

# **Vorlesung: Analysis 2 für Ingenieure**

Wintersemester 09/10

Michael Karow

*Thema: Numerische Optimierung*

## Problemstellung:

Wie findet man den minimalen Wert einer differenzierbaren Funktion

$$f : \mathbb{R}^n \supset G \rightarrow \mathbb{R}$$

## Lösungsverfahren:

- 1) Bestimme die Stellen  $\vec{x} \in G$  für die  $\text{grad}_{\vec{x}} f = \vec{0}$  ist.
- 2) Berechne die Funktionswerte  $f(\vec{x})$  an den gefundenen Stellen  $\vec{x}$ .  
Sei  $f(\vec{x}_*)$  der kleinste dieser Werte.
- 3) Wenn  $G$  kompakt, ist und  $f$  auf dem Rand von  $G$  keine kleineren Werte als  $f(\vec{x}_*)$  annimmt, dann ist  $f(\vec{x}_*)$  das globale Minimum. Falls  $G$  nicht kompakt ist, muss man noch prüfen, wie sich  $f$  in der Nähe des Randes verhält (vielleicht gibt es dann gar kein Minimum. Beispiel:  $f(x) = -1/x$ ,  $G = \{x \in \mathbb{R} \mid x > 0\}$ ).

Schon Schritt 1 ist durch Handrechnung nicht immer durchführbar.

Wenn dies der Fall ist, braucht man **Numerische Methoden**.

## Das Nullstellenproblem für vektorwertige Funktionen

Angenommen, wir wollen das Minimum der Funktion

$$(x, y) \mapsto f(x, y) = \sin^2(x^2 + y) + \cos(xy) + x^2$$

finden, indem wir zunächst die Nullstellen des Gradienten berechnen.  
Dann müssen wir die Gleichung

$$\vec{F}(x, y) := \text{grad}_{(x,y)} f = \begin{bmatrix} 4 \sin(x^2 + y) \cos(x^2 + y) x - \sin(xy)y + 2x \\ 2 \sin(x^2 + y) \cos(x^2 + y) - \sin(xy)y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

lösen. Dies ist durch Umformen nicht machbar.

**Verallgemeinerung des Problems:** Sei  $\vec{F} : \mathbb{R}^n \supset G \rightarrow \mathbb{R}^n$ . Finde  $\vec{x}$ , so dass

$$\vec{F}(\vec{x}) = \vec{0}.$$

Wir suchen also Nullstellen einer vektorwertigen Funktion  $\vec{F}$ .

## Das Newton-Verfahren zur Berechnung einer Nullstelle

Angenommen, die Funktion  $\vec{F} : \mathbb{R}^n \supset G \rightarrow \mathbb{R}^n$  hat die Nullstelle  $\vec{x}_* \in G$ , d.h. es ist  $\vec{F}(\vec{x}_*) = \vec{0}$ . Wie können wir  $\vec{x}_*$  finden, wenn eine Näherung  $\vec{x}_0$  für  $\vec{x}_*$  bekannt ist?

### Herleitung des Newton-Verfahrens:

(1) Es ist  $\vec{x}_* = \vec{x}_0 + \Delta\vec{x}$  mit  $\Delta\vec{x} = \vec{x}_* - \vec{x}_0$ .

(2) Taylor-Entwicklung um  $\vec{x}_0$ :  $\vec{0} = \vec{F}(\vec{x}_*) = \vec{F}(\vec{x}_0) + \vec{F}'(\vec{x}_0) \Delta\vec{x} + \vec{R}(\Delta\vec{x}) \quad (*)$

(3) Wenn wir (\*) nach  $\Delta\vec{x}$  auflösen (umstellen) könnten, dann könnten wir mit (1) die Nullstelle  $\vec{x}_*$  berechnen. Wegen des unbekannten Restglieds  $\vec{R}(\Delta\vec{x})$  ist dies aber nicht möglich. Wir hoffen, dass das Restglied klein ist und bestimmen eine Näherung  $\vec{x}_1$  durch Lösen der Gleichung

$$\vec{0} = \vec{F}(\vec{x}_1) = \vec{F}(\vec{x}_0) + \vec{F}'(\vec{x}_0) \Delta\vec{x} \quad \text{wobei} \quad \vec{x}_1 = \vec{x}_0 + \Delta\vec{x}.$$

Es ist

$$\vec{x}_1 = \vec{x}_0 - \vec{F}'(\vec{x}_0)^{-1} \vec{F}(\vec{x}_0).$$

$\vec{x}_1$  ist hoffentlich eine bessere Näherung als  $\vec{x}_0$ .

(4) Wiederhole das Verfahren mit  $\vec{x}_1$  statt mit  $\vec{x}_0$ . Fahre so fort. Das ergibt die rekursiv definierte Folge

$$\vec{x}_{k+1} = \vec{x}_k - \vec{F}'(\vec{x}_k)^{-1} \vec{F}(\vec{x}_k).$$

## Die Methode des steilsten Abstiegs (Gradientenverfahren)

Sei  $f : \mathbb{R}^n \supset G \rightarrow \mathbb{R}$  eine stetig differenzierbare Funktion.

Wir wissen:

$\text{grad}_{\vec{x}} f$  zeigt in die Richtung des stärksten Anstiegs von  $f$ .

$-\text{grad}_{\vec{x}} f$  zeigt in die Richtung des stärksten Abstiegs von  $f$ .

Eine Methode zur Bestimmung eines relativen Minimums von  $f$  ist die **Methode des steilsten Abstiegs**:

Gehe in jedem Schritt ein Stück in Richtung  $-\text{grad}_{\vec{x}} f$ . Definiere also eine Folge

$$\vec{x}_{k+1} = \vec{x}_k - t_k \text{grad}_{\vec{x}} f, \quad t_k > 0.$$

Dabei muss  $t_k$  so klein gewählt werden, dass  $f(\vec{x}_{k+1}) < f(\vec{x}_k)$ . Dies ist stets möglich.

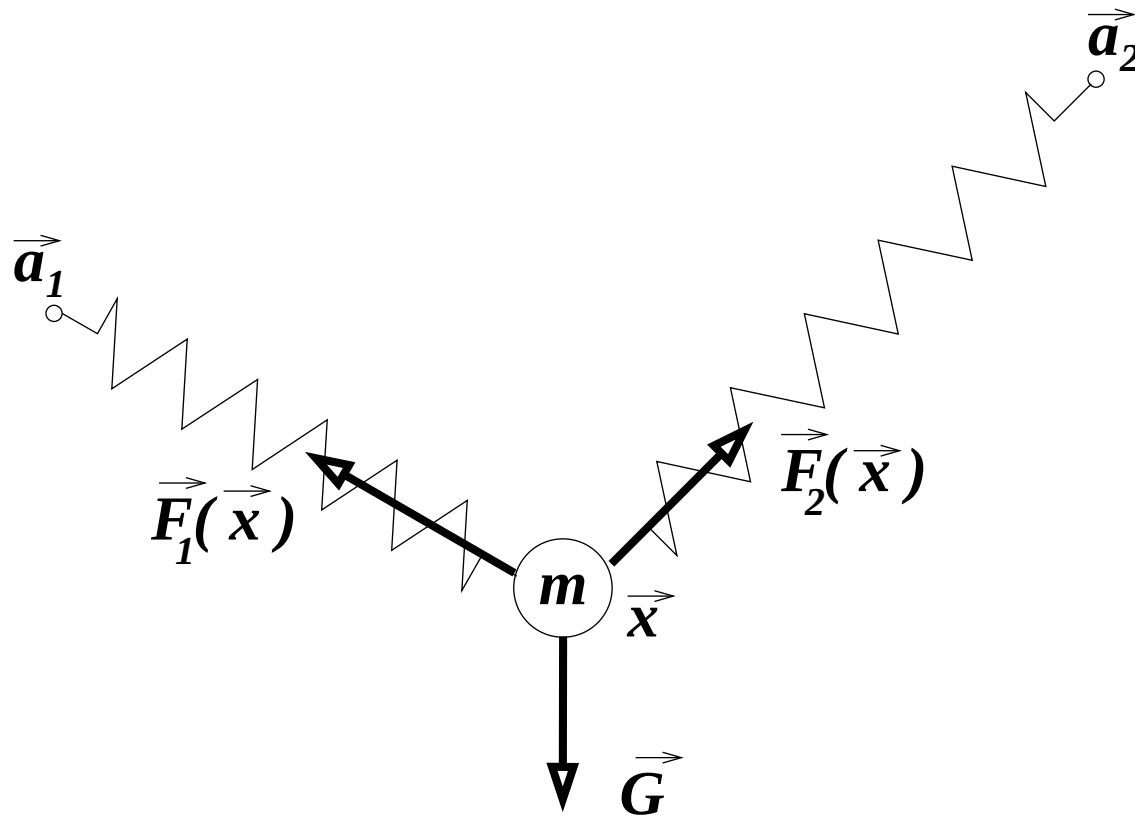
## Beispielproblem: Gleichgewichtspunkt in einem Kraftfeld

Situation: Eine Masse  $m$  hängt an 2 Federn.

Auf diese Masse wirken dann die Federkräfte und die Gravitation.

Die Summe dieser Kräfte  $\vec{F}$  hängt vom Ort  $\vec{x}$  ab,  
an dem die Masse sich gerade befindet:  $\vec{F} = \vec{F}(\vec{x})$ .

Gesucht: Ein Punkt  $\vec{x}$  mit  $\vec{F}(\vec{x}) = \vec{0}$ .



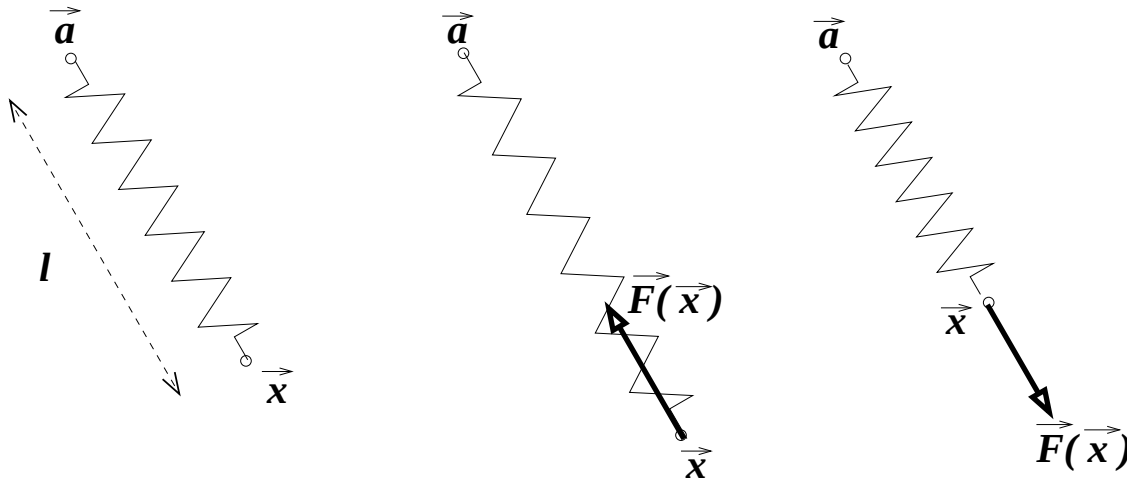
## Die Federkraft

Die Feder mit (entspannter) Länge  $\ell$  und Steifigkeit  $s$  sei am Punkt  $\vec{a}$  frei drehbar aufgehängt. Dann übt die Feder auf ihren zweiten Endpunkt  $\vec{x}$  folgende Kraft aus:

$$\vec{F}(\vec{x}) = -s (|\vec{x} - \vec{a}| - \ell) \frac{\vec{x} - \vec{a}}{|\vec{x} - \vec{a}|}. \quad (\text{Hooksches Gesetz})$$

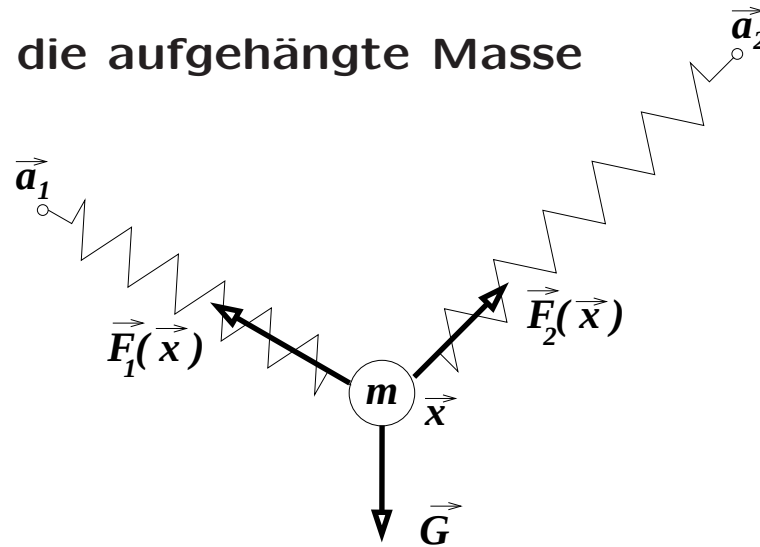
Es ist

$$\vec{F}(\vec{x}) = -\text{grad}_{\vec{x}} E_{el} \quad \text{mit} \quad E_{el}(\vec{x}) = \frac{s}{2} (|\vec{x} - \vec{a}| - \ell)^2 \quad (\text{elastische Energie}).$$



Links: entspannte Feder mit Länge  $\ell$ . Mitte: gedehnte Feder. Rechts: gestauchte Feder.

## Die Kraft auf die aufgehängte Masse



**Gesamtkraft:** 
$$\vec{F}(\vec{x}) = -s_1 (|\vec{x} - \vec{a}_1| - \ell_1) \frac{\vec{x} - \vec{a}_1}{|\vec{x} - \vec{a}_1|} - s_2 (|\vec{x} - \vec{a}_2| - \ell_2) \frac{\vec{x} - \vec{a}_2}{|\vec{x} - \vec{a}_2|} - \begin{bmatrix} 0 \\ gm \end{bmatrix}.$$

Der letzte Summand ist die Gravitationskraft.  $g$  = Fallbeschleunigung. Es ist

$$\vec{F}(\vec{x}) = -\text{grad}_{\vec{x}} V \quad \text{mit} \quad V(\vec{x}) = \frac{s_1}{2} (|\vec{x} - \vec{a}_1| - \ell_1)^2 + \frac{s_2}{2} (|\vec{x} - \vec{a}_2| - \ell_2)^2 + m g y.$$

$V(\vec{x})$  ist die potentielle Energie (**Potential** von  $\vec{F}$ ).

**Bewegungsgleichung** für die Masse:  $m \ddot{\vec{x}}(t) = \vec{F}(\vec{x}(t)).$

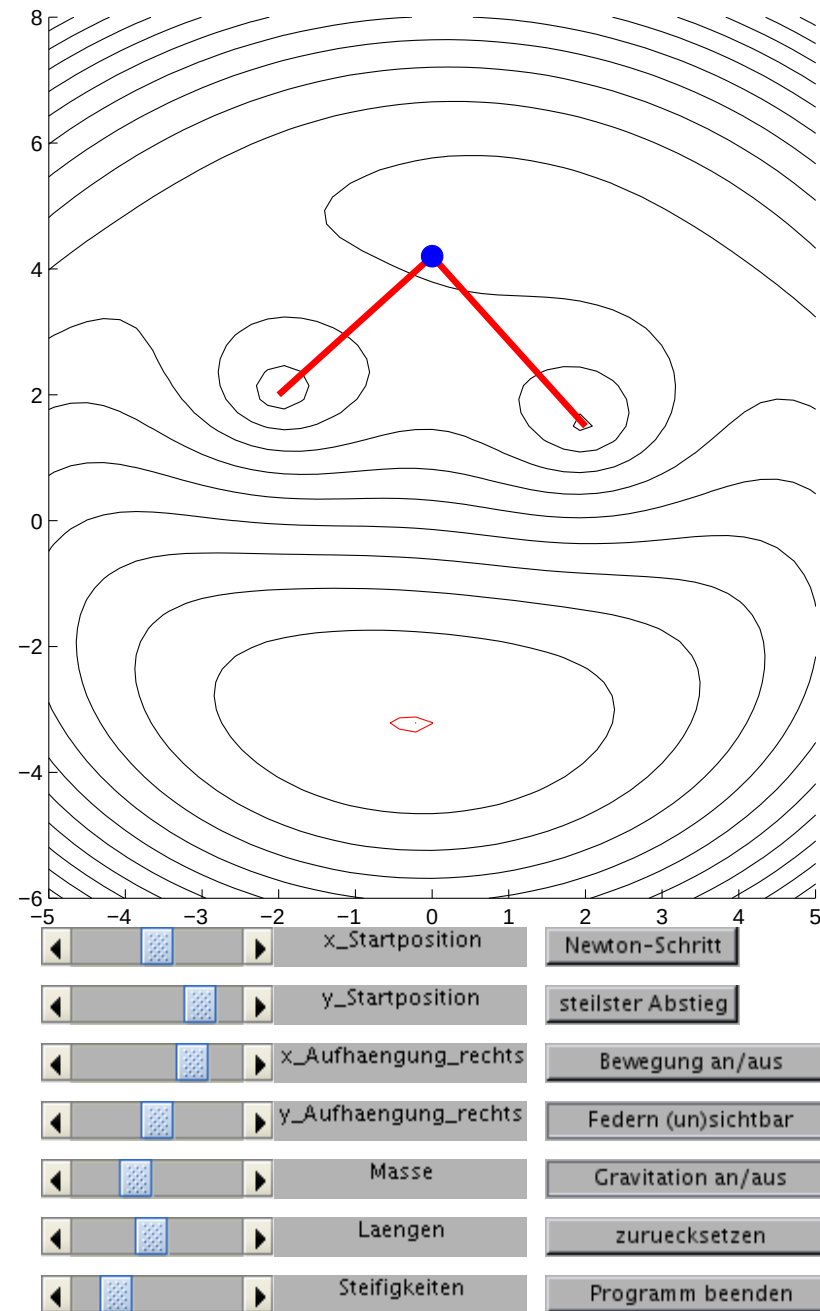
Ein Punkt  $\vec{x}$  heißt **Gleichgewichtspunkt**, wenn  $\vec{0} = \vec{F}(\vec{x}) = -\text{grad}_{\vec{x}} V.$

An einem Gleichgewichtspunkt erfährt die Masse keine Beschleunigung.

Eine Minimalstelle des Potentials  $V$  ist stets Gleichgewichtspunkt.



# Demo-Programm zum Federproblem (siehe Vorlesung)



## Ausgleichsprobleme I

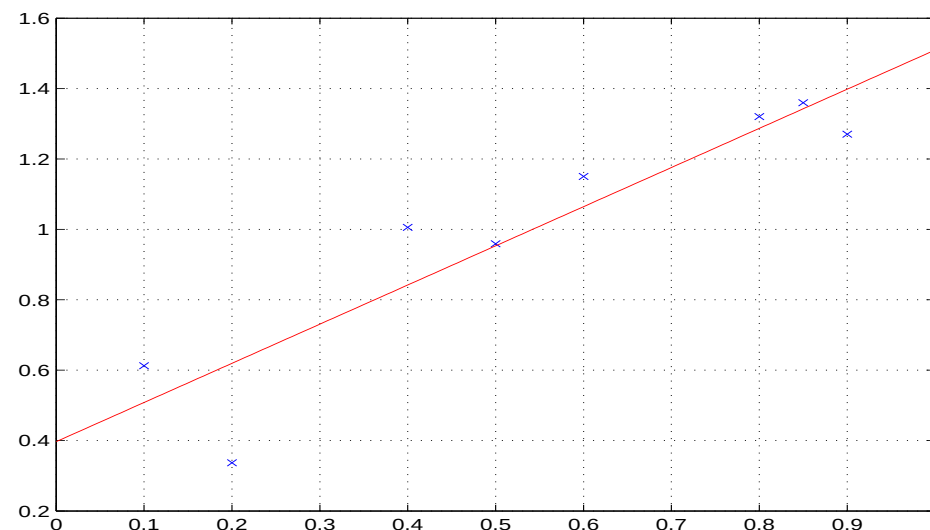
**Beispielproblem 1:** Zwischen zwei (z.B. physikalischen) Größen  $t$  und  $y$  wird (z.B. aufgrund theoretischer Überlegungen) ein linearer Zusammenhang der Form

$$y = y_{a,b}(x) = a + bt \quad (*)$$

angenommen. Die unbekannten Parameter  $a, b$  sollen durch Messungen bestimmt werden. Die Messungen ergeben die Wertepaare

$$(t_1, y_1), \quad (t_2, y_2), \quad (t_3, y_3), \quad \dots, (t_m, y_m).$$

Aufgrund von Messfehlern oder weil die Ausgangshypothese  $(*)$  nicht ganz korrekt ist, liegen die Messpunkte nicht auf einer Geraden. Was sind die besten Werte für  $a, b$ , die man in dieser Situation angeben kann?



## Ausgleichsprobleme II

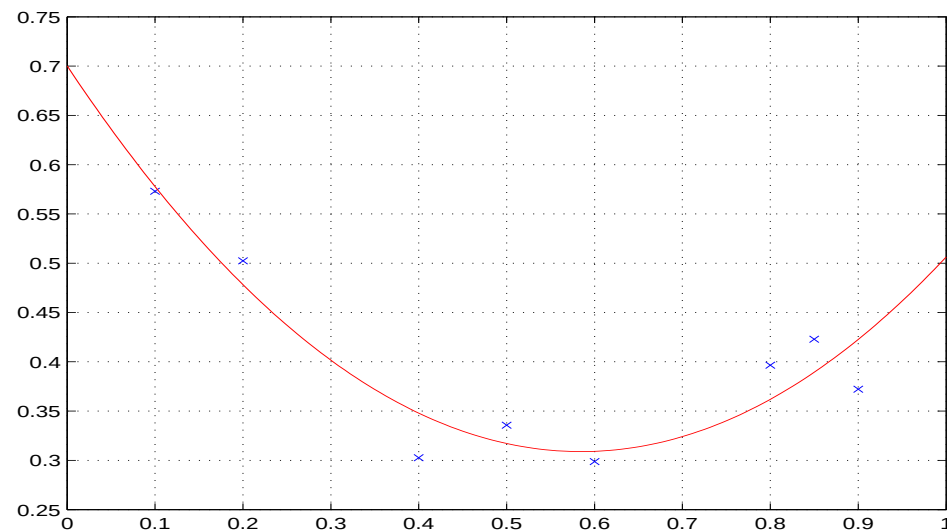
**Beispielproblem 2:** Zwischen zwei Größen  $t$  und  $y$  wird ein quadratischer Zusammenhang der Form

$$y = y_{a,b,c}(t) = a + bt + ct^2$$

angenommen. Die unbekannten Parameter  $a, b, c$  sollen durch Messungen bestimmt werden. Die Messungen ergeben die Wertepaare

$$(t_1, y_1), \quad (t_2, y_2), \quad (t_3, y_3), \quad \dots, (t_m, y_m).$$

Was sind die besten Werte für  $a, b, c$  ?



## Ausgleichsprobleme III

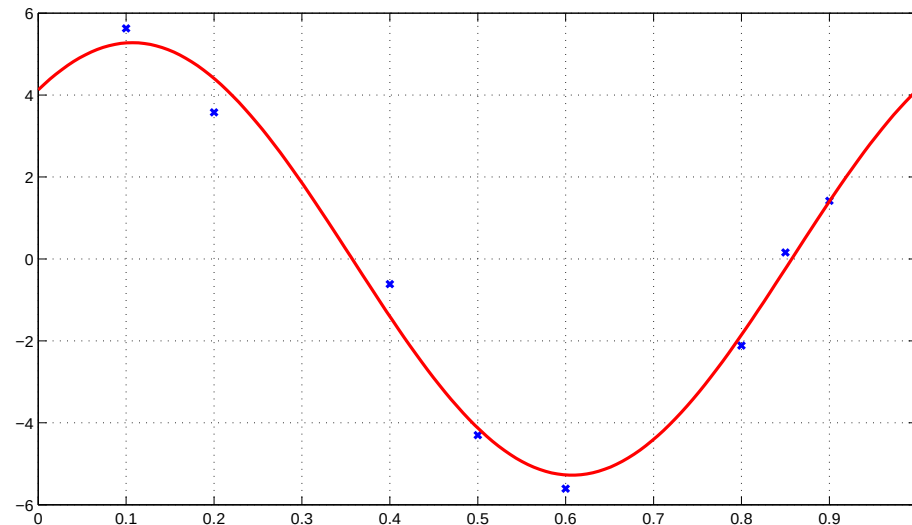
**Beispielproblem 3:** Zwischen zwei Größen  $t$  und  $y$  wird ein Zusammenhang der Form

$$y = y_{a,b}(t) = a \sin(t + b)$$

angenommen. Die unbekannten Parameter  $a, b$  sollen durch Messungen bestimmt werden. Die Messungen ergeben die Wertepaare

$$(t_1, y_1), \quad (t_2, y_2), \quad (t_3, y_3), \quad \dots (t_m, y_m).$$

Was sind die besten Werte für  $a, b$  ?



## Ausgleichsprobleme IV

### Allgemeines Ausgleichsproblem (für 2 Parameter):

Zwischen den Größen  $t$  und  $y$  wird ein Zusammenhang der Form

$$y = y_{a,b}(t) = y(a, b, t) \quad (*)$$

angenommen. Gegeben sind die Daten

$$(t_1, y_1), \quad (t_2, y_2), \quad (t_3, y_3), \quad \dots (t_m, y_m),$$

wobei  $m \geq 2$ . Wie muss man die freien Konstanten  $a, b$  wählen, so dass die Funktion  $(*)$  die Daten optimal approximiert (annähert)?

**Frage:** Was heißt eigentlich 'optimal approximiert' ?

Auf diese Frage kann man verschiedene Antworten geben. In dieser VL besprechen wir die (in der Praxis am häufigsten vorkommende)

**Optimalität im Sinne der kleinsten Fehlerquadrate**

## Ausgleichsprobleme V

Messwert zu  $t_j$ :  $y_j$

Funktionswert zu  $t_j$ :  $y_{a,b}(t_j)$

Fehlerquadrat:  $(y_{a,b}(t_j) - y_j)^2$

Summe der Fehlerquadrate:

$$f(a, b) = \sum_{j=1}^m (y_{a,b}(t_j) - y_j)^2.$$

Methode der kleinsten Fehlerquadrate (least squares method):

Bestimme die Parameter  $a, b$ , so, dass  $f(a, b)$  minimal wird.

$$f(a, b) \rightarrow \min$$

Dieses Optimierungsproblem kann z.B. mit der Methode des steilsten Abstiegs (Gradientenverfahren) oder mit dem Newton-Verfahren (für  $\text{grad } f$ ) gelöst werden.

# Ausgleichsprobleme VI

Demo-Programm aus der Vorlesung:

