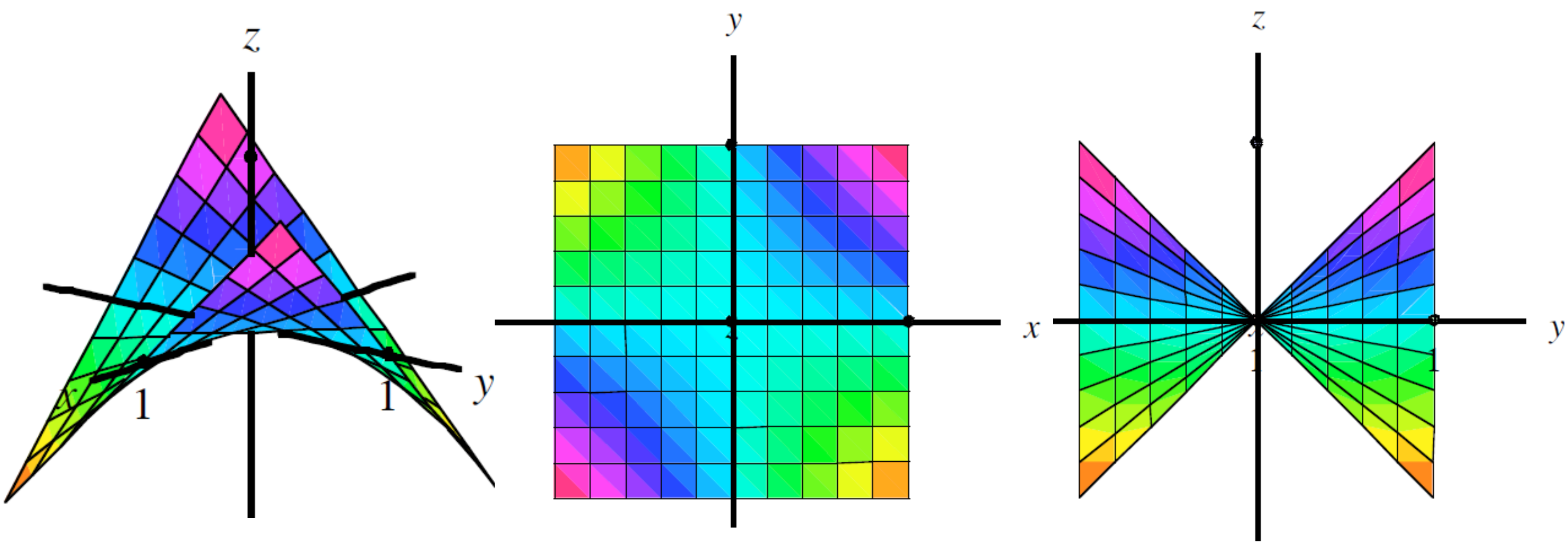
$f(x, y) = x^2 + y^2$ beschreibt ein **Paraboloid**. Im Grundriss sind die Niveaulinien Kreise, im Aufriss horizontale Geraden.



Paraboloid: Schrägbild, Grundriss, Aufriss

2. Beispiel: Sattelfläche

Beispiel: Die Funktion $\mathbf{f(x, y) = x \cdot y}$ heißt **Sattelfläche**. Sie beschreibt z. B. die Fläche eines Rechtecks.



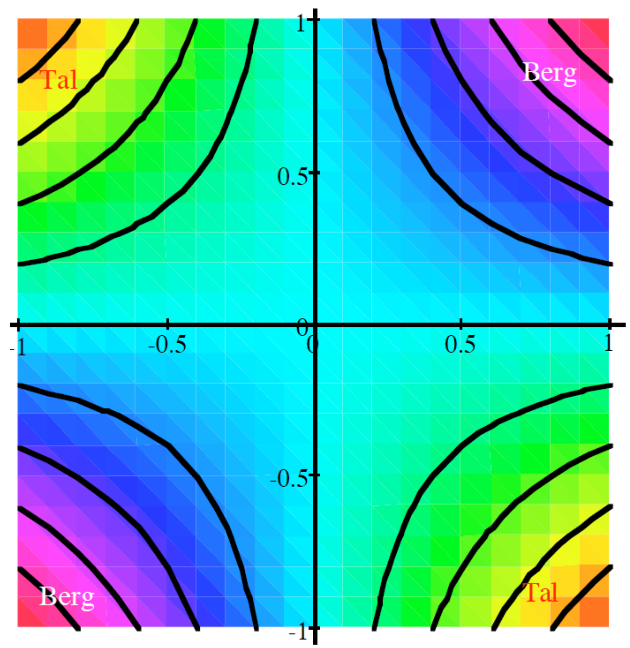
Ansicht, Sicht von oben, Sicht von vorne

Niveaulinien für die Niveaus -1, -0.9, …, 0.9, 1

Für die Niveaulinien der Sattelfläche gilt:

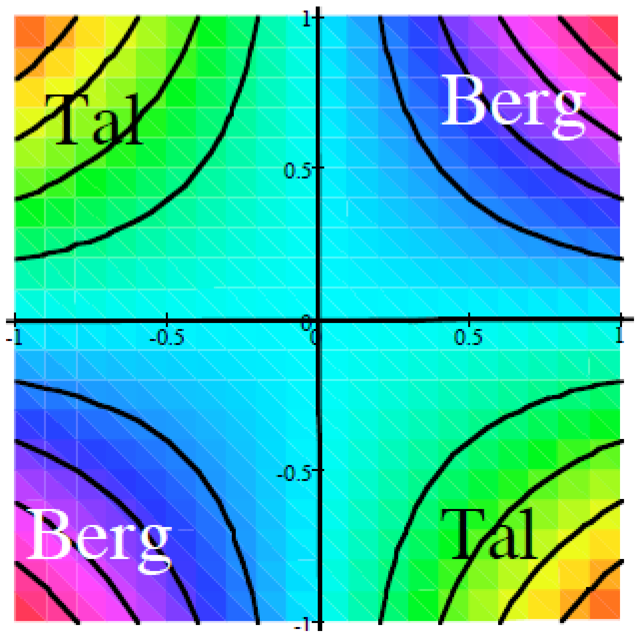
Niveau 0: $f(x, y) = x \cdot y = 0$ → $x = 0$ oder $y = 0$ (Achsen)

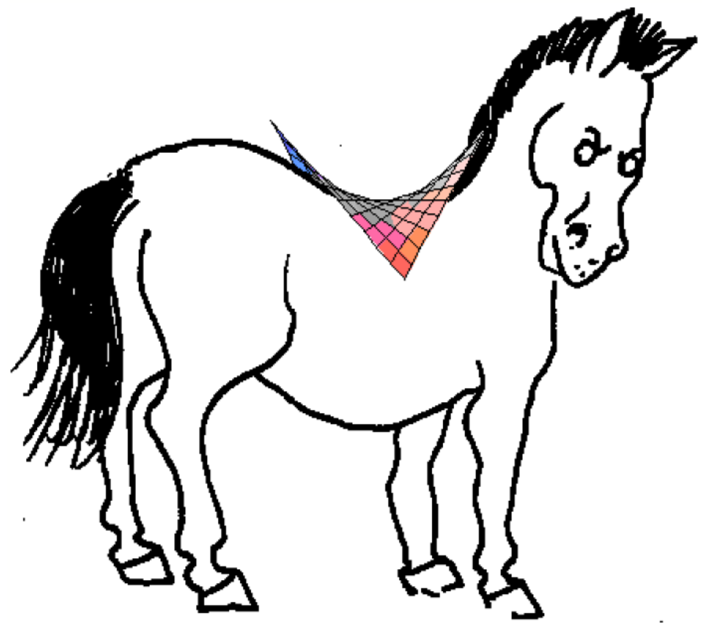
Niveau c: $f(x, y) = x \cdot y = c$ → $y = c/x$ (Hyperbeln)



Sattelfläche: Täler, Berge und ein Sattelpunkt

Im 1. und 3. Quadranten haben wir positive Funktionswerte, im 2. und 4. Quadranten negative Funktionswerte. Der Nullpunkt ist ein **Sattelpunkt**: In zwei Richtungen geht es hinauf, in zwei Richtungen hinunter.





5.2 Partielle Ableitungen

Partielle Ableitungen

Trick: $y = y_0$ setzen und „einfrieren"!

Ableitung von $f(x, y_0)$ bezüglich x an der Stelle (x_0, y_0):

Partielle Ableitung nach x (Ableitung in x-Richtung):

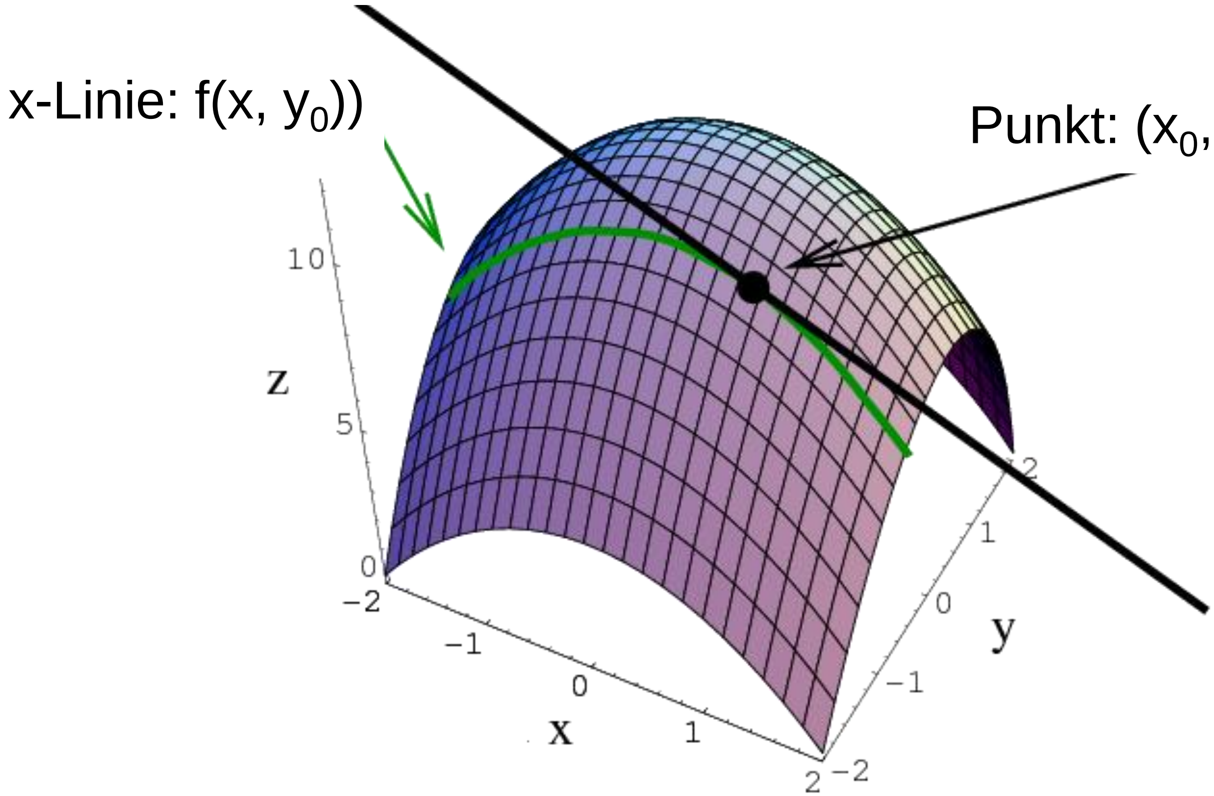
$$f_x(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0) = \lim_{x \to x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0}$$

Analog: **Partielle Ableitung nach y** (Ableitung in y-Richtung):

$$f_y(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = \lim_{y \to y_0} \frac{f(x_0, y) - f(x_0, y_0)}{y - y_0}$$

Geometrische Interpretation der partiellen Ableitung

$f_x(x_0, y_0)$ = (Tangenten-) Steigung der x-Linie



Beispiel und Übung

Beispiele:

(a) $f(x, y) = x^2 + y^2 \Rightarrow$ $f_x(x, y) = 2x$

$f_y(x, y) = 2y$

(b) $f(x, y) = x^3 - 3xy^2 \Rightarrow$ $f_x(x, y) = 3x^2 - 3y^2$

$f_y(x, y) = -6xy$

Übungen:

(c) $f(x, y) = x \cdot y \Rightarrow$ $f_x(x, y) =$

$f_y(x, y) =$

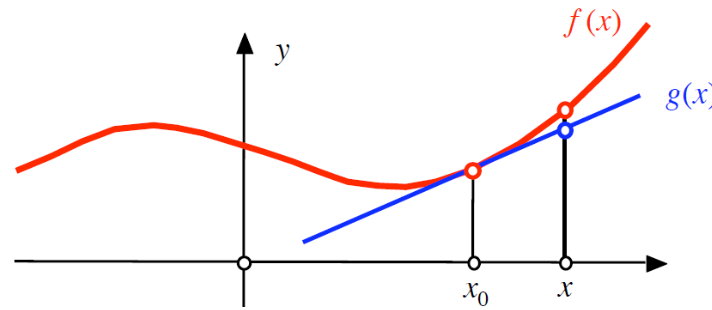
(d) $f(x, y, z) = x \cdot y^2 \cdot z^3 \Rightarrow$ $f_x(x, y, z) =$

$f_y(x, y, z) =$

$f_z(x, y, z) =$

Tangentialebenen

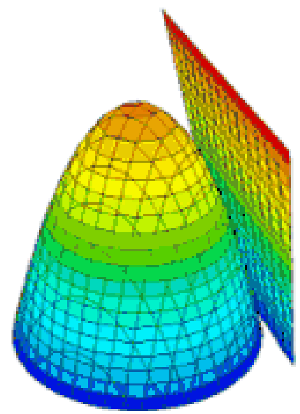
Erinnerung: Tangente an Funktion f(x)



$$g(x) = f(x_0) + f'(x_0)(x - x_0)$$

Analog: **Tangentialebene:**

$$g(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$



Beispiel

Paraboloid: $f(x, y) = x^2 + y^2$ mit $f_x(x, y) = 2x$ und $f_y(x, y) = 2y$

Tangentialebene an der Stelle (0.6, 0.8):

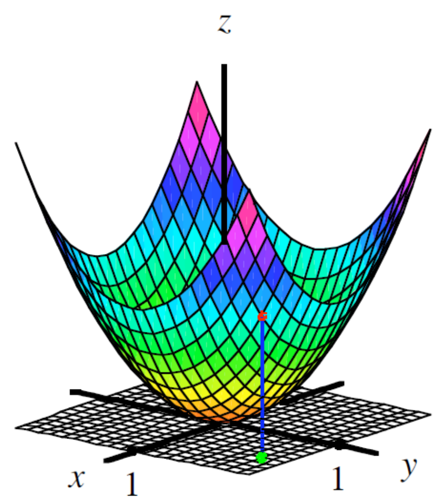
$$\mathbf{g(x, y)} = f(0.6, 0.8) + f_x(0.6, 0.8)(x - 0.6) + f_y(0.6, 0.8)(y - 0.8)$$

$$= 0.6^2 + 0.8^2 + 2 \cdot 0.6 \cdot (x - 0.6) + 2 \cdot 0.8 \cdot (y - 0.8)$$

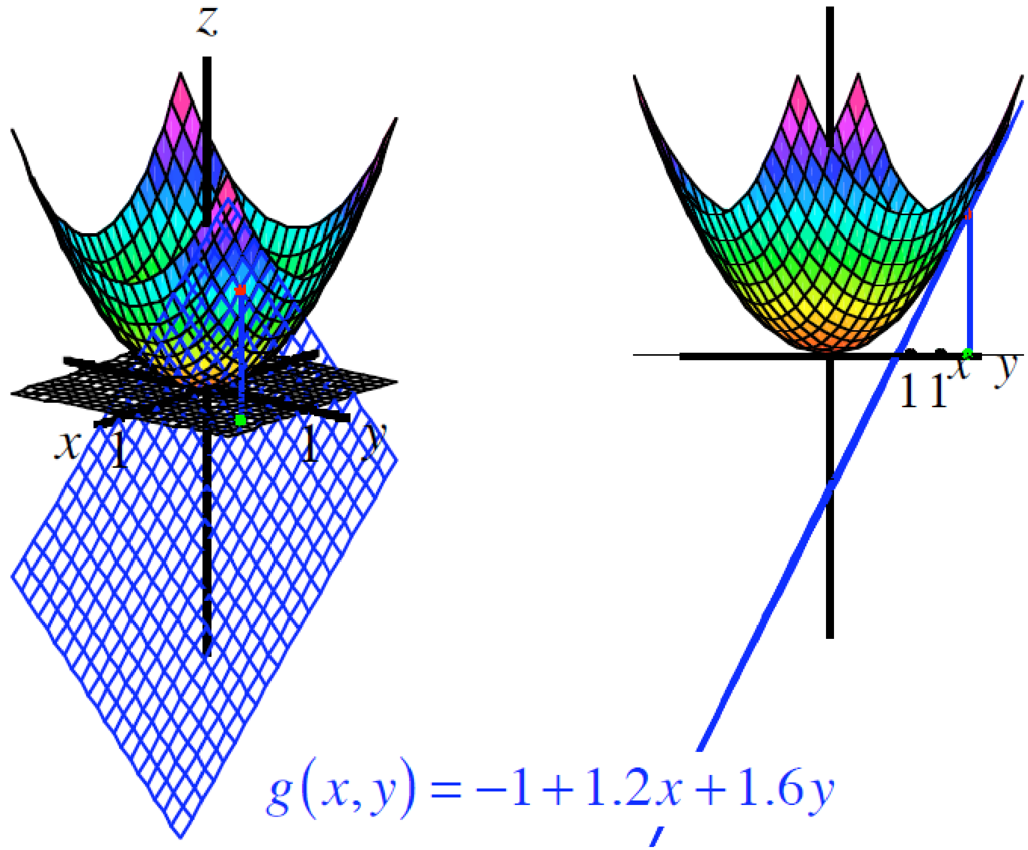
$$= 0.36 + 0.64 + 1.2 \cdot (x - 0.6) + 1.6 \cdot (y - 0.8)$$

$$= 1 + 1.2x - 0.72 + 1.6y - 1.28$$

$$\mathbf{= -1 + 1.2x + 1.6y}$$



Beispiel (Fortsetzung)

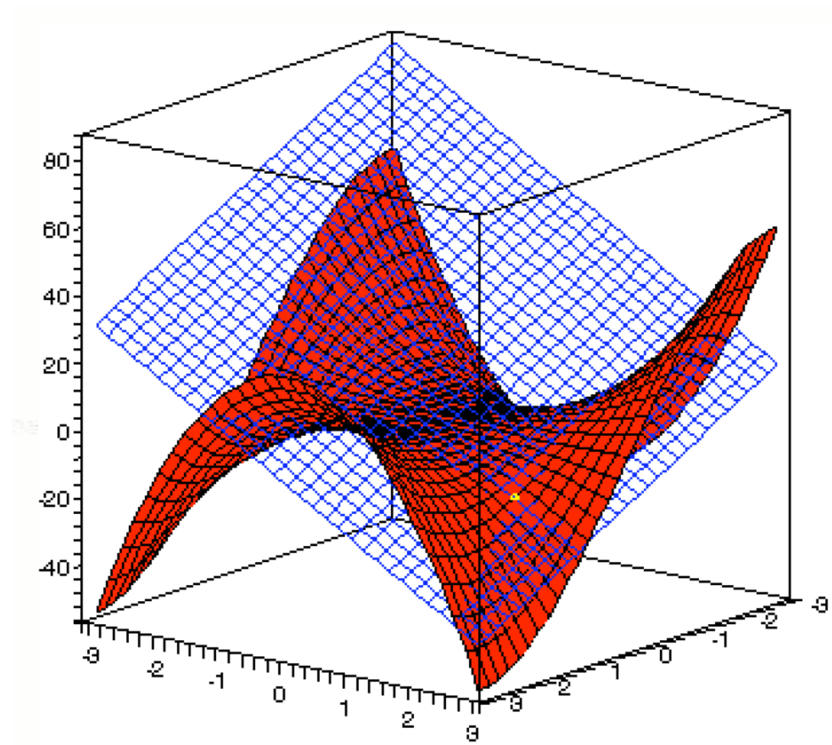


$$g(x,y) = -1 + 1.2x + 1.6y$$

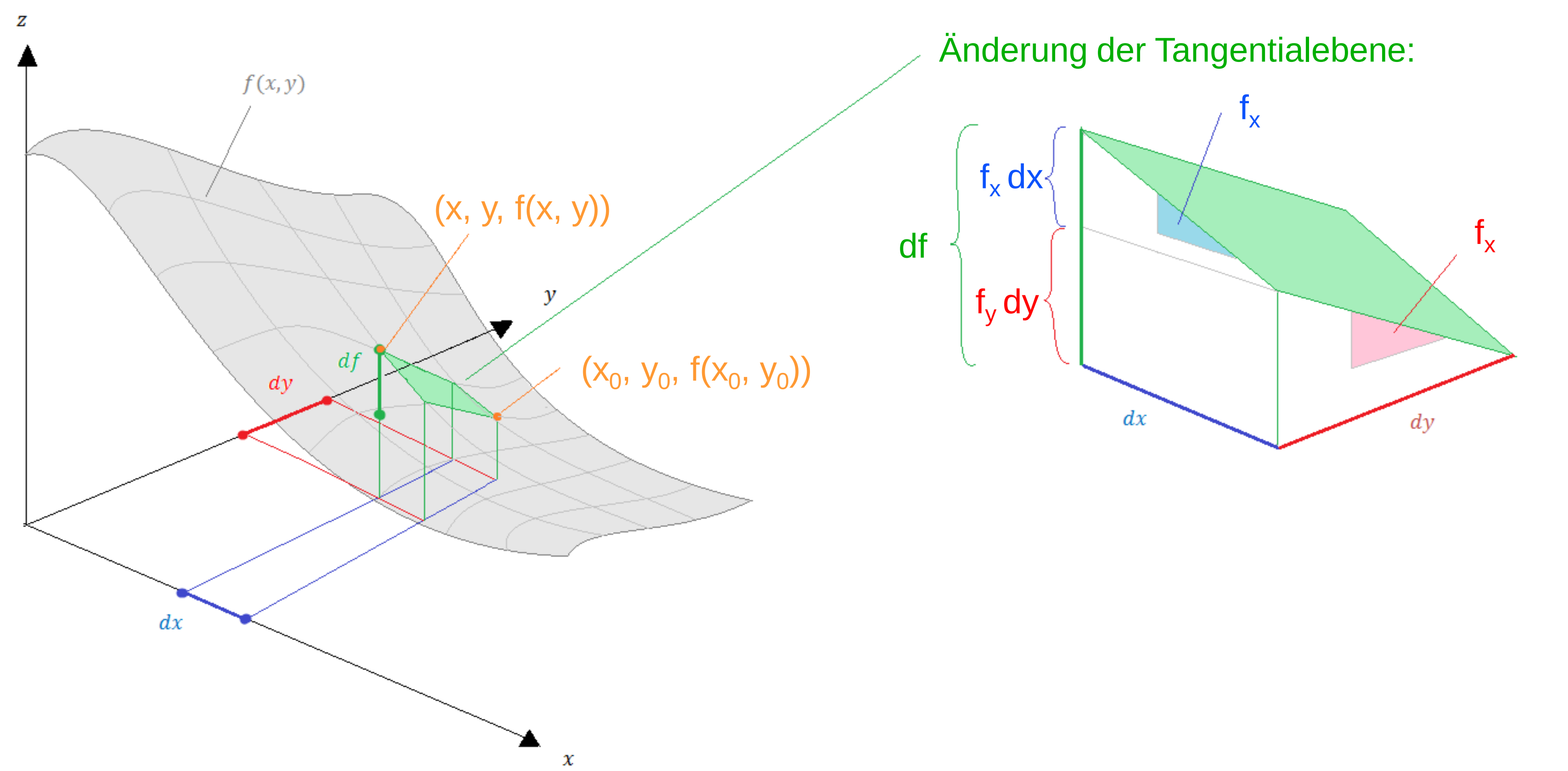
Fläche mit Tangentialebene, verschiedene Ansichten

Übung

Die Funktion $f(x, y) = x^3 - 3xy^2$ heißt **Affensattel**. Bestimmen Sie für f die Tangentialebene an der Stelle (1, 2).



Partielle Ableitungen und Tangentialebene



Gradient

Der **Gradient** einer Funktion f(x, y) ist der Vektor ihrer partiellen Ableitungen:

$$\operatorname{grad}(f) = \begin{pmatrix} f_x \\ f_y \end{pmatrix}$$

Statt grad(f) schreibt man auch: **∇f** (**Nabla-Operator**).

Beispiel: Die Funktion $f(x, y) = x^3 - 3xy^2$ (Affensattel) hat den Gradient

$$\operatorname{grad}(f) = \begin{pmatrix} f_x \\ f_y \end{pmatrix} = \begin{pmatrix} 3x^2 - 3y^2 \\ -6xy \end{pmatrix}$$

An der Stelle (1, 1) ergibt sich der Gradient: $\operatorname{grad}(f)(1,1) = \begin{pmatrix} 0 \\ -6 \end{pmatrix}$

Übung

Berechnen Sie folgende Gradienten an der Stelle (1, 1):

(a) $\mathrm{grad}(3x^2 - 7y)$

(b) $\mathrm{grad}(21x^2y)$

(c) $\mathrm{grad}(x^y)$

Richtungsableitungen

Wir kennen bereits

- die Ableitung in x-Richtung
  (Steigung einer x-Linie): $f_x = \frac{\partial f}{\partial x}$
- die Ableitung in y-Richtung
  (Steigung einer y-Linie): $f_y = \frac{\partial f}{\partial y}$

Fragen:

- Wie steil ist es *schräg* zu diesen Richtungen?

- In welcher Richtung ist die Steigung *am größten*?

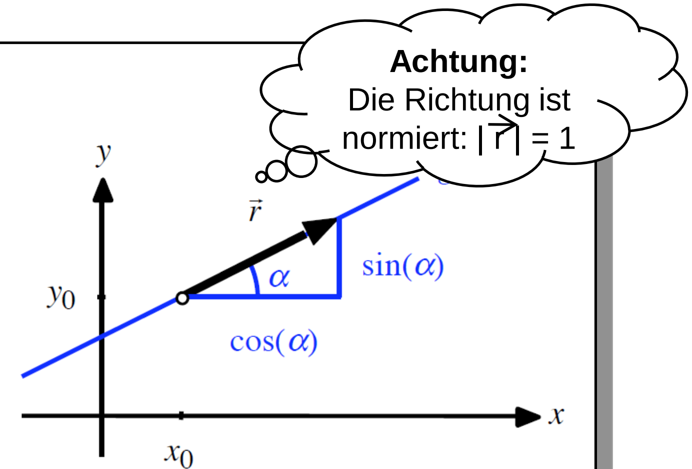
Richtungsableitung

Wir beschreiben eine **beliebige Richtung**

durch einen Vektor $\vec{r} = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}$ mit $|\vec{r}| = 1.$

Die **Richtungsableitung** einer Funktion f(x, y)

in Richtung $\vec{r}$ wird mit $\dfrac{\partial f}{\partial \vec{r}}$ bezeichnet.



Sie ergibt sich als Skalarprodukt der beiden Vektoren grad(f) und $\vec{r}$:

Satz: Für die Richtungsableitung gilt:

$$\frac{\partial f}{\partial \vec{r}} = \operatorname{grad}(f) \cdot \vec{r} = \begin{pmatrix} f_x \\ f_y \end{pmatrix} \cdot \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix} = f_x \cdot \cos(\alpha) + f_y \cdot \sin(\alpha).$$

Übung

Berechnen Sie für die Funktion

$$f(x, y) = x^3 \sin(y) + e^x y^2$$

im Punkt (1, 0, f(1, 0)) die Richtungsableitung in Richtung von $\vec{r} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$.

Folgerungen

Sei γ der Winkel zwischen den Vektoren grad(f) und $\vec{r}$. Dann gilt:

$$\frac{\partial f}{\partial \vec{r}} = \text{grad}(f) \cdot \vec{r} = |\,\text{grad}(f)\,| \cdot |\,\vec{r}\,| \cdot \cos(\gamma)$$

(1.) Für den Fall $\gamma = 0$ folgt: $\gamma = 0 \quad \Rightarrow \quad \cos(\gamma) = 1 \quad \Rightarrow \quad \frac{\partial f}{\partial \vec{r}}$ maximal

D. h.: Die Richtungsableitung ist maximal, wenn $\gamma = 0$ ist, wenn also $\vec{r}$ in die Richtung von grad(f) zeigt.

Folgerung 1: Der Gradient zeigt in Richtung des steilsten Anstiegs.

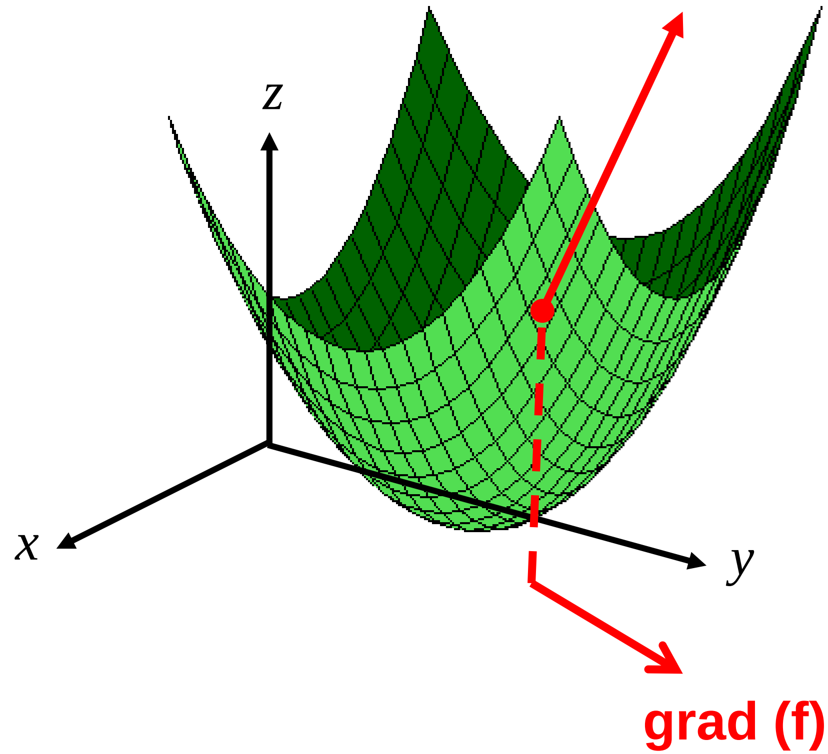
(2.) Für den Fall $\gamma = \pi/2$ (= 90°) folgt: $\gamma = \pi/2 \quad \Rightarrow \quad \cos(\gamma) = 0 \quad \Rightarrow \quad \frac{\partial f}{\partial \vec{r}} = 0$

Folgerung 2: Der Gradient steht senkrecht auf den Niveaulinien.

Eigenschaften des Gradienten

(1.) Der Gradient zeigt in Richtung des steilsten Anstiegs.
Sein Betrag ist die Steigung in dieser (steilsten) Richtung.

(2.) Der Gradient steht senkrecht auf den Niveaulinien.



Beispiel: f(x, y) = xy

$f(x, y) = x \cdot y$

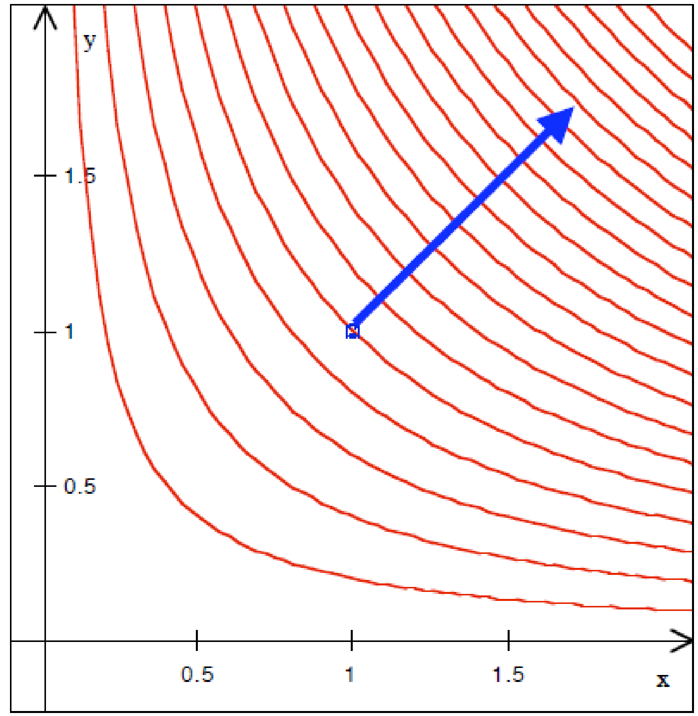
Gradient:

$$\mathrm{grad}(f) = \begin{pmatrix} f_x \\ f_y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$$

Im Punkt (1, 1):

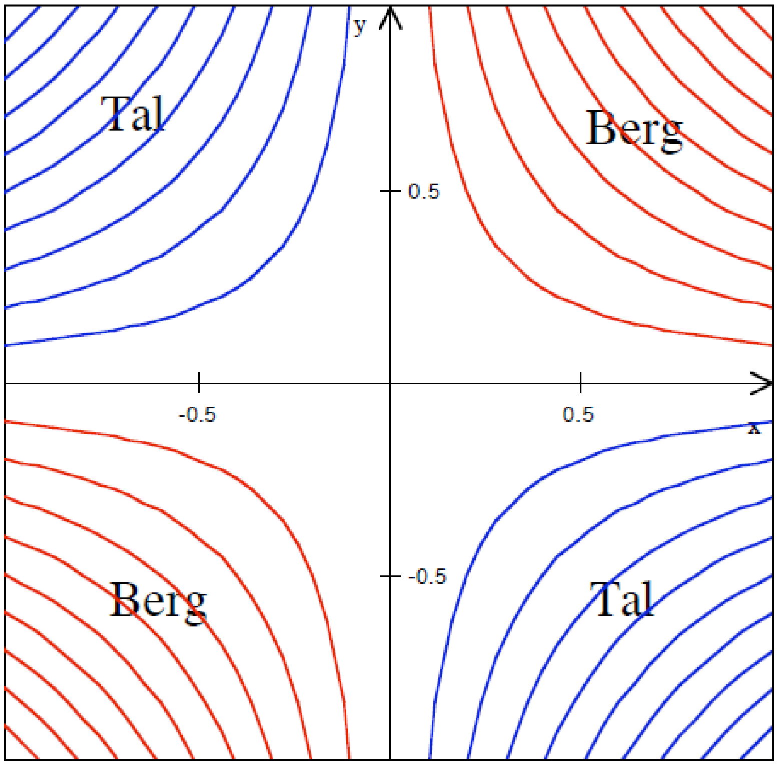
$$\mathrm{grad}(f)(1,1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

- zeigt in Richtung des steilsten Anstiegs, Steigung in dieser Richtung = $\sqrt{2}$
- steht senkrecht auf der Niveaulinie

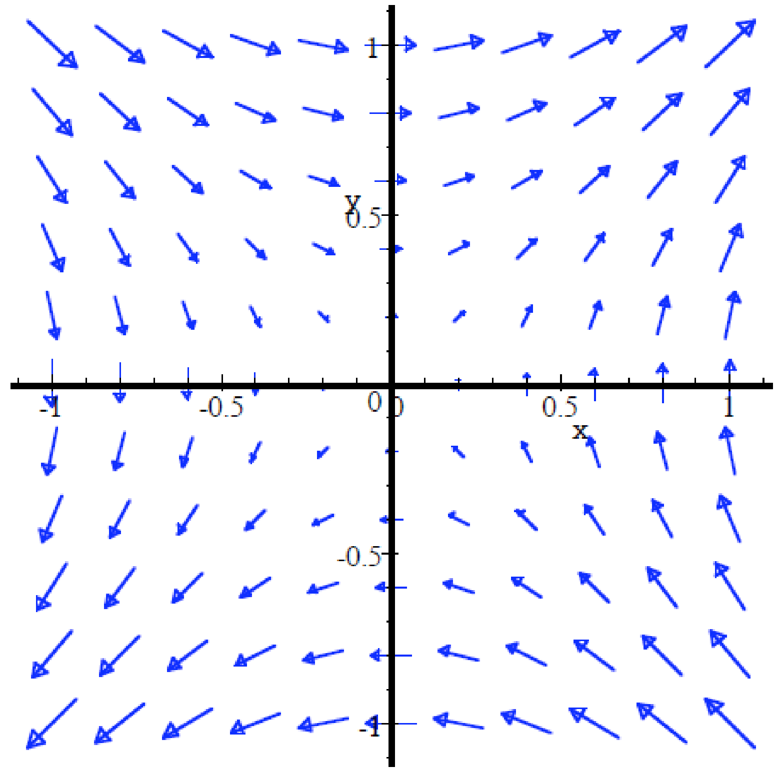


Niveaulinien von $f(x, y) = xy$

Gradientenfeld von f(x, y) = xy



Niveaulinien



Gradientenrichtung

Übung

Berechnen Sie die Richtung des größten Anstiegs der Funktion

$$f(x, y) = x^2 - y^2$$

im Punkt (1, 1, f(1, 1)) und die Steigung in dieser Richtung.

Höhere partielle Ableitungen

Die partiellen Ableitungen f_x und f_y können erneut nach x oder y abgeleitet werden.

Zweimalige partielle Ableitung nach x: $$f_{xx} = \frac{\partial^2 f}{\partial x^2}$$

Einmal nach x und einmal nach y abgeleitet: $$f_{xy} = \frac{\partial^2 f}{\partial x \partial y}$$

Zweimalige partielle Ableitung nach y: $$f_{yy} = \frac{\partial^2 f}{\partial y^2}$$

Beispiel

$$f(x,y) = x^4 y^7$$

$$f_x(x,y) = 4x^3y^7 \qquad f_y(x,y) = 7x^4y^6$$

$$f_{xx}(x,y) = 12x^2y^7 \qquad f_{xy}(x,y) = 28x^3y^6 \qquad f_{yy}(x,y) = 42x^4y^5$$

Satz: Die Reihenfolge der partiellen Ableitungen spielt keine Rolle.

$$f_{xxx}(x,y) = 24xy^7 \qquad f_{xxy}(x,y) = 84x^2y^6 \qquad f_{xyy}(x,y) = 168x^3y^5 \qquad f_{yyy}(x,y) = 210x^4y^4$$

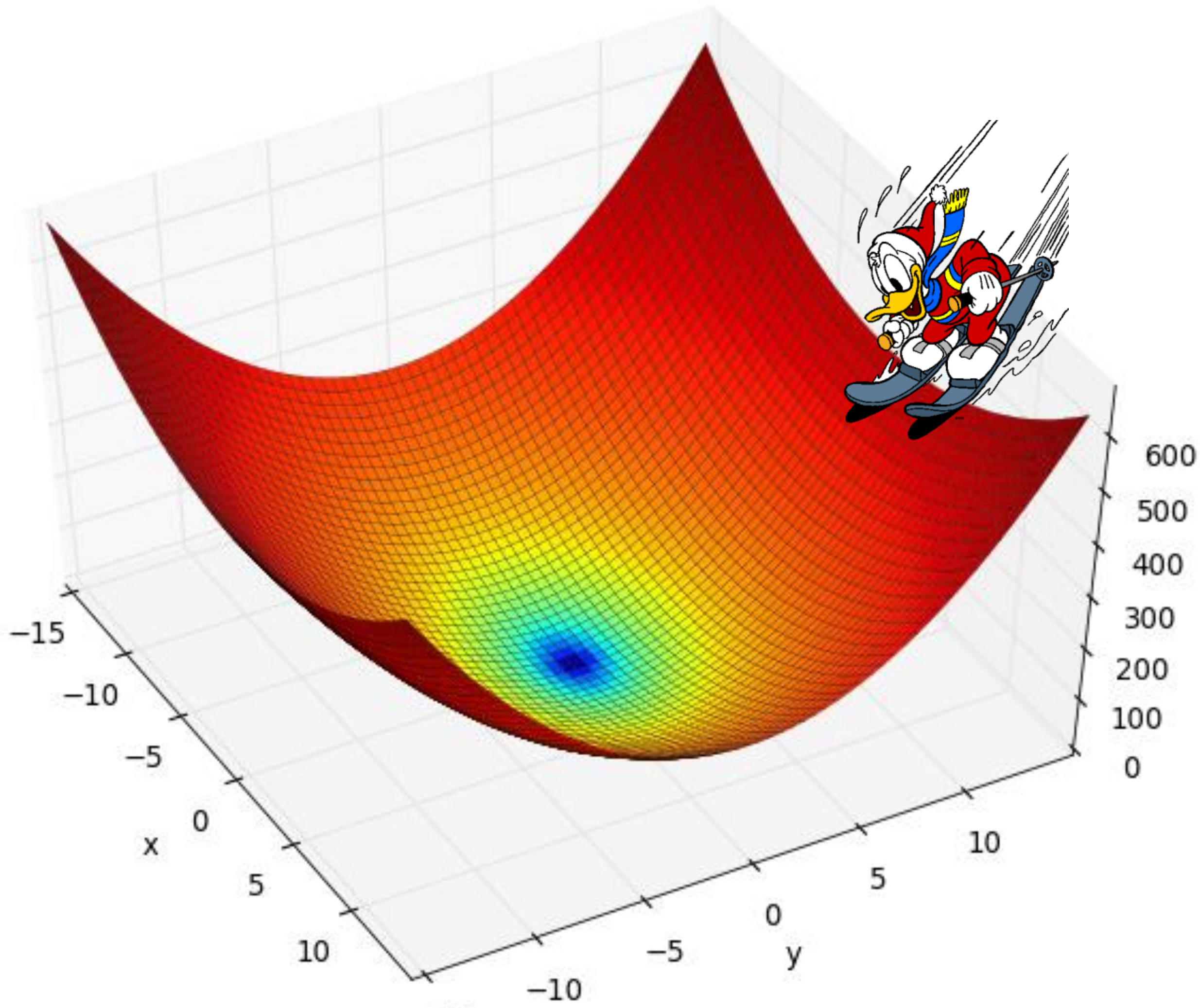
Übung

Bestimmen Sie f_x, f_y, f_{xy}, f_{yx}:

(a) $f(x, y) = x^2y + xy^3 + 5xy$

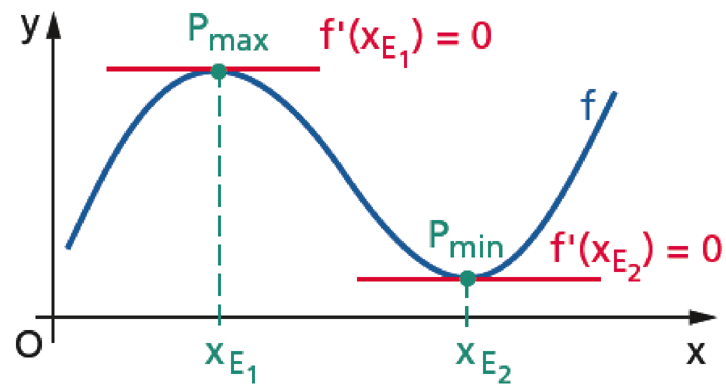
(b) $f(x, y) = \cos(2x) \cdot e^{5y}$

5.3 Extrema



Wdh.: Extrema von f(x)

Erinnerung an Funktionen f(x)
***einer* Variablen:**



Kriterium für lokale Extrema:

- Notwendig:

  $f'(x_E) = 0$ („waagerechte Tangente")

- Hinreichend:

  $f''(x_E) > 0$ („linksgekrümmt") $\Rightarrow$ Minimum in x_E

  $f''(x_E) < 0$ („rechtsgekrümmt") $\Rightarrow$ Maximum in x_E

Extrema von f(x, y): notwendig

Für Funktionen f(x, y) von *zwei* Variablen gelten analoge Kriterien.

Notwendiges Kriterium für lokale Extrema:

Wenn eine Funktion f(x, y) an der Stelle (x_0, y_0) ein lokales Extremum hat, dann gilt an dieser Stelle $f_x = 0$ und $f_y = 0$, also

$$\operatorname{grad}(f) = \vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Anstelle von f' müssen jetzt also *beide* ersten Ableitungen f_x und f_y gleich Null sein. Um mögliche Extrempunkte zu finden, muss daher ein Gleichungssystem mit zwei Gleichungen gelöst werden.

Hesse-Matrix

Anstelle von f'' benötigen wir für das hinreichende Kriterium die *vier* zweiten Ableitungen: f_{xx}, f_{xy}, f_{yx} und f_{yy}.

Diese werden in der **Hesse-Matrix** zusammengefasst:

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

Entscheidend für die Art des Extremums ist die **Determinante** der Hesse-Matrix:

$$\Delta = \det \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = f_{xx} \cdot f_{yy} - f_{xy}^2$$

Übung

Bestimmen Sie die Determinante der Hesse-Matrix von $f(x, y) = x^2y^3$.

$f_x =$ $\Rightarrow$ $f_{xx} =$, $f_{xy} =$

$f_y =$ $\Rightarrow$ $f_{yy} =$, $f_{yx} =$

$$\Delta = \det\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = f_{xx} \cdot f_{yy} - f_{xy}^2$$

Extrema von f(x, y): hinreichend

Hinreichendes Kriterium für lokale Extrema:

Wenn für eine Funktion f(x, y) an der Stelle (x_0, y_0) gilt grad(f) = 0, dann liegt in (x_0, y_0) unter folgenden Bedingungen folgendes vor:

$\Delta > 0$ und $f_{xx} > 0$ $\Rightarrow$ isoliertes Minimum

$\Delta > 0$ und $f_{xx} < 0$ $\Rightarrow$ isoliertes Maximum

$\Delta < 0$ $\Rightarrow$ Sattelpunkt

$\Delta = 0$ $\Rightarrow$ keine Aussage möglich

Vorgehen bei der Extremwertsuche: Zunächst werden die Punkte bestimmt, an denen grad(f) = 0 ist (Gleichungssystem lösen). Anschließend werden diese Punkte in Δ und ggf. in f_{xx} eingesetzt.

1. Beispiel

1. Beispiel: **$f(x, y) = x \cdot y$ (Sattelfläche)**

Bestimmung der Punkte mit grad(f) = 0, also mit $f_x = 0$ und $f_y = 0$:

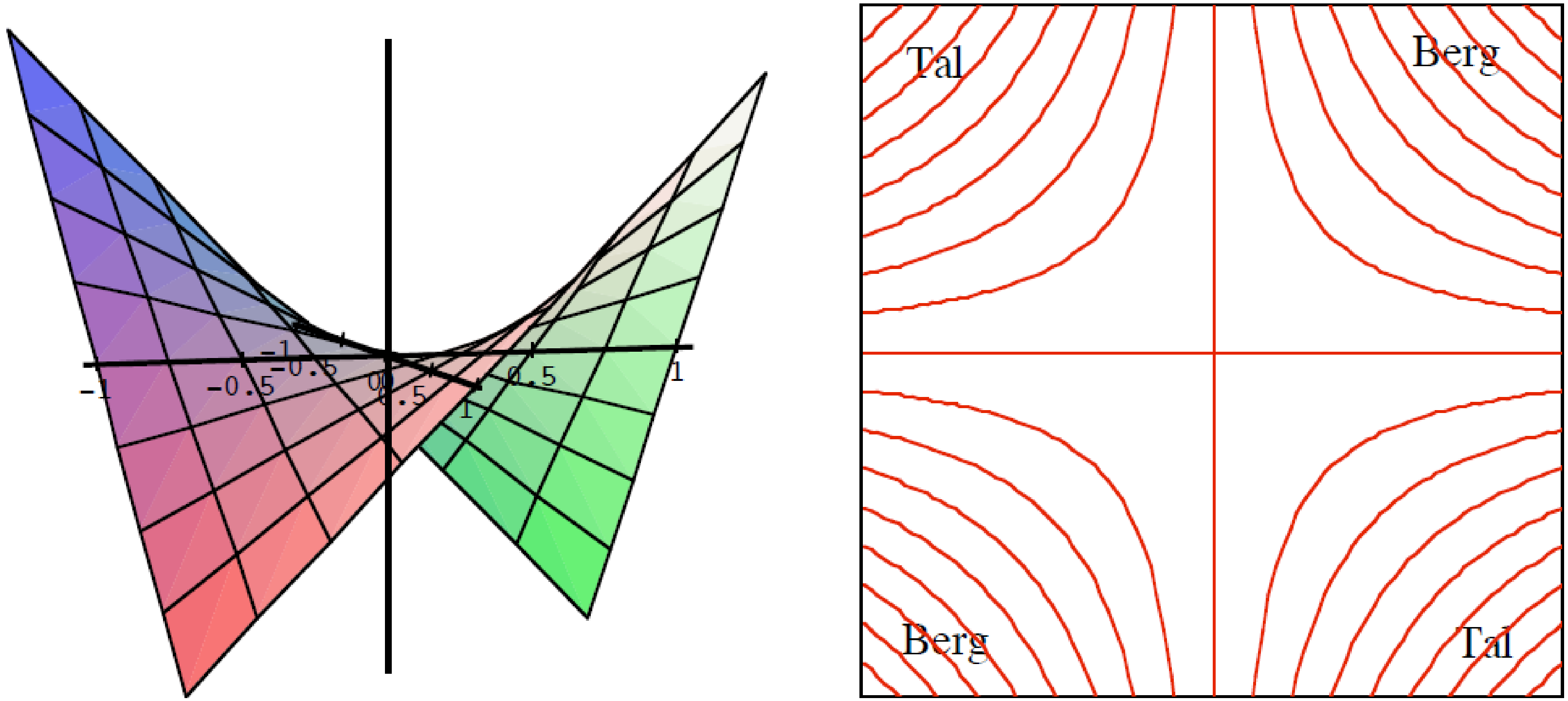
$$\left.\begin{aligned} f_x &= y = 0 \\ f_y &= x = 0 \end{aligned}\right\} \quad \Rightarrow \quad \text{Punkt: } (0, 0)$$

Untersuchen des Punkts (0, 0) mit Δ (und ggf. f_{xx}):

$$\left.\begin{aligned} f_{xx} &= 0 \qquad & f_{xy} &= 1 \\ f_{yx} &= 1 \qquad & f_{yy} &= 0 \end{aligned}\right\} \quad \Delta = \det\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = 0 \cdot 0 - 1^2 = -1$$

Wegen $\Delta < 0$ liegt im Ursprung (0, 0) ein **Sattelpunkt** vor.

1. Beispiel



$f(x,y) = xy$, **Sattelpunkt im Ursprung**

2. Beispiel

2. Beispiel: $f(x, y) = x^2 + y^2$ **(Paraboloid)**

Bestimmung der Punkte mit grad(f) = 0, also mit $f_x = 0$ und $f_y = 0$:

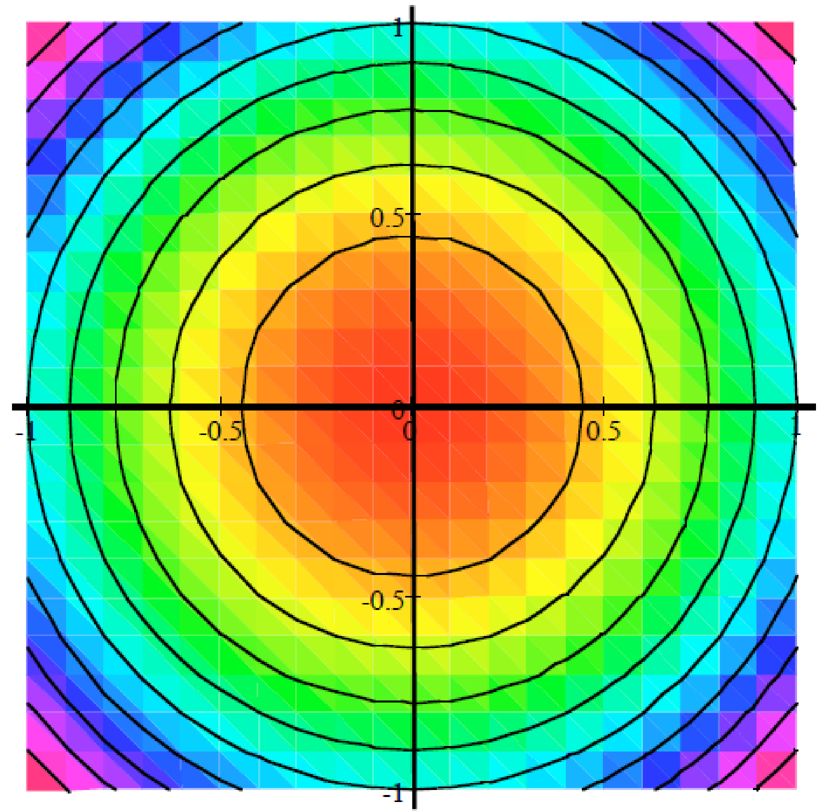
$$\left.\begin{array}{lll} f_x = 2x = 0 & \Rightarrow & x = 0 \\ f_y = 2y = 0 & \Rightarrow & y = 0 \end{array}\right\} \Rightarrow \text{ Punkt: } (0, 0)$$

Untersuchen des Punkts (0, 0) mit Δ (und ggf. f_{xx}):

$$\left.\begin{array}{ll} f_{xx} = 2 & f_{xy} = 0 \\ f_{yx} = 0 & f_{yy} = 2 \end{array}\right\} \quad \Delta = \det\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 2 \cdot 2 - 0^2 = 4 > 0$$

Wegen $\Delta > 0$ und außerdem $f_{xx}(0, 0) = 2 > 0$ befindet sich in (0, 0) ein isoliertes **Minimum**.

2. Beispiel



$f(x,y) = x^2 + y^2$, **isoliertes Minimum im Ursprung**

Übung

Untersuchen Sie folgende Funktion auf Extrema und Sattelpunkte:

$$f(x, y) = 1/3\ x^3 - x^2 + y^3 - 12y$$

Gradientenabstiegsverfahren

Probleme in der Praxis:

- Gleichungen lösen, höhere Ableitungen berechnen: oft schwierig, aufwändig, unmöglich
- Kriterium liefert manchmal keine Aussage

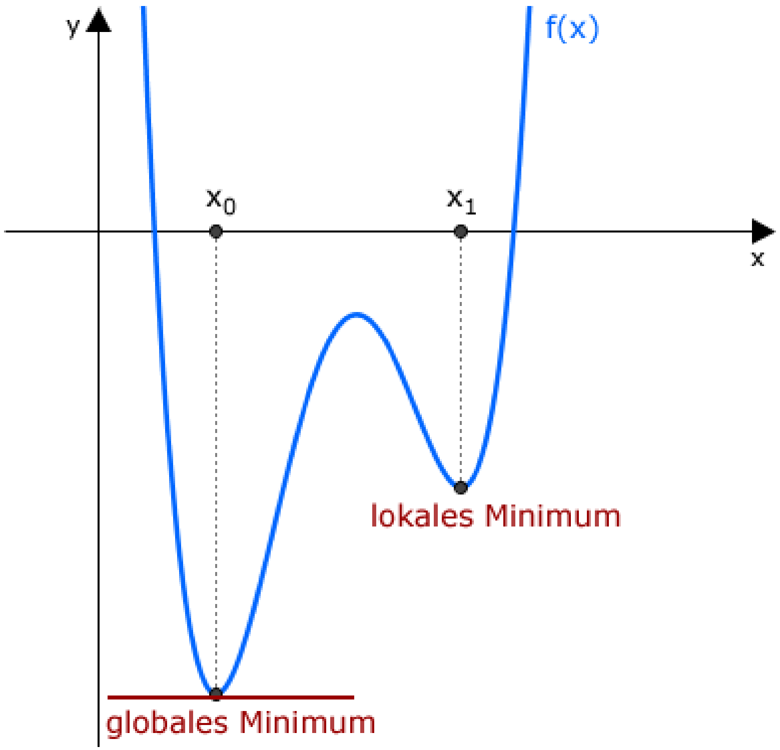
→ Ausweg: Numerische Näherung!

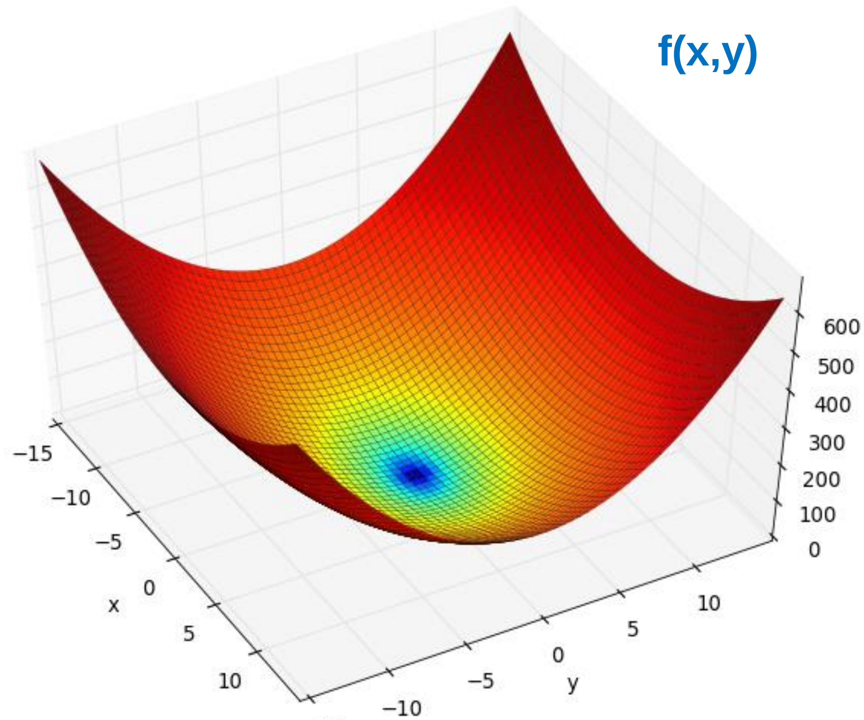
Gradientenabstiegsverfahren:

- auch: „Verfahren des steilsten Abstiegs“, „steepest descent“
- fundamentales Optimierungsverfahren, 1847 von Cauchy vorgestellt
- moderne Anwendungen, u. a. in der künstlichen Intelligenz

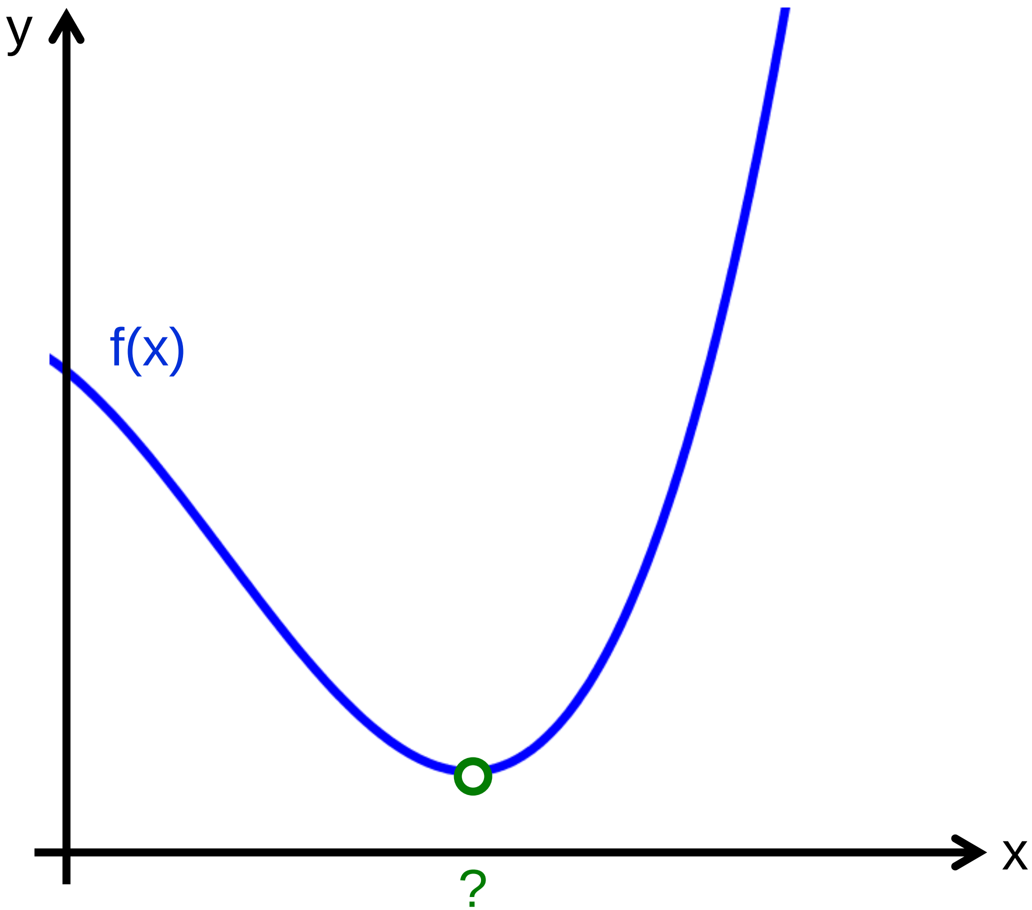
Gradientenabstiegsverfahren

Ziel: numerische Berechnung eines lokalen Minimums

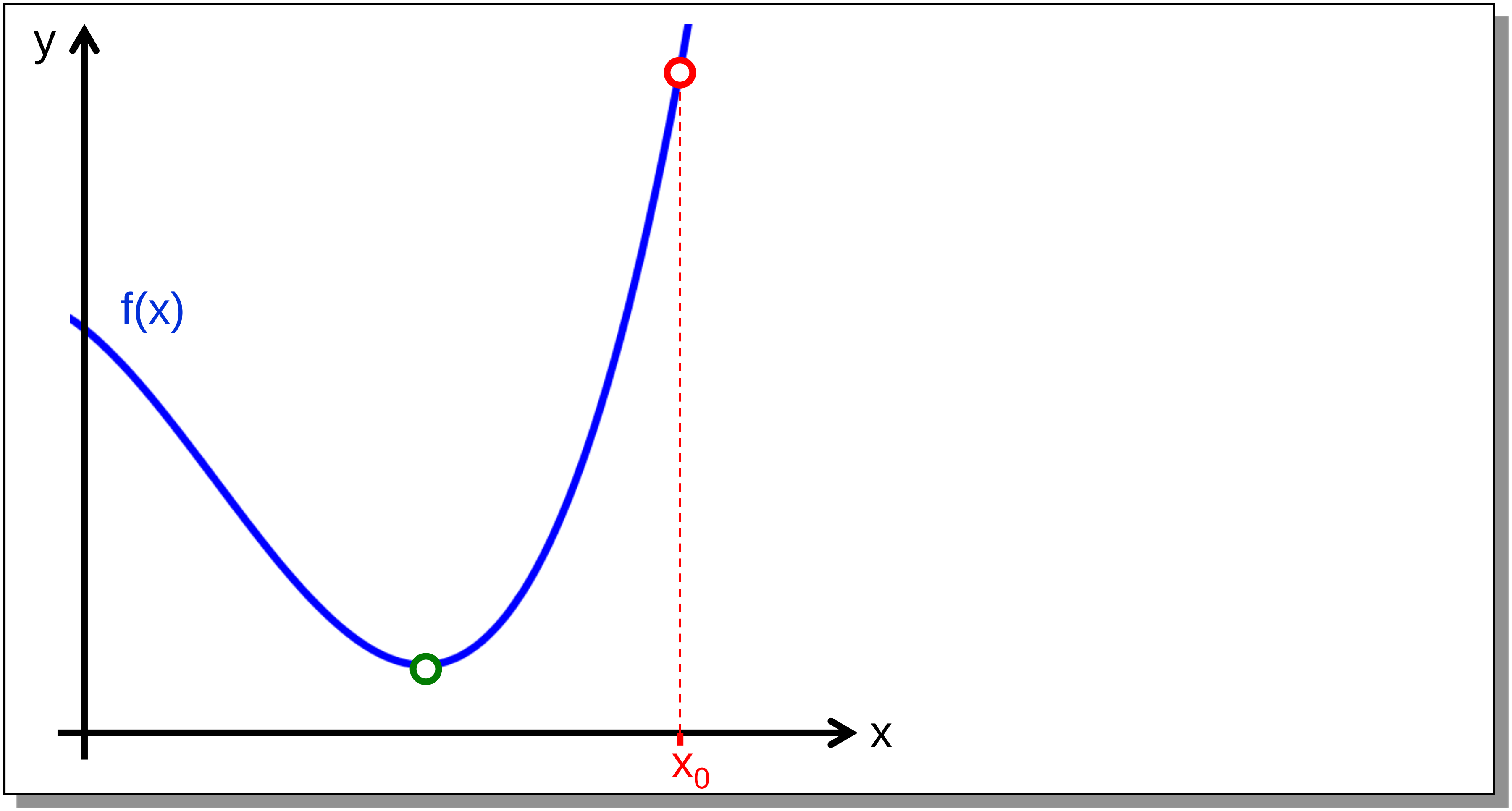




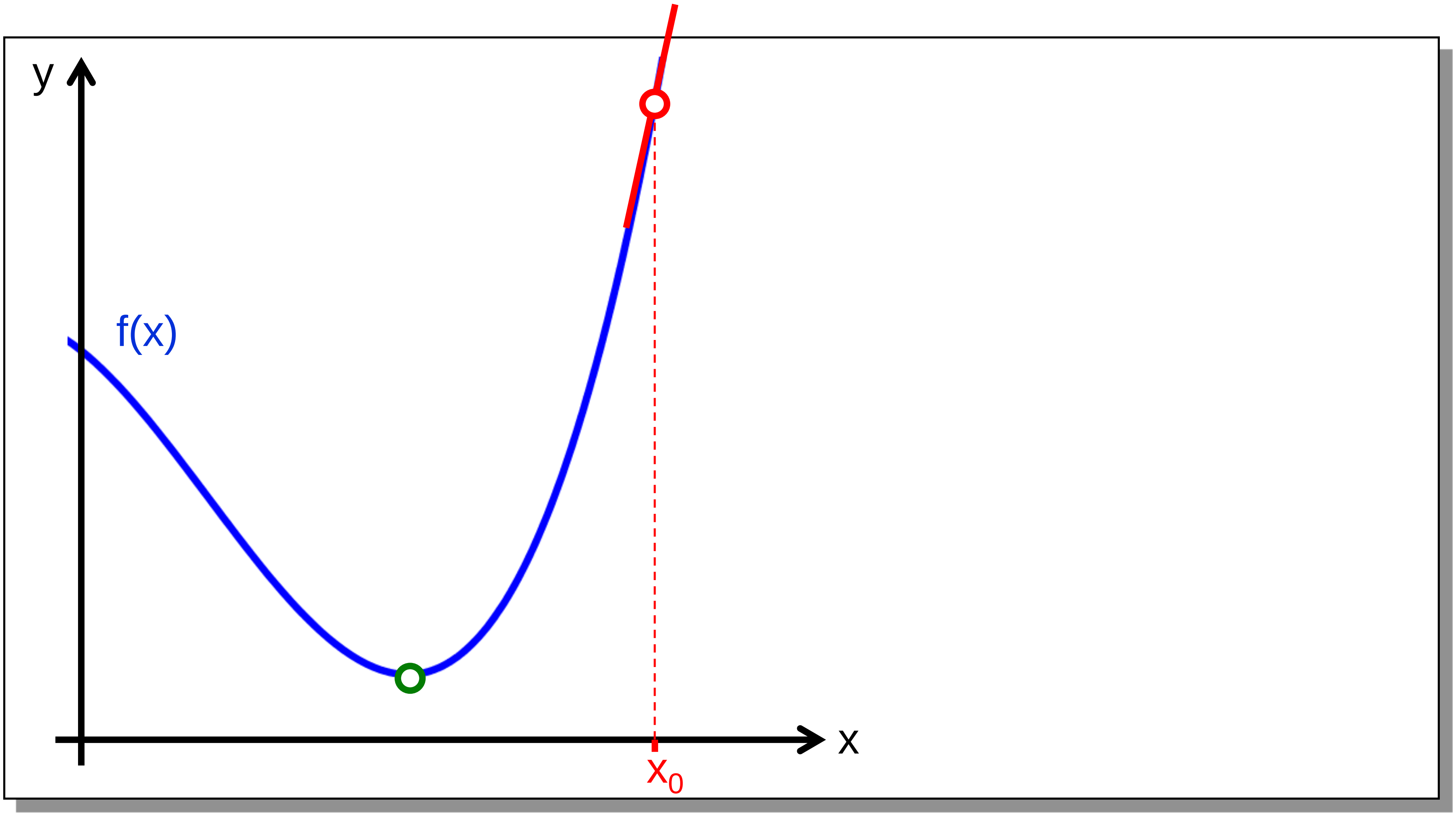
Gesucht: Minimum!



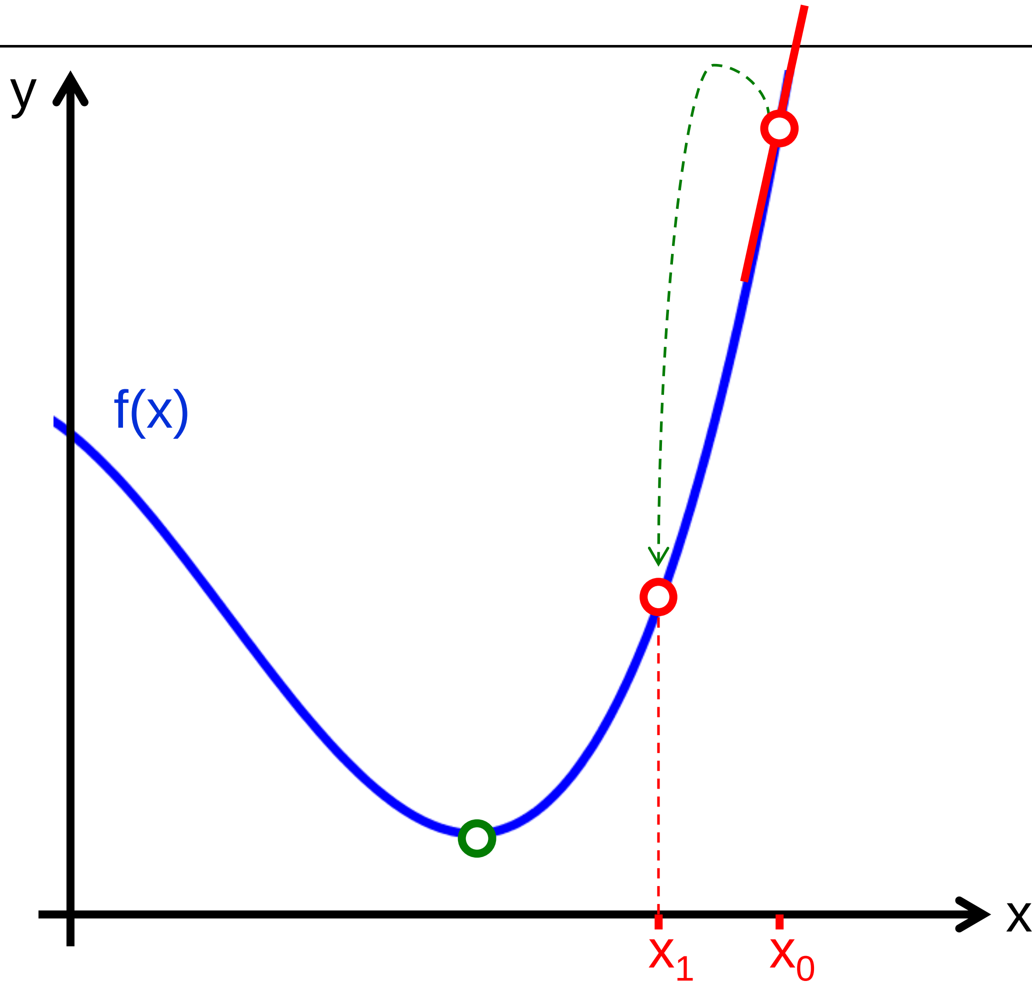
Startstelle wählen



Steigung berechnen



Richtung und Schrittweite?



Steigung positiv

→ nach links gehen

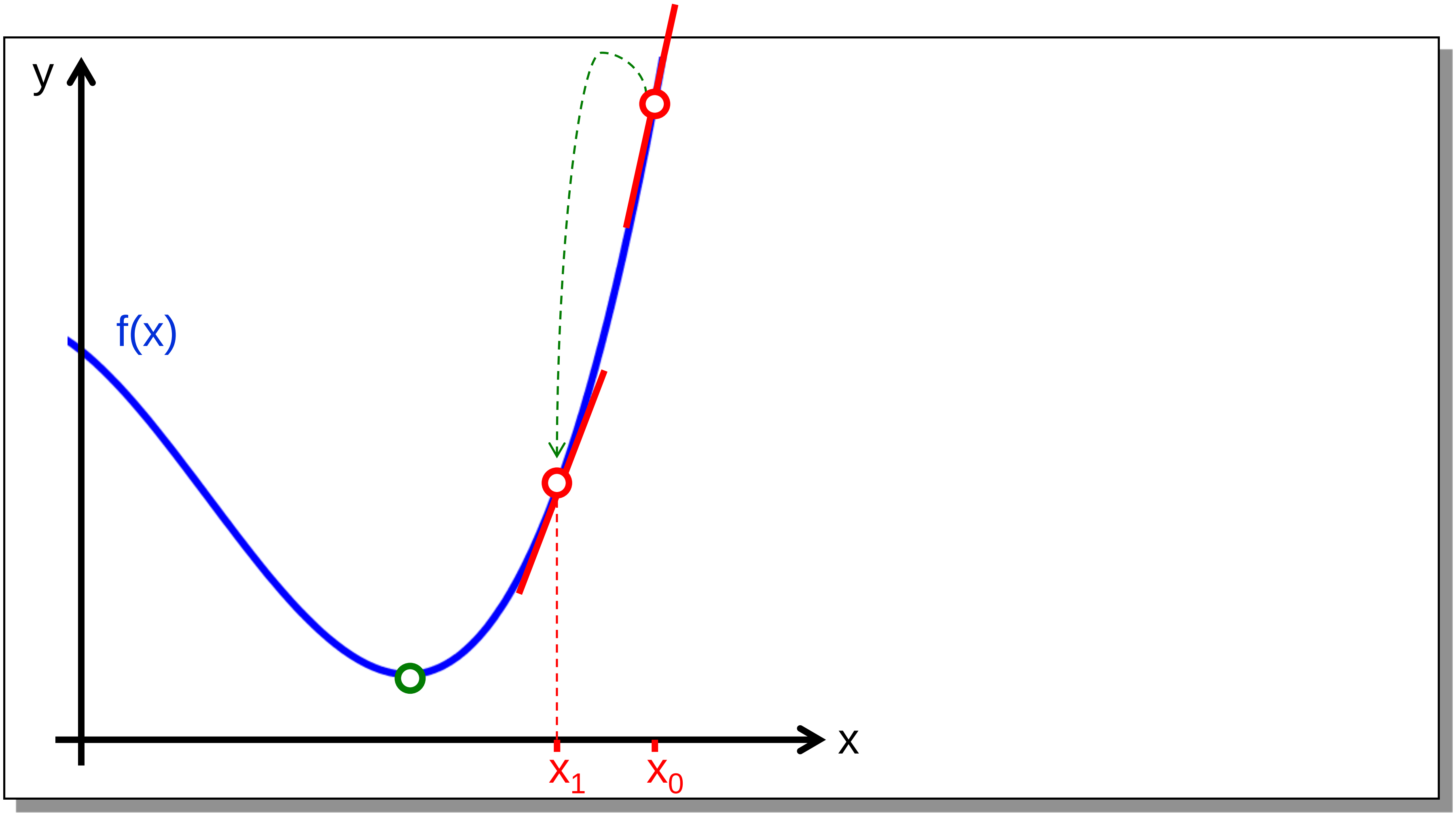
→ $x_{n+1} = x_n - \ldots$

Große Steigung

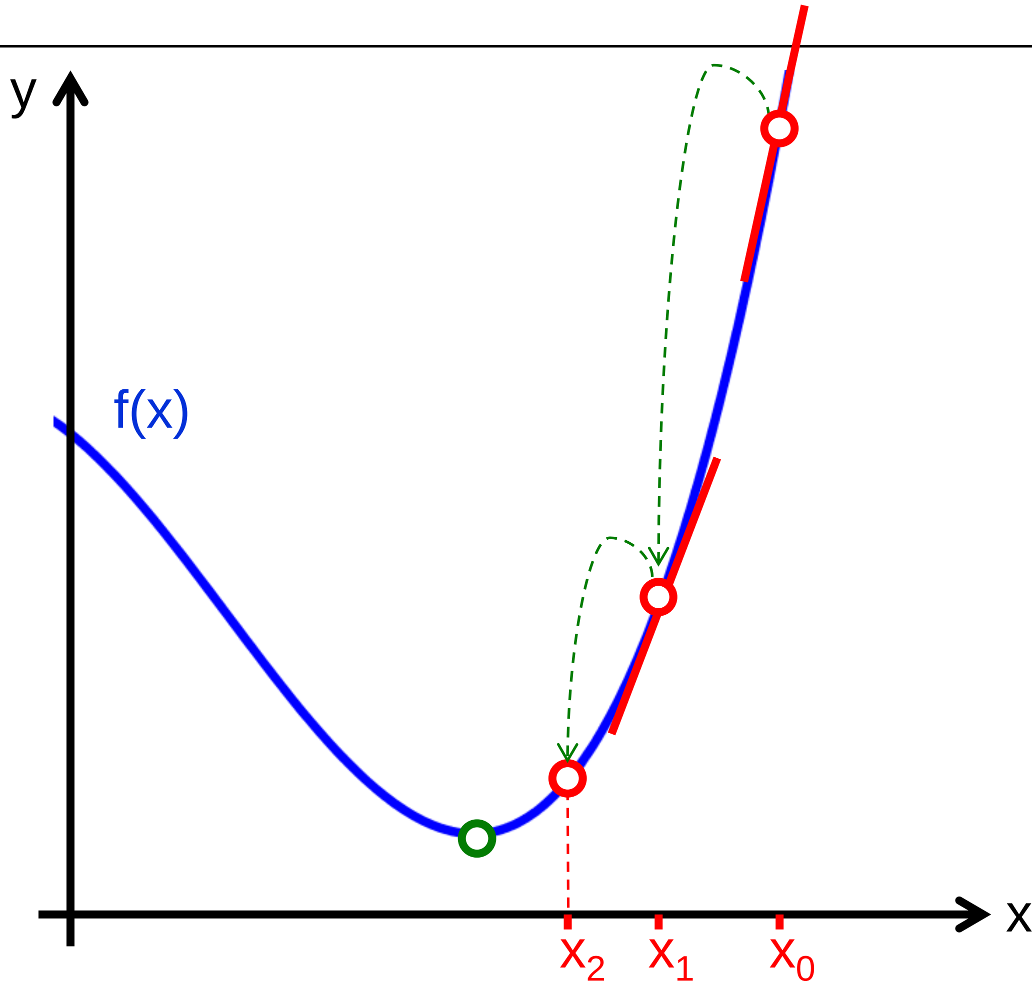
→ Min. weit entfernt

→ große Schritte

Und wieder Steigung berechnen …



Richtung und Schrittweite?



Steigung positiv

→ nach links gehen

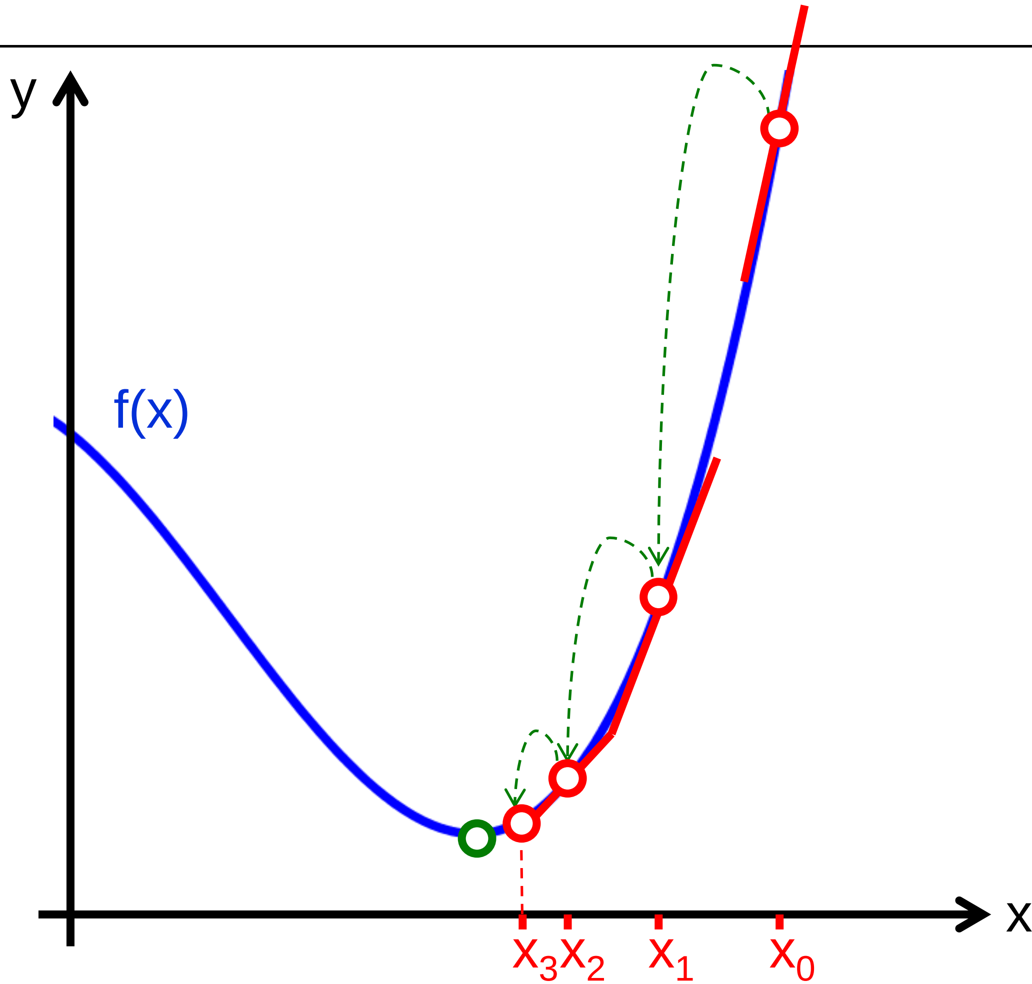
→ $x_{n+1} = x_n - \ldots$

Kleinere Steigung

→ Min. näher

→ kleinere Schritte

Usw.



Steigung positiv

→ nach links gehen

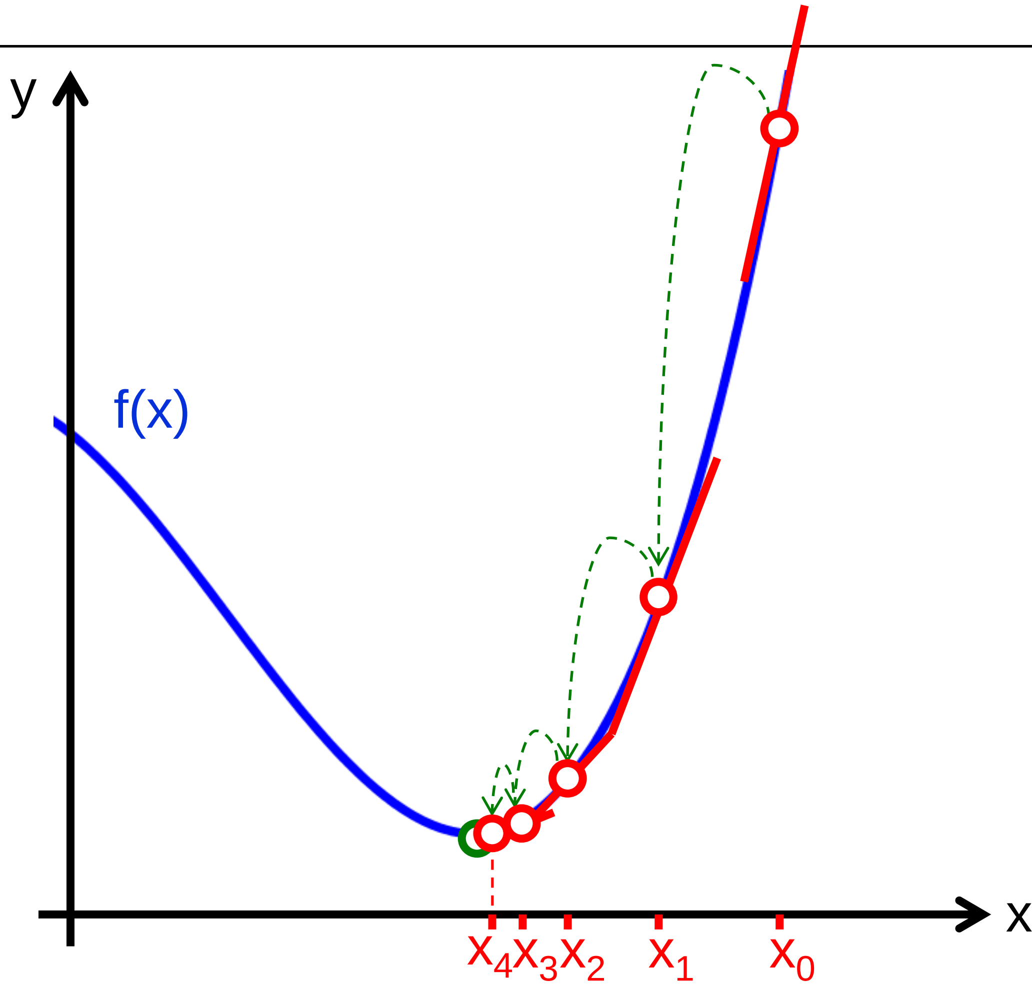
→ $x_{n+1} = x_n - \ldots$

Kleinere Steigung

→ Min. näher

→ kleinere Schritte

Usw. bis Steigung klein genug



Steigung positiv

→ nach links gehen

→ $x_{n+1} = x_n - \ldots$

Kleine Steigung

→ Min. in der Nähe

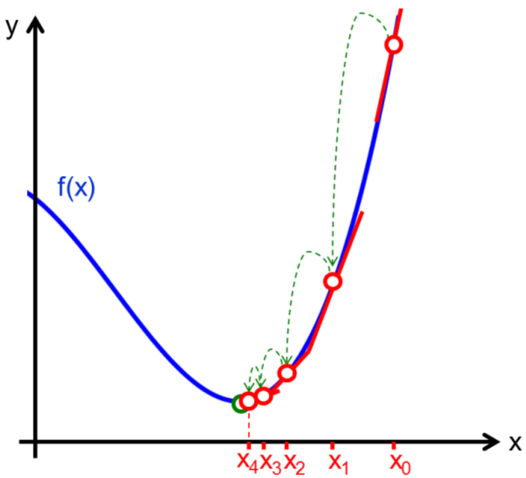
→ kleine Schritte

Gradientenabstiegsverfahren (eindimensional)

- **Startwert** x_0 beliebig wählen
- **Richtung?**

  Steigung positiv → nach links: $x_{n+1} = x_n - \dots$

  Steigung negativ → nach rechts: $x_{n+1} = x_n + \dots$



- **Schrittweite?**

  Große Steigung → Min. weit entfernt → große Schritte

  Kleine Steigung → Min. in der Nähe → kleine Schritte

  → Schrittweite proportional zur Steigung!

- **Iterationsvorschrift:** $\mathbf{x_{n+1} = x_n - s \cdot f'(x_n)}$

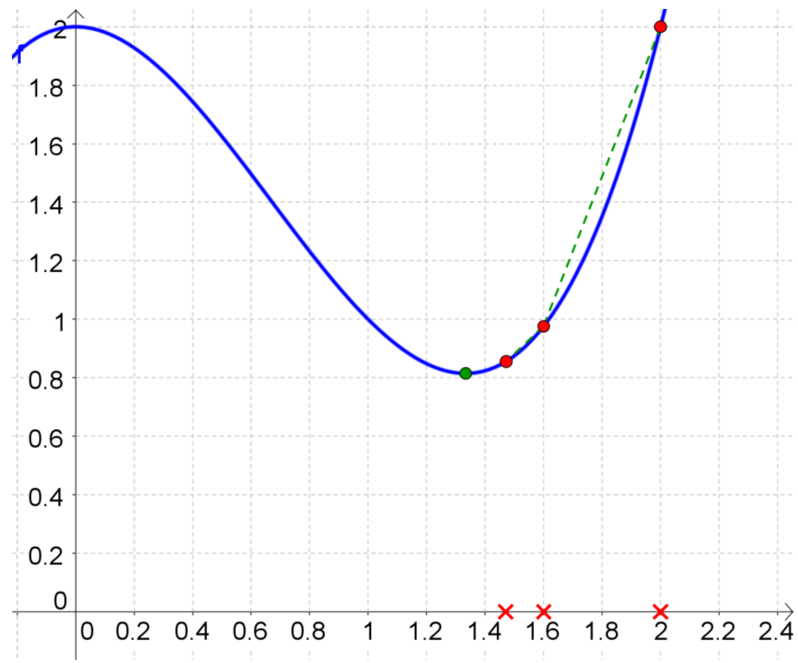
Beispiel

Funktion: $f(x) = x^3 - 2x^2 + 2$

Ableitung: $f'(x) = 3x^2 - 4x$

Wähle: Startwert: $x_0 = 2$

Schrittweite: $s = 0.1$



Gradientenabstieg:

$$x_{n+1} = x_n - s \cdot f'(x_n)$$
$$= x_n - 0.1 \cdot (3x_n^2 - 4x_n)$$

Ergebnis der ersten Iterationen:

$x_0 = 2$

$x_1 = 2 - 0.1 \cdot (3 \cdot 2^2 - 4 \cdot 2) = 1.6$

$x_2 = 1.6 - 0.1 \cdot (3 \cdot 1.6^2 - 4 \cdot 1.6) = 1.472$

Taschenrechner-Trick

Der TR speichert x_n im ANS-Speicher!

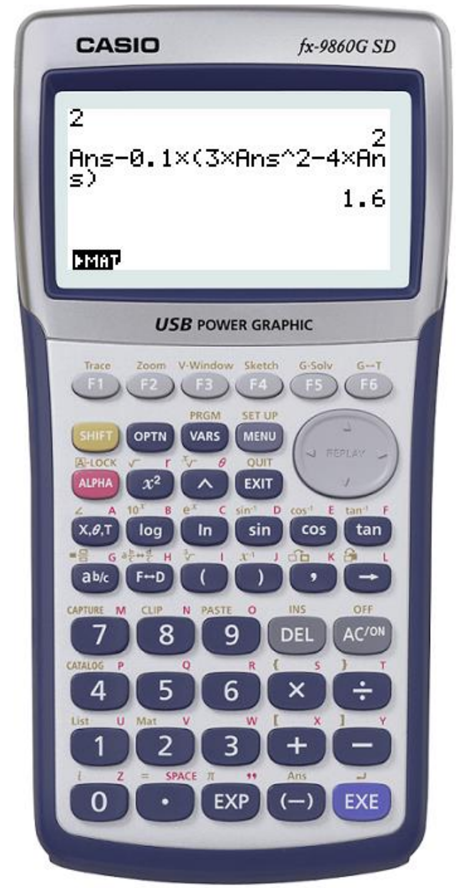
$$x_{n+1} = x_n - 0.1 \cdot (3x_n^2 - 4x_n)$$

$$= [ANS] - 0.1 \cdot (3 \cdot [ANS]^2 - 4 \cdot [ANS])$$

TR-Eingabe:

`2 =`

`Ans - 0.1 × ( 3 × Ans ^ 2 - 4 × Ans ) =`



Übung

Berechnen Sie ein Minimum der Funktion

$$f(x) = x^4 - 4x^3 + 2x^2 + 4x + 1$$

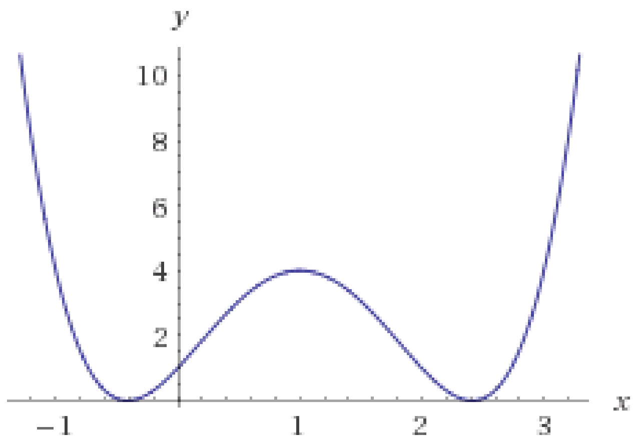
mit dem Gradientenabstiegsverfahren (mit $x_0 = 3$ und $s = 0.1$).

$$x_{n+1} = x_n - s \cdot f'(x_n)$$

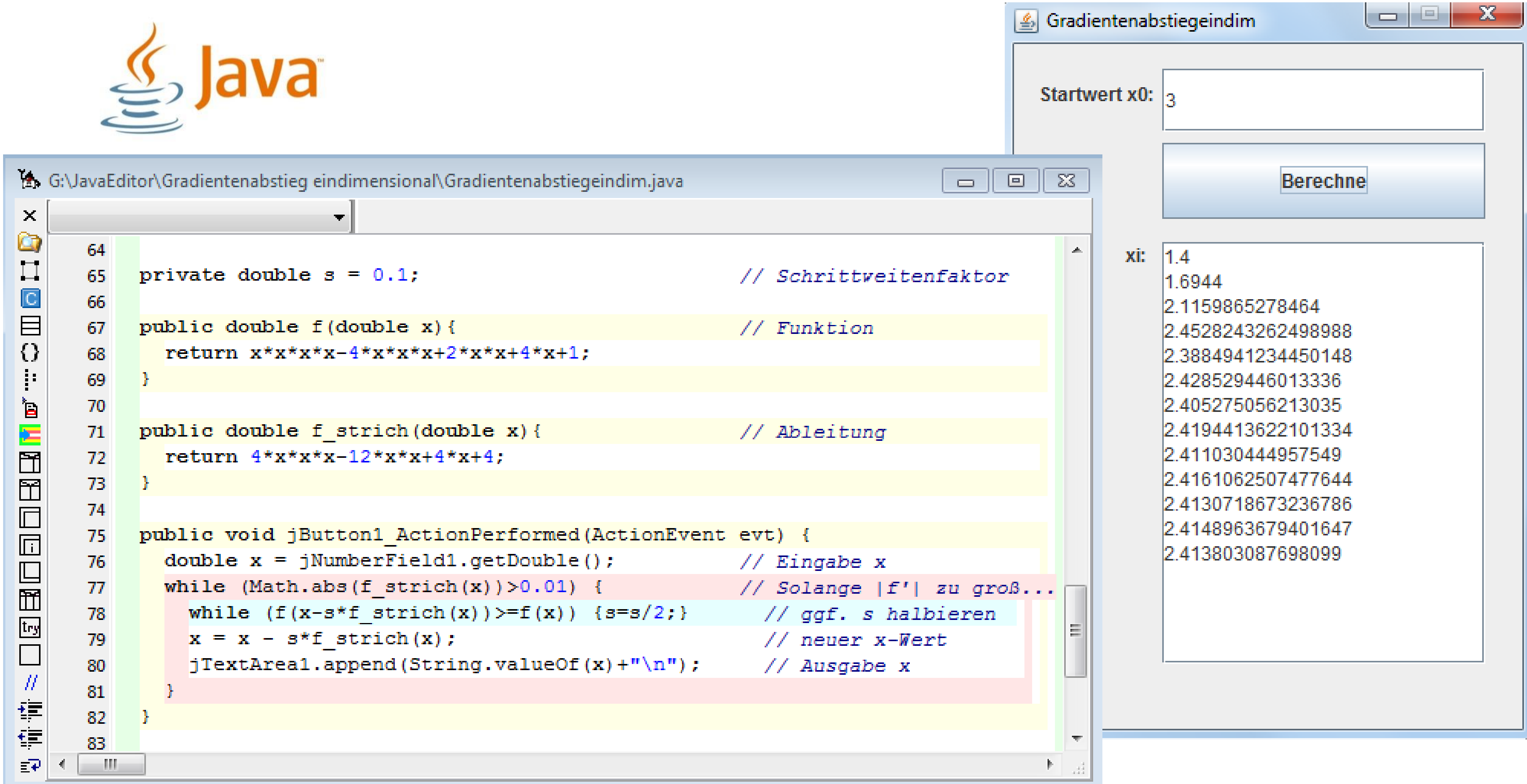
$$= x_n - 0.1 \cdot (4x_n^3 - 12x_n^2 + 4x_n + 4)$$

Ergebnis: $x = 2.4142$

Bei Startwert $x_0 = 0$: $x = -0.41421$

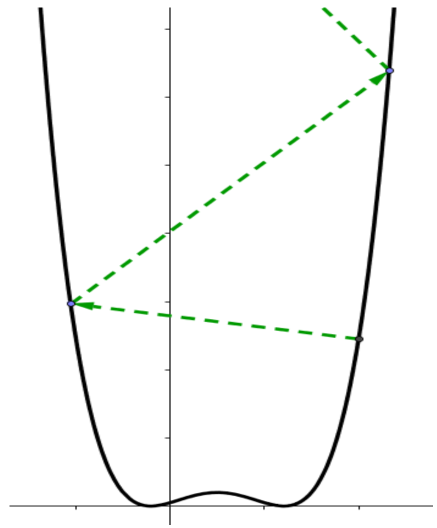


Programmierung des Gradientenabstiegsverfahrens



Steuerung der Schrittweite

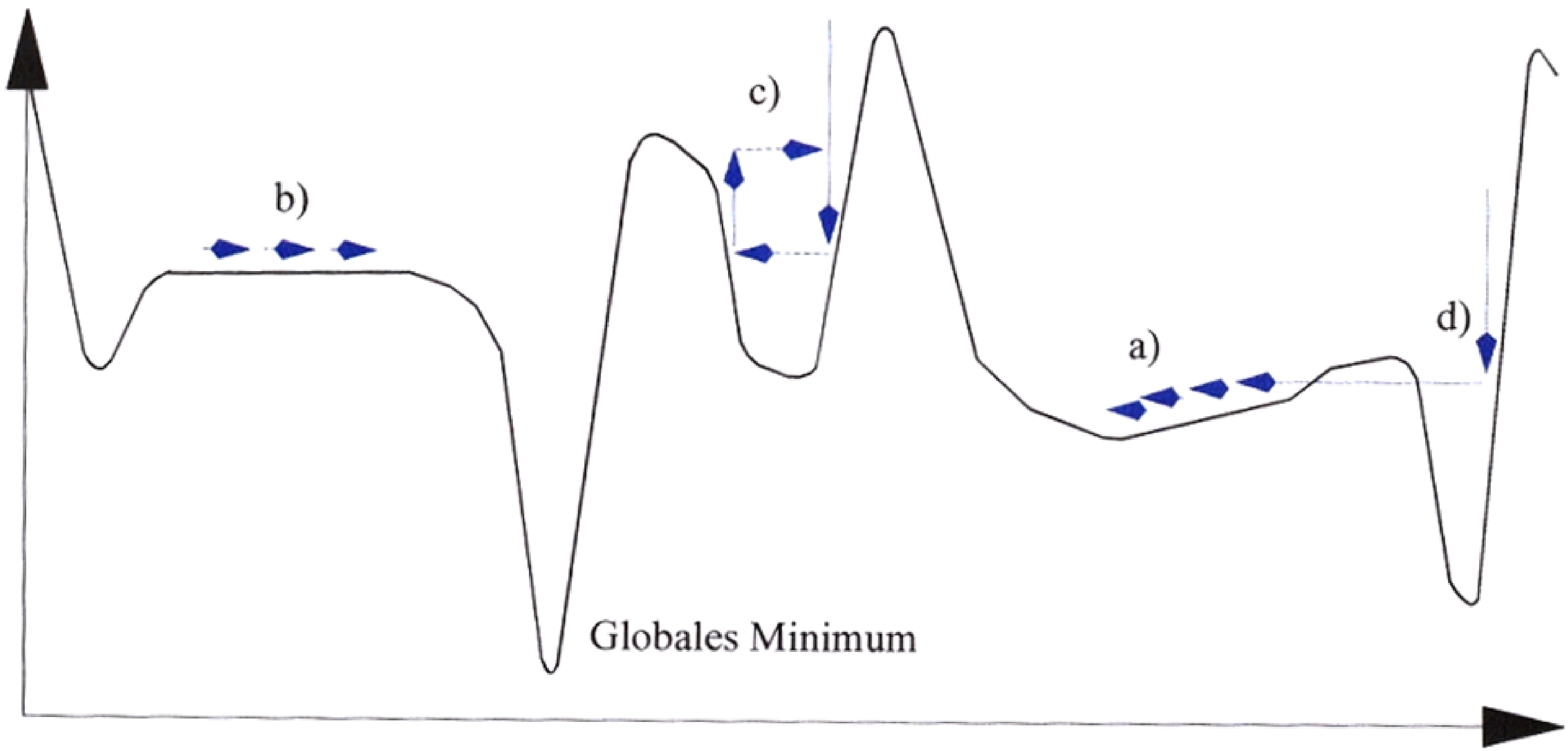
Problem: Divergenz bei zu großer Schrittweite:



Lösung: Solange $f(x_{n+1}) \geq f(x_n)$ Schrittweitenfaktor **s verkleinern**:

```
while (f(x - s*f_strich(x)) >= f(x)) {s = s/2}
```

Weitere mögliche Probleme



a) Finden schlechter Minima

b) Stillstand bei flachen Plateaus

c) Oszillation in Schluchten

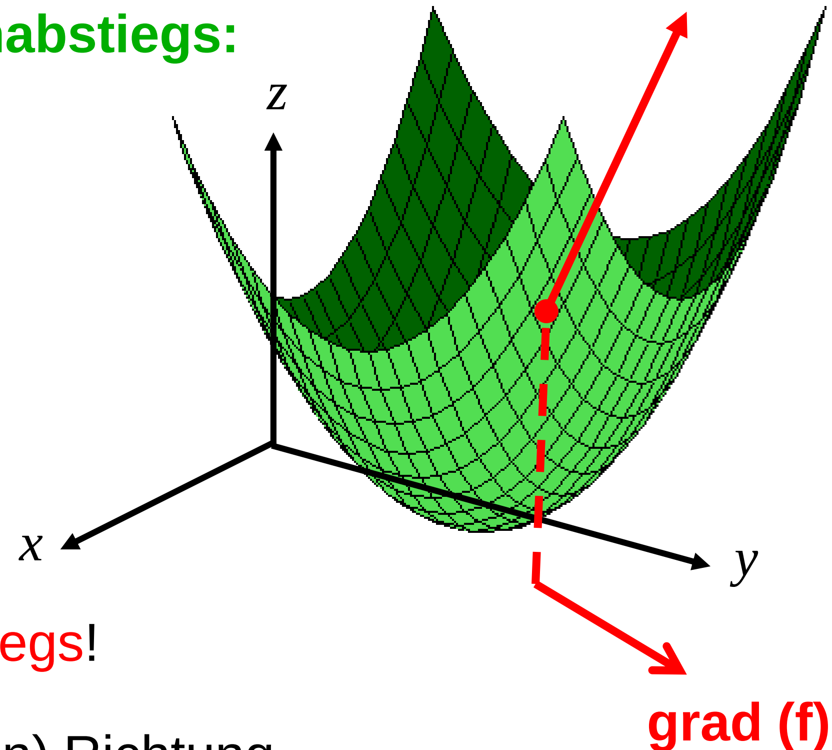
d) Verlassen guter Minima

Wie findet der Skifahrer im Nebel zurück ins Tal?

Minima mehrdimensionaler Funktionen

Verallgemeinerung des Gradientenabstiegs:

Wie finden wir ein Minimum bei einer *mehrdimensionalen* Funktion?



Erinnerung: Gradient

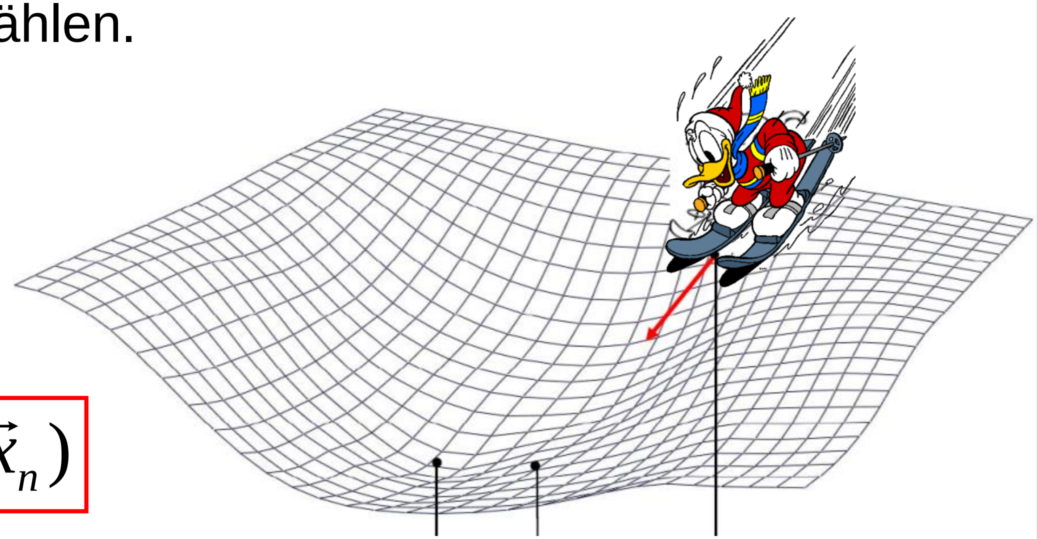
= Vektor der partiellen Ableitungen

- zeigt in Richtung des steilsten Anstiegs!
- Betrag = Steigung in dieser (steilsten) Richtung

Gradientenabstiegsverfahren (mehrdimensional)

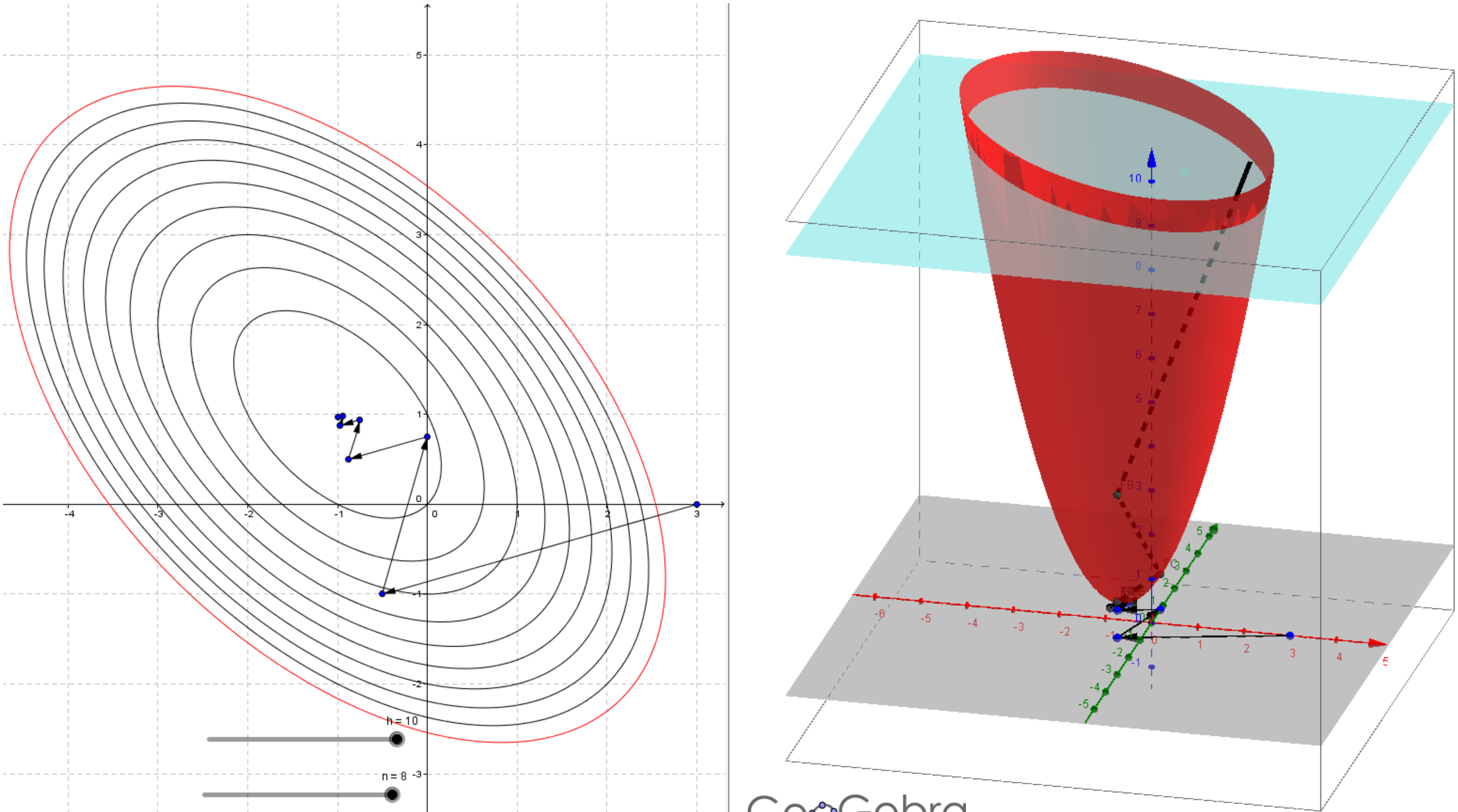
- **Startpunkt** $\vec{x}_0 = (x_0, y_0)$ wählen.
- In **Richtung** des negativen Gradienten fortschreiten (größter Abstieg!):

$$\boxed{\vec{x}_{n+1} = \vec{x}_n - s \cdot grad\ f(\vec{x}_n)}$$



- **Wiederholung** bis keine numerische Verbesserung.

Visualisierung mit GeoGebra



Programmierung (mehrdimensional)

G:\Javaeditor\Gradientenabstieg Mehrdimensional\Gradientenabstiegzweidim.java

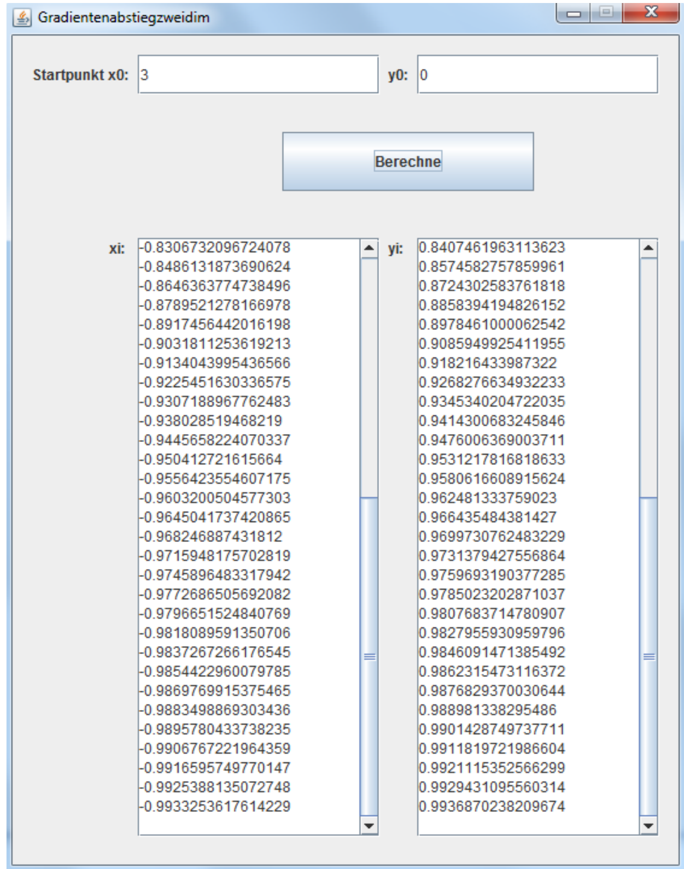
```java
public double f(double x, double y) {            // Funktion
  return x*x + y*y + x*y + x - y + 1;
}

public double f_x(double x, double y){          // part. Abl. nach x
  return 2*x + y + 1;
}

public double f_y(double x, double y){          // part. Abl. nach y
  return 2*y + x - 1;
}

public double betrag(double x, double y){       // Betrag eines Vektors
  return Math.sqrt(x*x+y*y);
}

public void jButton1_ActionPerformed(ActionEvent evt) {
  double x = jNumberField1.getDouble();         // Eingaben x und y
  double y = jNumberField2.getDouble();
  double s = 0.1;                               // Startschrittweite
  while (betrag(f_x(x,y), f_y(x,y))>0.001) {    // Solange |grad| zu groß...
    while (f(x-s*f_x(x,y), y-s*f_y(x,y)) >= f(x,y)) {  // ggf. s halbieren
      s=s/2;
    }
    x = x - s*f_x(x,y);                         // neues x und y
    y = y - s*f_y(x,y);
    jTextArea1.append(String.valueOf(x)+"\n");  // Ausgabe x und y
    jTextArea2.append(String.valueOf(y)+"\n");
  }
}
```



Anwendungen

- **Optimierung** von differenzierbaren, **mehrdimensionalen Funktionen** ohne Nebenbedingungen:

$$f(x, y, \ldots) = \min!$$

- **Lösen von nichtlinearen Gleichungssystemen:** Die Lösung von

$$g_1(x, y, \ldots) = 0$$
$$g_2(x, y, \ldots) = 0$$
$$\ldots.$$

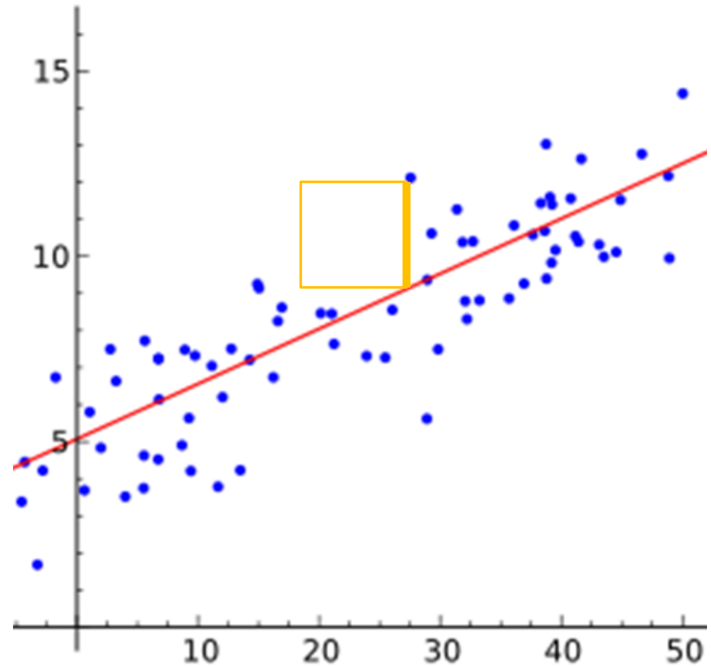
ist äquivalent zum Minimum von $f(x, y, \ldots) = \sum_i g_i^2(x, y, \ldots)$.

Anwendung: Regression

Beispiel:

- Welche Funktion $y = m\,x + b$ beschreibt die Messpunkte (x_i, y_i) am besten?
- Summe der Quadrate der Abstände soll minimal werden:

$$f(m, b) = \sum_i (y_i - (m\,x_i + b))^2$$



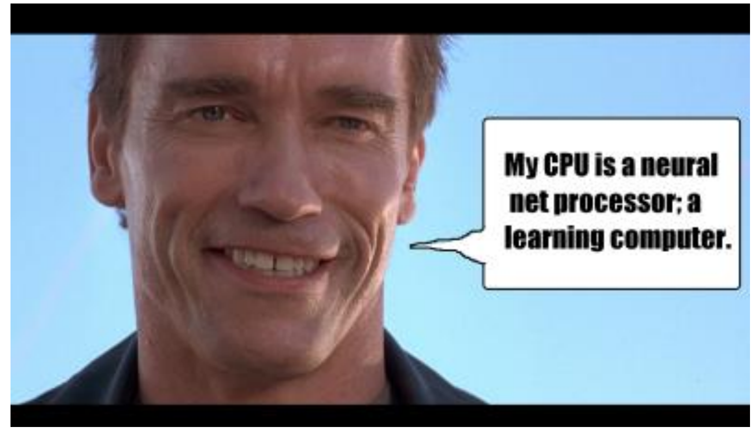
- Minimierung der „Fehlerfläche“ f(m, b) durch Gradientenabstieg!

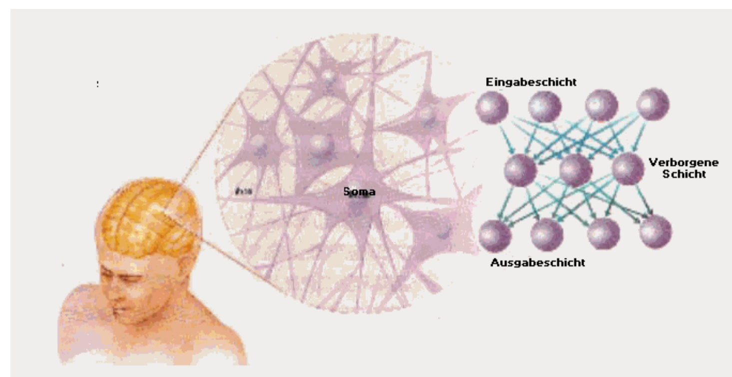
Wie können Maschinen lernen?

Anwendung: Lernen bei neuronalen Netzen

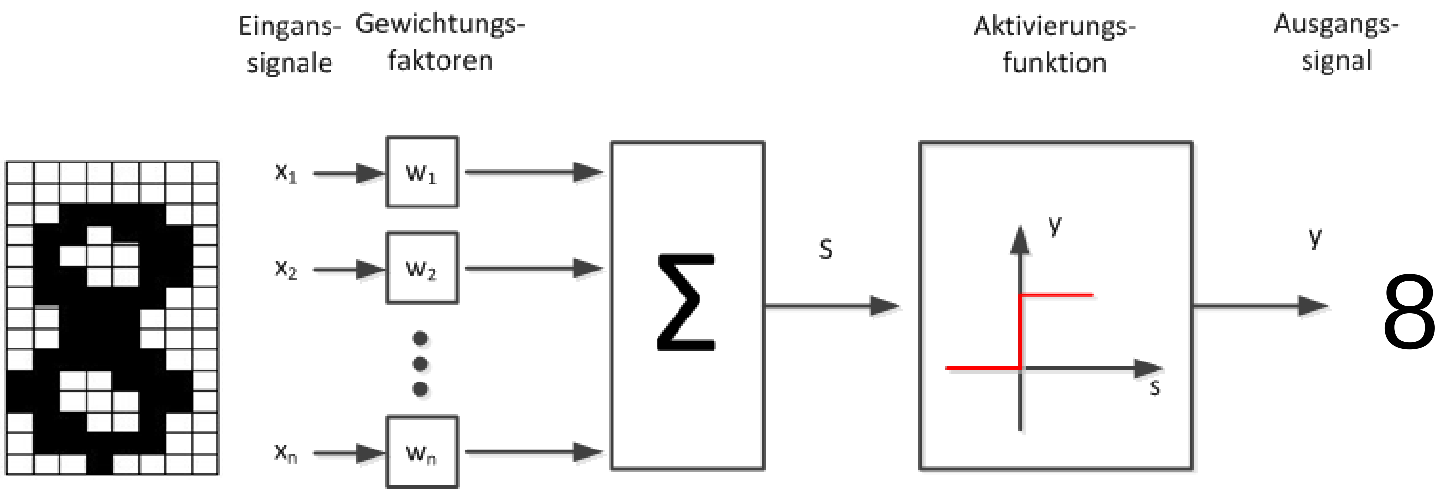
Neuronale Netze :

- von der Natur inspirierte Methode, Computer lernfähig zu machen.
- Anwendungen:
  - Sprach- und Schrifterkennung
  - Optimierungen in der Industrie
  - Künstliche Intelligenz
  - …
- Mathematisch: Gerichteter Graph, Knoten = künstliche Neuronen





Künstliches Neuron



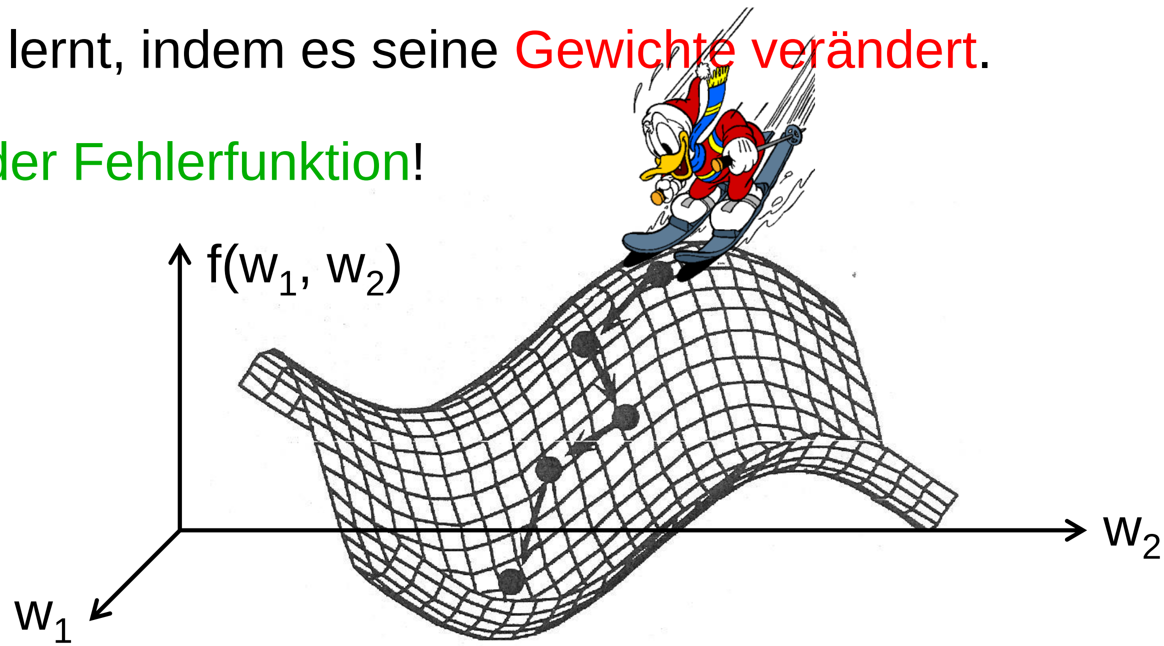
- **Trainingsphase:** Trainingseingaben → Vergleich mit erwarteten Ausgaben. Dabei passieren noch Fehler.
- **Ausführungsphase:** Lösung unbekannter Probleme (z.B: geschriebene Ziffern erkennen und diese als Output ausgeben).

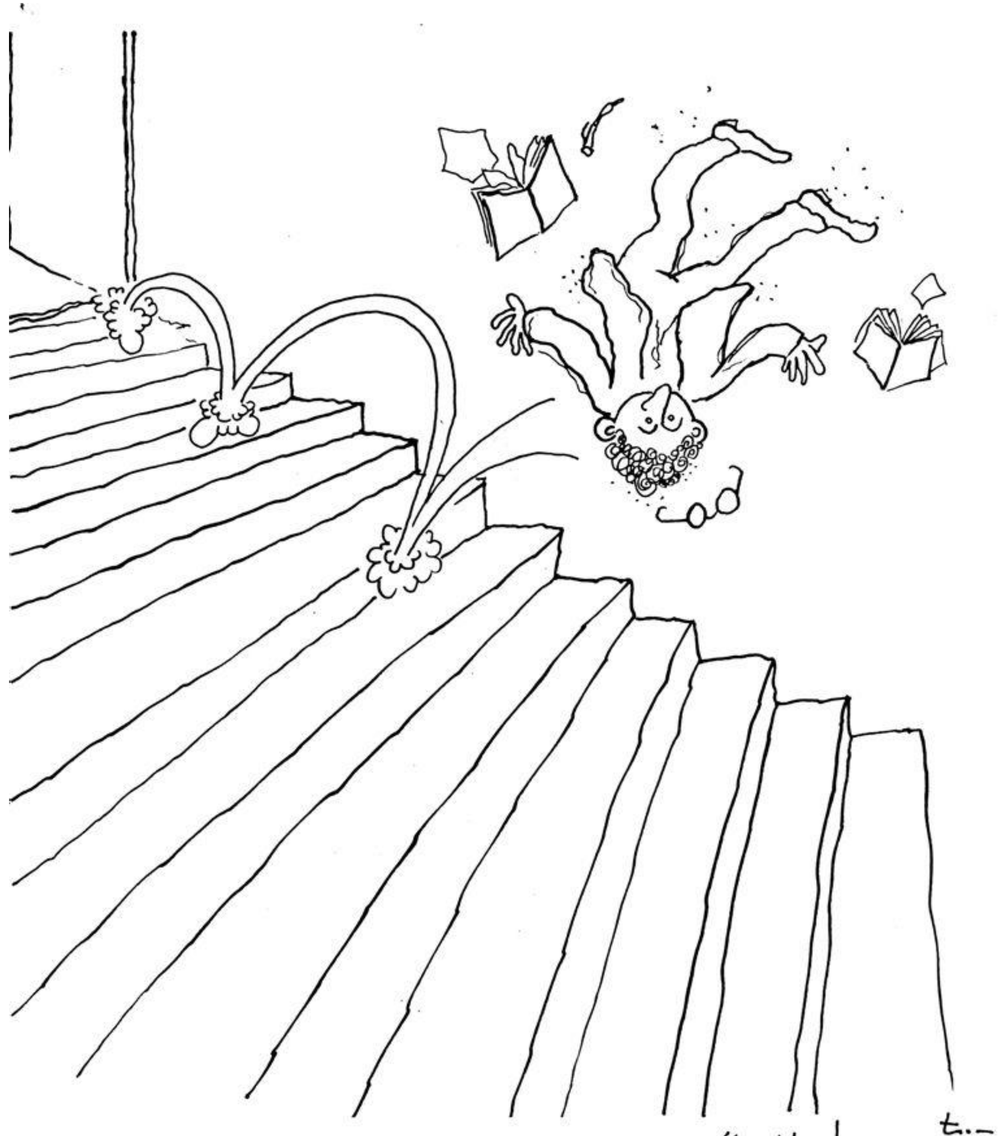
Lernen durch Fehlerminimierung

- Fehler des Neuron-Outputs ist Funktion der Gewichte:

$$f(w_1, w_2, \ldots)$$

- Das neuronale Netz lernt, indem es seine Gewichte verändert.
- Gesucht: Minimum der Fehlerfunktion!
- Gradientenabstieg auf Fehlerfläche!





Just after learning the "Steepest Descent" method
in optimization class...