

# Optimierung

Die Suche nach dem besten Weg

Schülerseminar an der Universität Mannheim  
vom 6.–9. September 2000

Wolfgang K. Seiler

**Nichtlineare Optimierung**

KAPITEL II: NICHTLINEARE OPTIMIERUNG .....	1
§0: Nichtlineare Optimierung im Eindimensionalen .....	1
§1: Funktionen mehrerer Veränderlicher .....	5
§2: Differenzierbare Funktionen .....	9
§3: Optimierung ohne Nebenbedingungen .....	21
§4: Optimierung zur Schätzung von Parametern .....	25
§5: Maxima und Minima unter Nebenbedingungen .....	33
§6: Numerische Verfahren .....	46
a) Die Gradientenmethode .....	46
b) Der Metropolis-Algorithmus .....	53
§7: Zusammenfassung .....	58

WOLFGANG K. SEILER  
Mathematisches Institut der  
Universität Mannheim  
68131 Mannheim

seiler@math.uni-mannheim.de

von mehreren Intervallen; wir wollen uns daher auf den wesentlichen Teil der Aufgabe konzentrieren, nämlich darauf, das Maximum oder Minimum von  $f$  in einem Intervall  $[a, b]$  zu finden.

Auch hier sind wir beileibe noch nicht auf dem wohlvertrauten Terrain der Kurvendiskussion, wo Extrema entweder an einer Intervallgrenze angenommen werden oder aber bei Nullstellen der Ableitung: Schließlich war nirgends davon die Rede, daß  $f$  überall in  $[a, b]$  differenzierbar sein soll. Die in Abbildung 1 zu sehende Funktion

$$f(x) = 1 - |\sin x|$$

etwa hat ihre Maxima bei den Vielfachen von  $\pi$ , und genau dort ist die Funktion nicht differenzierbar. Die Links- und Rechtsableitungen existieren zwar beide, sind aber  $\pm 1$  und damit deutlich von null verschieden. In einem Fall wie diesem muß man also das Intervall unterteilen in kleinere Intervalle, in deren Inneren  $f$  differenzierbar ist.

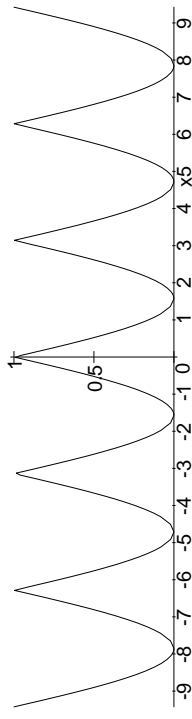


Abb. 1: Graph der Funktion  $f(x) = 1 - |\sin x|$

Selbst das geht aber nicht immer: Angenommen, wir gehen aus von der konstanten Funktion  $f_0(x) = 0$  auf dem Einheitsintervall  $I = [0, 1]$ ; ihr Graph ist einfach eine waagrechte Strecke. Daraus konstruieren wir eine weitere Funktion  $f_1$ , indem wir im Graph von  $f_0$  das mittlere Drittel um  $1/3$  anheben; wir betrachten also die Funktion

$$f_1: I \rightarrow \mathbb{R}; \quad x \mapsto \begin{cases} \frac{1}{3} & \text{falls } \frac{1}{3} \leq x \leq \frac{2}{3} \\ 0 & \text{sonst} \end{cases}$$

Mit dieser verfahren wir genauso: Wir ersetzen jede waagrechte Strecke im Graphen durch eine, deren mittleres Drittel angehoben ist, jetzt aber

## Kapitel 2 Nichtlineare Optimierung

Gestern haben wir gesehen, daß es in der *linearen* Optimierung darum geht, eine lineare Funktion  $f(x_1, \dots, x_n)$  zu maximieren oder minimieren unter Nebenbedingungen der Form

$$g_1(x_1, \dots, x_n) \leq 0, \quad g_2(x_1, \dots, x_n) \leq 0, \dots, \quad g_r(x_1, \dots, x_n) \leq 0$$

mit *linearen* Funktionen  $f, g_1, \dots, g_r$ . Damit ist ziemlich klar, daß es in der *nichtlinearen* Optimierung im wesentlichen um dieselbe Aufgabe geht, nur daß die genannten Funktionen nun auch *nichtlinear* sein können.

In dieser extrem allgemeinen Form ist das Problem allerdings weit von einer effizienten Lösbarkeit entfernt – nicht nur, wenn man sich auf Methoden beschränkt, die im Rahmen dieses dreitägigen Seminars behandelt werden können, sondern auch nach dem derzeitigen Stand der mathematischen Forschung.

Um zu sehen, wo die Probleme liegen, betrachten wir zunächst den (für Anwendungen nur selten interessanten) eindimensionalen Fall.

### §0: Nichtlineare Optimierung im Eindimensionalen

Hier haben wir eine Funktion  $f$  einer reellen Variablen  $x$  und wollen wissen, wo diese ihr Maximum oder Minimum annimmt unter Nebenbedingungen der Form  $g_1(x) \leq 0, \dots, g_r(x) \leq 0$ .

Für völlig beliebige Funktionen  $g_i$  haben wir kaum eine Chance, auch nur die Nebenbedingungen auszuwerten. Tatsächlich geht es aber meist darum, die Extremwerte in einem Intervall zu finden oder auch in einem

nur um ein Neuntel; dies führt auf die Funktion

$$f_2: I \rightarrow \mathbb{R}; \quad x \mapsto \begin{cases} \frac{1}{3} + \frac{1}{9} & \text{für } x \in [\frac{4}{9}, \frac{5}{9}] \\ \frac{1}{3} & \text{für } x \in [\frac{1}{3}, \frac{4}{9}) \cup (\frac{5}{9}, \frac{2}{3}] \\ \frac{1}{9} & \text{für } x \in [\frac{1}{9}, \frac{2}{9}) \cup [\frac{7}{9}, \frac{8}{9}] \\ 0 & \text{sonst} \end{cases}$$

Daraus konstruieren wir eine Funktion  $f_3$ , indem wir wiederum in jeder waagrechten Strecke des Graphen das mittlere Drittel anheben, jetzt aber nur noch um  $\frac{1}{3^3} = \frac{1}{27}$  und so weiter; Abbildung 2 zeigt die Funktion  $f_9$ .

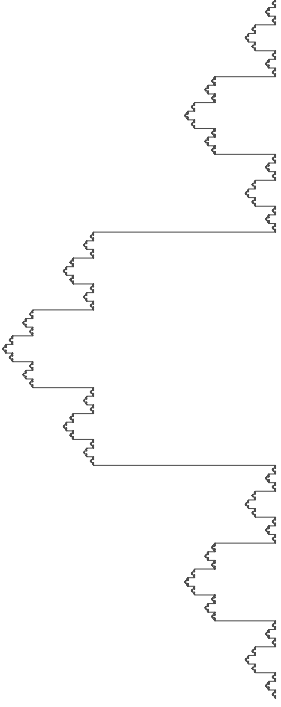


Abb. 2: Graph der Funktion  $f_9(x)$

Wenn wir dies unbegrenzt fortführen, erhalten wir schließlich eine Funktion  $f_\infty$ , deren Graph sich im Rahmen der Druckauflösung dieses Manuskripts nicht mehr von Abbildung 2 unterscheiden läßt, und die analytisch folgendermaßen beschrieben werden kann: Für  $x = 1$  ist  $f_\infty(x) = 0$ . Für alle anderen Argumente  $x$  schreiben wir  $x$  im Dreiersystem als

$$x = 0, a_1 a_2 a_3 \dots, \quad \text{d.h. } x = \sum_{j=1}^{\infty} a_j 3^{-j} \quad \text{mit } a_j \in \{0, 1, 2\},$$

und definieren dazu die Indexmenge

$$J(x) = \{j \in \mathbb{N} \mid a_j = 1\}.$$

Dann ist

$$f_\infty(x) = \sum_{j \in J(x)} 3^{-j}.$$

Die so konstruierte Funktion  $f_\infty$  ist offensichtlich nirgends stetig und somit erst recht in keinem Intervall differenzierbar. Trotzdem ist unmittelbar einsichtig, daß sie ihr einziges Maximum in dem Punkt annimmt, in dem  $J(x) = \mathbb{N}$  ist, d.h.

$$x = f(x) = 0,1111 \dots = \sum_{j=1}^{\infty} 3^{-j} = \frac{1}{2}.$$

Durch die Beschränkung auf den differenzierbaren Fall verlieren wir also Funktionen, für die das Optimierungsproblem durchaus lösbar ist; wir müssen uns überlegen, ob wir das in Kauf nehmen wollen.

Als erstes sollten wir uns darüber im klaren sein, daß für die Beschreibung realer Probleme sowohl differenzierbare Funktionen als auch Funktionen wie  $f_\infty$  stets Idealisierungen sein müssen: In der Realität sind Preise ganzzahlige Vielfache einer kleinsten Verrechnungseinheit, viele Güter können nicht zerteilt werden, ohne ihren Charakter wesentlich zu verändern (Wem nützt schon ein halber Blumentopf?), und selbst bei gut teilbaren Stoffen wie Flüssigkeiten oder Gasen kommen wir irgendwann einmal zu Molekülen.

Andererseits arbeiten wir bei der mathematischen Beschreibung eines realen Systems notgedrungen immer mit Modellen, die nur einen Teil der Wirklichkeit darstellen. Wie die Erfahrung zeigt, ist die Mathematik trotzdem (eher wohl gerade deshalb) sehr erfolgreich bei der Behandlung von Problemen aus Naturwissenschaft, Wirtschaft und Technik. Wichtig ist nur, daß man sich der Grenzen der mathematischen Behandlung bewußt ist und ein Modell wählt, bei dem die Idealisierung keine wesentlichen Aspekte des Systems verändert – was leider sehr viel einfacher gesagt als getan ist. Immerhin stellt die Mathematik auch Methoden zur Validierung von Modellen zur Verfügung, mit denen man zumindest grobe Fehler erkennen kann.

Unabhängig davon sollte klar sein, daß wir beispielsweise Funktionen nicht als differenzierbar annehmen können, wenn ihre Diskrettheit ein wesentlicher Aspekt des Problems ist: Wenn wir ganzzahlige Lösungen suchen, die zwischen eins und zehn liegen, ist die Modellierung mit differenzierbaren Funktionen sicherlich nicht sinnvoll; dann muß man mit

den aufwendigeren Methoden der diskreten Optimierung arbeiten, die Thema der morgigen Sitzung sind. Wenn wir aber ganzzahlige Lösungen in der Größenordnung mehrerer Milliarden suchen, werden andere Aspekte der Idealisierung im allgemeinen zu deutlich größeren Fehlern führen als die Beschreibung durch eine kontinuierliche Variable, so daß die Modellierung durch eine stückweise differenzierbare Funktion meist vertretbar ist.

Heute wollen wir stets annehmen, daß wir es mit Problemen zu tun haben, die einer Modellierung mit zumindest stückweise differenzierbaren Funktionen zugänglich sind. Im eindimensionalen Fall sind wir damit im wohlvertrauten Gebiet der Kurvendiskussion und müssen drei Arten von Punkten untersuchen: die Nullstellen der Ableitung, die Randpunkte und die Punkte, in denen die Funktion nicht differenzierbar ist.

Thema der heutigen Sitzung wird sein, diesen Ansatz auf Funktionen mehrerer Veränderlicher zu verallgemeinern.

## §1: Funktionen mehrerer Veränderlicher

Eine Funktion  $f$  einer Veränderlicher ordnet jeder Zahl  $x$  aus einer gewissen Teilmenge  $D \subseteq \mathbb{R}$  der reellen Zahlen, dem Definitionsbereich von  $f$ , einen Funktionswert  $f(x) \in \mathbb{R}$  zu; wir schreiben die Funktion als

$$f: \begin{cases} D \rightarrow \mathbb{R} \\ x \mapsto f(x) \end{cases} .$$

Genauso erklären wir Funktionen mehrerer Veränderlicher: Eine Funktion von  $n$  Veränderlichen hat eine Teilmenge  $D \subseteq \mathbb{R}^n$  als Definitionsbereich und ordnet jedem Punkt  $(x_1, \dots, x_n) \in D$  einen Funktionswert  $f(x_1, \dots, x_n)$  zu; wir schreiben

$$f: \begin{cases} D \rightarrow \mathbb{R} \\ (x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n) \end{cases} .$$

Im Falle von zwei oder drei Variablen schreiben wir natürlich meist  $f(x, y)$  bzw.  $f(x, y, z)$  anstelle von  $f(x_1, x_2)$  bzw.  $f(x_1, x_2, x_3)$ . Einfa-

che Beispiele solchen Funktionen sind etwa die Funktion

$$g: \begin{cases} \mathbb{R}^3 \rightarrow \mathbb{R} \\ (x, y, z) \mapsto \sqrt{x^2 + y^2 + z^2} \end{cases} ,$$

die jedem Punkt des  $\mathbb{R}^3$  seinen Abstand vom Nullpunkt zuordnet, oder die Funktion

$$h: \begin{cases} \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x, y) \mapsto \sin(x+y) \cdot \cos(x-y) \end{cases} .$$

Zum besseren Verständnis einer gegebenen Funktion  $f: D \rightarrow \mathbb{R}$  einer Veränderlichen wird man fast immer deren Graph

$$\Gamma_f \stackrel{\text{def}}{=} \{(x, y) \in D \times \mathbb{R} \mid y = f(x)\}$$

betrachten, und hier gibt es bereits die erste Schwierigkeit für Funktionen mehrerer Veränderlicher: Für eine Funktion  $f: D \rightarrow \mathbb{R}$  mit  $D \subseteq \mathbb{R}^n$  von  $n$  Veränderlichen können wir den Graphen

$$\Gamma_f \stackrel{\text{def}}{=} \{(x_1, \dots, x_n, y) \in D \times \mathbb{R} \mid y = f(x_1, \dots, x_n)\}$$

zwar problemlos als Menge definieren, aber bereits für  $n = 2$  liegt er im dreidimensionalen Raum, wo wir nicht mehr zeichnen können. Für  $n > 2$  liegt er gar in einem mindestens vierdimensionalen Raum, den wir uns nicht einmal mehr vorstellen können.

Trotzdem ist der Graph zumindest für Funktionen zweier Veränderlicher ein nützliches Hilfsmittel zur Veranschaulichung; Abbildung 3 etwa zeigt recht gut das Verhalten der oben definierten Funktion  $h$ , besonders wenn man sich den Graphen am Computer unter verschiedenen Perspektiven anschaut.

Wer mit Graphen von Funktionen zweier Veränderlicher genügend Erfahrung gesammelt hat, wird im Laufe der Zeit die für einen Mathematiker extrem wichtige „abstrakte Anschauung“ entwickeln, die ihm erlaubt, anhand analoger Phänomene im dreidimensionalen Raum auch ein Gefühl dafür zu entwickeln, was in höherdimensionalen Räumen gelten sollte und was nicht. In einfachen Fällen wie der obigen Funktion  $g$ ,

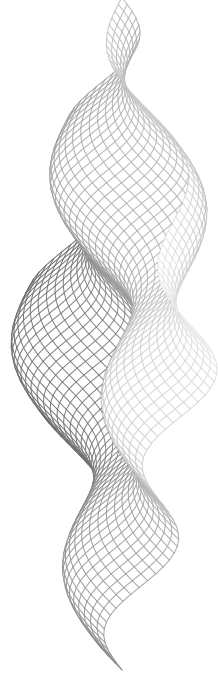


Abb. 3: Graph der Funktion  $h(x, y) = \sin(x + y) \cdot \cos(x - y)$

die den Abstand vom Nullpunkt beschreibt, gibt die zweidimensionale Variante

$$\tilde{g}: \begin{cases} \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x, y) \mapsto \sqrt{x^2 + y^2} \end{cases}$$

einen ersten Eindruck; für komplizierte Funktionen muß man nach und nach die notwendige Erfahrung gewinnen, um einen Blick für das Wesentliche zu bekommen.

Eine andere Möglichkeit zur Veranschaulichung von Funktionen zweier Veränderlicher ist von topographischen Karten her bekannt: Dort wird die Höhe über dem Meeresspiegel, eine Funktion der beiden Ebenenkoordinaten, dargestellt durch *Höhenlinien*. Entsprechend können wir für eine beliebige Funktion  $f: D \rightarrow \mathbb{R}$  mit  $D \subseteq \mathbb{R}^2$  und jeden Wert  $c \in \mathbb{R}$  die *Niveaulinie*

$$N_c(f) \stackrel{\text{def}}{=} \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = c\}$$

definieren; sie muß natürlich keine „Linie“ sein, sondern kann auch nur aus einigen Punkten bestehen, leer sein oder – im Falle einer konstanten Funktion – für einen Wert  $c$  aus dem gesamten Definitionsbereich  $D$  der Funktion bestehen.

Im Falle der obigen Funktion  $h$  etwa ist  $N_c(h)$  für  $|c| > 1$  die leere Menge; für  $c = \pm 1$  besteht sie nur aus isolierten Punkten, für  $c = 0$  aus einem Netz von Geraden, und für  $0 < |c| < 1$  aus geschlossenen Kurven.

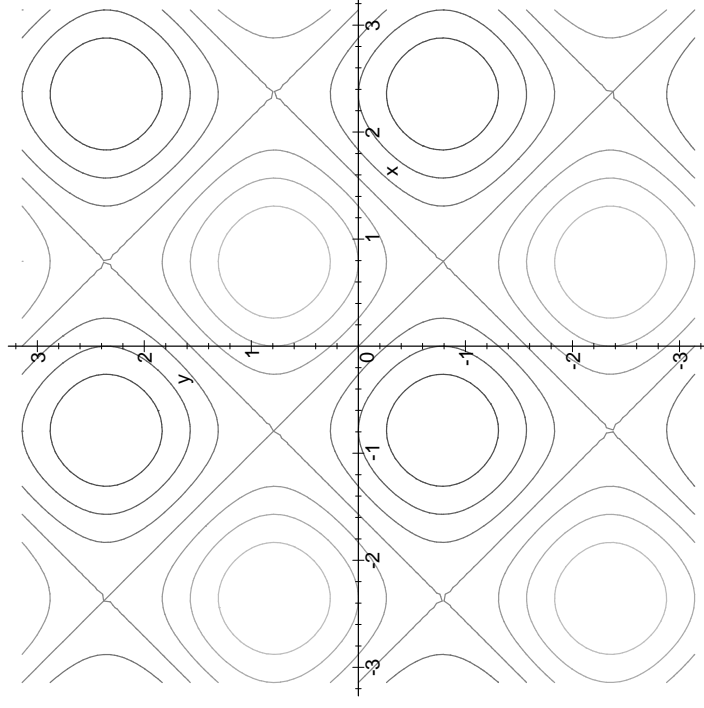


Abb. 4: Niveaulinien der Funktion  $h(x, y) = \sin(x + y) \cdot \cos(x - y)$

Für Funktionen von mehr als zwei Veränderlichen ist die Visualisierung naturgemäß schwieriger; wir können Graphen und Niveaulinien, -flächen usw. zwar problemlos definieren, aber nicht mehr zeichnen – es sei denn, es handelt sich um sehr einfache Niveauflächen im  $\mathbb{R}^3$ .

Ansonsten gibt es für Funktionen von mehr als zwei Veränderlichen im allgemeinen keine andere Möglichkeiten, als einen Teil der Variablen auf interessanten Werten festzuhalten und die so eingeschränkte Funktion darzustellen. Dies liefert natürlich kein vollständiges Bild der Funktion.

### §2: Differenzierbare Funktionen

Zum besseren Verständnis der Differenzierbarkeit von Funktionen mehrerer Veränderlicher empfiehlt es sich, zunächst noch einmal den wohl-bekanntesten Fall von Funktion *einer* Veränderlichen zu betrachten.

Die Idee hinter der Differenzierbarkeit besteht darin, daß man eine gegebene Funktion durch eine lineare Funktion annähern will oder – geometrisch gesprochen – ihren Graphen durch eine Gerade. Diese Annäherung kann man sich auch dadurch veranschaulichen, daß man Ausschnittsvergrößerungen des Graphen betrachtet: Wird der Ausschnitt immer kleiner vergrößern wir immer mehr, wird das Bild im Falle einer differenzierbaren Funktion immer ähnlicher zu einer Geraden. Die Abbildungen 5 bis 8 zeigen dies für  $f(x) = \sin x$  in der Umgebung von  $x = 1$ , wobei diese Umgebung von einer Abbildung zur nächsten jeweils um den Faktor fünf vergrößert wird.

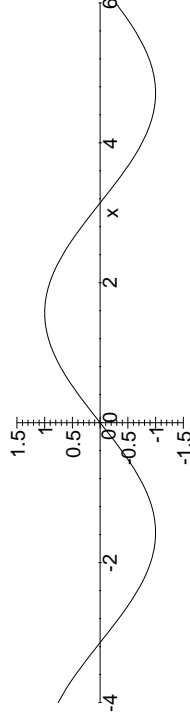


Abb. 5: Graph der Funktion  $y = \sin x$

Bei einer nichtdifferenzierbaren Funktion könnte sich ein völlig anderes Bild ergeben; beim Graph der Funktion  $f_\infty$  etwa sieht jede Ausschnittsvergrößerung wieder ungefähr so aus wie Abbildung 2.

Algebraisch betrachtet bedeutet die Annäherbarkeit durch lineare Funktionen, daß die Funktion in einer hinreichend kleinen Umgebung eines Punktes  $x$  ungefähr gleich einer linearen Funktion ist:

$$f(x+h) \approx f(x) + a \cdot h \quad \text{mit} \quad a = f'(x)$$

für kleine Werte von  $h$ .

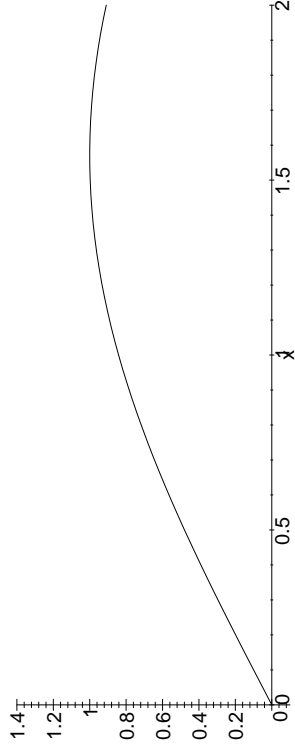


Abb. 6: Vergrößerung um den Faktor fünf um  $(1, \sin(1))$

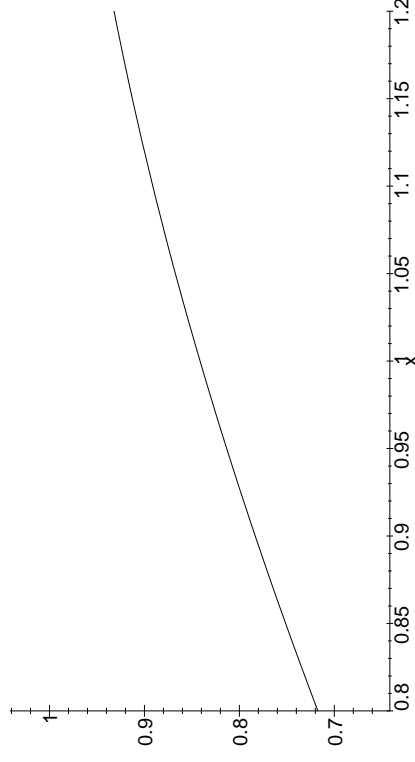


Abb. 7: Nochmalige Vergrößerung um den Faktor fünf um  $(1, \sin(1))$

Anstelle des etwas vagen Zeichens „ $\approx$ “ wollen wir in Zukunft genauer schreiben

$$f(x+h) \approx f(x) + a \cdot h + o(h),$$

wobei  $o(h)$  für *irgendeine* uns nicht weiter interessierende Funktion von  $h$  stehen soll, die für  $h \rightarrow 0$  schneller gegen null geht als  $h$  selbst; die „Gleichung“

$$\varphi(h) = \psi(h) + o(h)$$

soll also eine Abkürzung dafür sein, daß die Differenz  $\varphi(h) - \psi(h)$

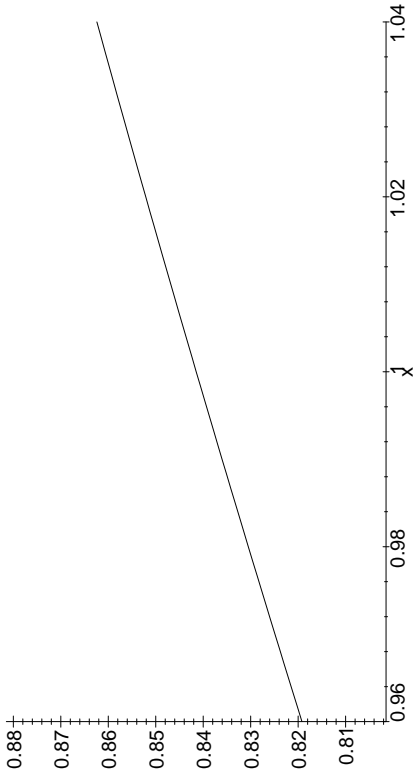


Abb. 8: Nach noch einer Vergrößerung sieht die Funktion praktisch linear aus

schneller als  $h$  gegen null geht; formal ausgedrückt:

$$\lim_{h \rightarrow 0} \frac{\varphi(h) - \psi(h)}{h} = 0.$$

Beispielsweise ist also

$$(1+h)^2 = 1 + 2h + o(h) \quad \text{und} \quad \sin h = h + o(h),$$

denn

$$(1+h)^2 = 1 + 2h + h^2 \quad \text{mit} \quad \lim_{h \rightarrow 0} \frac{h^2}{h} = \lim_{h \rightarrow 0} h = 0$$

und

$$\lim_{h \rightarrow 0} \frac{\sin h - h}{h} = \lim_{h \rightarrow 0} \frac{\sin h}{h} - 1 = 1 - 1 = 0.$$

Nach dem deutschen Mathematiker EDMUND LANDAU, der diese Schreibweise einführte, spricht man hier vom LANDAUSCHEN kleinen  $o$ . Anstelle von  $o(h)$  können wir auch allgemeinere Ausdrücke wie  $o(h^n)$  betrachten:

$$\varphi(h) = \psi(h) + o(h^n) \quad \text{besagt, daß} \quad \lim_{h \rightarrow 0} \frac{\varphi(h) - \psi(h)}{h^n} = 0$$

ist, und ganz allgemein soll für eine beliebige Funktion  $\omega(h)$  gelten

$$\varphi(h) = \psi(h) + o(\omega(h)) \quad \text{genau dann, wenn} \quad \lim_{h \rightarrow 0} \frac{\varphi(h) - \psi(h)}{\omega(h)} = 0$$

ist.



EDMUND GEORG HERMANN LANDAU (1877–1938) wurde in Berlin geboren und studierte an der dortigen Universität, wo er auch von 1899 bis 1909 lehrte. Dann bekam er einen Ruf an die damals führende deutsche Mathematikfakultät in Göttingen. 1933 verlor er seinen dortigen Lehrstuhl, denn die Studenten boykottierten seine Vorlesungen, da sie meinten, sie könnten Mathematik nur von einem Professor ihrer eigenen Rasse lernen. LANDAU zahlreiche Publikationen beschäftigten sich vor allem mit der Zahlentheorie, über die er auch ein bedeutendes Lehrbuch schrieb; sehr bekannt sind seine Arbeiten über die Verteilung von Primzahlen.

Mit LANDAUS  $o$ -Notation können wir auch sagen, die Funktion  $f$  sei genau dann differenzierbar in  $x$  mit Ableitung  $f'(x) = a$ , wenn

$$f(x+h) = f(x) + a \cdot h + o(h)$$

ist, denn das bedeutet gerade, daß

$$\lim_{h \rightarrow 0} \frac{f(x+h) - (f(x) + a \cdot h)}{h} = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} - a = 0$$

ist, d.h.

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = a.$$

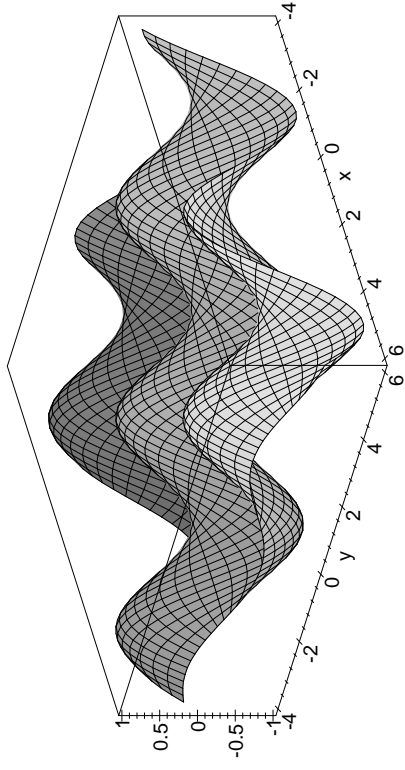
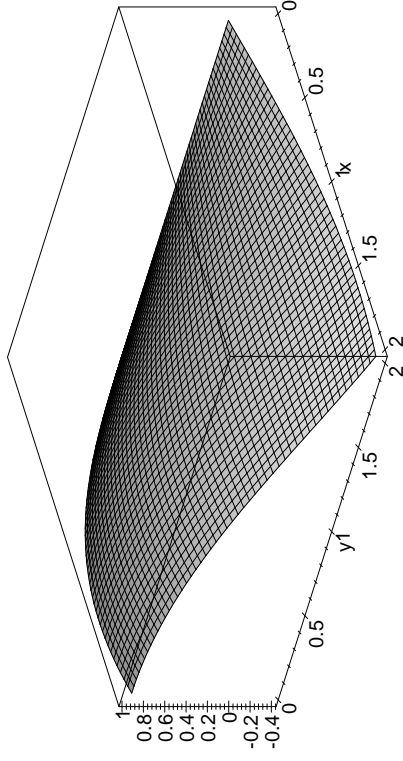
Im Zweidimensionalen können wir ganz entsprechend vorgehen: Wenn wir etwa die Funktion  $f(x, y) = \sin x \cos y$  auf  $\mathbb{R}^2$  in der Umgebung des Punktes  $(1, 1)$  betrachten und diese wieder sukzessive um den Faktor fünf vergrößern, erhalten wir die Abbildungen 9 bis 12, die zeigen, daß sich der Graph der Funktion in einer hinreichend kleinen Umgebung des betrachteten Punktes nur wenig von einer Ebenen unterscheidet, d.h. die Funktion ist „im Kleinen“ annähernd linear.

Eine lineare Funktion zweier Veränderlicher hat die Form

$$L(x, y) = ax + by + c \quad \text{mit} \quad a, b, c \in \mathbb{R},$$

also ist

$$L(x+h, y+k) = a(x+h) + b(y+k) + c = (ax + by + c) + ah + bk$$

Abb. 9: Graph der Funktion  $z = \sin x \cos y$ Abb. 10: Vergrößerung um den Faktor fünf um  $(1, 1, \sin(1) \cos(1))$ 

eine Approximation für  $f$  in der Umgebung des betrachteten Punkts  $(x, y)$ , im obigen Beispiel also des Punkts  $(1, 1)$ . Im Punkt  $(x, y)$  sollte

$$L(x, y) = a + bx + cy$$

natürlich mit  $f(x, y)$  übereinstimmen, und für  $(h, k) \neq (0, 0)$  sollte der Unterschied zwischen  $f$  und  $L$  schneller gegen Null gehen als  $h$  und  $k$ .

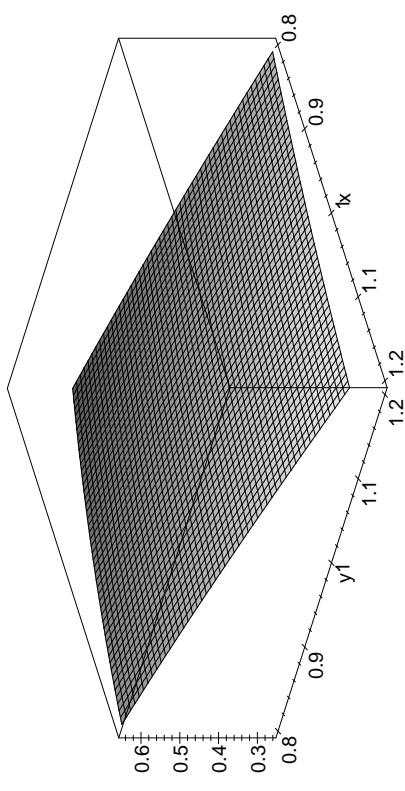
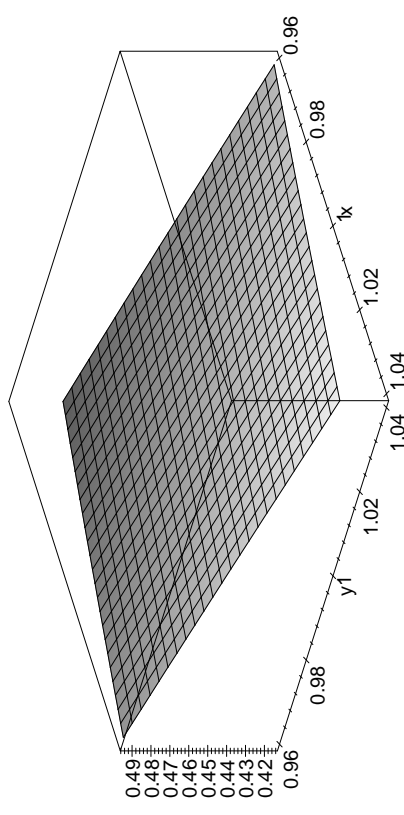
Abb. 11: Nochmalige Vergrößerung um den Faktor fünf um  $(1, 1, \sin(1) \cos(1))$ 

Abb. 12: Nach noch einer Vergrößerung sieht die Funktion praktisch linear aus

Mit dem LANDAUSchen kleinen  $o$  können wir dies schreiben als

$$f(x + h, y + k) = f(x, y) + ah + bk + o(\max(|h|, |k|)),$$

was, in Übereinstimmung mit der allgemeinen Konvention, bedeuten



soll, daß

$$\lim_{\substack{h \rightarrow 0 \\ k \rightarrow 0}} \frac{f(x+h, y+k) - f(x, y) + ah + bk}{\max(|h|, |k|)} = 0$$

ist. Selbstverständlich gilt das nicht für *jede* Funktion  $f$ ; diejenigen Funktionen zweier Veränderlicher, zu denen es solche Zahlen  $a, b$  gibt, bezeichnen wir als *differenzierbar*, was sich wie im Eindimensionalen wieder so interpretieren läßt, daß die Funktion „im Kleinen“ linear ist.

**Definition:** Eine Funktion  $f: D \rightarrow \mathbb{R}$  mit  $D \subseteq \mathbb{R}^2$  heißt *differenzierbar* im Punkt  $(x, y) \in D$ , wenn es reelle Zahlen  $(a, b)$  gibt, so daß für hinreichend kleine Zahlen  $h, k \in \mathbb{R}$  gilt

$$f(x+h, y+k) = f(x, y) + ah + bk + o\left(\max(|h|, |k|)\right).$$

Mit dieser Definition gibt es noch ein Problem: Je nachdem, wie die Menge  $D$  aussieht, liegt der Punkt  $(x+h, y+k)$  möglicherweise selbst für sehr kleine Werte von  $h$  und  $k$  gar nicht mehr in  $D$ ; dies wäre beispielsweise dann der Fall, wenn wir eine Funktion auf dem Gitter  $D = \mathbb{Z}^2 \subset \mathbb{R}^2$  betrachten. Für eine solche Funktion ist es aber ganz offensichtlich unsinnig, von Tangentialebenen und Ableitungen zu sprechen; wir müssen also voraussetzen, daß  $D$  mit  $(x, y)$  zumindest für hinreichend kleine Werte von  $h$  und  $k$  auch noch die Punkte  $(x+h, y+k)$  enthält. Eine Menge  $D$ , in der dies für jeden Punkt  $(x, y) \in D$  gilt, bezeichnet man als *offen*:

**Definition:** Eine Menge  $D \subseteq \mathbb{R}^2$  heißt *offen*, wenn es für jeden Punkt  $(x, y) \in D$  eine positive reelle Zahl  $\varepsilon > 0$  gibt, so daß für alle  $h, k \in \mathbb{R}$  mit  $|h| < \varepsilon$  und  $|k| < \varepsilon$  auch der Punkt  $(x+h, y+k)$  in  $D$  liegt.

Von differenzierbaren Funktionen können wir somit nur dann sinnvoll reden, wenn der Definitionsbereich  $D$  offen ist.

Beispiele offener Mengen sind etwa die offene Kreisscheibe

$$K_r = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < r^2\}$$

oder das offene Quadrat

$$Q_a = \{(x, y) \in \mathbb{R}^2 \mid |x| < a, |y| < a\}.$$

Nicht offen sind dagegen die entsprechenden Mengen, bei denen das Kleinerzeichen durch  $\leq$  ersetzt wird, denn im Falle etwa des abgeschlossenen Quadrats

$$\overline{Q}_a = \{(x, y) \in \mathbb{R}^2 \mid |x| \leq a, |y| \leq a\}$$

kann man vom Punkt  $(0, a)$  aus nicht mehr weiter nach rechts gehen, d.h. für kein  $k > 0$  liegt  $(0, a+k)$  in  $\overline{Q}_a$ .

**Bemerkung:** Einige Leser werden wohl eine andere Definition einer offenen Menge im  $\mathbb{R}^2$  gewohnt sein, nämlich daß für alle  $(h, k) \in \mathbb{R}^2$  aus einem Kreis mit Radius  $\varepsilon$  um den Nullpunkt, d.h. für alle  $(h, k)$  mit  $h^2 + k^2 < \varepsilon^2$ , der Punkt  $(x+h, y+k)$  in  $D$  liegen soll. Obige Definition dagegen besagt, daß ein Quadrat mit Mittelpunkt  $(x, y)$  und Kantenlänge  $2\varepsilon$  in  $D$  liegen soll. Da in beiden Fällen nichts über die Größe von  $\varepsilon$  ausgesagt wird, sind diese beiden Definitionen äquivalent, denn jeder Kreis enthält ein Quadrat und umgekehrt.

Damit wissen wir, wann  $f$  differenzierbar ist; was aber ist die *Ableitung* von  $f$ ?

Im eindimensionalen Fall sind die differenzierbaren Funktionen genau die, die sich durch Geraden annähern lassen, und als Ableitung  $f'(x)$  bezeichnen wir die Steigung derjenigen Geraden, die den Graphen von  $f$  im Punkt  $(x, f(x))$  am besten approximiert. Rechnerisch ist dann

$$f(x+h) = f(x) + h \cdot f'(x).$$

Im zweidimensionalen Fall läßt sich eine differenzierbare Funktion  $f(x, y)$  in jedem Punkt durch eine Ebene annähern, was wir rechnerisch durch die Formel

$$f(x+h, y+h) = f(x, y) + ah + bk + o\left(\max(|h|, |k|)\right)$$

ausdrücken können. Als Ableitung oder, wie man meistens sagt, *Gradient* von  $f$ , bezeichnen wir daher den Vektor  $\begin{pmatrix} a \\ b \end{pmatrix}$ :

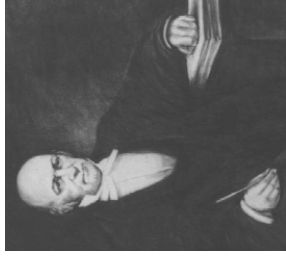
**Definition:** Ist die Funktion  $f: D \rightarrow \mathbb{R}$  mit  $D \subseteq \mathbb{R}^2$  differenzierbar und ist im Punkt  $(x, y) \in D$

$$f(x+h, y+h) = f(x, y) + ah + bk + o\left(\max(|h|, |k|)\right),$$

so bezeichnen wir als *Ableitung* oder *Gradient* von  $f$  im Punkt  $(x, y)$  den Vektor

$$\text{grad } f(x, y) = \nabla f(x, y) = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Das Symbol  $\nabla$  sieht zwar aus wie ein griechischer Buchstabe, ist aber keiner; es ist ein auf den Kopf gestelltes großes Delta ( $\Delta$ ).  $\nabla f$  wird „Nabla  $f$ “ ausgesprochen nach dem griechischen Wort  $\nu\alpha\beta\lambda\alpha$  = Leiter; die Bezeichnung wurde eingeführt von dem irischen Mathematiker WILLIAM ROWEN HAMILTON, den die Form von  $\nabla$  an eine Leiter erinnerte.



WILLIAM ROWEN HAMILTON (1805–1865) wurde in Dublin geboren; bereits mit fünf Jahren sprach er Latein, Griechisch und Hebräisch. Mit dreizehn begann er, mathematische Literatur zu lesen, mit 21 wurde er, noch als Student, Professor der Astronomie am Trinity College in Dublin. Schon bald verlor er sein Interesse für Astronomie, arbeitete aber weiterhin auf dem Gebiet der Mathematik und Physik. Am bekanntesten ist seine Entdeckung der Quaternionen 1843; vorher publizierte er bedeutende Arbeiten aus so unterschiedlichen Gebieten wie der Optik, der Dynamik und der Algebra.

Bei mehr als zwei Veränderlichen verläßt uns zwar die geometrische Anschauung, aber der Formalismus läßt sich problemlos auf beliebige Variablenzahl verallgemeinern:

**Definition:** a) Eine Teilmenge  $D \subseteq \mathbb{R}^n$  heißt offen, wenn es zu jedem Punkt  $(x_1, \dots, x_n) \in D$  ein  $\varepsilon > 0$  gibt, so daß für alle reellen Zahlen  $h_i$  mit  $|h_i| < \varepsilon$  auch der Punkt  $(x_1 + h_1, \dots, x_n + h_n)$  in  $D$  liegt.

b) Eine Funktion  $f: D \rightarrow \mathbb{R}$  auf der offenen Teilmenge  $D \subseteq \mathbb{R}^n$  heißt *differenzierbar* im Punkt  $(x_1, \dots, x_n) \in D$ , wenn es reelle Zahlen  $a_1, \dots, a_n$  gibt, so daß

$$f(x_1 + h_1, \dots, x_n + h_n) = f(x_1, \dots, x_n) + a_1 \cdot h_1 + \dots + a_n \cdot h_n + o(\max |h_i|)$$

ist. Alsdann bezeichnen wir den Vektor

$$\text{grad } f(x_1, \dots, x_n) = \nabla f(x_1, \dots, x_n) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n$$

als *Ableitung* oder *Gradient* von  $f$  im Punkt  $(x_1, \dots, x_n)$ .

Mit dieser Definition haben wir etwas ähnliches wie die Definition

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

für Funktionen einer Veränderlicher: Damit ist die Ableitung zwar *definiert*, aber sie wird – außerhalb von Anfängervorlesungen – praktisch nie so ausgerechnet. Entsprechend sollten wir auch für Funktionen mehrerer Veränderlicher eine effizientere Methode der Differentiation finden.

Auch dazu hilft uns die Geometrie weiter: Für eine Funktion zweier Veränderlicher gibt die Ableitung an, durch welche Ebene der Graph der Funktion am besten angenähert wird. Eine Ebene ist aber bestimmt durch zwei auf ihr liegende Geraden, und dazu können wir beispielsweise die beiden Geraden nehmen, entlang derer jeweils eine der beiden Variablen konstant ist.

Betrachten wir etwa die Gerade  $y = v$  auf der  $y$  gleich einer Konstanten  $v$  ist. Diese Gerade ist der Schnitt des Graphen von  $f$  mit der Ebenen  $y = v$  im dreidimensionalen Raum und ist damit auf dieser Ebenen der Graph der Funktion  $z = g_1(x) \stackrel{\text{def}}{=} f(x, v)$ , d.h. der Graph einer Funktion einer Veränderlichen. Anschaulich ist ziemlich klar, daß die Tangente an den Graphen dieser Funktion gleich dem Schnitt der Tangentialebenen an den Graphen von  $f$  mit der Ebenen  $y = v$  sein sollte, und bei rein formaler Betrachtung ist dies in der Tat völlig trivial: Wir wissen, daß

$$f(x+h, y+k) = f(x, y) + ah + bk + o(\max(|h|, |k|))$$

ist, und wenn wir in dieser Formel  $y = v$  konstant halten, ist  $k = 0$ , d.h.

$$g_1(x+h) = f(x+h, v) = f(x, v) + ah + o(h) = g_1(x) + ah + o(h).$$

Somit ist  $a$  gleich der Ableitung von  $g_1$  im Punkt  $x$ , d.h.

$$a = g_1'(x).$$

Genauso können wir auch  $x$  auf einem konstanten Wert  $u$  festhalten und die Funktion  $z = g_2(y) \stackrel{\text{def}}{=} f(u, y)$  betrachten; dieselbe Überlegung wie oben zeigt, daß dann

$$b = g_2'(y)$$

ist. Somit ist

$$\text{grad } f(x, y) = \begin{pmatrix} g'_1(x) \\ g'_2(y) \end{pmatrix}.$$

Wir bezeichnen die beiden Ableitungen  $g'_1(x)$  und  $g'_2(y)$  als die *partiellen* Ableitungen von  $f$  nach  $x$  und  $y$ ; in Zeichen

$$f_x(x, y) = \frac{\partial f}{\partial x}(x, y) = \underset{\text{def}}{=} g'_1(x) \quad \text{und} \quad f_y(x, y) = \frac{\partial f}{\partial y}(x, y) = \underset{\text{def}}{=} g'_2(y)$$

Das Zeichen  $\partial$  ist wieder etwas, was wie ein griechischer Buchstabe aussieht, aber keiner ist; einige Leute bezeichnen es als das „kleine nablá“, andere als „def“ wie *delta*. Meist sagt man aber einfach etwas wie „ $f$  partiell abgeleitet nach  $x$ “.

Für Funktionen von mehr als zwei Veränderlichen können wir zwar nicht mehr anschaulich mit Ebenen argumentieren, aber die formale Argumentation mit LANDAUS kleinem  $o$  bleibt gültig: Wenn die Funktion  $f: D \rightarrow \mathbb{R}$  mit  $D \subseteq \mathbb{R}^n$  im Punkt  $(x_1, \dots, x_n) \in D$  differenzierbar ist, ist für alle hinreichend kleinen reellen Zahlen  $h_i$

$$f(x_1 + h_1, \dots, x_n + h_n) = f(x_1, \dots, x_n) + a_1 h_1 + \dots + a_n h_n + o(\max_i |h_i|).$$

Setzen wir darin alle Variablen  $x_j$  *außer* der  $i$ -ten auf geeignete Konstanten  $u_j$  und betrachten wir  $x_i$  in der Umgebung der reellen Zahl  $u_i$ , so verschwinden alle  $h_j$  außer dem  $i$ -ten und

$$f(u_1, \dots, u_i + h_i, \dots, u_n) = f(u_1, \dots, u_n) + a_i h_i + o(h_i).$$

Definieren wir eine Funktion

$$g_i(x_i) \stackrel{\text{def}}{=} f(u_1, \dots, x_i, \dots, u_n),$$

so ist

$$g_i(u_i + h_i) = g_i(u_i) + a_i h_i + o(h_i),$$

d.h.  $g_i$  ist differenzierbar im Punkt  $u_i$  mit Ableitung  $g'_i(u_i) = a_i$ .

Auch hier schreiben wir wieder

$$f_{x_i}(x_1, \dots, x_n) = \frac{\partial f}{\partial x_i}(x_1, \dots, x_n) \stackrel{\text{def}}{=} g'_i(u_i)$$

und bezeichnen diese Funktion als die *partielle Ableitung* von  $f$  nach  $x_i$ ; alsdann ist

$$\text{grad } f(x_1, \dots, x_n) = \nabla f(x_1, \dots, x_n) = \begin{pmatrix} f_{x_1}(x_1, \dots, x_n) \\ \vdots \\ f_{x_n}(x_1, \dots, x_n) \end{pmatrix}.$$

Häufig wird der Gradient durch diese Formel *definiert*; dann folgt aber, wie das nachfolgende Beispiel zeigt, aus der Existenz des Gradienten leider nicht notwendigerweise die Differenzierbarkeit der Funktion:

Die Funktion

$$f: \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R} \\ (x, y) & \mapsto \begin{cases} 1 & \text{falls } xy \neq 0 \\ 0 & \text{falls } xy = 0 \end{cases} \end{cases}$$

ist Null auf der  $x$ -Achse und der  $y$ -Achse, und eins überall sonst. Damit existieren im Punkt  $(0, 0)$  beide partielle Ableitungen und sind identisch Null. Trotzdem ist  $f$  selbstverständlich nicht differenzierbar in  $(0, 0)$ , denn da  $f(h, k)$  für jeden Punkt, der nicht auf einer der beiden Koordinatenachsen liegt, gleich eins ist, kann es keine lineare Funktion  $L(h, k) = bh + ck$  geben, so daß

$$f(h, k) = f(0, 0) + L(h, k) + o(\max(|h|, |k|)) = o(\max(|h|, |k|))$$

ist, denn  $L(0, 0) = f(0, 0) = 0$  und  $f(h, k) = 1$  für  $hk \neq 0$ .

Nun wird natürlich jeder vernünftige Mensch einwenden, daß dieses Beispiel sehr künstlich ist, und in der Tat verhalten sich „gutartige“ Funktionen nicht so:

**Lemma:** Falls  $f: D \rightarrow \mathbb{R}$  auf der offenen Teilmenge  $D \subseteq \mathbb{R}^n$  stetig ist und auch alle partiellen Ableitungen von  $f$  dort existieren und stetig sind, ist  $f$  in  $D$  differenzierbar und

$$\text{grad } f(x_1, \dots, x_n) = \nabla f(x_1, \dots, x_n) = \begin{pmatrix} f_{x_1}(x_1, \dots, x_n) \\ \vdots \\ f_{x_n}(x_1, \dots, x_n) \end{pmatrix}.$$

Der Beweis dieses Lemmas und auch die für sein Verständnis wesentliche Definition der Stetigkeit von Funktionen mehrerer Veränderlicher würde uns zu weit vom Thema dieses Seminars wegführen; daher sei darauf verzichtet.

### §3: Optimierung ohne Nebenbedingungen

Nachdem wir wissen, was die Ableitung einer Funktion mehrerer Veränderlichen ist, sollten wir – falls die Erfahrungen aus dem Fall einer Veränderlichen verallgemeinerbar sind – in der Lage sein, eine notwendige Bedingung für ein Extremum anzugeben. Nachdem wir trotz der sehr begrenzten Zeit, die für dieses Seminar zur Verfügung steht, uns so lange mit dem Gradienten beschäftigt haben, ist eigentlich klar, daß – in Analogie zum eindimensionalen Fall – die folgende Aussage gelten sollte:

**Lemma:** Wenn die differenzierbare Funktion  $f: D \rightarrow \mathbb{R}$  im Punkt  $x_1, \dots, x_n \in D$  ein lokales Extremum hat, ist dort ihr Gradient gleich dem Nullvektor, d.h. alle partiellen Ableitungen verschwinden.

*Beweis:* Anschaulich ist die Behauptung ziemlich klar, denn wenn die lineare Funktion, durch die wir  $f$  approximieren, nicht konstant ist, gibt es Richtungen, in denen sie ansteigt, und solche in denen sie fällt; der Punkt kann also weder Maximum noch Minimum sein.

Auch formal läßt sich das leicht einsehen, denn nach Definition der Differenzierbarkeit ist für kleine Werte  $h_i$

$$f(x_1 + h_1, \dots, x_n + h_n) = f(x_1, \dots, x_n) + h_1 f_{x_1}(x_1, \dots, x_n) + \dots + h_n f_{x_n}(x_1, \dots, x_n) + o(\max_i |h_i|).$$

Falls auch nur eine der partiellen Ableitungen nicht verschwindet, lassen sich Zahlenwerte für die  $h_i$  finden, so daß die Summe

$$h_1 f_{x_1}(x_1, \dots, x_n) + \dots + h_n f_{x_n}(x_1, \dots, x_n)$$

positiv ist, und auch solche, für die sie negativ ist. Daran ändert sich auch dann nichts, wenn wir alle  $h_i$  mit einer kleinen Konstanten  $\lambda > 0$

multiplizieren, und wenn wir diese hinreichend klein wählen, kann der durch das kleine  $\lambda$  abgeschätzte Term vernachlässigt werden. Also gibt es in einer Umgebung des betrachteten Punktes sowohl größere als auch kleinere Funktionswerte, so daß dort kein Extremum sein kann. ■

Die Frage, welche Nullstellen des Gradienten wirklich Extremwerten entsprechen, ist – wie auch schon im Eindimensionalen – schwieriger; in der Praxis wird es meist am einfachsten sein, sich die Umgebung des betreffenden Punktes mit irgendwelchen *ad hoc*-Methoden genauer anzusehen und dann zu entscheiden.

Bei der Funktion

$$f: \begin{cases} \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x, y) \mapsto (x+y)^2 + \cos^2(x-y) \end{cases}$$

etwa ist

$$\nabla f(x, y) = \begin{pmatrix} 2(x+y) - 2\cos(x-y)\sin(x-y) \\ 2(x+y) + 2\cos(x-y)\sin(x-y) \end{pmatrix}$$

und für jedes lokale Extremum müssen beide Komponenten dieses Vektors verschwinden. Wir müssen also das nichtlineare Gleichungssystem

$$2(x+y) - 2\cos(x-y)\sin(x-y) = 0$$

$$2(x+y) + 2\cos(x-y)\sin(x-y) = 0$$

lösen. Einzelnen betrachtet sieht zwar keine der beiden Gleichungen sonderlich angenehm aus, als System sind sie aber durchaus handhabbar: Das Gleichungssystem ist genau dann erfüllt, wenn die Summe und die Differenz der beiden Gleichungen null ist, und das führt, wenn wir die hier irrelevanten Vorfaktoren  $\pm 4$  gleich weglassen, auf das neue System

$$x+y=0 \quad \text{und} \quad \cos(x-y)\sin(x-y)=0.$$

Die erste Gleichung sagt, daß  $y = -x$  sein muß, und wenn wir das in die zweite einsetzen, folgt

$$\cos 2x \sin 2x = 0.$$

Der Sinus verschwindet genau bei den ganzzahligen Vielfachen von  $\pi$ , der Cosinus genau zwischen diesen Werten; also verschwindet das Produkt aus Sinus und Cosinus genau bei den ganzzahligen Vielfachen

von  $\frac{\pi}{4}$  und die Funktion  $\cos 2x \sin 2x$  bei denen von  $\frac{\pi}{4}$ . Kandidaten für lokale Extreme sind also die Punkte

$$\left( \frac{k\pi}{4}, -\frac{k\pi}{4} \right) \quad \text{mit } k \in \mathbb{Z}.$$

Die Funktionswerte dort sind

$$\begin{aligned} f\left(\frac{k\pi}{4}, -\frac{k\pi}{4}\right) &= \left(\frac{k\pi}{4} - \frac{k\pi}{4}\right)^2 + \cos^2\left(\frac{k\pi}{4} + \frac{k\pi}{4}\right) = \cos^2 \frac{k\pi}{2} \\ &= \begin{cases} 0 & \text{für ungerade } k \\ 1 & \text{für gerade } k \end{cases}. \end{aligned}$$

Da die Funktion  $f$  als Summe zweier Quadrate keine negativen Werte annimmt, ist klar, daß wir für ungerade  $k$  nicht nur lokale sondern sogar globale Minima haben.

Für gerades  $k$  sollte man nach den Erfahrungen mit dem eindimensionalen Fall (lokale) Maxima erwarten, denn bekanntlich liegt bei einer stetigen Funktion einer Veränderlichen zwischen zwei Minima stets ein Maximum. Schon im zweidimensionalen Fall ist allerdings nicht mehr ganz klar, was das Wort „zwischen“ bedeuten soll: In der Ebenen gibt es schließlich viele Wege, die von einem gegebenen Punkt zu einem anderen führen. Um von einem Minimum zum nächsten zu kommen, müssen wir zwar nach oben klettern und dann wieder absteigen, so daß wir zwischendrin irgendwo einen höchsten Punkt erreichen; aber zum Glück für jeden Bergwanderer muß dieser höchste Punkt des Weges nicht unbedingt ein Gipfel sein.

Betrachten wir daher den Fall  $k = 0$  etwas genauer; wir befinden uns dann also im Nullpunkt. Wenn wir uns auf der zweiten Winkelhalbierenden bewegen, d.h. auf der Geraden  $y = -x$ , befinden wir uns dort in der Tat in einem Maximum, denn

$$f(x, -x) = (x - x)^2 + \cos^2 2x = \cos^2 2x$$

hat bei  $x = 0$  ein (sogar absolutes) Maximum.

Bewegen wir uns aber auf der ersten Winkelhalbierenden, ist

$$f(x, x) = (x + x)^2 + \cos^2(x - x) = 4x^2 + 1,$$

und jetzt ist der Punkt  $x = 0$  zum (absoluten) Minimum geworden.

Damit ist klar, daß der Punkt  $x = 0$  kein Maximum der Funktion  $f(x, y)$  sein kann, obwohl der Gradient von  $f$  dort verschwindet. Das geometrische Bild ist ähnlich einer Paßhöhe im Gebirge: Die Paßstraße hat ihren höchsten Punkt erreicht und geht nun wieder abwärts, links und rechts geht es aber höher in Richtung Gipfel.

Der mathematische Begriff für solche Punkte orientiert sich allerdings nicht an Paßhöhen, sondern an einem (Pferde-)Sattel: Dieser geht rechts und links nach unten, vorne und hinten aber nach oben. In Analogie dazu bezeichnet man einen Punkt, in dem der Gradient verschwindet, der aber weder ein relatives Maximum noch ein relatives Minimum ist, als Sattelpunkt.

Abbildung 13 zeigt das einfachste Beispiel eines Sattelpunkts, den Nullpunkt auf dem Graph der Funktion  $g(x, y) = x^2 - y^2$ .

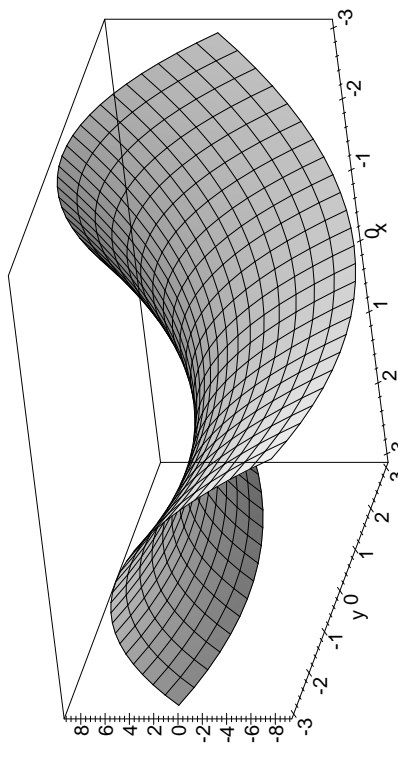


Abb. 13: Graph der Funktion  $g(x, y) = x^2 - y^2$

Im obigen Beispiel zeigt die gleiche Überlegung wie oben im Fall  $k = 0$ , daß jeder der Punkte  $(\frac{k\pi}{4}, -\frac{k\pi}{4})$  mit geradem  $k$ , d.h. jeder Punkt  $(\frac{\ell\pi}{2}, -\frac{\ell\pi}{2})$  mit ganzzahligem  $\ell$ , ein Sattelpunkt ist. Abbildung 14 zeigt

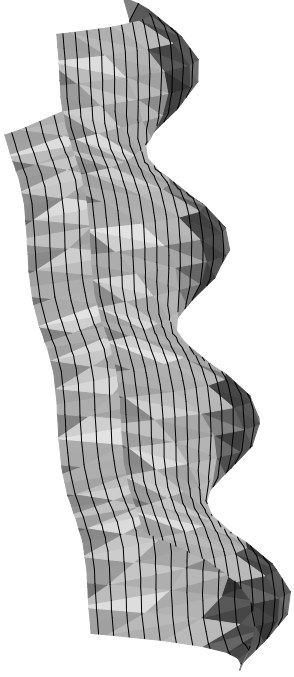


Abb. 14: Graph der Funktion  $f(x, y) = (x + y)^2 + \cos^2(x - y)$

einige der Minima und Sattelpunkte; damit man diese besser sieht, sind im Graphen alle Funktionswerte größer zwei ausgeblendet.

Einen fast noch klareren Eindruck von  $f$  vermitteln die Niveaulinien in Abbildung 15: Für Funktionswerte, die größer null, aber kleiner als eins sind, bestehen die Niveaulinien aus unendlich vielen nicht miteinander zusammenhängenden geschlossenen Kurven, die um die Minima herumführen. Für das Niveau eins haben wir eine verzopfte Kurve, die sich in den Sattelpunkten selbst schneidet, und für größere Funktionswerte bestehen die Niveaulinien aus jeweils zwei wellenförmigen Kurven, deren Amplitude mit zunehmendem Niveau abnimmt.

#### §4: Optimierung zur Schätzung von Parametern

Oftmals ist zu gegebenen Beobachtungsdaten grundsätzlich bekannt, welcher Art von Gesetz sie genügen sollten, und das Problem besteht „nur“ noch darin, die in diesem Gesetz vorkommenden Parameter zu bestimmen. Im einfachsten Fall könnte man etwa an einen Widerstand denken, der dadurch gemessen wird, daß man verschiedene Spannungen  $U_i$  anlegt und die zugehörigen Stromstärken  $I_i$  mißt. Nach dem Ohmschen Gesetz ist dann  $U_i = R \cdot I_i$ , aber aufgrund der unvermeidlichen Meßfehler werden die verschiedenen Quotienten  $U_i/I_i$  natürlich nicht alle gleich sein. Die Lösung dieses Problems ist klar: Man nimmt

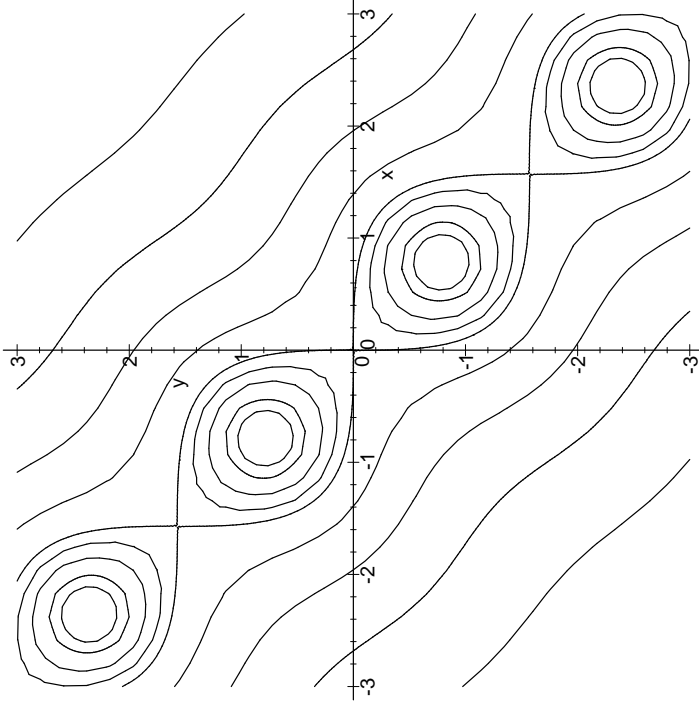


Abb. 15: Niveaulinien der Funktion  $f(x, y) = (x + y)^2 + \cos^2(x - y)$

den Mittelwert der Quotienten. Schwieriger wird es, wenn mehrere Parameter ins Spiel kommen.

Der allgemeine Ansatz ist folgender: Sei  $y = f(a, b, \dots; x)$  der von den Parametern  $a, b, \dots$  abhängige Zusammenhang zwischen den Größen  $x$  und  $y$ , und setze  $(x_1, y_1), \dots, (x_n, y_n)$  gegebene (Meß-)Werte, an die die Parameter angepaßt werden sollen.

Nehmen wir an, es gäbe Parameter  $a, b, \dots$ , für die  $y = f(a, b, \dots; x)$  den exakten Zusammenhang zwischen den Größen  $y$  und  $x$  wiedergibt,

daß aber die Werte  $y_i$  aufgrund von Beobachtungsfehlern von den Funktionswerten  $f(a, b, \dots; x_i)$  abweichen – das ist eine Situation, mit denen man es bei einfacheren Problemen in den Naturwissenschaften und in der Technik oft zu tun hat.

In anderen Gebieten wissen wir möglicherweise genau, daß die Funktion  $y_i = f(a, b, \dots; x_i)$  eine drastische Vereinfachung der Realität ist, wir wollen sie aber trotzdem, mangels einer besseren Alternative, optimal anpassen – dies ist üblicherweise bei komplexen Systemen der Fall – egal ob diese aus der Technik, den Natur-, den Sozial- oder den Wirtschaftswissenschaften kommen.

Betrachten wir etwa ein Beispiel aus der Medizin: Nach dem *Merck Manual of Diagnosis and Therapy*, Rahway NH, <sup>15</sup>1987 (hier zitiert nach

E.K. YEARGES, R.W. SCHINKWILER, J.V. HEROD: *An introduction to the mathematics of biology: with computer algebra models*, Birkhäuser, 1996, S. 19),

hängt die durchschnittliche Körpergröße von Kindern folgendermaßen mit dem Alter zusammen:

Alter in Jahren	1	3	5	7	9	11	13
Größe in cm	75	92	108	121	130	142	155

Wie Abbildung 16 zeigt, liegen die Punkte ungefähr auf einer Geraden.

Nun ist es natürlich absurd anzunehmen, die Körpergröße sei grundsätzlich eine lineare Funktion des Alters: Achtzigjährige sind im Durchschnitt nicht größer wie Zwanzigjährige, eher im Gegenteil. Trotzdem können wir versuchen, eine lineare Funktion

$$\text{Körpergröße} = m \cdot \text{Alter} + b$$

zu finden, die den Zusammenhang *im angegebenen Bereich* möglichst gut beschreibt.

Unter „möglichst gut“ können wir dabei vieles verstehen: Wir könnten zum Beispiel verlangen, daß die *Summe* aller Abweichungen möglichst gering ist. Die hat dann allerdings den meist nicht erwünschten Effekt,

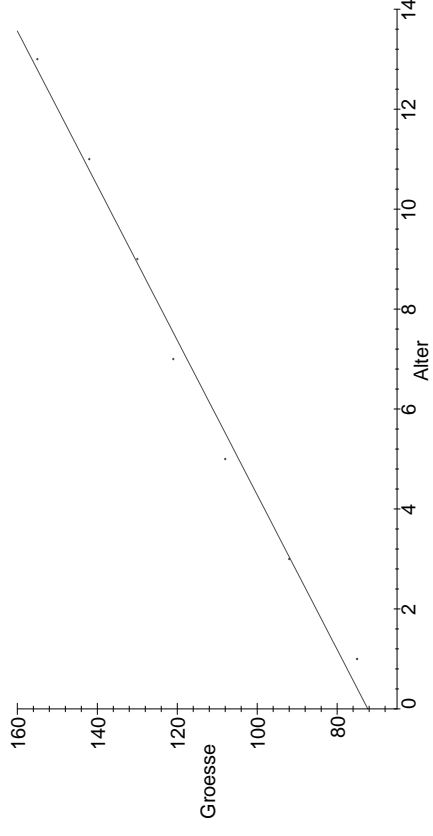


Abb. 16: Abhängigkeit der Körpergröße vom Alter

daß sich große positive und große negative Abweichungen gegenseitig aufheben können.

Besser wäre es wohl zu verlangen, daß die *maximale Abweichung* möglichst kleinen Betrag hat. Diese Forderung ist oftmals sinnvoll, beispielsweise, wenn man in der numerischen Mathematik ein Polynom sucht, das eine gegebene Funktion in einem gewissen Intervall möglichst gut approximiert. Im vorgegebenen Beispiel wissen wir allerdings, daß die Körpergröße nur in einem beschränkten Intervall einermäßen durch eine lineare Funktion beschreibbar ist, und wenn das an einem oder auch beiden Enden des Intervalls nicht mehr so gut klappt, sollten wir diesem Umstand keine allzugroße Bedeutung zumessen.

Ein dritter Ansatz könnte sein, daß wir die mittlere betragsmäßige Abweichung möglichst klein halten wollen; dies klingt vernünftig, wird aber selten praktiziert. Stattdessen verwendet man die mittlere *quadratische* Abweichung, versucht also, die Größe

$$\frac{1}{N} \sum_{i=1}^N (y_i - f(a, b, \dots; x_i))^2$$

zu minimieren. (Der Vorfaktor  $1/N$  ist dabei natürlich irrelevant.)

Für diese von CARL FRIEDRICH GAUSS eingeführte *Methode der kleinsten Quadrate* spricht sehr vieles (allein GAUSS gab im Laufe seines Lebens drei verschiedene Begründungen dafür):

- Durch Quadrieren des Fehlers werden große Abweichungen stärker „bestraft“.
- Wenn wir zur Bestimmung der Parameter ein Minimum suchen, ist eine Summe von Quadraten sehr viel angenehmer als eine Summe von Beträgen: Die Ableitung einer Summe von Quadraten ist linear, die einer Summe von Beträgen ist dort, wo sie überhaupt existiert, eine Treppenfunktion.
- Wenn wir das Tupel der theoretischen Werte und das der empirischen Werte als Vektoren im  $\mathbb{R}^N$  auffassen, minimieren wir nach dieser Methode den EUKLIDISCHEN ABSTAND dieser Vektoren.
- Das *Fehlerfortpflanzungsgesetz* besagt, daß sich Fehler quadratisch aufsummieren.
- Statistische Überlegungen nach der sogenannten *maximum likelihood Methode* zeigen, daß bei den so ermittelten Parameter die korrekten Parameter die größte Wahrscheinlichkeit haben, in der Nähe der berechneten zu liegen.



CARL FRIEDRICH GAUSS (1777–1855) leistete wesentliche Beiträge zur Zahlentheorie, zur nichteuklidischen Geometrie, zur Differentialgeometrie und Kartographie, zur Fehlerrechnung und Statistik, zur Astronomie und Geophysik usw.; vielen Lesern wird er auch durch den GAUSS-Algorithmus zur Lösung linearer Gleichungssysteme bekannt sein. Als Direktor der Göttinger Sternwarte baute er zusammen mit dem Physiker WEBER den ersten Telegraphen. Er leitete die erste Vermessung und Kartierung des Königreichs Hannover und zeitweise auch den Witwenfond der Universität Göttingen; seine hierbei gewonnene Erfahrung benutzte er für erfolgreiche Spekulationen mit Aktien.

Im vorliegenden Beispiel suchen wir eine lineare Funktion

$$u = f(a, b; t) = a \cdot t + b,$$

gesucht sind also reelle Zahlen  $a$  und  $b$ , bei denen die Funktion

$$F(a, b) = \sum_{i=1}^N (y_i - a \cdot x_i - b)^2$$

minimal wird. Dort müssen dann insbesondere die beiden partiellen Ableitungen  $F_a(a, b)$  und  $F_b(a, b)$  verschwinden. Bevor wir diese ausrechnen, empfiehlt es sich aber, die Funktion zunächst durch Ausmultiplizieren nach Termen in  $a$  und  $b$  zu sortieren:

$$\begin{aligned} F(a, b) &= \sum_{i=1}^N y_i^2 + a^2 \cdot \sum_{i=1}^N x_i^2 + N \cdot b^2 - 2 \cdot a \cdot \sum_{i=1}^N y_i \cdot x_i \\ &\quad - 2 \cdot b \cdot \sum_{i=1}^N y_i + 2 \cdot a \cdot b \cdot \sum_{i=1}^N x_i. \end{aligned}$$

Damit ist

$$F_a(a, b) = 2 \cdot a \cdot \sum_{i=1}^N x_i^2 - 2 \cdot \sum_{i=1}^N y_i \cdot x_i + 2 \cdot b \cdot \sum_{i=1}^N x_i$$

und

$$F_b(a, b) = 2 \cdot n \cdot b - 2 \cdot \sum_{i=1}^N y_i + 2 \cdot a \cdot \sum_{i=1}^N x_i.$$

Division durch zwei ergibt die beiden Gleichungen für das Verschwinden dieser partiellen Ableitungen:

$$\begin{aligned} a \cdot \sum_{i=1}^N x_i^2 + b \cdot \sum_{i=1}^N x_i &= \sum_{i=1}^N x_i \cdot y_i \\ \text{und} \quad a \cdot \sum_{i=1}^N x_i + b \cdot N &= \sum_{i=1}^N y_i. \end{aligned}$$

Damit haben wir die Bestimmung von  $a$  und  $b$  zurückgeführt auf das Lösen eines linearen Gleichungssystems und müssen nur noch für das



vorliegende Beispiel die Koeffizienten dieses Systems berechnen:

$$\sum_{i=1}^N x_i^2 = 1^2 + 3^2 + \dots + 13^2 = 455$$

$$\sum_{i=1}^N x_i = 1 + 3 + \dots + 13 = 49$$

$$\sum_{i=1}^N x_i y_i = 1 \cdot 75 + 32 + \dots + 13 \cdot 155 = 6485$$

$$\sum_{i=1}^N y_i = 75 + 92 + \dots + 155 = 823.$$

Wir müssen also das lineare Gleichungssystem

$$455a + 49b = 6485$$

$$49a + 7b = 823$$

lösen, wobei es hoffentlich niemandem große Schwierigkeiten bereitet, die einzige Lösung

$$a = \frac{181}{28} \approx 6,464 \quad \text{und} \quad b = \frac{2025}{28} \approx 72,321$$

zu finden. Unser lineares Modell stimmt also dann am besten mit der Daten überein, wenn wir von einer Geburtsgröße von etwa 72 cm und einem jährlichen Wachstum von etwa  $6\frac{1}{2}$  cm ausgehen. In Abbildung 16 ist die entsprechende Gerade zusammen mit den Datenpunkten eingezeichnet; wie zu erwarten war, wird die Wachstumsrate bei Kleinstkindern unterschätzt, bei älteren dagegen überschätzt, aber im großen und ganzen ist die Übereinstimmung recht gut.

Natürlich hätten wir das Gleichungssystem auch lösen können, bevor wir für die Koeffizienten Zahlenwerte eingesetzt haben; für die Behandlung des vorliegenden Beispiels wäre das zwar erheblich umständlicher gewesen, aber wenn man viele Beispiele behandeln muß oder möchte, hat auch eine allgemeine Formel durchaus ihre Vorteile, so daß wir sie hier

kurz herleiten wollen. Wir gehen aus vom linearen Gleichungssystem

$$a \cdot \sum_{i=1}^N x_i^2 + b \cdot \sum_{i=1}^N x_i = \sum_{i=1}^N x_i \cdot y_i$$

$$\text{und} \quad a \cdot \sum_{i=1}^N x_i + b \cdot N = \sum_{i=1}^N y_i.$$

Um  $b$  zu eliminieren, multiplizieren wir die zweite Gleichung zunächst mit  $\frac{1}{N} \sum_{i=1}^N x_i$ :

$$a \cdot \frac{1}{N} \left( \sum_{i=1}^N x_i \right)^2 + b \cdot \sum_{i=1}^N x_i = \frac{1}{N} \left( \sum_{i=1}^N x_i \right) \cdot \left( \sum_{i=1}^N y_i \right)$$

Dies subtrahieren wir dann von der ersten Gleichung:

$$a \cdot \left( \sum_{i=1}^N x_i^2 - \frac{1}{N} \left( \sum_{i=1}^N x_i \right)^2 \right) = \sum_{i=1}^N x_i \cdot y_i - \frac{1}{N} \left( \sum_{i=1}^N x_i \right) \cdot \left( \sum_{i=1}^N y_i \right),$$

woraus sofort folgt, daß

$$a = \frac{\sum_{i=1}^N x_i \cdot y_i - \frac{1}{N} \left( \sum_{i=1}^N x_i \right) \cdot \left( \sum_{i=1}^N y_i \right)}{\sum_{i=1}^N x_i^2 - \frac{1}{N} \left( \sum_{i=1}^N x_i \right)^2}$$

ist. Entsprechend läßt sich  $a$  eliminieren, wenn wir die erste Gleichung mit  $\sum x_i$  und die zweite mit  $\sum x_i^2$  multiplizieren.; wir erhalten

$$b = \frac{\left( \sum_{i=1}^N x_i \right) \cdot \left( \sum_{i=1}^N x_i \cdot y_i \right) - \left( \sum_{i=1}^N x_i^2 \right) \cdot \left( \sum_{i=1}^N y_i \right)}{\left( \sum_{i=1}^N x_i \right)^2 - N \sum_{i=1}^N x_i^2}.$$

In der Statistik bezeichnet man die Anwendung der Methode der kleinsten Quadrate auf dieses Problem als *lineare Regression*.

(Auf genau dieselbe Weise wie im folgenden lassen sich auch alle anderen Probleme, bei denen die Funktion  $f$  linear von ihren Parametern abhängt, auf das Lösen linearer Gleichungssysteme zurückführen.)

### §5: Maxima und Minima unter Nebenbedingungen

In der Realität können sich Variablen selten frei im gesamten  $\mathbb{R}^n$  bewegen; von den Gegebenheiten her sinnvoll ist meist nur eine beschränkte Teilmenge. Im Gegensatz zur Dimension eins, wo diese Teilmenge praktisch immer ein Intervall ist, gibt es aber im Mehrdimensionalen keinen Grund, warum diese Teilmenge offen oder zumindest der Abschluß einer offenen Teilmenge sein sollte: Im  $\mathbb{R}^3$  kann man sich beispielsweise auch interessieren für das Maximum oder Minimum der Ladungsdichte auf einer Kugeloberfläche oder die Temperaturverteilung auf der Innenhaut eines Reaktordruckbehälters.

Diese Maxima oder Minima sind im allgemeinen keine lokalen Maxima oder Minima der betrachteten Funktion: Wenn man die jeweilige Fläche verläßt, läßt sich der Funktionswert für einen solchen Extremwert meist noch – je nach Richtung – sowohl vergrößern als auch verkleinern. Dementsprechend können die Methoden, die wir in §3 diskutiert haben, solche Extremwerte üblicherweise nicht finden; wir brauchen weitere Werkzeuge, die in diesem Paragraphen bereitgestellt werden sollen.

Wir beginnen mit einem speziellen, aber dennoch typischen Fall:

Gegeben sei eine Funktion  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ , möglicherweise auch nur auf einer Teilmenge  $D \subset \mathbb{R}^n$  definiert, deren (lokale) Extremwerte auf einer Teilmenge gesucht werden, die durch das Verschwinden einer weiteren Funktion  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  gegeben ist. Falls wir uns für Extremwerte auf einer Kugel vom Radius  $r$  um den Nullpunkt interessieren, wäre dies etwa die Funktion

$$g: \begin{cases} \mathbb{R}^3 & \rightarrow \mathbb{R} \\ (x, y, z) & \mapsto x^2 + y^2 + z^2 - r^2. \end{cases}$$

Was unter einem lokalen Extremwert unter der Nebenbedingung  $g = 0$  zu verstehen ist, dürfte wohl jedem klar sein; der Vollständigkeit halber sei es trotzdem noch als formale Definition hingeschrieben:

**Definition:** Wir sagen, die Funktion  $f: D \rightarrow \mathbb{R}$  auf einer Teilmenge  $D \subseteq \mathbb{R}^n$  habe im Punkt  $(a_1, \dots, a_n) \in D$  ein lokales  $\begin{cases} \text{Maximum} \\ \text{Minimum} \end{cases}$  unter der Nebenbedingung  $g = 0$ , wobei  $g: D \rightarrow \mathbb{R}$  eine weitere Funktion ist, wenn  $g(a) = 0$  ist und es eine Umgebung  $U$  von  $a$  gibt, so daß für alle  $x_1, \dots, x_n \in U$  gilt:

$$\text{Ist } g(x_1, \dots, x_n) = 0, \text{ so ist } f(x_1, \dots, x_n) \begin{cases} \leq \\ \geq \end{cases} f(a_1, \dots, a_n).$$

Eine mögliche Strategie zur Lösung solcher Probleme besteht darin, die Gleichung  $g = 0$  nach einer der Variablen aufzulösen, diese dann in  $f$  einzusetzen und sodann eine gewöhnliche Extremwertaufgabe zu lösen. Diese Auflösung ist *explizit* nur in sehr einfachen Fällen möglich, aber selbst wenn wir nur wissen, daß eine solche Auflösung *existiert*, können wir doch damit argumentieren und Kriterien ableiten.

Als Einstiegsbeispiel betrachten wir eine beliebige Schulbuchaufgabe zur Minimumsbestimmung: Eine Konservendose soll bei einem vorgegebenen Volumen von  $100 \text{ cm}^3$  möglichst wenig Blech benötigen, d.h. ihre Oberfläche soll minimal sein.

Die Oberfläche eines Zylinders der Höhe  $h$  mit einer Grundfläche vom Radius  $r$  ist

$$f(r, h) = 2\pi r^2 + 2\pi r \cdot h;$$

die Nebenbedingung für das Volumen  $V = \pi r^2 h$  besagt, daß

$$g(r, h) = \pi r^2 h - 100 = 0$$

sein soll.

Hier läßt sich natürlich die Nebenbedingung sofort nach  $h$  auflösen:

$$h = \frac{100}{\pi r^2},$$

und wir müssen nur noch die Funktion

$$F(r) = f\left(r, \frac{100}{\pi r^2}\right) = 2\pi r^2 + \frac{200}{r}$$

minimieren. Für diese ist

$$F'(r) = 4\pi r - \frac{200}{r^2},$$

und dies verschwindet genau dann, wenn

$$4\pi r^3 = 200 \quad \text{oder} \quad r = \sqrt[3]{\frac{50}{\pi}}$$

ist.

In diesem einfachen Fall kann man solche Aufgaben also zurückführen auf gewöhnliche Extremwertaufgaben, indem man die Nebenbedingung nach einer der Variablen auflöst und diese dann in  $f$  einsetzt; in anderen Fällen kann man gelegentlich die Nebenbedingung durch geeignete Parameterwahl oder Wahl eines angepaßten Koordinatensystems berücksichtigen. Im allgemeinen wird aber beides nicht möglich sein, so daß wir andere Methoden brauchen.

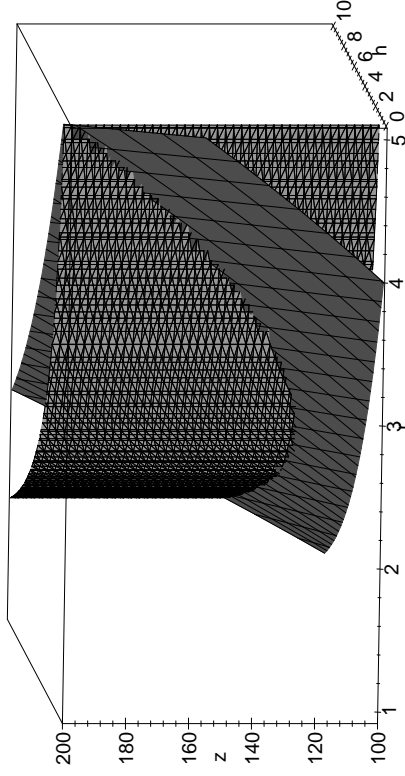


Abb. 17: Oberfläche einer Konservendose mit festem Volumen

Unser bisherige Theorie für lokale Extrema ist in dieser Situation nicht anwendbar, denn die lokalen Extrema von  $f$  werden nur in den seltensten Fällen die Nebenbedingung  $g = 0$  erfüllen; im obigen Beispiel zeigt Abbildung 17 die Nebenbedingung als eng schraffierte Fläche dargestellt und der Graph von  $f$  als weiter schraffierte; wie man sieht, läßt sich der Wert von  $f$  problemlos verkleinern, wenn man nur die Fläche  $g = 0$  verläßt, und in der Tat ist auch ohne jede Mathematik sofort klar, daß man

mit weniger Blech auskommt, wenn man die Konservendose einfach schmälert oder kürzer macht.

Die Grundidee für ein alternatives Verfahren wird klar bei der Betrachtung der Niveaulinien in Abbildung 18: Die Niveaulinie für  $g = 0$  ist gestrichelt eingezeichnet, verschiedene Niveaulinien von  $f$  als durchgezogene Kurven.

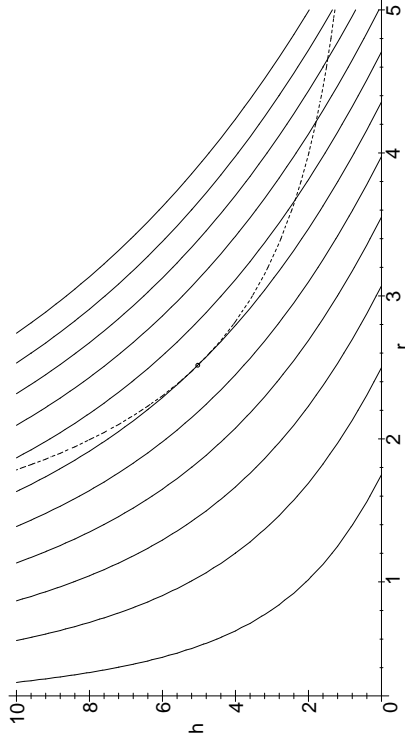


Abb. 18: Niveaulinien für Oberfläche und Volumen

Wie man sieht, schneiden einige dieser Niveaulinien die gestrichelte Kurve überhaupt nicht: Wenn man zu wenig Blech hat, kann man keine Dose mit  $100 \text{ cm}^3$  Inhalt zusammenlöten. Wenn es dagegen genug Blech gibt, gibt es gleich zwei Schnittpunkte: Die Dose kann entweder eher höher oder eher breiter gemacht werden. In einem solchen Fall kann man die Niveaulinie durch eine zu einem etwas niedrigeren Niveau ersetzen, die im allgemeinen auch wieder Schnittpunkte haben wird, so daß das Niveau noch nicht minimal sein kann. Erst wenn man im Minimum ist, fallen die beiden Schnittpunkte zusammen; wenn man nun das Niveau noch weiter erniedrigt, gibt es keine Schnittpunkte mehr.

Da somit im Minimum zwei Schnittpunkte zusammenfallen, berühren sich dort die Niveaulinien von  $f$  und von  $g$ , d.h. sie haben eine gemeinsame Tangente. Da der Gradient, wie wir wissen, senkrecht auf der

Tangenten der Niveaulinien steht (die Richtungsableitung entlang einer Niveaulinie ist schließlich null), sind somit die Gradienten von  $f$  und  $g$  im Minimum zueinander parallel, d.h. der eine ist ein Vielfaches des anderen.

Dies gilt nicht nur im vorliegenden Beispiel, sondern allgemein:

**Satz:**  $D \subseteq \mathbb{R}^n$  sei eine offene Menge und  $f, g \in C^1(D, \mathbb{R})$  seien stetig differenzierbare Funktionen auf  $D$ . Falls  $f$  im Punkt  $(a_1, \dots, a_n) \in D$  ein Extremum hat unter der Nebenbedingung  $g(x_1, \dots, x_n) = 0$ , so sind die Vektoren

$$\text{grad } f(a_1, \dots, a_n) \quad \text{und} \quad \text{grad } g(a_1, \dots, a_n)$$

proportional.

*Beweis:* Die Grundidee ist einfach: Auch wenn wir die Nebenbedingung nicht *explizit* nach einer der Variablen auflösen können, läßt sich in vielen Fällen relativ leicht nachweisen, daß zumindest lokal eine Auflösung zumindest *existiert*. Diese Auflösung kennen wir zwar nicht, aber wir können mit ihr argumentieren und, zumindest formal, auch rechnen.

Falls  $\text{grad } g(a_1, \dots, a_n)$  der Nullvektor ist, gibt es nichts mehr zu beweisen, denn der Nullvektor ist proportional zu jedem anderen Vektor.

Wir können daher annehmen, daß mindestens eine Komponente nicht verschwindet, und durch Ummummerierung der Koordinaten können wir o.B.d.A. annehmen, daß dies die  $n$ -te Komponente ist, d.h.  $g_{x_n}(a) \neq 0$ .

Dann gibt es nach dem sogenannten Satz über implizite Funktionen (für dessen Beweis wir leider keine Zeit haben) eine Umgebung  $U$  von  $(a_1, \dots, a_{n-1})$  und eine Funktion  $h: U \rightarrow \mathbb{R}$  mit  $h(a_1, \dots, a_{n-1}) = a_n$ , so daß

$$g(x_1, \dots, x_{n-1}, h(x_1, \dots, x_{n-1})) = 0 \quad \text{für alle } (x_1, \dots, x_{n-1}) \in U.$$

Nachdem  $f$  in  $a$  ein lokales Extremum unter der Nebenbedingung  $g = 0$  hat, nimmt die Funktion

$$F(x_1, \dots, x_{n-1}) \stackrel{\text{def}}{=} f(x_1, \dots, x_{n-1}, h(x_1, \dots, x_{n-1}))$$

in  $(a_1, \dots, a_{n-1})$  ein lokales Extremum im üblichen Sinne an, d.h. der Gradient von  $F$  verschwindet dort.

Die Kettenregel, die hier im Mehrdimensionalen fast genauso aussieht, wie man es vom Eindimensionalen her gewohnt ist, sagt für  $i = 1, \dots, n-1$ , daß  $F_{x_i}(a_1, \dots, a_{n-1})$

$$= f_{x_i}(a_1, \dots, a_n) + f_{x_n}(a_1, \dots, a_n) \cdot h_{x_i}(a_1, \dots, a_{n-1})$$

ist, und nach dem Satz über implizite Funktionen gibt uns die partielle Ableitung von  $h$  nach  $x_i$  als

$$h_{x_i} = -\frac{g_{x_i}}{g_{x_n}},$$

d.h.

$$F_{x_i}(a_1, \dots, a_{n-1}) = f_{x_i}(a_1, \dots, a_n) - f_{x_n}(a_1, \dots, a_n) \frac{g_{x_i}(a_1, \dots, a_n)}{g_{x_n}(a_1, \dots, a_n)}.$$

Da die linke Seite verschwindet, gilt dasselbe auch für die rechte. Die rechte Seite ist im Gegensatz zur linken auch für  $i = n$  definiert und verschwindet aus trivialen Gründen; also verschwindet die rechte Seite für alle  $i$ , d.h.

$$\text{grad } f(a_1, \dots, a_n) - \frac{f_{x_n}(a_1, \dots, a_n)}{g_{x_n}(a_1, \dots, a_n)} \text{grad } g((a_1, \dots, a_n)) = \vec{0}.$$

Also sind die beiden Gradienten in der Tat proportional. ■

Falls der Gradient von  $g$  im Punkt  $a$  nicht verschwindet, gibt es somit eine Zahl  $\lambda \in \mathbb{R}$ , so daß

$$\text{grad } f(a) + \lambda \text{ grad } g(a) = \vec{0}$$

ist, nämlich

$$\lambda = -\frac{f_{x_n}(a)}{g_{x_n}(a)}.$$

Diese Zahl heißt als LAGRANGEScher Multiplikator; mit ihrer inhaltlichen Interpretation werden wir uns in Kürze beschäftigen.

Zur praktischen Bestimmung von Extremwerten unter Nebenbedingungen geht man wie folgt vor: Über die Punkte, in denen der Gradient

von  $g$  verschwindet, macht obiger Satz keine verwertbare Aussage; diese Punkte müssen also vorab berechnet und untersucht werden.

Danach müssen die Punkte gefunden werden, in denen es ein  $\lambda \in \mathbb{R}$  gibt, so daß

$$\begin{aligned} f_{x_1}(x_1, \dots, x_n) + \lambda g_{x_1}(x_1, \dots, x_n) &= 0 \\ &\vdots \\ f_{x_n}(x_1, \dots, x_n) + \lambda g_{x_n}(x_1, \dots, x_n) &= 0 \\ g(x_1, \dots, x_n) &= 0 \end{aligned}$$

ist. Dies ist ein System von  $n + 1$  Gleichungen für die  $n + 1$  Unbekannten, allerdings ist dieses Gleichungssystem nur selten linear und damit oft nicht mit bekannten Methoden lösbar. Manchmal kann man das Gleichungssystem durch geeignete Umformungen und Fallunterscheidungen vollständig lösen, in anderen Fällen helfen numerische Verfahren wie etwa das Verfahren von NEWTON-RAPHSON.

Falls alle Gleichungen Polynomgleichungen sind (oder durch Einführung geeigneter zusätzlicher Variablen auf Polynomgleichungen zurückgeführt werden können), kann man im Falle einer endlichen Lösungsmenge diese auch exakt bestimmen: Genau wie der GAUSS-Algorithmus zur Lösung eines linearen Gleichungssystems dieses auf eine Treppengestalt bringt, aus der man die Lösungen einfach ermitteln kann, gibt es in der Computeralgebra einen Algorithmus, der dasselbe für beliebige Systeme von Polynomgleichungen versucht; die Gleichungen, die dieser Algorithmus liefert, bezeichnet man als GRÖBNER-Basis oder Standardbasis. Zum Verständnis dieses Algorithmus, den man als eine Art Synthese aus EUKLIDISCHEN Algorithmus und GAUSS-Algorithmus ansehen kann, sind Kenntnisse der kommutativen Algebra erforderlich, für die die Zeit in dieser Vorlesung nicht ausreicht; bei einigen Implementierungen werden zusätzlich auch noch Algorithmen aus der Informatik eingesetzt, die typischerweise nicht in Grundvorlesungen behandelt werden. Deshalb sei hier nur darauf hingewiesen, daß die gängigen universellen Computeralgebrasysteme wie Maple, Mathematica, MuePad allesamt entsprechende Routinen enthalten, mit denen man auch dann experimentieren kann, wenn man die dahinterstehende Theorie nicht versteht.

Als Beispiel, wie gelegentlich auch ein nichtlineares Gleichungssystem elementar gelöst werden kann, betrachten wir eine Anwendung aus den Wirtschaftswissenschaften: Die Gesamtproduktion eines Unternehmens oder eines Staates in Abhängigkeit von  $n$  eingesetzten Ressourcen

$x_1, \dots, x_n$  wird oft modelliert durch eine sogenannte COBB-DOUGLAS-Funktion der Form

$$P(x_1, \dots, x_n) = \alpha x_1^{\epsilon_1} \dots x_n^{\epsilon_n},$$

benannt nach den beiden Wissenschaftlern, die dieses Modell 1928 für die amerikanische Gesamtproduktion in Abhängigkeit von Kapital und Arbeit in den Jahren 1899 bis 1922 entwickelten. (Sie fanden  $P \approx 1,01A^{3/4}K^{1/4}$  mit  $A =$  Anzahl der Beschäftigten und  $K =$  Kapitaleinsatz.)

Betrachten wir stattdessen die Produktion eines Wirtschaftsguts aus zwei Ressourcen  $x, y$  gemäß der Funktion

$$f(x, y) = P(x, y) = x^{1/2}y^{1/4}.$$

Falls wir der Einfachheit halber annehmen, daß die Kosten pro Einheit für  $x$  und  $y$  gleich sind und die Gesamtkosten höchstens gleich zwölf sein dürfen, müssen wir  $f$  maximieren unter der Nebenbedingung

$$x + y \leq 12.$$

Nun ist aber  $f$  eine monoton wachsende Funktion sowohl von  $x$  als auch von  $y$ , d.h. die maximale Produktion wird sicherlich erreicht in einem Punkt, für den  $x + y = 12$  ist, denn für jeden anderen Punkt  $(x, y)$  mit  $x + y < 12$  ist  $f(x, y) < f(x, 12 - x)$ . Daher können wir die Nebenbedingung in der gewohnten Form

$$g(x, y) = x + y - 12 = 0$$

schreiben. Diese Nebenbedingung sowie die zu maximierende Funktion sind in Abbildung 19 dargestellt.

Ableitung beider Funktionen zeigt, daß

$$\text{grad } g = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad \text{grad } f = \begin{pmatrix} y^{1/4}/2x^{1/2} \\ x^{1/2}/4y^{3/4} \end{pmatrix}$$

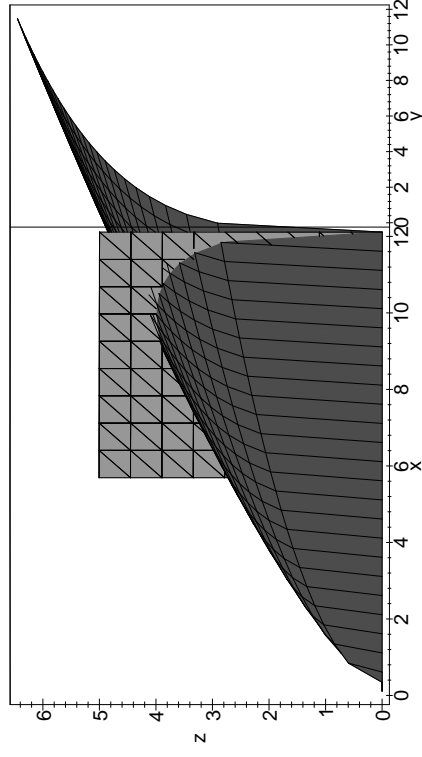


Abb. 19: Maximierung einer Produktionsfunktion bei festem Kapitaleinsatz

ist; das zu lösende Gleichungssystem wird also zu

$$\frac{y^{1/4}}{2x^{1/2}} + \lambda = 0$$

$$\frac{x^{1/2}}{4y^{3/4}} + \lambda = 0$$

$$x + y - 12 = 0$$

(Die Nenner brauchen uns nicht zu stören, denn da  $f(0, y) = f(x, 0) = 0$  ist, kommen Lösungen mit  $x = 0$  oder  $y = 0$  für das Maximum ohnehin nicht in Frage; wir können sie also getrost ausschließen.)

Als Ansatz zu einer möglichen Lösung können wir ausnutzen, daß  $\lambda$  in den beiden ersten Gleichungen isoliert steht; wenn wir danach auflösen und gleichsetzen, erhalten wir die Gleichung

$$\frac{y^{1/4}}{2x^{1/2}} = \frac{x^{1/2}}{4y^{3/4}}$$

Multiplikation mit dem Hauptnenner macht daraus

$$4y^{1/4}y^{3/4} = 2x^{1/2}x^{1/2} \quad \text{oder} \quad 2y = x.$$

Einsetzen in die dritte Gleichung ergibt  $3y = 12$ , also ist

$$y = 4 \quad \text{und} \quad x = 8,$$

und der Maximalwert von  $f$  ist

$$f(8, 4) = 8^{1/2} \cdot 4^{1/4} = 2\sqrt{2} \cdot \sqrt{2} = 4.$$

Auch den LAGRANGESchen Multiplikator  $\lambda$  können wir noch ausrechnen:

$$\lambda = \frac{y^{1/4}}{2x^{1/2}} = \frac{4^{1/4}}{2 \cdot 8^{1/2}} = \frac{\sqrt{2}}{2 \cdot 2\sqrt{2}} = \frac{1}{4}.$$

Die Berechnung von  $\lambda$  war für die Bestimmung des Optimums eigentlich überflüssig;  $\lambda$  ist nur eine Hilfsgröße zur Berechnung des Extremums. Wir wollen uns als nächstes überlegen, daß wir  $\lambda$  auch inhaltlich interpretieren können: Dazu betrachten wir eine Nebenbedingung

$$g(x_1, \dots, x_n) = c$$

mit *variabler* rechter Seite  $c$  und ein Extremum der Funktion

$$f(x_1, \dots, x_n).$$

Dieses Extremum wird natürlich von  $c$  abhängen; wir schreiben es in der Form

$$(x_1(c), \dots, x_n(c))$$

und nehmen an, daß die Funktionen  $x_i(c)$  stetig differenzierbar seien. (Ein interessierter Leser kann sich anhand des Satzes über implizite Funktionen überlegen, welche Bedingungen  $f$  und  $g$  erfüllen müssen, damit dies garantiert ist.) Der Optimalwert von  $f$  in Abhängigkeit von  $c$  ist dann

$$F(c) \stackrel{\text{def}}{=} f(x_1(c), \dots, x_n(c)).$$

Nach der Kettenregel ist

$$F'(c) = \sum_{i=1}^n f_{x_i} x'_i(c).$$

Genauso können wir

$$G(c) \stackrel{\text{def}}{=} g(x_1(c), \dots, x_n(c))$$

betrachten und erhalten

$$G'(c) = \sum_{i=1}^n g_{x_i} x'_i(c).$$

Da  $(x_1(c), \dots, x_n(c))$  ein Optimum ist, sind dort die Gradienten von  $f$  und  $g - c$  proportional mit Proportionalitätsfaktor  $\lambda$ . Da wir bei der Gradientenbildung nur nach den  $x_i$  ableiten, von denen die rechte Seite  $c$  nicht abhängt, ist der Gradient von  $g - c$  gleich dem von  $g$  selbst, d.h.

$$f_{x_i} = \lambda g_{x_i} \quad \text{für alle } i.$$

Somit ist  $F'(c) = \lambda G'(c)$ . Da der Punkt  $(x_1(c), \dots, x_n(c))$  die Nebenbedingung mit rechter Seite  $c$  erfüllt, ist aber  $G(c) = c$  und damit  $G'(c) \equiv 1$ . Also ist  $\lambda = F'(c)$  die Wachstumsrate für das Optimum bei Änderung der rechten Seite der Nebenbedingung.

Im obigen Beispiel steigt also die Maximalmenge  $f(x, y)$ , die mit Kapitaleinsatz 12 produziert werden kann, für kleines  $h$  ungefähr um  $h/4$ , wenn wir den Kapitaleinsatz auf  $12 + h$  erhöhen. Die Erhöhung des Kapitaleinsatzes lohnt sich, wenn für das fertige Produkt ein Preis pro Einheit erzielt werden kann, der größer ist als vier.

Als letztes wollen wir uns noch überlegen, was passiert, wenn wir nicht nur eine, sondern mehrere Nebenbedingungen erfüllen müssen. Es geht also wieder darum, eine Funktion  $f(x_1, \dots, x_n)$  zu optimieren, jetzt aber unter den Nebenbedingungen

$$g_1(x_1, \dots, x_n) \geq 0, \quad \dots \quad g_r(x_1, \dots, x_n) \geq 0.$$

(Es genügt, Bedingungen mit  $\geq$  zu betrachten, denn durch Multiplikation mit minus Eins kann man jede Ungleichung mit  $\leq$  in eine mit  $\geq$  überführen. Auch Gleichungen  $g_i = 0$  kann man zumindest formal durch die beiden Ungleichungen  $g_i \geq 0$  und  $-g_i \geq 0$  ausdrücken.)

Die wichtigsten Beispiele solcher Optimierungsaufgaben sind die Fälle mit linearen Funktionen  $f$  und  $g_i$ ; hier redet man von *linearen Programmen*. (Das Wort *Programme* in diesem Zusammenhang hat natürlich nichts mit Computerprogrammen zu tun.) Das wichtigste Verfahren zur Lösung solcher Aufgaben, der Simplex-Algorithmus, wurde gestern behandelt; heute geht es um die *nichtlineare Programmierung*.

Im linearen Fall definieren die Nebenbedingungen ein (endliches oder unendliches) Polyeder im  $\mathbb{R}^n$ , und eine lineare Funktion, so sie ein

endliches Maximum oder Minimum hat, nimmt dieses auf dem Rand dieses Polyeders an, sogar in einer Ecke. Man muß daher „nur“ die Ecken dieses Polyeders untersuchen.

Bei nichtlinearen Funktionen ist die Situation komplizierter, denn nun kann es auch im Innern Extrema geben: Die Funktion

$$f(x, y) = e^{-x^2 - y^2} \quad \text{mit der Nebenbedingung} \quad x^2 + y^2 \leq 1$$

etwa nimmt ihr Maximum im Punkt  $(0, 0)$  an; auf dem Rand des Einheitskreises liegen nur die Minima. Im allgemeinen Fall eines nichtlinearen Programms kann ein Optimum also entweder ganz im Innern liegen oder aber eine beliebige Teilmenge der Nebenbedingungen exakt erfüllen.

Falls wir es mit inneren Punkte zu tun haben, sind diese lokale Maxima oder Minima ohne Nebenbedingungen, und wir haben uns bereits in §3 überlegt, wie man diese bestimmt: In jedem solchen Punkt verschwindet der Gradient der zu optimierenden Funktion.

Im Falle einer einzigen *Gleichung* als Nebenbedingung sind der Gradient von  $f$  und der Gradient der Nebenbedingung proportional; da der Nullvektor zu jedem anderen Vektor proportional ist, schließt dies auch den Fall der Optima bei inneren Punkten mit ein.

Bei mehreren Nebenbedingungen müssen wir die Proportionalität durch eine allgemeinere Eigenschaft ersetzen, die *lineare Abhängigkeit*:

**Definition:**  $r + 1$  Vektoren  $\vec{v}_0, \dots, \vec{v}_r$  heißen *linear abhängig*, wenn es reelle Zahlen  $\lambda_0, \dots, \lambda_r \in \mathbb{R}$  gibt, die nicht allesamt verschwinden, so daß

$$\lambda_0 \vec{v}_0 + \dots + \lambda_r \vec{v}_r = \vec{0}$$

ist.

Mit dieser Definition gilt dann

Die naheliegende Verallgemeinerung auf den Fall mehrerer Nebenbedingungen ist daher der

**Satz:** Die Funktion  $f: D \rightarrow \mathbb{R}$  auf  $D \subseteq \mathbb{R}^n$  habe im Punkt  $\mathbf{a} = (a_1, \dots, a_n) \in D$

ein Extremum unter den Nebenbedingungen

$$g_1(\mathbf{a}) \geq 0, \quad g_2(\mathbf{a}) \geq 0, \quad \dots, \quad g_r(\mathbf{a}).$$

Dann sind die  $r + 1$  Vektoren

$$\text{grad } f(\mathbf{a}), \quad \text{grad } g_1(\mathbf{a}), \quad \text{grad } g_2(\mathbf{a}), \quad \dots, \quad \text{grad } g_r(\mathbf{a})$$

linear abhängig.

Der *Beweis* erfordert keine wesentlich neuen Ideen gegenüber dem Fall einer einzigen Nebenbedingung und sei daher nur kurz skizziert: Falls die Gradienten der  $g_i$  im Punkt  $\mathbf{a}$  bereits untereinander linear abhängig sind, gibt es nichts mehr zu beweisen; nehmen wir also an, sie seien linear unabhängig. Dann gibt es (mindestens)  $r$  verschiedene Variablen  $x_{j_1}$  bis  $x_{j_r}$ , so daß

$$(g_i)_{x_{j_i}}(\mathbf{a}) \neq 0$$

ist. Also kann nach dem Satz über implizite Funktionen jede Nebenbedingung zur Elimination einer anderen Variablen benutzt werden, und im wesentlichen dieselbe Rechnung wie im Fall einer Nebenbedingung zeigt die Behauptung. ■

Die lineare Abhängigkeit der Vektoren

$$\text{grad } f(\mathbf{a}), \quad \text{grad } g_1(\mathbf{a}), \quad \text{grad } g_2(\mathbf{a}), \quad \dots, \quad \text{grad } g_r(\mathbf{a})$$

bezeichnet man als KUHN-TUCKER-Bedingung; sie ist eine offensichtliche Verallgemeinerung der Bedingung von LAGRANGE, ist allerdings deutlich jünger: Sie erschien 1951 in einer gemeinsamen Arbeit von H.W. KUHN und A.W. TUCKER, vier Jahre, nachdem G. DANTZIG den Simplex-Algorithmus entwickelt hatte, und fast zweihundert Jahre, nachdem LAGRANGE seine Multiplikatoren zur Bestimmung von Extrema unter einer Nebenbedingung eingeführt hatte.

Die KUHN-TUCKER-Bedingung ist nicht das Ende, sondern der Ausgangspunkt für die nichtlineare Programmierung: Die vielfältigen linearen Abhängigkeiten, die zwischen den Gradienten der verschiedenen Nebenbedingungen möglich sind, führen zu so vielen Fallunterscheidungen, daß die Lösung für große Werte von  $r$  sehr schnell sehr aufwendig wird.

In der Praxis versucht man daher, sich soweit wie möglich auf spezielle Funktionsklassen zu beschränken, bei denen die Situation einfacher ist. Durch Konvexitätsbedingungen kann man etwa erreichen, daß Optima eindeutig festgelegt sind, und wenn man sich auf quadratische Funktionen beschränkt, führen die KUHN-TUCKER-Bedingungen auf etwas, womit man mit einem leicht modifizierten Simplex-Verfahren arbeiten kann.

Da die Zeit aber schon fast zu Ende ist, können wir uns mit diesen Verfahren, durch die die nichtlineare Optimierung erst richtig interessant und anwendbar wird, leider nicht mehr beschäftigen. Stattdessen sollen die letzten Minuten der Frage gewidmet sein, wie man sich bei komplizierten Problemen Näherungslösungen verschafft.

## §6: Numerische Verfahren

Wie wir gesehen haben, führt die Methode der LAGRANGESchen Multiplikatoren im allgemeinen auf nichtlineare Gleichungssysteme, die nur in einfachen Fällen explizit lösbar sind. In allen anderen Fällen muß man mit numerischen Methoden arbeiten, und da bietet sich an, das Problem von vornherein ohne den Umweg über LAGRANGESche Multiplikatoren Extrema numerisch zu bearbeiten.

### a) Die Gradientenmethode

Bei der Diskussion der Differenzierbarkeit in §2 haben wir gesehen, daß für eine differenzierbare Funktion  $f$  von zwei Variablen gilt

$$f(x+h, y+k) = f(x, y) + f_x(x, y) \cdot h + f_y(x, y) \cdot k + o(\max(|h|, |k|));$$

entsprechende Formeln lassen sich auch für Funktionen von  $n$  Variablen hinschreiben.

Wenn wir ein Maximum (oder Minimum) von  $f$  ansteuern wollen, liegt es daher nahe,  $h$  und  $k$  so zu wählen, daß sich der Funktionswert möglichst stark vergrößert (oder verkleinert).

Der für kleine Werte von  $h$  und  $k$  dominante Teil

$$f_x(x, y) \cdot h + f_y(x, y) \cdot k$$



der Veränderung des Funktionswerts läßt sich auch als Skalarprodukt

$$\begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix} \cdot \begin{pmatrix} h \\ k \end{pmatrix}$$

zweier Vektoren schreiben, von denen der erste gerade der Gradient von  $f$  im betrachteten Punkt ist.

Das Skalarprodukt kann bekanntlich auch berechnet werden als Produkt der Längen der beiden Vektoren mit dem Cosinus des eingeschlossenen Winkels und ist somit, bei festgehaltener Länge des Vektors  $\begin{pmatrix} h \\ k \end{pmatrix}$  genau dann maximal, wenn dieser Cosinus eins ist, der Winkel also  $0^\circ$ , und genau dann minimal, wenn der Cosinus minus eins ist, der Winkel also  $180^\circ$ . Wenn wir also ein Maximum anstreben, sollten wir immer in Richtung des Gradienten gehen; für ein Minimum immer in Gegenrichtung.

In beliebiger Dimension ist nicht mehr so ganz klar, was ein Winkel ist (obwohl zwei Vektoren auch dort immer eine Ebene aufspannen, in der es eigentlich klar sein sollte); aber auch dort ist klar, daß das Skalarprodukt zweier zueinander proportionaler Vektoren  $\vec{v}$  und  $\vec{w}$  bei gleicher Richtung von  $\vec{v}$  und  $\vec{w}$  gleich dem Produkt  $\|\vec{v}\| \cdot \|\vec{w}\|$  der Längen ist, und bei entgegengesetzter Richtung gleich  $-\|\vec{v}\| \cdot \|\vec{w}\|$ . Das dies bei festgehaltenen Längen der maximal mögliche Wert ist, zeigt die

**Cauchy-Schwarzsche Ungleichung:** Für zwei Vektoren

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad \text{und} \quad \vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

ist  $|\vec{v} \cdot \vec{w}| \leq \|\vec{v}\| \cdot \|\vec{w}\|$ .

Der Beweis beruht auf folgendem Trick: Da Längen von Vektoren nie negativ sind, ist für beliebige reelle Zahlen  $\lambda, \mu \in \mathbb{R}$

$$\begin{aligned} (\lambda \vec{v} + \mu \vec{w}) \cdot (\lambda \vec{v} + \mu \vec{w}) &= \lambda^2 \vec{v} \cdot \vec{v} + \lambda \mu \vec{v} \cdot \vec{w} + \mu \lambda \vec{w} \cdot \vec{v} + \mu^2 \vec{w} \cdot \vec{w} \\ &= \lambda^2 \vec{v} \cdot \vec{v} + 2\lambda \mu \vec{v} \cdot \vec{w} + \mu^2 \vec{w} \cdot \vec{w} \end{aligned}$$

stets größer oder gleich null. Speziell für  $\lambda = (\vec{w} \cdot \vec{w})$  und  $\mu = -(\vec{v} \cdot \vec{w})$  erhalten wir

$$(\vec{w} \cdot \vec{w})^2 (\vec{v} \cdot \vec{v}) - 2(\vec{w} \cdot \vec{w})(\vec{v} \cdot \vec{w}) + (\vec{v} \cdot \vec{w})(\vec{v} \cdot \vec{w})(\vec{w} \cdot \vec{w}) \geq 0,$$

also

$$(\vec{w} \cdot \vec{w})(\vec{w} \cdot \vec{w})(\vec{v} \cdot \vec{v}) - |\vec{v} \cdot \vec{w}|^2 \geq 0.$$

Ist hier  $\vec{w} \cdot \vec{w} \neq 0$ , können wir durch diese Zahl dividieren, und die Behauptung ist bewiesen. Andernfalls ist  $\vec{w}$  der Nullvektor, und die Behauptung stimmt, da dann beide Seiten verschwinden. ■



Baron AUGUSTIN LOUIS CAUCHY (1789–1857) stellte als erster durch die exakte Definition von Begriffen wie *Konvergenz* und *Stetigkeit* die Analysis auf ein solches Fundament. In insgesamt 789 Arbeiten beschäftigte er sich u.a. auch mit komplexer Analysis, Variationsrechnung, Differentialgleichungen, FOURIER-Analyse, Permutationsgruppen, der Diagonalisierung von Matrizen und der theoretischen Mechanik. Als überzeugter Royalist hatte er häufig Schwierigkeiten mit den damaligen Regierungen; er lebte daher mehrere Jahre im Exil in Turin und später in Prag, wo er (mit sehr mäßigen Erfolg) den französischen Thronfolger unterrichtete.



Der deutsche Mathematiker KARL HERMAN AMANDUS SCHWARZ (1843–1921) beschäftigte sich hauptsächlich mit konformen Abbildungen und mit sogenannten Minimalflächen, d.h. Flächen mit vorgegebenen Eigenschaften, deren Flächeninhalt minimal ist. Im Rahmen einer entsprechenden Arbeit für die WEIERSTRASS-Festschrift von 1885 (im Falle eines durch Doppelintegrale definierten Skalarprodukts) bewies er die obige Ungleichung; CAUCHY hatte sie bereits in seinem Analysislehrbuch von 1821 für endlichdimensionale Vektoren bewiesen. SCHWARZ lehrte nacheinander in Halle, Zürich, Göttingen und Berlin.

Damit bietet sich für Funktionen zweier Veränderlicher (die Verallgemeinerung auf  $n$  Veränderliche ist offensichtlich und soll uns daher nicht weiter beschäftigen) folgende Strategie an: Wir wählen irgendeinen Ausgangspunkt  $(x_0, y_0)$  und berechnen dort den Gradienten  $\nabla f(x_0, y_0)$ . Weiter geben uns eine Länge  $\ell_0$  für den Vektor  $\begin{pmatrix} h \\ k \end{pmatrix}$  vor, die von der Länge des Gradienten abhängen kann oder auch nicht. Dann setzen wir bei der Suche nach einem Maximum

$$\begin{pmatrix} h_0 \\ k_0 \end{pmatrix} = \frac{\ell_0}{\|\nabla f(x_0, y_0)\|} \nabla f(x_0, y_0);$$

bei der Suche nach Minima nehmen wir das Negative davon.

Als nächstes betrachten wir den Punkt

$$(x_1, y_1) \stackrel{\text{def}}{=} (x_0 + h_0, y_0 + k_0),$$

berechnen dort den Gradienten  $\nabla f(x_0, y_0)$ , setzen mit geeignetem  $\ell_1$

$$\begin{pmatrix} h_1 \\ k_1 \end{pmatrix} = \pm \frac{\ell_1}{\|\nabla f(x_1, y_1)\|} \nabla f(x_1, y_1)$$

(+ für Maxima, – für Minima) zur Definition des nächsten Punktes

$$(x_2, y_2) \stackrel{\text{def}}{=} (x_1 + h_1, y_1 + k_1),$$

und so weiter. In jedem Schritt erhöhen (oder erniedrigen) wir den Funktionswert soweit wie es mit der vorgegebenen Länge  $\ell_i$  nur möglich ist in der Hoffnung, so irgendwann auf ein Maximum (oder Minimum) zu stoßen. Dieses können wir erreichen, wenn wir am Rand des Definitionsbereichs von  $f$  angelangt sind, oder aber wenn wir in einem Punkt sind, in dem der Gradient verschwindet: Von dort aus geht es mit diesem Verfahren nicht mehr weiter.

Da wir mit einem numerischen Verfahren nur ein verschwindend geringe Chance haben, exakt in einem Extremum zu enden, zeigt sich hier auch die Notwendigkeit einer intelligenten Wahl der Schrittweiten  $\ell_i$ : Wenn diese zu groß sind, kann es passieren, daß wir endlos um ein Extremum herumoszillieren.

Theoretisch ist auch möglich, daß wir in einem Sattelpunkt landen, aber wenn man sich überlegt, wie die Gradienten in der Umgebung eines Sattelpunktes aussehen, wird schnell klar, daß dies nur sehr selten passiert.

Abbildung 20 zeigt ein einfaches Beispiel für einen mit der Gradientenmethode zurückgelegten Weg; hier wurde in jedem Schritt

$$\begin{pmatrix} h_i \\ k_i \end{pmatrix} = 0,1 \cdot \nabla f(x_i, y_i)$$

gesetzt. Der Weg geht offensichtlich recht zielstrebig auf das Maximum zu.

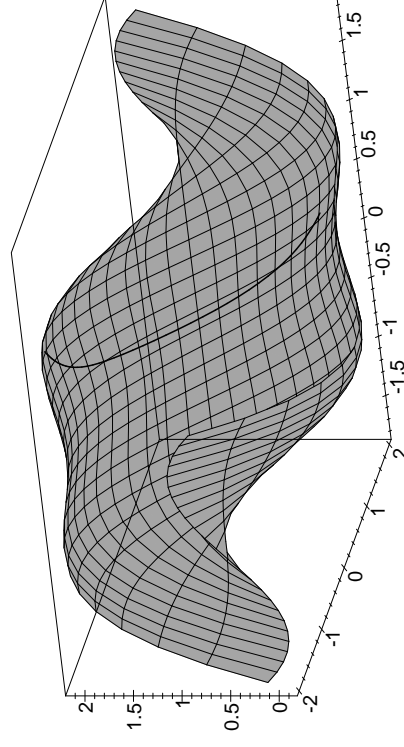


Abb. 20: Eine Anwendung der Gradientenmethode

Abbildung 21 zeigt dasselbe Bild in einen etwas größeren Zusammenhang; hier sehen wir, daß unser Streben nach kurzfristigen Gewinnen langfristig wohl doch nicht so erfolgreich war: Wenn wir vom Startpunkt aus nach rechts in die kleine Mulde abgestiegen wären, hätten wir auf dem gegenüberliegenden Hang deutlich größere Funktionswerte erreicht als im lokalen Maximum, in dem wir schließlich gelandet sind.

Dies ist ein grundsätzliches Problem von Gradientenverfahren: Falls man sie in der Nähe des (absoluten) Optimums starten läßt, führen sie schnell und zuverlässig ans Ziel, ansonsten aber ist die Gefahr sehr groß, daß man in einem nur lokalen Optimum steckenbleibt.

Um von dort wieder weiterzukommen, gibt es verschiedene Strategien. Eine anschaulich recht klare ist die sogenannte „Tunnelung“. Der Name entstand aus der Betrachtung von Minimierungsproblemen; nehmen wir also an, wir wollen das Minimum der Funktion  $f(x, y)$  in einem gewissen Bereich finden und ein Gradientenverfahren hat uns in einen Punkt  $(x_M, y_M)$  geführt, von dem aus es nicht mehr weiterkommt. Um zu sehen, ob  $z_M = f(x_M, y_M)$  wirklich der kleinste Wert ist, den  $f$  im betrachteten Bereich annehmen kann, versuchen wir, eine weitere Lösung der Gleichung

$$f(x, y) = z_M$$

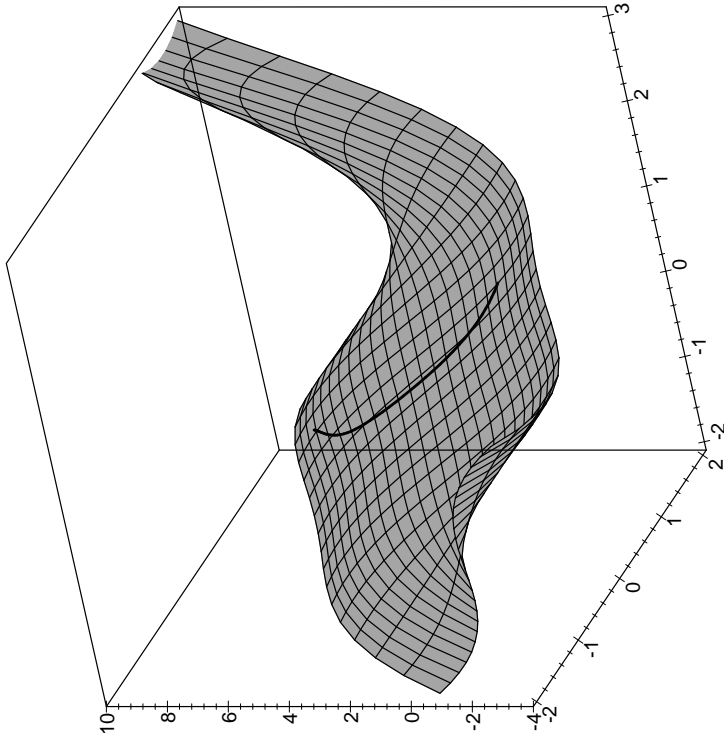


Abb. 21: Der Weg aus Abb. 20 aus einem weiteren Blickwinkel

zu finden. Dafür gibt es eine ganze Reihe numerischer Verfahren, z.B. das Verfahren von NEWTON-RAPHSON, mit denen sich zumindest ein solcher Punkt leicht finden läßt. Leider könnte dieser Punkt unser Ausgangspunkt  $(x_M, y_M)$  sein; deshalb sucht man tatsächlich nicht nach Lösungen der Gleichung  $f(x, y) = z_M$ , sondern nach Lösungen einer leicht abgewandelten Gleichung der Form

$$\tilde{f}(x, y) = z_M,$$

wobei  $\tilde{f}$  dadurch aus  $f$  entsteht, daß man die Funktionswerte in der

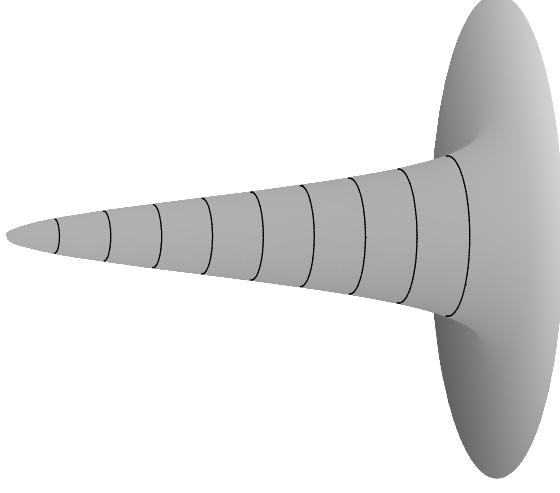
unmittelbaren Umgebung von  $(x_M, y_M)$  stark anhebt, so daß das dortige Minimum verschwindet. Dazu kann man beispielsweise eine Funktion der Form

$$G(x, y) = a e^{\frac{(x-x_M)^2 + (y-y_M)^2}{b}}$$

mit geeigneten Parametern  $a, b$  wählen, wie sie in Abbildung 22 zu sehen ist, und

$$\tilde{f}(x, y) = f(x, y) + G(x, y)$$

setzen.

Abb. 22:  $G(x, y) = e^{-3(x^2+y^2)}$

Dies bringt das Minimum im Punkt  $(x_M, y_M)$  zum Verschwinden und verändert die Funktion praktisch nicht, wenn man nur hinreichend weit entfernt ist von  $(x_M, y_M)$ . (Je kleiner  $b$  ist, umso lokalisierter ist die Veränderung.) Eine Lösung der Gleichung

$$\tilde{f}(x, y) = z_M,$$

so es eine gibt, liegt also nicht in der unmittelbaren Umgebung von  $(x_M, y_M)$  und ist daher ein guter Ausgangspunkt, um dort die Gradientenmethode noch einmal zu starten bis zum nächsten lokalen Minimum und so weiter. Sobald die Gleichung nicht mehr lösbar ist, können wir ziemlich sicher sein, daß  $z_M$  das globale Minimum ist – es sei denn, wir hätten die Parameter  $a$  und  $b$  sehr dumm gewählt.

Tunnelung ist auch ein wichtiges Konzept in der Physik: Dort versucht ein System bekanntlich stets, sein Energieminimum zu erreichen. Dies kann jedoch daran scheitern, daß es sich in einem lokalen Minimum befindet und nicht genügend Energie aufbringen kann, um den Energie-wall zu überwinden, der es vom absoluten Minimum trennt. Zumindest im Bereich der Quantentheorie gibt es dann auch den sogenannten *Tunnelleffekt*, der es einzelnen Teilchen erlaubt, diesen Wall zu tunneln und auf diese Weise einen Zustand niedrigerer Energie zu erreichen.

Im obigen Beispiel geht es nicht um ein Minimum, sondern um ein Maximum, da die Suche danach graphisch besser darstellbar ist. Also graben wir auch keinen Tunnel, sondern spannen ein Hochseil, das irgendwo auf der eingezeichneten Ebenen liegt und uns vom erreichten Zwischenhoch zur Startposition für einen weiteren Anstieg bringt. (Tatsächlich ist die Ebene etwas zu tief eingezeichnet, damit man das alte Maximum noch erkennen kann; das Seil muß also etwas höher hängen.)

## b) Der Metropolis-Algorithmus

Eine weitere Idee zur Vermeidung von Zwischenhochs hat ebenfalls viel mit Physik zu tun: Ein Gas erreicht seinen Zustand minimaler Energie dann, wenn die Bewegungsenergie  $\frac{1}{2}mv^2$  eines jeden Teilchens gleich null ist, wenn sich also nichts mehr bewegt. Dies geschieht aber höchstens am absoluten Nullpunkt; bei positiven Temperaturen werden die meisten Teilchen positive kinetische Energie haben. Nach LUDWIG

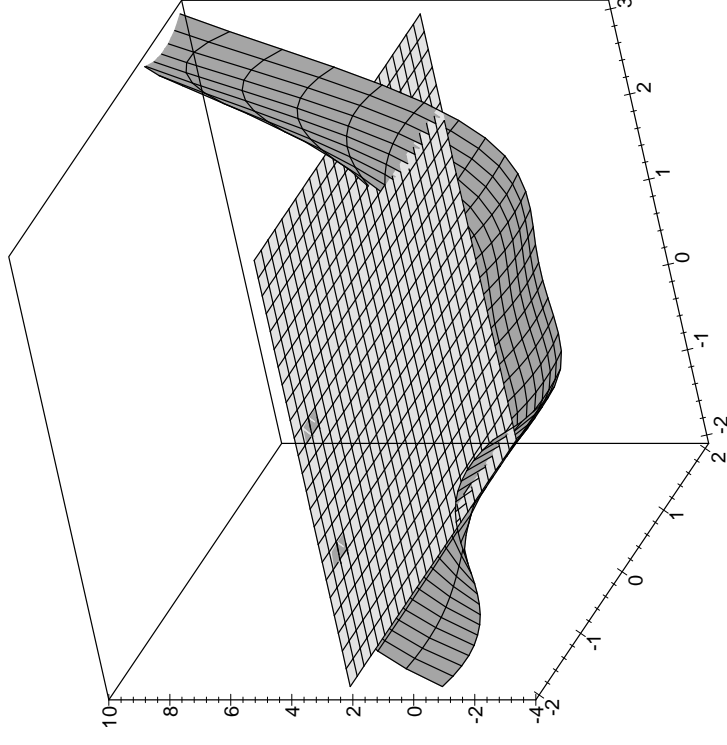


Abb. 23: „Tunnelung“ für Maxima

BOLTZMANN ist dabei die Wahrscheinlichkeit dafür, daß ein Teilchen die Energie  $E = \frac{1}{2}mv^2$  hat, bei Temperatur  $T$  proportional zu

$$e^{-\frac{E}{kT}},$$

mit einer Konstanten  $k \approx 1,38066 \cdot 10^{-23} \text{ J/K}$ , die heute als BOLTZMANN-Konstante bezeichnet wird.

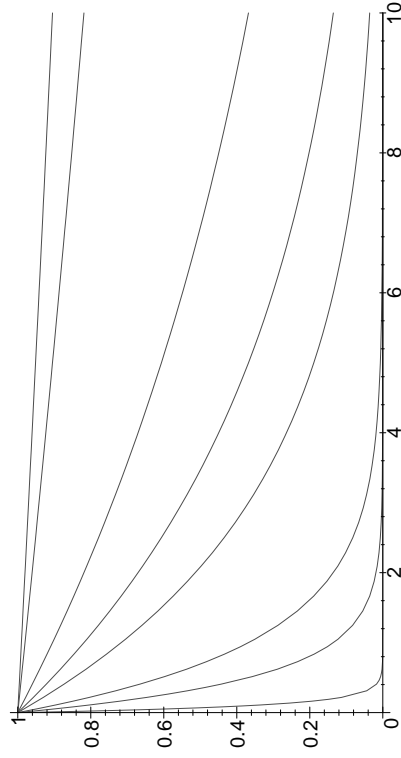


Abb. 24:  $e^{-E/kT}$  für  $kT = 0, 1, 0, 5, 1, 3, 5, 20, 40, 100$



LUDWIG BOLTZMANN (1844–1906) wuchs auf und studierte in Wien; danach lehrte er in Graz, Heidelberg, Berlin, Graz, Wien, Graz, Wien, Leipzig und Wien. Er war jeweils Professor für Theoretische Physik, für Mathematik oder für Experimentalphysik. Auf seiner letzten Stelle in Wien hielt er auch eine so erfolgreiche Philosophievorlesung, daß in Kaiser Franz Josef in den Palast einlud. Am bekanntesten ist er für die Begründung der statistischen Mechanik, eine damals sehr umstrittene Theorie. Ob die damit verbundenen Anfeindungen zu seinem Selbstmord führten, ist unbekannt.

Bei der *simulierten Abkühlung* oder *BOLTZMANN-Maschine* ahmt man dies nach, indem man mit einer hohen Temperatur startet und der Richtung, in der man weitergeht, einer dieser Temperatur entsprechende Freiheit läßt. Man geht also nicht mehr unbedingt in Richtung des Gradienten, sondern geht zufällig in eine von endlich vielen vorgegebenen Richtungen. Die Wahrscheinlichkeit für den Richtungsvektor  $\binom{h_j}{k_j}$  soll dabei analog zur BOLTZMANN-Verteilung festgelegt werden, d.h. wir ordnen ihm eine „Energie“

$$E_j = \pm (f(x + h_j, y + k_j) - f(x, y))$$

zu (positiv bei der Suche nach einem Minimum, negativ bei der Suche nach einem Maximum) und die Wahrscheinlichkeit dafür, daß wir in

Richtung  $\binom{h_j}{k_j}$  gehen, soll proportional sein zu  $e^{-E_j/kT}$ . Sie ist also, falls  $N$  Richtungen zur Verfügung stehen, gleich

$$p_j \stackrel{\text{def}}{=} e^{-E_j/kT} / \sum_{\ell=1}^N e^{-E_\ell/kT}$$

Zur Wahl einer Richtung erzeugen wir uns daher eine Zufallszahl  $Z$  zwischen null und eins und gehen in Richtung  $\binom{h_j}{k_j}$ , wenn

$$\sum_{\ell=1}^{j-1} p_\ell < Z \leq \sum_{\ell=1}^j p_\ell$$

ist. (Die Frage, wie lang die Richtungsvektoren im wievielten Schritt sein sollen, wollen wir hier ausklammern.)

Bie hohen Temperaturen ist damit die Richtung fast vollständig zufallsbedingt gewählt, während in der Nähe des absoluten Nullpunkts praktisch nur noch die optimale Richtung eine Chance hat. Falls wir bei hoher Temperatur in einem Zwischenextremum landen, sorgt dies also mit recht hoher Wahrscheinlichkeit dafür, daß wir dort nicht steckenbleiben.

Am Ende wollen wir allerdings im absoluten Optimum steckenbleiben, d.h. wir müssen die Temperatur im Verlauf der Rechnung immer weiter senken – daher der Name *simulated annealing* = simulierte Abkühlung. Bei der Anwendung auf Optimierungsprobleme bezeichnet man diese Vorgehensweise als den METROPOLIS-Algorithmus. In welcher Weise man die Temperatur am besten senkt, ist immer noch ein Gebiet aktiver Forschung. Man kann zeigen, daß man statistisch betrachtet praktisch immer im Optimum landet, wenn man mit einer hinreichend hohen Ausgangstemperatur  $T_1$  startet und im  $r$ -ten Schritt mit Temperatur  $T_1/\log(r+1)$  arbeitet, aber bei einer derart langsamen Abkühlung braucht der Algorithmus viel zu lange, um ans Ziel zu kommen.

In Abbildung 25 sieht man, wie sich der Algorithmus bei einer Abkühlungsregel verhält, die im  $r$ -ten Schritt mit Temperatur  $T_1/r$  arbeitet: Zumindest im gezeigten Fall funktioniert das recht gut.

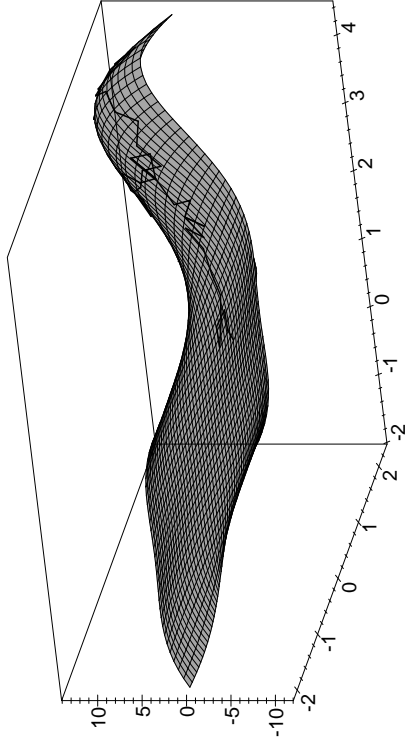


Abb. 25: Der METROPOLIS-Algorithmus für obiges Problem

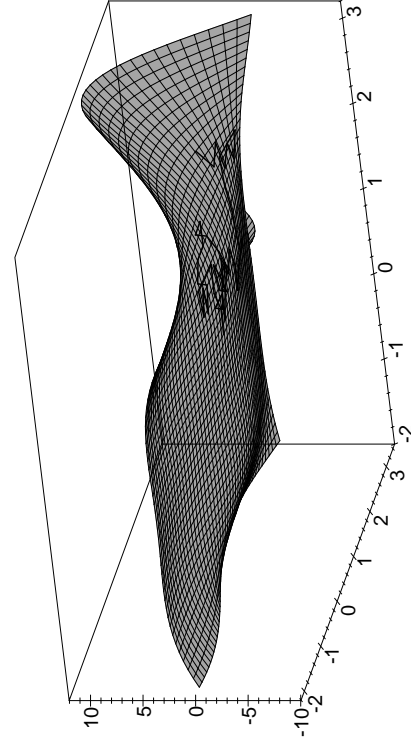


Abb. 26: Ditto mit anderen Zufallszahlen

In anderen Fällen (d.h., wenn andere Zufallszahlen gezogen werden) bleibt man damit aber auch gelegentlich ziemlich lange im Tal hängen; ein Beispiel dafür zeigt Abbildung 26.

Auch hier kommt man aber immerhin in eine gute Startposition, und oft wird es am besten sein, nach hinreichend vielen METROPOLIS-Schritten einfach ein gewöhnliches Gradientenverfahren zu starten.

Zusammenfassend läßt sich sagen, daß der METROPOLIS-Algorithmus und verwandte Verfahren (die sogenannten Monte-Carlo-Methoden) sehr nützliche Hilfsmittel zur Optimierung sind, falls man so gut wie nichts über die zu optimierende Funktion weiß. Sie funktionieren nicht nur bei kontinuierlichen Problemen, wie den hier betrachteten, sondern auch für diskrete und kombinatorische Optimierungsprobleme.

Sie haben allerdings den Nachteil, daß man nie garantieren kann, daß man ein Optimum erreichen wird, und selbst wenn man eines erreicht, kann die Methode dies nicht erkennen. (Es gibt alternative numerische Methoden, die das können.)

## §7: Zusammenfassung

Die nichtlineare Optimierung ist ein sehr weites Feld, von dem wir heute nur einen kleinen Ausschnitt gesehen haben. Dieser Ausschnitt besteht nicht aus den für die Praxis wichtigsten Verfahren, sondern aus denen, die ausgehend von der Schulmathematik der beginnenden Oberstufe am leichtesten zugänglich sind; sie sind zwar (in Kombination mit dem gestern behandelten Simplex-Verfahren) die Grundbausteine, aus denen die meisten praktisch relevanten Verfahren zusammengesetzt sind, aber für die vielen kleinen Abwandlungen, die dazu führen, daß man ein Problem wirklich effizient lösen kann, sind drei Stunden Seminar zu kurz: Dafür muß man eine entsprechende Vorlesung aus dem Bereich Mathematik oder Operations Research besuchen.