
Kapitel IV

Analysis 2

§1 Reihen von Zahlen und Funktionen

Sei a_0, a_1, a_2, \dots eine Folge von (reellen oder komplexen) Zahlen. Mit

$$S_N := \sum_{n=0}^N a_n$$

bezeichnet man die N -te *Partialsumme* der a_n . Die **Folge** (S_N) der Partialsummen nennt man eine *unendliche Reihe* und bezeichnet sie mit

$$\boxed{\sum_{n=0}^{\infty} a_n.}$$

Die Reihe heißt *konvergent* (bzw. *divergent*), falls die Folge (S_N) konvergent (bzw. divergent) ist.

Der Grenzwert wird – wenn er existiert – ebenfalls mit dem Symbol

$$\sum_{n=0}^{\infty} a_n \text{ bezeichnet!}$$

Einige **Beispiele** kennen wir schon:

1. Ist $0 < q < 1$, so konvergiert die *geometrische Reihe* $\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}$.

2. Die *Eulersche Zahl* kann durch eine Reihe dargestellt werden:

$$\sum_{n=0}^{\infty} \frac{1}{n!} = e.$$

Aus den Regeln für die Konvergenz von Folgen ergeben sich analoge Regeln für Reihen:

1. Konvergieren die Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ gegen a bzw. b , so konvergiert auch

$$\sum_{n=0}^{\infty} (a_n + b_n), \text{ und zwar gegen } a + b.$$

2. Ist c eine feste Zahl, so konvergiert $\sum_{n=0}^{\infty} (c \cdot a_n)$ gegen $c \cdot a$.

Ein erstes Kriterium für die Konvergenz einer Reihe ist schnell gefunden:

Notwendiges Kriterium für die Konvergenz einer Reihe

Ist $\sum_{n=0}^{\infty} a_n$ konvergent, so ist (a_n) eine Nullfolge (d.h. eine gegen 0 konvergente Folge)

BEWEIS: Die Folgen S_N und $T_N := S_{N-1}$ konvergieren beide gegen den gleichen Grenzwert, eine Zahl a . Aber dann konvergiert $a_N := S_N - T_N$ gegen $a - a = 0$. \square

Daß dieses Kriterium nicht hinreicht, zeigt das Beispiel der „harmonischen Reihe“:

Beispiel:

Konvergiert die Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$? Die Anschauung hilft hier nicht viel weiter.

Es ist aber leicht, die Partialsummen nach unten abzuschätzen:

$$\begin{aligned} S_{2^N} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \cdots + \left(\frac{1}{2^{N-1}+1} + \cdots + \frac{1}{2^N}\right) \\ &\geq 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \cdots + \frac{1}{8}\right) + \cdots + \left(\frac{1}{2^N} + \cdots + \frac{1}{2^N}\right) \\ &= 1 + \frac{1}{2} + \frac{2}{4} + \frac{4}{8} + \cdots + \frac{2^{N-1}}{2^N} \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2} \quad (1 + N \text{ Summanden}) \\ &= 1 + \frac{N}{2} \rightarrow +\infty. \end{aligned}$$

Also ist die Reihe divergent!

In einem besonderen Spezialfall kommt man fast mit dem notwendigen Kriterium aus:

Leibniz-Kriterium

(a_n) sei eine **monoton fallende** Nullfolge reeller Zahlen.

Dann ist die „alternierende Reihe“ $\sum_{n=0}^{\infty} (-1)^n a_n$ konvergent.

BEWEIS: Wir betrachten die Folgen $u_N := S_{2N-1}$ und $v_N := S_{2N}$. Dann gilt:

$$\begin{aligned} u_{N+1} &= S_{2N+1} \\ &= S_{2N-1} + a_{2N} - a_{2N+1} \\ &\geq S_{2N-1} = u_N. \end{aligned}$$

$$\begin{aligned} v_{N+1} &= S_{2N+2} \\ &= S_{2N} - a_{2N+1} + a_{2N+2} \\ &\leq S_{2N} = v_N. \end{aligned}$$

Weiter ist $v_N = S_{2N} = S_{2N-1} + a_{2N} \geq u_N$, denn die a_n müssen alle ≥ 0 sein. Zusammen ergibt das die folgende Ungleichungskette:

$$\dots \leq u_N \leq u_{N+1} \leq \dots \leq v_{N+1} \leq v_N \leq \dots$$

Nach dem Satz von der monotonen Konvergenz strebt also u_N gegen eine Zahl u und v_N gegen eine Zahl v . Da außerdem $v_N - u_N = a_{2N}$ gegen Null konvergiert, muß $u = v$ sein. Es ist klar, daß dann auch S_N gegen diese Zahl konvergiert. \square

Beispiel:

Die *alternierende harmonische Reihe* $\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n}$ konvergiert! Über den Grenzwert können wir allerdings im Augenblick noch nichts aussagen.

Es sei (a_n) eine konvergente Zahlenfolge. Dann gibt es ein $a \in \mathbb{R}$, so daß folgendes gilt:

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N}, \text{ s.d. } \forall n \geq n_0 \text{ gilt: } |a_n - a| < \varepsilon.$$

Da die Folgenglieder a_n dem Grenzwert beliebig nahe kommen, erwarten wir, daß sich die Folgenglieder auch untereinander beliebig nahe kommen. Das ist in der Tat der Fall:

Sei ein $\varepsilon > 0$ beliebig klein vorgegeben. Dann gibt es ein $N_0 \in \mathbb{N}$, so daß $|a_n - a| < \frac{\varepsilon}{2}$ für $n \geq N_0$ ist. Aber dann gilt für $n, m \geq N_0$:

$$\begin{aligned} |a_n - a_m| &= |(a_n - a) + (a - a_m)| \\ &\leq |a_n - a| + |a_m - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Dieses Verhalten ist so typisch, daß es eine besondere Bezeichnung dafür gibt:

Definition:

Eine Folge (a_n) heißt *Cauchyfolge*, falls gilt:

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N}, \text{ s.d. } \forall n, m \geq n_0 : |a_n - a_m| < \varepsilon.$$

Wir haben oben gerade eingesehen, daß jede konvergente Folge eine Cauchyfolge ist. Das war nicht so überraschend. Es gilt aber für reelle Folgen auch die Umkehrung:

Cauchysches Konvergenzkriterium

Jede (reelle) Cauchyfolge besitzt einen Grenzwert.

BEWEIS: Daß dieser Satz gilt, ist keineswegs selbstverständlich. Das Hauptproblem besteht darin, einen Kandidaten für den Grenzwert zu finden.

1) *Cauchyfolgen sind beschränkt:*

Es sei $n_0 \in \mathbb{N}$ so gewählt, daß $|a_n - a_m| \leq 1$ für $n, m \geq n_0$ ist. Dann ist

$$C := \max\{|a_n| : n = 1, 2, \dots, n_0\}$$

eine endliche reelle Zahl, und es gilt:

Für $n \leq n_0$ ist $|a_n| \leq C < C + 1$, und für $n > n_0$ ist

$$|a_n| = |a_{n_0} + (a_n - a_{n_0})| \leq |a_{n_0}| + |a_n - a_{n_0}| \leq C + 1.$$

Also ist die Folge durch $C + 1$ beschränkt.

2) Nach dem Satz von Bolzano und Weierstraß (der ziemlich direkt aus der Vollständigkeit der reellen Zahlen folgt) besitzt jede beschränkte Folge einen Häufungspunkt. Sei also a ein Häufungspunkt der Cauchyfolge (a_n) . Dann kann man eine Teilfolge $(a_{n(k)})$ von (a_n) finden, die gegen a konvergiert. Demnach ist a ein guter Kandidat für den Grenzwert.

3) Um nun zu zeigen, daß (a_n) selbst schon gegen a konvergiert, geben wir uns ein $\varepsilon > 0$ vor. Da (a_n) Cauchyfolge ist, gibt es ein $n_1 \in \mathbb{N}$, so daß $|a_n - a_m| < \frac{\varepsilon}{2}$ für $n, m \geq n_1$ ist. Und weil $(a_{n(k)})$ gegen a konvergiert, gibt es ein $k_0 \in \mathbb{N}$, so daß $n_0 := n(k_0) \geq n_1$ und $|a_{n_0} - a| < \frac{\varepsilon}{2}$ ist. Für $n \geq n_0$ gilt dann:

$$\begin{aligned} |a_n - a| &= |(a_{n_0} - a) + (a_n - a_{n_0})| \\ &\leq |a_{n_0} - a| + |a_n - a_{n_0}| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Das bedeutet, daß (a_n) gegen a konvergiert. □

Besonders bei den Reihen spielt das Cauchy-Kriterium eine wichtige Rolle:

Cauchy-Kriterium für Reihen

Die Reihe (reeller oder komplexer Zahlen) $\sum_{n=0}^{\infty} a_n$ konvergiert genau dann, wenn gilt:

$$\forall \varepsilon > 0 \exists N_0, \text{ so daß } \forall N \geq N_0 \text{ gilt: } \left| \sum_{n=N_0+1}^N a_n \right| < \varepsilon.$$

BEWEIS: Da $\sum_{n=N_0+1}^N a_n = S_N - S_{N_0}$ ist, folgt das Cauchy-Kriterium für Reihen unmittelbar aus dem für Folgen. □

Der Vorteil des Kriteriums besteht darin, daß man es mit **endlichen** Summen zu tun hat!

Definition:

Eine Reihe (reeller oder komplexer Zahlen) $\sum_{n=0}^{\infty} a_n$ heißt *absolut konvergent*, falls die Reihe $\sum_{n=0}^{\infty} |a_n|$ konvergiert.

Absolute Konvergenz bedingt gewöhnliche Konvergenz

Eine absolut konvergente Reihe konvergiert auch im gewöhnlichen Sinne.

Zum BEWEIS verwendet man das Cauchy-Kriterium. Es ist

$$\left| \sum_{n=N_0+1}^N a_n \right| \leq \sum_{n=N_0+1}^N |a_n|.$$

Konvergiert die Reihe der Absolutbeträge, so wird die rechte Seite bei großem N_0 beliebig klein, und das gilt dann erst recht für die linke Seite. \square

Man beachte: Der Grenzwert einer absolut konvergenten Reihe ist immer noch der Grenzwert im gewöhnlichen Sinne. Die absolute Konvergenz besagt nur, daß die Reihe mindestens so gut konvergiert wie die Reihe der Absolutbeträge.

Besonders häufig wird das folgende Vergleichskriterium benutzt:

Majoranten-Kriterium

Ist $\sum_{n=0}^{\infty} a_n$ eine konvergente Reihe nicht-negativer reeller Zahlen, und ist (c_n) eine Folge reeller oder komplexer Zahlen, so daß $|c_n| \leq a_n$ für fast alle $n \in \mathbb{N}$ gilt, so konvergiert die Reihe $\sum_{n=0}^{\infty} c_n$ absolut!

Auch hier wird der *Beweis* mit Hilfe des Cauchy-Kriteriums geführt.

Wenn nun eine Reihe nicht zufällig das Leibniz-Kriterium erfüllt, wird man i.a. versuchen, die Konvergenz mit Hilfe des Majoranten-Kriteriums auf die Konvergenz einer Vergleichsreihe mit positiven Gliedern zurückzuführen. Für letztere gibt es zahlreiche Untersuchungsmethoden, von denen wir hier nur die populärste betrachten können:

Quotienten-Kriterium

Es sei (a_n) eine Folge von (reellen oder komplexen) Zahlen $\neq 0$. Außerdem gebe es eine reelle Zahl q mit $0 < q < 1$, so daß

$$\left| \frac{a_{n+1}}{a_n} \right| \leq q < 1 \quad \text{für fast alle } n \text{ gilt.}$$

Dann ist die Reihe $\sum_{n=0}^{\infty} a_n$ absolut konvergent.

BEWEIS: „Für fast alle“ bedeutet:

$$\exists N_0 \in \mathbb{N}, \text{ so daß für } n \geq N_0 \text{ gilt: } \left| \frac{a_{n+1}}{a_n} \right| \leq q < 1.$$

Dann ist

$$|a_{N_0+k}| \leq q \cdot |a_{N_0+k-1}| \leq \dots \leq q^k \cdot |a_{N_0}|.$$

Also ist $\sum_{n=0}^{\infty} q^n \cdot |a_{N_0}|$ eine Majorante der Reihe $\sum_{n=0}^{\infty} a_{N_0+n}$. Die erstere konvergiert, es handelt sich ja um eine geometrische Reihe. Nach dem Majorantenkriterium konvergiert dann die zweite Reihe absolut, und damit auch die Ausgangsreihe, die lediglich ein paar Anfangsterme mehr besitzt. \square

Bemerkung: In der Praxis arbeitet man meist mit folgenden Kriterien:

1. Ist $a_n \neq 0$ für fast alle n und $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1$, so konvergiert $\sum_{n=0}^{\infty} a_n$ absolut.

Das ist aus folgendem Grund richtig: Wenn die Quotienten $\left| \frac{a_{n+1}}{a_n} \right|$ gegen ein $q^* < 1$ konvergieren, so gibt es ein $\varepsilon > 0$ und ein $n_0 \in \mathbb{N}$, so daß gilt:

$$\left| \frac{a_{n+1}}{a_n} \right| \leq q^* + \varepsilon < 1 \text{ für } n \geq n_0.$$

Setzt man nun $q := q^* + \varepsilon$, so ist das Quotientenkriterium mit diesem q erfüllt.

2. Ist $a_n \neq 0$ und $\left| \frac{a_{n+1}}{a_n} \right| \geq 1$ für fast alle n , so divergiert die Reihe $\sum_{n=0}^{\infty} a_n$.

Klar, die Glieder a_n bilden keine Nullfolge.

3. (ACHTUNG) Ist $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = 1$, so kann man – zumindest mit dem Quotientenkriterium – keine Aussage machen!

Beispiele:

1. Bei der harmonischen Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ ist

$$a_n = \frac{1}{n}, \text{ also } \left| \frac{a_{n+1}}{a_n} \right| = \frac{n}{n+1},$$

und dieser Ausdruck konvergiert gegen 1. Man kann das Quotientenkriterium nicht anwenden. Tatsächlich wissen wir schon, daß die Reihe divergent ist.

2. Wie steht es mit $\sum_{n=0}^{\infty} \frac{1}{n^2}$?

Der Quotient

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{n^2}{(n+1)^2} = \frac{n^2}{n^2 + 2n + 1} = \frac{1}{1 + \frac{2}{n} + \frac{1}{n^2}}$$

konvergiert gegen 1, also sagt auch hier das Quotientenkriterium nichts aus. Die Situation ist aber anders als bei der harmonischen Reihe. Man kann nämlich wie folgt abschätzen:

$$\begin{aligned} \sum_{n=1}^N \frac{1}{n^2} &\leq 1 + \sum_{n=2}^N \frac{1}{n(n-1)} \\ &= 1 + \sum_{n=2}^N \left(\frac{1}{n-1} - \frac{1}{n} \right) \\ &= 1 + \left(1 + \frac{1}{2} + \cdots + \frac{1}{N-1} \right) - \left(\frac{1}{2} + \cdots + \frac{1}{N} \right) \\ &= 1 + 1 - \frac{1}{N} \leq 2. \end{aligned}$$

Die Folge der Partialsummen ist monoton wachsend und beschränkt, also konvergent. Den Grenzwert können wir leider noch nicht bestimmen.

3. Bei der Reihe $\sum_{n=0}^{\infty} \frac{n^2}{2^n}$ hilft das Quotientenkriterium: Für $n \geq 3$ ist

$$\begin{aligned} \left| \frac{a_{n+1}}{a_n} \right| &= \frac{(n+1)^2 \cdot 2^n}{n^2 \cdot 2^{n+1}} \\ &= \frac{1}{2} \left(1 + \frac{1}{n} \right)^2, \end{aligned}$$

und dieser Ausdruck konvergiert gegen $\frac{1}{2}$. Also ist die Reihe konvergent.

4. Betrachten wir noch die (schon bekannte) Reihe $\sum_{n=0}^{\infty} \frac{1}{n!}$. Hier ist $a_n = \frac{1}{n!}$, und

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{n!}{(n+1)!} = \frac{1}{n+1}$$

konvergiert gegen Null. Also konvergiert die Reihe! Das wußten wir schon, aber der vorliegende Beweis ist erheblich einfacher als der früher angegebene.

5.

Absolut konvergente Reihen verhalten sich sehr gutartig, die anderen Reihen hingegen ausgesprochen böse. Das ist der Inhalt des folgenden Satzes, den wir ohne Beweis zitieren:

Umordnungssatz

1. Ist die Reihe $\sum_{n=0}^{\infty} a_n$ absolut konvergent, so konvergiert jede Umordnung der Reihe gegen den gleichen Grenzwert wie die ursprüngliche Reihe.
2. Ist die Reihe $\sum_{n=0}^{\infty} a_n$ **nicht** absolut konvergent, so gibt es zu jedem $x \in \mathbb{R}$ eine Umordnung der Reihe, die gegen x konvergiert.

Probleme bereitet oft das Produkt von Reihen:

Produktsatz für Reihen

Die Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ seien absolut konvergent und im gewöhnlichen Sinne jeweils gegen a bzw. b konvergent.

Für $n \in \mathbb{N}_0$ sei $c_n := \sum_{i+j=n} a_i b_j = a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0$. Dann ist die Reihe

$\sum_{n=0}^{\infty} c_n$ absolut konvergent und im gewöhnlichen Sinne konvergent gegen $a \cdot b$.

Der BEWEIS ist etwas technisch:

Nach Voraussetzung konvergiert $A_N := \sum_{n=0}^N a_n$ gegen a und $B_N := \sum_{n=0}^N b_n$ gegen b . Wir setzen noch $C_N := \sum_{n=0}^N c_n$ und $a^* := \sum_{n=0}^{\infty} |a_n|$. Hier nutzen wir die absolute Konvergenz aus.

Setzen wir noch $\beta_N := B_N - b$, so ist $B_N = b + \beta_N$, und es gilt:

$$\begin{aligned} C_N &= a_0 b_0 + (a_0 b_1 + a_1 b_0) + \cdots + (a_0 b_N + \cdots + a_N b_0) \\ &= a_0 B_N + a_1 B_{N-1} + \cdots + a_N B_0 \\ &= a_0 (b + \beta_N) + \cdots + a_N (b + \beta_0) \\ &= A_N \cdot b + (a_0 \beta_N + \cdots + a_N \beta_0). \end{aligned}$$

Wir wollen zeigen, daß (C_N) gegen $a \cdot b$ konvergiert. Das ist offensichtlich der Fall, wenn

$$\gamma_N := a_0 \beta_N + \cdots + a_N \beta_0$$

gegen Null konvergiert. Letzteres können wir tatsächlich beweisen:

Sei $\varepsilon > 0$ vorgegeben. Wir wählen ein δ mit $0 < \delta < \frac{\varepsilon}{2a^*}$. Es gibt dann ein N_0 , so daß $|\beta_N| \leq \delta$ für $N \geq N_0$ ist (denn $\beta_N = B_N - b$ konvergiert ja gegen 0). Dieses N_0 halten wir fest. Außerdem wählen wir ein $C > 0$, so daß $|\beta_N| \leq C$ für alle N ist. Dann gilt für $N \geq N_0$:

$$\begin{aligned} |\gamma_N| &\leq |\beta_0 a_N + \cdots + \beta_{N_0} a_{N-N_0}| + |\beta_{N_0+1} a_{N-N_0-1} + \cdots + \beta_N a_0| \\ &\leq C \cdot (|a_N| + \cdots + |a_{N-N_0}|) + \delta \cdot a^* \\ &< C \cdot (|a_N| + \cdots + |a_{N-N_0}|) + \frac{\varepsilon}{2}. \end{aligned}$$

Der linke Summand wird bei festem N_0 und wachsendem N beliebig klein (Cauchy-Kriterium für die absolute Konvergenz der Reihe über die a_n). Also ist $|\gamma_N|$ bei hinreichend großem N kleiner als ε . Das war zu zeigen.

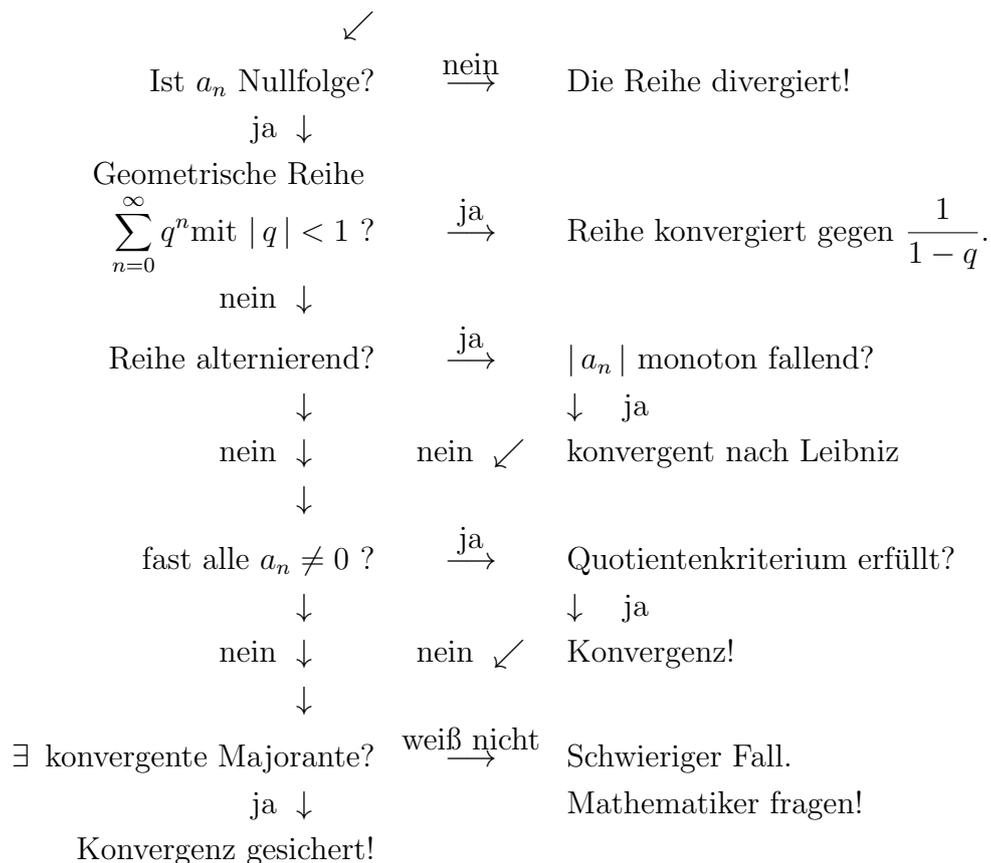
Für die absolute Konvergenz der Produktreihe benutzt man die Abschätzung

$$\sum_{n=0}^N |c_n| \leq \sum_{n=0}^N \sum_{i+j=n} |a_i| \cdot |b_j| \leq \left(\sum_{i=0}^N |a_i| \right) \cdot \left(\sum_{j=0}^N |b_j| \right).$$

Die rechte Seite ist durch das Produkt der absoluten Reihen beschränkt. □

Zum Schluß ein Überblick zur Untersuchung von Reihen:

Es sei eine Reihe $\sum_{n=0}^{\infty} a_n$ gegeben.



Ist die Konvergenz gesichert, so muß man – außer im Falle der geometrischen Reihe – noch nach dem Grenzwert suchen. Da bieten sich folgende Alternativen an:

1. Nachdenken!

(a) Wurde die Reihe in der Vorlesung behandelt?

Wenn ja, im Skript nachschlagen!

(b) Kann man den Grenzwert erraten? (etwa durch Vergleich mit geeigneten Minoranten und Majoranten)

Wenn ja, Konvergenz gegen den mutmaßlichen Grenzwert beweisen!

2. Im Bronstein nachschlagen!

3. Einen Experten fragen!

Ist $M \subset \mathbb{R}^n$ eine beliebige Teilmenge und (f_n) eine Folge von reell- oder komplexwertigen Funktionen auf M , so kann man auch die Folge der Partialsummen $F_N := \sum_{n=0}^N f_n$ betrachten. Diese Folge nennt man eine *Reihe von Funktionen* und schreibt dafür: $\sum_{n=0}^{\infty} f_n$.

Wie bei den Funktionenfolgen unterscheidet man auch bei den Funktionenreihen zwischen punktweiser und gleichmäßiger Konvergenz. Besonders nützlich ist folgendes Konvergenzkriterium:

Weierstraß-Kriterium

Es sei $\sum_{n=0}^{\infty} a_n$ eine konvergente Reihe nicht-negativer reeller Zahlen und (f_n) eine Folge von Funktionen auf einer Menge $M \subset \mathbb{R}^n$, so daß gilt:

$$|f_n(x)| \leq a_n \quad \text{für fast alle } n \text{ und alle } x \in M.$$

Dann konvergiert $\sum_{n=0}^{\infty} f_n$ gleichmäßig auf M .

BEWEIS: Nach dem Majorantenkriterium ist $\sum_{n=0}^{\infty} f_n(x)$ in jedem $x \in M$ absolut konvergent. Also können wir (punktweise) die Grenzfunktion $F(x) := \sum_{n=0}^{\infty} f_n(x)$ bilden. Ist andererseits $F_N(x) := \sum_{n=0}^N f_n(x)$ die N-te Partialsumme, so ist

$$F(x) - F_N(x) = \sum_{n=N+1}^{\infty} f_n(x).$$

Für beliebiges, aber festes x und $N^* > N$ ist stets

$$\left| \sum_{n=N+1}^{N^*} f_n(x) \right| \leq \sum_{n=N+1}^{N^*} |f_n(x)| \leq \sum_{n=N+1}^{N^*} a_n.$$

Läßt man N^* gegen Unendlich gehen, so bleibt die Ungleichung erhalten:

$$|F_N(x) - F(x)| = \left| \sum_{n=N+1}^{\infty} f_n(x) \right| \leq \sum_{n=N+1}^{\infty} a_n.$$

Ist jetzt $\varepsilon > 0$ vorgegeben, so kann man N so groß wählen, daß $\sum_{n=N+1}^{\infty} a_n < \varepsilon$ ist. Für diese N und beliebiges $x \in M$ ist dann auch $|F_N(x) - F(x)| < \varepsilon$. Also konvergiert (F_N) auf M gleichmäßig gegen F . \square

§2 Uneigentliche Integrale

Wir beginnen mit einem Beispiel:

Sei $f(x) := \frac{1}{\sqrt{x}}$. Dann ist $\lim_{x \rightarrow 0^+} f(x) = +\infty$, also f über $[0, 1]$ nicht integrierbar.

Andererseits ist $F(x) := 2\sqrt{x}$ eine Stammfunktion von $f(x)$, und daher

$$\int_{\varepsilon}^1 f(x) dx = F(1) - F(\varepsilon) = 2(1 - \sqrt{\varepsilon}).$$

Niemand kann einen nun daran hindern, ε gegen Null gehen zu lassen. Und siehe da:

$$\lim_{\varepsilon \rightarrow 0^+} \int_{\varepsilon}^1 f(x) dx = 2.$$

Da dieser Grenzwert existiert, wollen wir ihn natürlich gerne als Integral von f über $[0, 1]$ auffassen. Damit das möglich ist, müssen wir den Integralbegriff erweitern. Das geschieht in Gestalt der „uneigentlichen Integrale“.

Definition:

Sei $f : [a, b) \rightarrow \mathbb{R}$ eine Regelfunktion.

Der Grenzwert

$$\int_a^b f(t) dt := \lim_{x \rightarrow b^-} \int_a^x f(t) dt$$

wird als *uneigentliches Integral* bezeichnet. Falls er existiert, nennt man das uneigentliche Integral *konvergent*, andernfalls *divergent*.

Analog erklärt man das uneigentliche Integral einer Regelfunktion $f : (a, b] \rightarrow \mathbb{R}$ durch den rechtsseitigen Limes, wie im obigen Beispiel.

Man kann zeigen, daß dieser Begriff nichts Neues bringt, wenn f schon eine Regelfunktion auf $[a, b]$ ist.

Ist f eine Regelfunktion auf einem offenen Intervall (a, b) , so bildet man das uneigentliche Integral über $[a, b]$, indem man einen Punkt $c \in (a, b)$ wählt und die uneigentlichen Integrale von f über $[a, c]$ und über $[c, b]$ bildet und dann addiert. Das Ergebnis hängt nicht von der Wahl des Punktes c ab, wichtig ist nur, daß man beide Grenzübergänge unabhängig voneinander durchführt!

Beispiel:

Wir betrachten $f(x) := \frac{1}{x^\alpha}$ auf $(0, b]$ für verschiedene α .

a) Ist $\alpha = 1$, so ist $F(x) := \ln(x)$ eine Stammfunktion für $f(x)$, und daher

$$\int_{\varepsilon}^b \frac{1}{x} dx = \ln(b) - \ln(\varepsilon) \longrightarrow +\infty \text{ für } \varepsilon \rightarrow 0.$$

Das uneigentliche Integral divergiert!

b) Ist $\alpha \neq 1$, so ist $F(x) := -\frac{1}{(\alpha-1)x^{\alpha-1}}$ Stammfunktion für f .

Wir betrachten zunächst den Fall $\alpha < 1$: dann ist

$$\int_{\varepsilon}^b \frac{1}{x^{\alpha}} dx = -\frac{1}{\alpha-1} \cdot \left(\frac{1}{b^{\alpha-1}} - \frac{1}{\varepsilon^{\alpha-1}} \right).$$

Da $1 - \alpha > 0$ ist, strebt $\frac{1}{\varepsilon^{\alpha-1}} = \varepsilon^{1-\alpha}$ gegen Null für $\varepsilon \rightarrow 0$.

Also existiert $\int_0^b \frac{1}{x^{\alpha}} dx = -\frac{1}{(\alpha-1)b^{\alpha-1}}$ für $\alpha < 1$.

Insbesondere ist $\int_0^1 \frac{1}{x^{\alpha}} dx = \frac{1}{1-\alpha}$, z.B. $\int_0^1 \frac{1}{\sqrt{x}} dx = 2$.

c) Ist $\alpha > 1$, so ist $\alpha - 1 > 0$, und $\frac{1}{\varepsilon^{\alpha-1}}$ strebt gegen $+\infty$ für $\varepsilon \rightarrow 0$. In diesem Fall divergiert das uneigentliche Integral.

Das trifft z.B. auf $\int_0^1 \frac{1}{x^2} dx$ zu.

Bisher haben wir nur über beschränkte Intervalle integriert. Manchmal möchte man jedoch die Integration auf ganz \mathbb{R} oder zumindest auf eine Halbachse ausdehnen. Dann behilft man sich mit einer anderen Sorte von uneigentlichen Integralen:

Definition:

Sei $f : [a, +\infty) \rightarrow \mathbb{R}$ eine Regelfunktion. Dann wird auch der Grenzwert

$$\int_a^{\infty} f(t) dt := \lim_{x \rightarrow \infty} \int_a^x f(t) dt$$

als *uneigentliches Integral* bezeichnet. Konvergenz und Divergenz des uneigentlichen Integrals erklärt man wie oben.

Analog definiert man das uneigentliche Integral über $(-\infty, b]$.

Beispiel:

Wir betrachten noch einmal $f(x) = \frac{1}{x^{\alpha}}$ für verschiedene α .

a) $\alpha = 1$: $\int_a^x \frac{1}{t} dt = \ln(x) - \ln(a)$ strebt für $x \rightarrow +\infty$ gegen $+\infty$.

Das uneigentliche Integral divergiert also auch hier.

b) Ist $\alpha < 1$, so ist $\int_a^x \frac{1}{t^\alpha} dt = -\frac{1}{\alpha-1} \cdot \left(\frac{1}{x^{\alpha-1}} - \frac{1}{a^{\alpha-1}} \right)$, wobei $\frac{1}{x^{\alpha-1}} = x^{1-\alpha}$ gegen $+\infty$ strebt, für $x \rightarrow \infty$. Auch dieses Integral divergiert.

c) Ist $\alpha > 1$, so konvergiert das uneigentliche Integral offensichtlich. Das bedeutet, daß insbesondere das Integral $\int_1^\infty \frac{1}{x^2} dx$ konvergiert, während $\int_1^\infty \frac{1}{\sqrt{x}} dx$ divergiert.

Z. B. ist $\int_1^\infty \frac{1}{x^2} dx = 1$.

Wir haben übrigens gesehen, daß $\int_0^\infty \frac{1}{x^\alpha} dx$ für **kein** α konvergiert!

Wie bei offenen Intervallen müssen uneigentliche Integrale über ganz \mathbb{R} durch zwei voneinander unabhängige Grenzprozesse berechnet werden:

Definition:

Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine Regelfunktion. Das *uneigentliche Integral* $\int_{-\infty}^{+\infty} f(t) dt$ konvergiert genau dann, wenn die beiden uneigentlichen Integrale

$$\int_{-\infty}^0 f(t) dt \text{ und } \int_0^{+\infty} f(t) dt$$

konvergieren. Der Wert des Integrals über ganz \mathbb{R} ist gleich der Summe der Werte der Teilintegrale, d.h. es ist

$$\int_{-\infty}^{+\infty} f(t) dt = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow +\infty}} \int_a^b f(t) dt.$$

Man beachte, daß in der letzten Formel a und b unabhängig voneinander gegen die Grenzen streben. Manchmal findet man auch noch den folgenden Grenzwert:

$$HW \int_{-\infty}^{+\infty} f(t) dt := \lim_{r \rightarrow \infty} \int_{-r}^{+r} f(t) dt.$$

Man spricht dann vom *Cauchy'schen Hauptwert*. Es kann passieren, daß dieser Hauptwert existiert, obwohl das uneigentliche Integral von f über \mathbb{R} divergiert.

Wenn keine explizite Stammfunktion gegeben ist, wird es schwierig mit dem Nachweis von Konvergenz oder Divergenz eines uneigentlichen Integrals. Für den Fall gibt es aber gewisse Konvergenzkriterien.

Cauchy Kriterium für uneigentliche Integrale

Sei $f : [a, +\infty) \rightarrow \mathbb{R}$ eine Regelfunktion. Das uneigentliche Integral $\int_a^\infty f(x) dx$ konvergiert genau dann, wenn gilt:

$$\forall \varepsilon > 0 \exists C = C(\varepsilon) \geq a, \text{ s.d. } \left| \int_{x_1}^{x_2} f(x) dx \right| < \varepsilon \text{ für } C < x_1 < x_2 \text{ ist.}$$

BEWEIS:

„ \implies “: Das uneigentliche Integral konvergiere gegen $A \in \mathbb{R}$. Ist $\varepsilon > 0$ vorgegeben, so wählen wir $C = C(\varepsilon)$ so, daß

$$\left| \int_a^r f(x) dx - A \right| < \frac{\varepsilon}{2} \text{ für } r \geq C \text{ ist.}$$

Für $C < x_1 < x_2$ gilt dann:

$$\begin{aligned} \left| \int_{x_1}^{x_2} f(x) dx \right| &= \left| \int_a^{x_2} f(x) dx - \int_a^{x_1} f(x) dx \right| \\ &= \left| \left(\int_a^{x_2} f(x) dx - A \right) + \left(A - \int_a^{x_1} f(x) dx \right) \right| \\ &\leq \left| \int_a^{x_2} f(x) dx - A \right| + \left| \int_a^{x_1} f(x) dx - A \right| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

„ \impliedby “: Nun sei das Kriterium erfüllt.

Wir setzen $x_n := a + n$ für $n \in \mathbb{N}$. Dann ist (x_n) eine monoton wachsende Folge mit $\lim_{n \rightarrow \infty} x_n = +\infty$.

Wir setzen außerdem $A_n := \int_a^{x_n} f(x) dx$. Wir wollen zeigen, daß die Folge (A_n) gegen eine reelle Zahl A konvergiert, und daß A gerade das uneigentliche Integral ist.

a) Sei $\varepsilon > 0$ vorgegeben. Wir wählen $C = C(\varepsilon)$ gemäß dem Kriterium. Außerdem wählen wir $n_0 \in \mathbb{N}$ so groß, daß $x_n \geq C$ für $n \geq n_0$ ist. Für solche n gilt:

$$\left| A_n - A_{n_0} \right| = \left| \int_{x_{n_0}}^{x_n} f(x) dx \right| < \varepsilon.$$

Daraus folgt, daß (A_n) eine Cauchyfolge ist, die dementsprechend gegen ein $A \in \mathbb{R}$ konvergieren muß.

Wir wollen zeigen, daß $\lim_{r \rightarrow \infty} \int_a^r f(x) dx = A$ ist. Dazu sei noch einmal ein $\varepsilon > 0$ vorgegeben.

Wir wählen $n_0 \in \mathbb{N}$ so, daß $\left| A_n - A \right| < \frac{\varepsilon}{2}$ und $\left| \int_{x_n}^r f(x) dx \right| < \frac{\varepsilon}{2}$ für $n \geq n_0$ und $r > x_n$ ist. Für solche r ist dann

$$\begin{aligned} \left| \int_a^r f(x) dx - A \right| &= \left| \int_a^{x_n} f(x) dx + \int_{x_n}^r f(x) dx - A \right| \\ &\leq \left| \int_{x_n}^r f(x) dx \right| + \left| A_n - A \right| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Damit ist alles gezeigt. □

Der Nutzen des Cauchyriteriums wird sich gleich zeigen. Zuvor brauchen wir jedoch noch einen weiteren Begriff:

Definition:

Das uneigentliche Integral $\int_a^\infty f(x) dx$ konvergiert absolut, falls $\int_a^\infty |f(x)| dx$ konvergiert.

Für andere Typen von uneigentlichen Integralen definiert man die absolute Konvergenz entsprechend.

Absolute Konvergenz impliziert gewöhnliche Konvergenz

Konvergiert ein uneigentliches Integral über f absolut, so auch im gewöhnlichen Sinne.

BEWEIS: Es ist

$$\left| \int_{x_1}^{x_2} f(x) dx \right| \leq \int_{x_1}^{x_2} |f(x)| dx.$$

Wenn also $|f(x)|$ das Cauchyriterium erfüllt, so erst recht $f(x)$ selbst. Daraus folgt die Behauptung. □

Die Umkehrung des Satzes ist falsch!

Nun haben wir ein Mittel an der Hand, die Konvergenz von Integralen durch den Vergleich mit bekannteren Integralen zu beweisen:

Majorantenkriterium für uneigentliche Integrale

Es seien f und g zwei Regelfunktionen über einem Intervall I , mit $|f| \leq g$. Konvergiert das uneigentliche Integral über g , so konvergiert das uneigentliche Integral über f absolut.

BEWEIS: Man verwende das Cauchyriterium und die Monotonie des Integrals. □

Beispiele:

1. $\int_1^\infty \frac{\cos x}{x^2} dx$ konvergiert, weil $\int_1^\infty \frac{1}{x^2} dx$ konvergiert.

2.

$$\begin{aligned} \text{Es ist} \quad \int_1^r e^{-x} dx &= -\int_1^r (e^{-x})' dx \\ &= -(e^{-r} - e^{-1}), \end{aligned}$$

und das konvergiert gegen $\frac{1}{e}$ für $r \rightarrow \infty$.

Also konvergiert das uneigentliche Integral $\int_1^{\infty} e^{-x} dx$.

Analog ist

$$\int_{-\infty}^{-1} e^x dx = \lim_{r \rightarrow \infty} \int_{-r}^{-1} (e^x)' dx = \lim_{r \rightarrow \infty} (e^{-1} - e^{-r}) = e^{-1}.$$

Für $x \geq 1$ ist $x^2 \geq x$, also $e^{-x^2} \leq e^{-x}$.

Für $x \leq -1$ ist $x^2 = |x|^2 \geq |x| = -x$, also $-x^2 \leq x$. Damit ist dort $e^{-x^2} \leq e^x$, und es folgt die Konvergenz des „Fehlerintegrals“ $\int_{-\infty}^{+\infty} e^{-x^2} dx$.

3. Die Funktion $f(x) := \frac{\sin(x)}{x}$ ist überall stetig, und es gilt:

$$\begin{aligned} \int_s^t \frac{\sin x}{x} dx &= - \int_s^t \frac{\cos' x}{x} dx \\ &= - \frac{\cos x}{x} \Big|_s^t - \int_s^t \frac{\cos x}{x^2} dx, \end{aligned}$$

also

$$\left| \int_s^t \frac{\sin x}{x} dx \right| \leq \frac{1}{s} + \frac{1}{t} + \int_s^t \frac{1}{x^2} dx.$$

Der Ausdruck auf der rechten Seite strebt für $0 < s < t$ und $s \rightarrow \infty$ gegen Null. Mit dem Cauchy Kriterium folgt nun, daß

$$\int_0^{\infty} \frac{\sin x}{x} dx$$

konvergiert.

Man kann jedoch zeigen, daß $\int_0^{\infty} \left| \frac{\sin x}{x} \right| dx$ divergiert!

4. Die Funktion $\Gamma : (0, \infty) \rightarrow \mathbb{R}$ sei definiert durch

$$\Gamma(x) := \int_0^{\infty} e^{-t} t^{x-1} dt.$$

Dabei ist $t^{x-1} = e^{(x-1) \ln t}$.

Das uneigentliche Integral $\Gamma(1) = \int_0^{\infty} e^{-t} dt = -(e^{-t}) \Big|_0^{\infty} = 1$ konvergiert offensichtlich.

Nun sei $x > 0$ und $x \neq 1$. Wir untersuchen zunächst die Konvergenz des uneigentlichen Integrals bei $t = 0$. Da wir dann $0 < t \leq 1$ voraussetzen können, ist $1 < e^t \leq e$, also $|e^{-t} t^{x-1}| < t^{x-1}$.

Ist $x > 1$, so strebt t^{x-1} für $t \rightarrow 0$ gegen Null und es gibt keine Probleme.

Für $0 < x < 1$ ist aber auch $0 < 1 - x < 1$, und das uneigentliche Integral über $t^{x-1} = \frac{1}{t^{1-x}}$ konvergiert.

Nun zeigen wir die Konvergenz des Integrals für $t \rightarrow \infty$. Das geht folgendermaßen:

Da die Exponentialfunktion stärker als jede Potenz wächst, strebt $e^{-t}t^{x+1}$ für $t \rightarrow \infty$ gegen Null. Also gibt es eine Zahl $C > 0$, so daß $e^{-t}t^{x+1} \leq C$ für alle t ist. Daraus folgt:

$$|e^{-t}t^{x-1}| \leq C \cdot \frac{1}{t^2} \text{ für } t \geq 1.$$

Mit dem Majorantenkriterium ergibt sich die Konvergenz des Integrals.

Die so eingeführte Funktion nennt man die *Gamma-Funktion*. Es gilt:

a) $\Gamma(1) = 1$.

b) $\Gamma(x+1) = x \cdot \Gamma(x)$.

Zum BEWEIS von (b):

Es ist $\Gamma(x+1) = \int_0^\infty e^{-t}t^x dt$. Mit partieller Integration erhält man:

$$\int_\varepsilon^r e^{-t}t^x dt = -e^{-t}t^x \Big|_\varepsilon^r + x \int_\varepsilon^r e^{-t}t^{x-1} dt.$$

Der erste Summand strebt gegen 0, der zweite gegen $x \cdot \Gamma(x)$.

Insbesondere folgt nun:

$$\Gamma(2) = \Gamma(1+1) = 1 \cdot \Gamma(1) = 1,$$

$$\Gamma(3) = \Gamma(2+1) = 2 \cdot \Gamma(2) = 2$$

$$\Gamma(4) = \Gamma(3+1) = 3 \cdot \Gamma(3) = 2 \cdot 3$$

$$\text{und allgemein } \Gamma(n+1) = n \cdot \Gamma(n) = n! \text{ für } n \in \mathbb{N}.$$

Zum Schluß wollen wir noch den engen Zusammenhang zwischen Reihen und uneigentlichen Integralen hervorheben:

Vergleichssatz

Sei $m \in \mathbb{N}$ und $f : [m, \infty) \rightarrow \mathbb{R}$ positiv und monoton fallend. Dann haben die Reihe $\sum_{k=m}^{\infty} f(k)$ und das uneigentliche Integral $\int_m^{\infty} f(x) dx$ das gleiche Konvergenzverhalten.

BEWEIS: Als monotone Funktion ist f eine Regelfunktion. Auf dem Intervall $[k, k+1]$ ist

$$f(k) \geq f(x) \geq f(k+1),$$

also auch

$$f(k) \geq \int_k^{k+1} f(x) dx \geq f(k+1)$$

und damit

$$\sum_{k=m}^N f(k) \geq \int_m^{N+1} f(x) dx \geq \sum_{k=m+1}^{N+1} f(k).$$

Daraus folgt die Behauptung. □

Beispiel :

Aus dem Vergleichssatz und unseren Kenntnissen über uneigentliche Integrale folgt sofort:

$$\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}} \text{ konvergiert genau dann, wenn } \alpha > 1 \text{ ist.}$$

So ist etwa $\sum_{n=1}^{\infty} \frac{1}{n}$ divergent und $\sum_{n=1}^{\infty} \frac{1}{n\sqrt{n}} = \sum_{n=1}^{\infty} \frac{1}{n^{1.5}}$ konvergent.

§3 Potenzreihen

Sei (c_n) eine Folge (reeller oder komplexer) Zahlen, $a \in \mathbb{C}$. Dann heißt

$$f(z) := \sum_{n=0}^{\infty} c_n (z - a)^n$$

eine *Potenzreihe* mit *Entwicklungspunkt* a . Die Zahlen c_n heißen die *Koeffizienten* der Potenzreihe.

Ist $a \in \mathbb{R}$ und sind alle Koeffizienten c_n reell, so spricht man von einer *reellen Potenzreihe* und schreibt die Variable auch reell:

$$f(x) = \sum_{n=0}^{\infty} c_n (x - a)^n.$$

Um nicht doppelte Arbeit machen zu müssen, betreiben wir die Theorie gleich im Komplexen. Der reelle Fall ist darin enthalten. Man beachte: $z \in \mathbb{C}$ ist genau dann reell, wenn $\bar{z} = z$ ist.

Die Betragsfunktion liefert auf \mathbb{C} eine Metrik. Wir können also mit den topologischen Begriffen arbeiten, die wir schon von den metrischen Räumen her kennen. Die metrischen „Kugeln“ sind im Falle der komplexen Ebene allerdings Kreisscheiben. Statt des Buchstabens B (für *ball*) verwenden wir den Buchstaben D (für *disk*).

$$D_\varepsilon(z_0) := \{z \in \mathbb{C} : |z - z_0| < \varepsilon\}$$

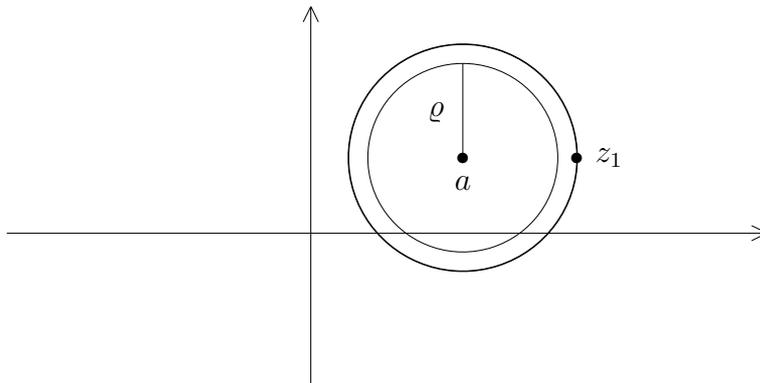
und $\overline{D_\varepsilon(z_0)} := \{z \in \mathbb{C} : |z - z_0| \leq \varepsilon\}.$

Über die gleichmäßige Konvergenz von Potenzreihen

Die Potenzreihe $f(z) = \sum_{n=0}^{\infty} c_n (z - a)^n$ konvergiere für ein $z_1 \in \mathbb{C}$, $z_1 \neq a$. Es sei $0 < \varrho < |z_1 - a|$.

Dann konvergiert $f(z)$ auf der abgeschlossenen Kreisscheibe $\overline{D_\varrho(a)}$ absolut und gleichmäßig,

und die Reihe $g(z) := \sum_{n=1}^{\infty} n \cdot c_n (z - a)^{n-1}$ konvergiert ebenfalls absolut und gleichmäßig auf $\overline{D_\varrho(a)}$.



BEWEIS: 1) Da $\sum_{n=0}^{\infty} c_n(z_1 - a)^n$ nach Voraussetzung konvergiert, gibt es eine Konstante $M > 0$, so daß $|c_n(z_1 - a)^n| \leq M$ für alle n ist. Und da $\rho < |z_1 - a|$ sein soll, ist $q := \frac{\rho}{|z_1 - a|} < 1$.

Für alle z mit $|z - a| \leq \rho$ gilt dann:

$$\begin{aligned} |c_n(z - a)^n| &= |c_n(z_1 - a)^n| \cdot \left| \frac{z - a}{z_1 - a} \right|^n \\ &\leq M \cdot q^n. \end{aligned}$$

Die geometrische Reihe $\sum_{n=0}^{\infty} M q^n$ konvergiert. Mit dem Majorantenkriterium folgt, daß $\sum_{n=0}^{\infty} c_n(z - a)^n$ absolut konvergiert, und mit dem Weierstraßkriterium folgt, daß die Reihe sogar gleichmäßig auf $\overline{D_\rho(a)}$ konvergiert.

2) Nach (1) ist $|n \cdot c_n(z - a)^{n-1}| \leq n \cdot M \cdot q^{n-1}$, und

$$\frac{(n+1)M \cdot q^n}{nM \cdot q^{n-1}} = \frac{n+1}{n} \cdot q$$

konvergiert gegen $q < 1$. Aus dem Quotientenkriterium folgt jetzt, daß $\sum_{n=0}^{\infty} n \cdot M \cdot q^{n-1}$ konvergiert, und wie oben kann man daraus schließen, daß $\sum_{n=0}^{\infty} n \cdot c_n(z - a)^{n-1}$ auf $\overline{D_\rho(a)}$ absolut und gleichmäßig konvergiert. \square

Der vorliegende Satz hat weitreichende Konsequenzen für das Konvergenzverhalten von Potenzreihen.

Definition:

Sei $f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n$ eine Potenzreihe. Die Zahl

$$R := \sup\{r \geq 0 \mid \exists z_1 \in \mathbb{C}, \text{ so daß } f(z) \text{ konvergent in } z_1 \text{ und } r = |z_1 - a| \text{ ist.}\}$$

heißt *Konvergenzradius* der Potenzreihe. Die Fälle $R = 0$ und $R = +\infty$ sind dabei auch zugelassen!

Der Kreis um a mit Radius R heißt der *Konvergenzkreis* der Reihe.

Konvergenzverhalten und Konvergenzradius

R sei der Konvergenzradius der Potenzreihe $f(z)$. Dann gilt:

1. Für $0 < r < R$ konvergiert $f(z)$ auf $\overline{D_r(a)}$ absolut und gleichmäßig.
2. Ist $|z_1 - a| > R$, so divergiert $f(z)$ in z_1 .

BEWEIS: 1) ist klar, wegen Satz 4.1.

2) Nach Definition von R kann $f(z)$ in einem Punkt z_1 mit $|z_1 - a| > R$ nicht mehr konvergieren. \square

Bemerkung: Bei **reellen** Potenzreihen geht alles genauso, man hat lediglich die Kreise durch Intervalle zu ersetzen.

Wir wissen jetzt, daß eine Potenzreihe im Innern ihres Konvergenzkreises konvergiert und außerhalb divergiert. Das Verhalten auf dem Rand des Kreises kann man nicht allgemein vorhersagen. Dafür sind gesonderte Betrachtungen erforderlich, die manchmal sehr schwierig werden können.

Auf jeden Fall ist es wichtig, den Konvergenzradius bestimmen zu können. In vielen Fällen gibt es dafür eine praktische Formel:

Quotientenformel für den Konvergenzradius

Sei (c_n) eine Folge von (reellen oder komplexen) Zahlen, $c_n \neq 0$ für fast alle n .

Wenn die Folge $\left| \frac{c_n}{c_{n+1}} \right|$ konvergiert, dann ist

$$R := \lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right|$$

der Konvergenzradius der Reihe $f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n$.

Man beachte, daß der Entwicklungspunkt a dabei keine Rolle spielt!

BEWEIS: Wir verwenden das Quotientenkriterium: Es ist

$$\left| \frac{c_{n+1}(z-a)^{n+1}}{c_n(z-a)^n} \right| = \left| \frac{c_{n+1}}{c_n} \right| \cdot |z-a|,$$

und dieser Ausdruck konvergiert (für festes z) gegen $\frac{1}{R} \cdot |z-a|$.

Ist $|z-a| < R$, also $\frac{1}{R} \cdot |z-a| < 1$, so konvergiert die Reihe. Ist $|z-a| > R$, so divergiert sie. Also muß R der Konvergenzradius sein!

Wir haben uns dabei übrigens nicht um gleichmäßige Konvergenz kümmern müssen. \square

Beispiele:

1. Sei $f(z) = \sum_{n=0}^{\infty} z^n$. Dann ist $a = 0$ und $c_n = 1$ für alle $n \in \mathbb{N}$. Das ergibt den Konvergenzradius $R = 1$.

Für $|z| < 1$ konvergiert die Reihe gegen $\frac{1}{1-z}$. Da alle Koeffizienten reell sind, kann man die Reihe auch reell auffassen. Tatsächlich nimmt die Grenzfunktion dann auf dem Konvergenzintervall $(-1, 1)$ nur reelle Werte an. An den Randpunkten $x = -1$ und $x = +1$ divergiert die Reihe.

2. Sei $f(z) = \sum_{n=1}^{\infty} \frac{z^n}{n}$. Hier ist $a = 0$ und $c_n = \frac{1}{n}$.

Da $\frac{c_n}{c_{n+1}} = \frac{n+1}{n}$ gegen 1 konvergiert, ist $R = 1$. An den Rändern des Konvergenzintervalls ist das Verhalten diesmal unterschiedlich:

Die harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ divergiert, die alternierende harmonische Reihe $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$ konvergiert.

3. Sei $f(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$. Wieder ist $a = 0$, und außerdem ist $c_n = \frac{1}{n!}$ für alle n , also

$$\frac{c_n}{c_{n+1}} = \frac{(n+1)!}{n!} = n+1 \text{ konvergent gegen } +\infty.$$

Diese Reihe konvergiert auf ganz \mathbb{C} .

Wir wollen nun einige Eigenschaften derjenigen Funktionen zusammenstellen, die als Grenzfunktionen von Potenzreihen auftreten:

Stetigkeit von Potenzreihen

Hat $f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n$ den Konvergenzradius R , so ist $f(z)$ im Innern des Konvergenzkreises $D_R(a) = \{z \in \mathbb{C} \mid |z-a| < R\}$ stetig.

BEWEIS: Die identische Funktion $z \mapsto z$ ist sicherlich auf ganz \mathbb{C} stetig, ebenso jede konstante Funktion. Da Summen und Produkte stetiger Funktionen wieder stetig sind, ist auch jedes komplexe Polynom auf ganz \mathbb{C} stetig.

Die Folge der Partialsummen der Potenzreihe ist eine Folge von komplexen Funktionen, und sie konvergiert auf jeder abgeschlossenen Kreisscheibe $\overline{D_\rho(a)}$, $0 \leq \rho < R$, gleichmäßig. Deshalb ist auch die Grenzfunktion f dort stetig. \square

Differenzierbarkeit und Integrierbarkeit von Potenzreihen

Sei $a \in \mathbb{R}$ und $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$ eine Potenzreihe mit reellen Koeffizienten um a , mit Konvergenzradius R .

Dann ist die Grenzfunktion $f(x)$ auf dem Konvergenzintervall $(a-R, a+R)$ stetig differenzierbar, und es gilt:

Die Reihe $\sum_{n=1}^{\infty} n \cdot c_n(x-a)^{n-1}$ konvergiert im Innern des Konvergenzintervalls absolut und gleichmäßig gegen die Ableitung $f'(x)$.

Umgekehrt konvergiert die Reihe $\sum_{n=0}^{\infty} \frac{c_n}{n+1} (x-a)^{n+1}$ im Innern des Konvergenzintervalls absolut und gleichmäßig gegen $\int_a^x f(t) dt$.

Potenzreihen können also gliedweise differenziert und integriert werden!

BEWEIS: Wir beginnen mit der gliedweisen Differenzierbarkeit.

Die Funktionenfolge $F_N(x) := \sum_{n=0}^N c_n(x-a)^n$ konvergiert auf $(a-R, a+R)$ punktweise gegen eine stetige Funktion f . Jede der Funktionen F_N ist stetig differenzierbar, und die Folge der Ableitungen F'_N konvergiert nach dem Satz über die gleichmäßige Konvergenz von Potenzreihen auf jedem abgeschlossenen Teilintervall gleichmäßig.

Also ist f stetig differenzierbar, und es gilt:

$$f'(x) = \lim_{N \rightarrow \infty} F'_N(x) = \sum_{n=1}^{\infty} n \cdot c_n(x-a)^{n-1}.$$

Nun kommen wir zur gliedweisen Integrierbarkeit:

Ist $0 < r < R$, so konvergiert (F_N) auf $(a-r, a+r)$ gleichmäßig gegen f .

Also konvergiert auch $\int_a^x F_N(t) dt$ für jedes feste $x \in (a - r, a + r)$ gegen $\int_a^x f(t) dt$. Es gilt aber:

$$\int_a^x F_N(t) dt = \sum_{n=0}^N c_n \int_a^x (t-a)^n dt = \sum_{n=0}^N c_n \cdot \frac{(t-a)^{n+1}}{n+1} \Big|_a^x.$$

Das bedeutet, daß die gewünschte Potenzreihe auf dem gesamten Konvergenzintervall $(a - R, a + R)$ punktweise gegen das Integral konvergiert. Aus dem schon angesprochenen Satz über die gleichmäßige Konvergenz von Potenzreihen folgt die Behauptung. \square

Da die Ableitung einer Potenzreihe wieder eine Potenzreihe ist, kann man das obige Argument wiederholen und erhält:

Folgerung

Eine (reelle) Potenzreihe $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$ ist im Konvergenzintervall beliebig oft differenzierbar, und es gilt:

$$c_n = \frac{f^{(n)}(a)}{n!}, \quad \text{für alle } n \geq 0.$$

BEWEIS: Wir müssen nur noch die Formel beweisen:

Offensichtlich ist

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1) \cdots (n-k+1) c_n (x-a)^{n-k}.$$

Setzt man $x = a$ ein, so erhält man:

$$f^{(k)}(a) = k(k-1) \cdots (k-k+1) c_k = k! c_k.$$

\square

Die Formel erinnert an die Taylorentwicklung:

Sei $I \subset \mathbb{R}$ ein Intervall, $f : I \rightarrow \mathbb{R}$ eine unendlich oft differenzierbare Funktion. Dann kann man von f Taylor-Polynome beliebig hohen Grades bilden: Sei $a \in I$ ein fest gewählter Entwicklungspunkt,

$$T_n f(x) = T_n f(x; a) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k$$

das n-te Taylorpolynom von f in a . Dann ist

$$f(x) = T_n f(x) + R_n(x),$$

wobei das Restglied durch

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!} \cdot (x-a)^{n+1}$$

gegeben ist, mit einem (von x abhängigen) Punkt c zwischen a und x .

Definition:

Ist $f : I \rightarrow \mathbb{R}$ beliebig oft differenzierbar und $a \in I$, so heißt

$$Tf(x; a) := \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n$$

die *Taylorreihe* von f in a .

Beispiele:

1. $f(x) = \sum_{n=0}^{\infty} c_n (x - a)^n$ sei eine konvergente Potenzreihe, mit Konvergenzradius R . Dann stimmt f auf $(a - R, a + R)$ offensichtlich mit seiner Taylor-Reihe überein.

2. Wir betrachten die Potenzreihe $\sum_{n=0}^{\infty} (-1)^n x^{2n}$. Dann ist $a = 0$ und

$$c_n = \begin{cases} (-1)^k & \text{falls } n = 2k \\ 0 & \text{sonst} \end{cases}.$$

Wir können die Formel für den Konvergenzradius nicht benutzen, aber da es sich um eine geometrische Reihe handelt, können wir direkt sehen, daß $R = 1$ ist. Als Grenzwert ergibt sich:

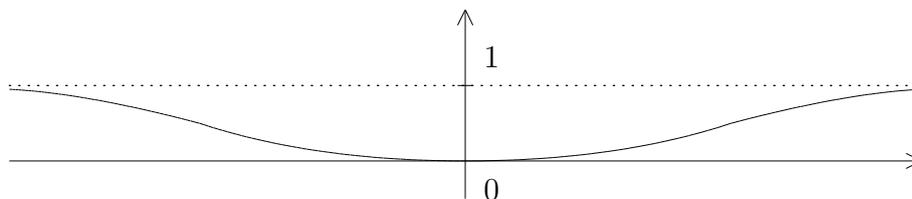
$$\sum_{n=0}^{\infty} (-1)^n x^{2n} = \sum_{n=0}^{\infty} (-x^2)^n = \frac{1}{1 + x^2}.$$

Die Grenzfunktion ist auf ganz \mathbb{R} definiert und beliebig oft differenzierbar, aber die Taylorreihe konvergiert nur auf $(-1, +1)$.

3. Noch verrückter verhält sich die Funktion

$$f(x) := \begin{cases} \exp\left(-\frac{1}{x^2}\right) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

Offensichtlich ist f stetig und für $x \neq 0$ beliebig oft differenzierbar. Außerdem ist $x = 0$ die einzige Nullstelle von f .



Behauptung: Es gibt rationale Funktionen $q_k(x)$, so daß

$$f^{(k)}(x) = q_k(x) \cdot e^{-1/x^2}$$

ist, für $x \neq 0$ und $k \in \mathbb{N}_0$.

BEWEIS dafür durch vollständige Induktion.

Offensichtlich können wir $q_0(x) := 1$ setzen. Ist nun $k \geq 1$ und die Behauptung für $k - 1$ schon bewiesen, so folgt:

$$\begin{aligned} f^{(k)}(x) &= f^{(k-1)'}(x) \\ &= q'_{k-1}(x) \cdot e^{-1/x^2} + q_{k-1}(x) \cdot 2x^{-3} \cdot e^{-1/x^2} \\ &= (q'_{k-1}(x) + 2q_{k-1}(x)x^{-3}) \cdot e^{-1/x^2}. \end{aligned}$$

Also hat $f^{(k)}$ die gewünschte Gestalt. □

Da die Exponentialfunktion stärker als jedes Polynom wächst, folgt:

$$\lim_{x \rightarrow 0} f^{(k)}(x) = \lim_{x \rightarrow 0} \frac{q_k(x)}{e^{1/x^2}} = \lim_{x \rightarrow \infty} \frac{q_k(1/x)}{e^{x^2}} = 0.$$

Also ist f auch im Nullpunkt beliebig oft differenzierbar, und $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}_0$.

Damit folgt, daß $Tf(x; 0) \equiv 0$ ist. Da die Funktion außerhalb des Nullpunktes stets $\neq 0$ ist, konvergiert die Taylorreihe nur im Entwicklungspunkt selbst gegen die Funktion.

Bevor wir weitere Beispiele betrachten, wollen wir die Formel für den Konvergenzradius noch etwas verallgemeinern:

Konvergenzradius für Potenzreihen mit Lücken

In der Potenzreihe $f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n$ sei $c_{2k} = 0$ und $c_{2k+1} \neq 0$ für fast alle k , und es existiere

$$c := \lim_{k \rightarrow \infty} \left| \frac{c_{2k+1}}{c_{2k+3}} \right|.$$

Dann ist $R := \sqrt{c}$ der Konvergenzradius.

Wenn $c_{2k+1} = 0$ und $c_{2k} \neq 0$ für fast alle k ist und der Grenzwert

$$c := \lim_{k \rightarrow \infty} \left| \frac{c_{2k}}{c_{2k+2}} \right|$$

existiert, so ist ebenfalls $R := \sqrt{c}$ der Konvergenzradius.

Der BEWEIS geht ähnlich wie bei der früher bewiesenen Formel. Wir betrachten nur den Fall $c_{2k+1} = 0$:

$$\left| \frac{c_{2k+2}z^{2k+2}}{c_{2k}z^{2k}} \right| = \left| \frac{c_{2k+2}}{c_{2k}} \right| \cdot |z|^2$$

konvergiert gegen $\frac{1}{c}|z|^2$, und dieser Ausdruck muß < 1 sein, damit die Reihe (nach Quotientenkriterium) konvergiert. Also muß $|z| < \sqrt{c}$ sein. \square

Beispiele :

1. Sei $f(x) := e^x$. Dann hat die Taylorreihe von f im Nullpunkt die Gestalt

$$E(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Von dieser Reihe wissen wir schon, daß sie den Konvergenzradius $R = +\infty$ hat. Also ist $E(x)$ eine auf ganz \mathbb{R} definierte unendlich oft differenzierbare Funktion. Wir wollen nun zeigen, daß $E(x) = e^x$ für alle $x \in \mathbb{R}$ ist.

(a) Es ist $E(0) = 1$.

(b) Nach dem Satz über die gliedweise Differenzierbarkeit von Reihen ist

$$E'(x) = \sum_{n=1}^{\infty} n \cdot \frac{x^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} = E(x).$$

Also erfüllt E auf ganz \mathbb{R} die Differentialgleichung $y' - y = 0$, mit der Anfangsbedingung $E(0) = 1$. Daraus folgt sofort $E(x) = e^x$.

Wir haben aber noch mehr herausbekommen:

$$E(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

ist eine auf ganz \mathbb{C} stetige Funktion (die *komplexe Exponentialfunktion*), deren Einschränkung auf die reelle Achse mit der reellen Exponentialfunktion e^x übereinstimmt. Wir schreiben wieder $\exp(z)$ statt $E(z)$.

2. Als nächstes wollen wir die Taylorreihen von $\sin(x)$ und $\cos(x)$ betrachten:

$$S(x) := \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} \quad \text{für den Sinus}$$

und $C(x) := \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} \quad \text{für den Cosinus.}$

Beide Reihen konvergieren auf ganz \mathbb{R} (sogar auf ganz \mathbb{C}), wie man aus der Formel für den Konvergenzradius von Reihen mit Lücken entnimmt. Weiter gilt:

(a) $S(0) = 0$ und $C(0) = 1$.

(b) $S'(x) = C(x)$ und $C'(x) = -S(x)$.

Daraus folgt: S ist Lösung der Differentialgleichung $y'' + y = 0$, mit den Anfangsbedingungen $S(0) = 0$ und $S'(0) = 1$.

Das zugehörige charakteristische Polynom $x^2 + 1$ hat die Nullstellen $x = \pm \mathbf{j}$. Also hat die allgemeine Lösung die Gestalt

$$y(x) = c_1 \cdot \cos(x) + c_2 \cdot \sin(x).$$

Setzt man die Anfangsbedingungen ein, so erhält man $c_1 = 0$ und $c_2 = 1$, also $S(x) = \sin(x)$ und $C(x) = \cos(x)$.

Auch hier sind die Funktionen S und C sogar auf ganz \mathbb{C} definiert und stetig. Diese Erweiterungen der reellen Winkelfunktionen werden wieder mit $\sin(z)$ und $\cos(z)$ bezeichnet.

Der Zusammenhang zwischen Exponentialfunktion und Winkelfunktionen wird nun deutlicher: es ist

$$\exp(\mathbf{j}z) = \sum_{n=0}^{\infty} \frac{(\mathbf{j}z)^n}{n!} = \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k}}{(2k)!} + \mathbf{j} \cdot \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k+1}}{(2k+1)!} = \cos(z) + \mathbf{j} \sin(z).$$

Für reelles z ist das die schon bekannte Eulersche Formel, aber die linke Seite der Gleichung hat nun a priori eine konkrete Bedeutung. Aus der Definition ist ein Satz geworden.

3. Ist $|x| < 1$, so ist $\sum_{n=0}^{\infty} (-x)^n = \frac{1}{1+x}$. Da man Potenzreihen gliedweise integrieren kann, folgt:

$$\begin{aligned} \ln(1+x) &= \ln(1+x) - \ln(1+0) \\ &= \int_0^x \frac{1}{1+t} dt = \sum_{n=0}^{\infty} \int_0^x (-t)^n dt \\ &= \sum_{n=0}^{\infty} (-1)^n \int_0^x t^n dt = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} t^{n+1} \Big|_0^x \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} x^{n+1}. \end{aligned}$$

Also ist

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n.$$

4. Aus der Darstellung $\sum_{n=0}^{\infty} (-x^2)^n = \frac{1}{1+x^2}$ folgt für $|x| < 1$:

$$\begin{aligned} \arctan(x) &= \int_0^x \frac{1}{1+t^2} dt \\ &= \sum_{n=0}^{\infty} (-1)^n \int_0^x t^{2n} dt \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}. \end{aligned}$$

Normalerweise kann man über das Konvergenzverhalten einer Potenzreihe auf dem Rand des Konvergenzkreises keine Aussage machen. Es gibt allerdings eine ganz kleine Ausnahme.

Abelscher Grenzwertsatz

Es sei $f(x) = \sum_{n=0}^{\infty} c_n x^n$ eine Potenzreihe mit reellen Koeffizienten und dem Konvergenzradius $R = 1$. Dann gilt:

$$\text{Ist } c := \sum_{n=0}^{\infty} c_n < \infty, \text{ so ist } \lim_{x \rightarrow 1} f(x) = c.$$

Der BEWEIS ist sicherlich nur für die Experten interessant:

Sei $s_{-1} := 0$ und $s_n := \sum_{i=0}^n c_i$, für $i \geq 0$. Dann kann man schreiben:

$$\begin{aligned} \sum_{n=0}^N c_n x^n &= \sum_{n=0}^N (s_n - s_{n-1}) x^n \\ &= \sum_{n=0}^N s_n x^n - \sum_{n=-1}^{N-1} s_n x^{n+1} \\ &= \sum_{n=0}^{N-1} s_n (1-x) x^n + s_N x^N - s_{-1}. \end{aligned}$$

Ist $|x| < 1$, so konvergiert die Potenzreihe in x absolut, und man kann beliebige Umformungen vornehmen. Daher folgt aus der gerade bewiesenen Gleichung für $N \rightarrow \infty$ die Beziehung

$$f(x) = (1-x) \cdot \sum_{n=0}^{\infty} s_n x^n.$$

Nun sei ein $\varepsilon > 0$ vorgegeben. Wir wählen ein n_0 , so daß für $n > n_0$ gilt:

$$|s_n - c| < \frac{\varepsilon}{2}.$$

Für $|x| < 1$ ist $(1-x) \cdot \sum_{n=0}^{\infty} x^n = 1$ und daher

$$\begin{aligned} |f(x) - c| &= \left| (1-x) \cdot \sum_{n=0}^{\infty} (s_n - c) x^n \right| \\ &= \left| (1-x) \cdot \sum_{n=0}^{n_0} (s_n - c) x^n + (1-x) \cdot \sum_{n=n_0+1}^{\infty} (s_n - c) x^n \right| \\ &\leq (1-x) \cdot \sum_{n=0}^{n_0} |s_n - c| \cdot |x|^n + \frac{\varepsilon}{2} < \varepsilon, \end{aligned}$$

wenn nur der erste Summand $< \frac{\varepsilon}{2}$ ist. Aber das ist sicher der Fall, wenn $1 > x > 1 - \delta$ ist, für genügend kleines $\delta > 0$. □

Beispiele :

1. Die Potenzreihe $\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n$ hat den Konvergenzradius 1, und die alternierende harmonische Reihe konvergiert. Also gilt:

$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n} = \lim_{x \rightarrow 1} \ln(1+x) = \ln(2).$$

Leider konvergiert die Reihe so schlecht, daß sie nicht zur Berechnung von $\ln(2)$ benutzt werden kann.

2. Die Reihe $\arctan(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$ hat auch den Konvergenzradius 1.

Die Reihe $\sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1}$ erfüllt das Leibniz-Kriterium und ist daher konvergent.

Also folgt:

$$\sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1} = \lim_{x \rightarrow 1} \arctan(x) = \frac{\pi}{4}.$$

Zum Schluß wollen wir noch eine besonders interessante Taylorentwicklung betrachten:

Wir beginnen mit dem Polynom $p(x) := (1+x)^n$. Dann ist

$$p^{(k)}(x) = n \cdot (n-1) \cdot \dots \cdot (n-k+1) \cdot (1+x)^{n-k} \quad \text{für } k = 1, \dots, n.$$

Insbesondere ist

$$\begin{aligned} p(0) &= 1, \\ p^{(k)}(0) &= n(n-1) \cdot \dots \cdot (n-k+1) \quad \text{für } 1 \leq k \leq n, \\ \text{insbesondere } p^{(n)}(0) &= n! \\ \text{und } p^{(k)}(0) &= 0 \quad \text{für } k > n. \end{aligned}$$

Also ist

$$(1+x)^n = \sum_{k=0}^n \binom{n}{k} x^k = \sum_{k=0}^n \frac{p^{(k)}(0)}{k!} x^k$$

die (in diesem Falle endliche) Taylorreihe von $p(x)$.

Das wollen wir jetzt verallgemeinern und die Funktion

$$f(x) := (1+x)^\alpha \quad \text{für } |x| < 1 \text{ und } \alpha \in \mathbb{R}$$

untersuchen. Hier ist

$$\frac{f^{(k)}(0)}{k!} = \frac{\alpha(\alpha-1) \cdot \dots \cdot (\alpha-k+1)}{k!} =: \binom{\alpha}{k}.$$

Die verallgemeinerten Binomialkoeffizienten $\binom{\alpha}{k}$ werden durch die obige Gleichung definiert. Insbesondere setzt man $\binom{\alpha}{0} := 1$. Man beachte, daß diese Zahlen weder ganz noch positiv zu sein brauchen. Es gilt aber z.B. die bekannte Formel

$$\binom{\alpha-1}{n} + \binom{\alpha-1}{n-1} = \binom{\alpha}{n}.$$

Die Konvergenz der Binomialreihe

Die Binomialreihe $B(x) := \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$ konvergiert für $|x| < 1$ gegen $(1+x)^\alpha$.

BEWEIS: 1) Den Konvergenzradius der Reihe bestimmen wir mit der Quotientenregel:
Es ist $c_n = \binom{\alpha}{n}$ und

$$\begin{aligned} \left| \frac{c_n}{c_{n+1}} \right| &= \left| \frac{\alpha(\alpha-1) \cdots (\alpha-n+1) \cdot (n+1)!}{\alpha(\alpha-1) \cdots (\alpha-n) \cdot n!} \right| \\ &= \left| \frac{n+1}{\alpha-n} \right| \\ &= \left| \frac{1 + \frac{1}{n}}{1 - \frac{\alpha}{n}} \right|, \end{aligned}$$

und dieser Ausdruck konvergiert gegen 1.

2) $B(x)$ ist also eine differenzierbare Funktion auf $(-1, +1)$. Es gilt:

$$\begin{aligned} B'(x) &= \sum_{n=1}^{\infty} n \cdot \binom{\alpha}{n} x^{n-1} = \sum_{n=0}^{\infty} (n+1) \cdot \binom{\alpha}{n+1} x^n \\ &= \sum_{n=0}^{\infty} \frac{(n+1)\alpha(\alpha-1) \cdots (\alpha-n)}{(n+1)!} x^n \\ &= \alpha \cdot \sum_{n=0}^{\infty} \frac{(\alpha-1)((\alpha-1)-1) \cdots ((\alpha-1)-n+1)}{n!} x^n \\ &= \alpha \cdot \sum_{n=0}^{\infty} \binom{\alpha-1}{n} x^n, \end{aligned}$$

also

$$\begin{aligned} (1+x)B'(x) &= \alpha \cdot \left(\sum_{n=0}^{\infty} \binom{\alpha-1}{n} x^n + \sum_{n=0}^{\infty} \binom{\alpha-1}{n} x^{n+1} \right) \\ &= \alpha \cdot \left(1 + \sum_{n=1}^{\infty} \left(\binom{\alpha-1}{n} + \binom{\alpha-1}{n-1} \right) x^n \right) \\ &= \alpha \cdot \left(1 + \sum_{n=1}^{\infty} \binom{\alpha}{n} x^n \right) \\ &= \alpha \cdot B(x). \end{aligned}$$

Setzen wir $h(x) := \frac{B(x)}{(1+x)^\alpha}$, so folgt:

$$\begin{aligned} h'(x) &= \frac{(1+x)^\alpha B'(x) - B(x)\alpha(1+x)^{\alpha-1}}{(1+x)^{2\alpha}} \\ &= \frac{(1+x)^{\alpha-1}}{(1+x)^{2\alpha}} \cdot ((1+x)B'(x) - \alpha B(x)) \equiv 0. \end{aligned}$$

Demnach muß $h(x) \equiv c$ konstant sein, also $B(x) = c(1+x)^\alpha$. Dabei ist $c = B(0) = 1$.

Das zeigt, daß $B(x) = (1+x)^\alpha$ für $|x| < 1$ ist. \square

Beispiel:

Es ist

$$\begin{aligned} \sqrt{1+x} &= (1+x)^{1/2} \\ &= \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} x^n \\ &= 1 + \binom{\frac{1}{2}}{1} x + \binom{\frac{1}{2}}{2} x^2 + \dots \\ &= 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 \pm \dots \end{aligned}$$

Diese Reihenentwicklung leistet gute Dienste, wenn man etwa ein elliptisches Integral näherungsweise auswerten will:

$$I := \int_0^{\frac{\pi}{2}} \sqrt{1 - k^2 \sin^2 \varphi} \, d\varphi, \quad \text{mit } 0 < k < 1.$$

Solche Integrale kann man nicht elementar berechnen. Also entwickelt man den Integranden in eine Taylorreihe und nutzt aus, daß man hier Integration und Reihenbildung miteinander vertauschen kann.

$$\begin{aligned} I &= \int_0^{\frac{\pi}{2}} (1 + (-k^2 \sin^2 \varphi)^{1/2}) \, d\varphi \\ &= \int_0^{\frac{\pi}{2}} \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-k^2 \sin^2 \varphi)^n \, d\varphi \\ &= \sum_{n=0}^{\infty} (-1)^n k^{2n} \binom{\frac{1}{2}}{n} \int_0^{\frac{\pi}{2}} \sin^{2n} \varphi \, d\varphi. \end{aligned}$$

Die Integrale

$$I(n) := \int_0^{\frac{\pi}{2}} \sin^{2n} \varphi \, d\varphi$$

kann man rekursiv berechnen:

Es ist $I(0) = \frac{\pi}{2}$ und $I(n) = \frac{2n-1}{2n} \cdot I(n-1)$, wie man leicht durch partielle Integration herausbekommt.

Also ist

$$I(1) = \frac{\pi}{4}, \quad I(2) = \frac{3\pi}{16}, \quad I(3) = \frac{5\pi}{32}, \dots$$

Das ergibt als Wert für das elliptische Integral:¹

$$I = \pi \left(\frac{1}{2} - \frac{k^2}{8} - \frac{3k^4}{128} - \frac{5k^6}{512} - \dots \right).$$

¹Die Binomialkoeffizienten werden folgendermaßen berechnet:

$$\begin{aligned} \binom{\frac{1}{2}}{1} &= \frac{\frac{1}{2}}{1} = \frac{1}{2}, \\ \binom{\frac{1}{2}}{2} &= \frac{\frac{1}{2}(\frac{1}{2}-1)}{2!} = -\frac{1}{8} \\ \binom{\frac{1}{2}}{3} &= \frac{\frac{1}{2}(\frac{1}{2}-1)(\frac{1}{2}-2)}{3!} = \frac{1}{16} \\ &\dots \text{ usw.} \end{aligned}$$

§4 Ergänzungen zur Topologie

Definition:

Sei $M \subset \mathbb{R}^n$ eine beliebige Teilmenge.

Ein Punkt $\mathbf{x}_0 \in \mathbb{R}^n$ heißt *innerer Punkt* von M , falls es ein $\varepsilon > 0$ gibt, so daß noch die ganze Kugel

$$B_\varepsilon(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{x}_0\| < \varepsilon\}$$

in M enthalten ist.

Eine Menge $M \subset \mathbb{R}^n$ ist definitionsgemäß genau dann *offen* im \mathbb{R}^n , wenn es zu **jedem** $\mathbf{x} \in M$ ein $\varepsilon > 0$ gibt, so daß $B_\varepsilon(\mathbf{x})$ ganz in M liegt. Also kann man auch sagen: M ist genau dann offen im \mathbb{R}^n , wenn M nur aus inneren Punkten besteht.

Wir erinnern uns: Der ganze \mathbb{R}^n und die leere Menge sind offen, endliche Durchschnitte und beliebige Vereinigungen von offenen Mengen sind wieder offen.

Definition:

Ist $M \subset \mathbb{R}^n$ eine beliebige Teilmenge, so nennt man die Menge

$$\overset{\circ}{M} := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \text{ ist innerer Punkt von } M\}$$

den *offenen Kern* von M .

Beispiele:

1. Ist M offen (z.B. $M = B_\varepsilon(\mathbf{x}_0)$), so ist $\overset{\circ}{M} = M$.
2. Die Begriffe „innerer Punkt“, „offen“, „offener Kern“ beziehen sich immer auf den betrachteten \mathbb{R}^n . Liegt eine Menge M ganz in einem niederdimensionalen affinen Unterraum des \mathbb{R}^n , so ist $\overset{\circ}{M} = \emptyset$ (im \mathbb{R}^n). Das trifft z.B. auf die folgende Menge zu:

$$M := \{(x_1, x_2) \in \mathbb{R}^2 : |x_1| < 1 \text{ und } x_2 = 0\}.$$

Man könnte aber auch sagen, daß M in \mathbb{R} liegt (indem man die Komponente x_2 vergißt, die ja sowieso $= 0$ ist). Und in \mathbb{R} ist $M = \{x_1 : |x_1| < 1\}$ offen, also ist **dort** $\overset{\circ}{M} = M$.

3. Besitzt M nur endlich viele Elemente, so ist $\overset{\circ}{M} = \emptyset$.
4. Ist \mathbf{x}_0 ein Randpunkt der Menge M , so enthält jede Umgebung von \mathbf{x}_0 wenigstens einen Punkt aus M und einen Punkt aus dem Komplement zu M . Also kann \mathbf{x}_0 kein innerer Punkt von M sein. Liegt dagegen \mathbf{x}_0 in M und ist kein Randpunkt von

M , so muß \mathbf{x}_0 ein innerer Punkt von M sein. Also gilt:

$$\overset{\circ}{M} = M \setminus \partial M.$$

Eigenschaften des offenen Kerns

1. $\overset{\circ}{M}$ ist die größte offene Menge, die in M enthalten ist.
2. Ist $N \subset M$, so ist auch $\overset{\circ}{N} \subset \overset{\circ}{M}$.
3. Es ist $(\overset{\circ}{M})^\circ = \overset{\circ}{M}$.
4. M offen $\iff \overset{\circ}{M} = M$.

BEWEIS: Es handelt sich um einfache Übungsaufgaben! Man beweise zunächst, daß $\overset{\circ}{M}$ stets offen ist. Daraus folgt sofort Eigenschaft (3) und auch sehr schnell Eigenschaft (1), und daraus dann Eigenschaft (4). (2) ist trivial. \square

Definition:

Sei $M \subset \mathbb{R}^n$ eine beliebige Teilmenge.

Ein Punkt $\mathbf{x}_0 \in \mathbb{R}^n$ heißt *Berührungspunkt* von M , falls für jede Umgebung $U = U(\mathbf{x}_0) \subset \mathbb{R}^n$ gilt: $U \cap M \neq \emptyset$.

Bemerkung: Folgende Aussagen sind äquivalent:

1. \mathbf{x}_0 ist Berührungspunkt von M .
2. $\forall \varepsilon > 0 \exists \mathbf{x} \in M$ mit $\|\mathbf{x} - \mathbf{x}_0\| < \varepsilon$.
3. Es gibt eine Folge (\mathbf{x}_n) von Punkten aus M mit $\lim_{n \rightarrow \infty} \mathbf{x}_n = \mathbf{x}_0$.

Definition:

Sei $M \subset \mathbb{R}^n$ eine beliebige Teilmenge.

Ein Punkt $\mathbf{x}_0 \in \mathbb{R}^n$ heißt *isolierter Punkt* von M , falls es eine Umgebung $U = U(\mathbf{x}_0)$ mit $U \cap M = \{\mathbf{x}_0\}$ gibt.

Ein Punkt $\mathbf{x}_0 \in \mathbb{R}^n$ heißt *Häufungspunkt* von M , falls in jeder Umgebung von \mathbf{x}_0 unendlich viele Punkte von M liegen.

Bemerkung: Ein isolierter Punkt von M gehört selbst zu M . Ein Häufungspunkt von M braucht nicht zu M zu gehören.

Einteilung der Berührungspunkte

Ein Punkt \mathbf{x}_0 ist genau dann ein Berührungspunkt der Menge M , wenn er ein isolierter Punkt oder ein Häufungspunkt von M ist.

Offensichtlich ist jeder Punkt von M auch ein Berührungspunkt von M . Ein Randpunkt von M , der nicht zu M gehört, muß zumindest ein Häufungspunkt von M sein, also auch ein Berührungspunkt. Ein Punkt aber, der weder zu M noch zu ∂M gehört, besitzt eine Umgebung, die M nicht trifft, kann also kein Berührungspunkt sein. Daraus folgt:

Charakterisierung der Berührungspunkte

Sei $M \subset \mathbb{R}^n$ eine beliebige Teilmenge. Dann stimmt die Menge

$$\overline{M} := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \text{ ist Berührungspunkt von } M\}$$

mit der früher definierten abgeschlossenen Hülle von M überein:

$$\overline{M} = M \cup \partial M.$$

Eigenschaften der abgeschlossenen Hülle

1. \overline{M} ist die kleinste abgeschlossene Menge, die M umfaßt.
2. Ist $N \subset M$, so ist auch $\overline{N} \subset \overline{M}$.
3. Es ist $\overline{\overline{M}} = \overline{M}$.
4. M abgeschlossen $\iff \overline{M} = M$.

BEWEIS: Auch hier handelt es sich wieder um eine einfache Übungsaufgabe. Wir wissen schon aus Kapitel II, daß M genau dann abgeschlossen ist, wenn $\partial M \subset M$ ist, bzw. wenn $\overline{M} = M$ ist. Damit ist (4) schon erledigt und (3) folgt trivial. Außerdem ist damit klar, daß \overline{M} eine abgeschlossene Menge ist, die M umfaßt. Der Rest von (1) und die Aussage (2) werden bewiesen, indem man isolierte Punkte und Häufungspunkte getrennt behandelt. \square

Folgerung

$$\partial M = \overline{M} \setminus \overset{\circ}{M}.$$

Beispiele :

1. Sei $M := \{\frac{1}{n} \mid n \in \mathbb{N}\} \subset \mathbb{R}$. Dann ist jeder Punkt von M ein isolierter Punkt und 0 ein Häufungspunkt, also $\overline{M} = M \cup \{0\}$.

2. Sei $M := \mathbb{Q} \cap [0, 1] \subset \mathbb{R}$. Dann besitzt M keinen isolierten Punkt, aber jeder Punkt von $[0, 1]$ ist Häufungspunkt von M . Also ist $\overline{M} = [0, 1]$.
3. Sei $M = B_r(\mathbf{0}) \setminus \{\mathbf{0}\}$. Dann ist $\mathbf{0}$ ein Häufungspunkt von M und ein isolierter Punkt von $\mathbb{R}^n \setminus M$. Es ist $\overline{M} = \overline{B_r(\mathbf{0})}$.

Wir betrachten jetzt ein abgeschlossenes Intervall $I = [a, b] \subset \mathbb{R}$ und eine Folge (x_n) von Zahlen aus I . Dann gilt:

1. (x_n) besitzt – als beschränkte Folge – einen Häufungspunkt x_0 . Das folgt aus dem Satz von Bolzano-Weierstraß.
2. Man kann nun eine Teilfolge $(x_{n(k)})$ auswählen, die gegen x_0 konvergiert.
3. Da die $x_{n(k)}$ in I liegen, ist x_0 auch ein Häufungspunkt der Menge I . Und da I abgeschlossen ist, x_0 schon selbst zu I .

Insgesamt haben wir herausbekommen:

Jede Folge in I besitzt eine konvergente Teilfolge, deren Grenzwert in I liegt.

Beim Beweis haben wir nur zwei Eigenschaften von I benutzt, nämlich daß I beschränkt und abgeschlossen ist. Wir werden sehen, daß die beschränkten und abgeschlossenen Teilmengen von \mathbb{R}^n genau diejenigen sind, bei denen jede Folge in der Menge auch eine in dieser Menge konvergente Teilfolge besitzt. Allerdings wollen wir das gleich im \mathbb{R}^n zeigen.

Definition:

Eine Menge $K \subset \mathbb{R}^n$ heißt *kompakt*, wenn jede Folge von Punkten aus K eine Teilfolge besitzt, die gegen einen Grenzwert in K konvergiert.

Bevor wir beweisen, daß die kompakten Mengen im \mathbb{R}^n genau die abgeschlossenen und beschränkten Mengen sind, wollen wir noch die Grundlagen für ein weiteres Kompaktheitskriterium besprechen.

Ein (endliches oder unendliches) System $(U_\iota)_{\iota \in I}$ von Teilmengen des \mathbb{R}^n heißt eine *offene Überdeckung* einer Menge $M \subset \mathbb{R}^n$, falls gilt:

1. Alle Mengen U_ι sind offen.
2. $M \subset \bigcup_{\iota \in I} U_\iota$.

Ist $J \subset I$ eine Teilmenge der Indexmenge, so nennt man das System $(U_\iota)_{\iota \in J}$ ein *Teilsystem*. Liegt M auch noch in der Vereinigung aller U_ι , $\iota \in J$, so spricht man von einer *Teilüberdeckung*. Ist die Menge J außerdem endlich, so spricht man von einer *endlichen Teilüberdeckung*. Im allgemeinen wird man jedoch gar keine Elemente der Überdeckung weglassen können, und schon gar nicht so viele, daß eine endliche Überdeckung übrigbleibt.

Beispiele :

1. Sei $M := \{n \mid n \in \mathbb{N}\} \subset \mathbb{R}$. Die Intervalle $U_n := (n - \frac{1}{2}, n + \frac{1}{2})$ bilden eine offene Überdeckung von M . Die Indexmenge I ist in diesem Falle die Menge \mathbb{N} der natürlichen Zahlen, also unendlich. Offensichtlich kann man keine der Mengen U_n weglassen.
2. Sei $M := \{0\} \cup \{\frac{1}{n} \mid n \in \mathbb{N}\}$, und $(U_\iota)_{\iota \in I}$ irgendeine offene Überdeckung von M . Dann gibt es ein $\iota_0 \in I$, so daß $0 \in U_{\iota_0}$ ist. Aber weil $\frac{1}{n}$ gegen 0 konvergiert, gibt es ein n_0 , so daß auch $\frac{1}{n} \in U_{\iota_0}$ ist, für $n > n_0$. Zu jeder der Zahlen $n = 1, 2, 3, \dots, n_0$ gibt es aber jeweils ein $\iota_n \in I$ mit $\frac{1}{n} \in U_{\iota_n}$. Also bilden die Mengen $U_{\iota_0}, U_{\iota_1}, \dots, U_{\iota_{n_0}}$ eine endliche Teilüberdeckung.

Wir benötigen auch noch einen Hilfssatz über offene Mengen in kartesischen Produkten.

Hilfssatz

Ein Punkt $(\mathbf{x}_0, \mathbf{y}_0)$ in einer Teilmenge $U \subset \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$ ist genau dann ein innerer Punkt, wenn es Umgebungen $V(\mathbf{x}_0) \subset \mathbb{R}^n$ und $W(\mathbf{y}_0) \subset \mathbb{R}^m$ mit $V \times W \subset U$ gibt.

BEWEIS: 1) Sei $(\mathbf{x}_0, \mathbf{y}_0)$ innerer Punkt von $U \subset \mathbb{R}^n \times \mathbb{R}^m$. Dann gibt es ein $\varepsilon > 0$, so daß die Kugel

$$B_\varepsilon(\mathbf{x}_0, \mathbf{y}_0) = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{n+m} : \|\mathbf{x} - \mathbf{x}_0\|^2 + \|\mathbf{y} - \mathbf{y}_0\|^2 < \varepsilon^2\}$$

ganz in U enthalten ist. Wir setzen dann $\delta := \frac{1}{\sqrt{2}} \cdot \varepsilon$.

$V := \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}_0\| < \delta\}$ und $W := \{\mathbf{y} : \|\mathbf{y} - \mathbf{y}_0\| < \delta\}$ sind Umgebungen von \mathbf{x}_0 im \mathbb{R}^n bzw. \mathbf{y}_0 im \mathbb{R}^m , und für $(\mathbf{x}, \mathbf{y}) \in V \times W$ ist

$$\|\mathbf{x} - \mathbf{x}_0\|^2 + \|\mathbf{y} - \mathbf{y}_0\|^2 < 2\delta^2 = \varepsilon^2.$$

Also liegt $V \times W$ in $B_\varepsilon(\mathbf{x}_0, \mathbf{y}_0)$ und damit in U .

2) Gibt es umgekehrt Umgebungen V bzw. W von \mathbf{x}_0 im \mathbb{R}^n bzw. von \mathbf{y}_0 im \mathbb{R}^m mit $V \times W \subset U$, so kann man o.B.d.A. annehmen, daß $V = B_\varepsilon(\mathbf{x}_0)$ und $W = B_\varepsilon(\mathbf{y}_0)$ ist, für ein geeignetes $\varepsilon > 0$. Aber dann gilt für $(\mathbf{x}, \mathbf{y}) \in B_\varepsilon(\mathbf{x}_0, \mathbf{y}_0) \subset \mathbb{R}^{n+m}$:

$$\|\mathbf{x} - \mathbf{x}_0\|^2 \leq \|\mathbf{x} - \mathbf{x}_0\|^2 + \|\mathbf{y} - \mathbf{y}_0\|^2 < \varepsilon^2,$$

und analog $\|\mathbf{y} - \mathbf{y}_0\|^2 < \varepsilon^2$, also $\mathbf{x} \in V$ und $\mathbf{y} \in W$. Das bedeutet, daß (\mathbf{x}, \mathbf{y}) in $V \times W$ und damit in U liegt. Also gehört ganz $B_\varepsilon(\mathbf{x}_0, \mathbf{y}_0)$ zu U , und $(\mathbf{x}_0, \mathbf{y}_0)$ ist innerer Punkt von U . □

Der Satz von Heine-Borel

Folgende Aussagen über eine Teilmenge $K \subset \mathbb{R}^n$ sind äquivalent:

1. Jede offene Überdeckung von K enthält eine endliche Teilüberdeckung.
2. K ist kompakt.
3. K ist abgeschlossen und beschränkt.

BEWEIS:

(1) \implies (2): Sei (\mathbf{x}_n) eine Folge von Punkten aus K ,

$$X := \{\mathbf{x}_n : n \in \mathbb{N}\}.$$

Ist X eine endliche Menge, so enthält (\mathbf{x}_n) eine konstante Teilfolge, die natürlich gegen ein Element von K konvergiert. Wir können also annehmen, daß X unendlich ist.

Behauptung: Die Menge X hat einen Häufungspunkt in K .

BEWEIS dafür: Wir nehmen an, die Behauptung sei falsch. Dann besitzt jeder Punkt $\mathbf{x} \in K$ eine offene Umgebung $U_x \subset \mathbb{R}^n$, in der nur endlich viele Elemente von X liegen. Das System der Umgebungen U_x bildet eine offene Überdeckung von K , und nach Voraussetzung kommt man dann sogar mit endlich vielen solcher Umgebungen aus. Aber dann muß X endlich sein. Das ist ein Widerspruch!

Sei $\mathbf{x}_0 \in K$ ein Häufungspunkt von X . Dann kann man zu jedem $k \in \mathbb{N}$ einen Punkt $\mathbf{x}_{n_k} \in X \cap B_{\frac{1}{k}}(\mathbf{x}_0)$ finden. Das liefert die gewünschte (gegen \mathbf{x}_0) konvergente Teilfolge.

(2) \implies (3): Sei K kompakt. Wäre K nicht beschränkt, so gäbe es zu jedem $n \in \mathbb{N}$ ein $\mathbf{x}_n \in K$ mit $\|\mathbf{x}_n\| > n$. Aber die Folge der \mathbf{x}_n könnte dann keine konvergente Teilfolge besitzen, und das ist nicht zugelassen.

Um zu zeigen, daß K auch abgeschlossen ist, betrachten wir einen beliebigen Berührungspunkt \mathbf{x}_0 von K . Definitionsgemäß gibt es zu jedem $n \in \mathbb{N}$ einen Punkt $\mathbf{x}_n \in K \cap B_{\frac{1}{n}}(\mathbf{x}_0)$. Nach Voraussetzung gibt es eine in K konvergente Teilfolge $(\mathbf{x}_{n(k)})$. Als Grenzwert dieser Folge kommt aber nur \mathbf{x}_0 in Frage. Also liegt \mathbf{x}_0 in K , und wir haben gezeigt, daß $\overline{K} = K$ ist.

(3) \implies (1): Das ist der schwierige Teil des Beweises. Wir führen ihn in mehreren Schritten. Dabei benutzen wir die Bezeichnung „ü-kompakt“ für die Überdeckungseigenschaft (1).

1. Schritt: Jedes abgeschlossene Intervall $[a, b]$ ist ü-kompakt in \mathbb{R} .

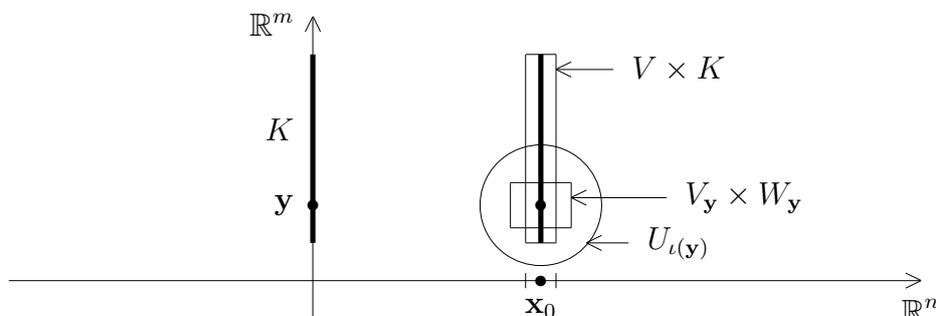
BEWEIS dafür: Sei $(U_\iota)_{\iota \in I}$ eine offene Überdeckung von $[a, b]$ in \mathbb{R} . Dann liegt speziell a in einem U_{ι_1} , und weil U_{ι_1} offen ist, gibt es sogar ein t mit $a < t < b$ und $[a, t] \subset U_{\iota_1}$.

Nun sei $M := \{t \in [a, b] : [a, t] \text{ kann durch endlich viele } U_\iota \text{ überdeckt werden}\}$. Die Menge M wird durch b nach oben beschränkt, daher existiert $t_0 := \sup M$, und es ist $a < t_0 \leq b$. Wir wollen zeigen, daß $t_0 = b$ ist und machen die Annahme, es sei $t_0 < b$.

Es gibt ein $\iota_0 \in I$, so daß t_0 in U_{ι_0} liegt. Wir können dann auch noch ein t^* mit $a < t^* < t_0$ finden, so daß $[t^*, t_0] \subset U_{\iota_0}$ ist. Aber dann muß t^* in M liegen, und nach Konstruktion gibt es Indizes $\iota_1, \dots, \iota_N \in I$, so daß $[a, t^*]$ durch $U_{\iota_1}, \dots, U_{\iota_N}$ überdeckt wird. Weil wir außerdem angenommen haben, daß $t_0 < b$ ist, gibt es auch ein t^{**} mit $t_0 < t^{**} \leq b$, so daß $[t_0, t^{**}] \subset U_{\iota_0}$ ist. Damit wird $[a, t^{**}]$ durch $U_{\iota_0}, U_{\iota_1}, \dots, U_{\iota_N}$ überdeckt, und t^{**} gehört zu M , im Widerspruch zur Supremumseigenschaft von t_0 . Die Annahme war falsch, es ist tatsächlich $t_0 = b$, und das ganze Intervall kann durch endlich viele U_{ι} überdeckt werden.

2. Schritt: Sei $\mathbf{x}_0 \in \mathbb{R}^n$ und $K \subset \mathbb{R}^m$ ü-kompakt. Dann gibt es zu jeder offenen Überdeckung $(U_{\iota})_{\iota \in I}$ von $\{\mathbf{x}_0\} \times K$ in \mathbb{R}^{n+m} eine endliche Teilmenge $J \subset I$ und eine offene Umgebung V von \mathbf{x}_0 im \mathbb{R}^n , so daß $(U_{\iota})_{\iota \in J}$ schon eine Überdeckung von $V \times K$ im \mathbb{R}^{n+m} ist.

Der BEWEIS dafür ist leicht: Zu jedem Punkt $\mathbf{y} \in K$ gibt es ein $\iota(\mathbf{y}) \in I$ mit $(\mathbf{x}_0, \mathbf{y}) \in U_{\iota(\mathbf{y})}$, und dann gibt es offene Umgebungen $V_{\mathbf{y}}$ von \mathbf{x}_0 im \mathbb{R}^n und $W_{\mathbf{y}}$ von \mathbf{y} im \mathbb{R}^m mit $V_{\mathbf{y}} \times W_{\mathbf{y}} \subset U_{\iota(\mathbf{y})}$.



Weil K als ü-kompakt vorausgesetzt wird, überdecken schon endlich viele Umgebungen $W_{\mathbf{y}_1}, \dots, W_{\mathbf{y}_N}$ die Menge K . Setzt man $V := V_{\mathbf{y}_1} \cap \dots \cap V_{\mathbf{y}_N}$, so wird $V \times K$ von den Mengen $U_{\iota(\mathbf{y}_\nu)}$ überdeckt, $\nu = 1, \dots, N$.

3. Schritt: Sind $A \subset \mathbb{R}^n$ und $B \subset \mathbb{R}^m$ ü-kompakt, so ist auch $A \times B \subset \mathbb{R}^{n+m}$ ü-kompakt.

BEWEIS dafür: Sei $(U_{\iota})_{\iota \in I}$ eine offene Überdeckung von $A \times B$. Zu jedem $\mathbf{x} \in A$ gibt es eine offene Umgebung $V_{\mathbf{x}}$, so daß $V_{\mathbf{x}} \times B$ schon von endlich vielen U_{ι} überdeckt wird. Da aber auch schon endlich viele $V_{\mathbf{x}}$ die Menge A überdecken, kommt man insgesamt bei $A \times B$ mit endlich vielen U_{ι} aus.

4. Schritt: Jeder „Quader“ $[a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n] \subset \mathbb{R}^n$ ist ü-kompakt.

Der BEWEIS ist eine einfache Induktion über n .

5. Schritt: Ist $K \subset \mathbb{R}^n$ abgeschlossen und beschränkt, so ist K ü-kompakt.

BEWEIS: Da K beschränkt ist, gibt es einen Quader K_0 mit $K \subset K_0$. Sei nun $(U_{\iota})_{\iota \in I}$ eine offene Überdeckung von K . Dann bilden die U_{ι} zusammen mit der offenen Menge $U^* := \mathbb{R}^n \setminus K$ eine offene Überdeckung von K_0 . Wir wissen aber schon, daß K_0 ü-kompakt ist. Also bilden endlich viele der U_{ι} , eventuell zusammen mit U^* , eine Überdeckung von K_0 . Das geht nur, wenn die endlich vielen U_{ι} bereits K überdecken. Das war's! \square

Es folgen nun einige Sätze über kompakte Mengen.

Abgeschlossene Teilmengen kompakter Mengen

Sei $K \subset \mathbb{R}^n$ kompakt und $A \subset K$ abgeschlossen. Dann ist auch A kompakt.

BEWEIS: Mit K ist auch A beschränkt. Weil A außerdem abgeschlossen ist, ist A sogar kompakt. \square

Das stetige Bild kompakter Mengen

Sei $K \subset \mathbb{R}^n$ kompakt und $f : K \rightarrow \mathbb{R}^m$ stetig. Dann ist auch die Bildmenge $f(K) \subset \mathbb{R}^m$ kompakt.

BEWEIS: Sei (\mathbf{y}_n) eine Folge in $f(K)$. Dann gibt es Punkte $\mathbf{x}_n \in K$ mit $f(\mathbf{x}_n) = \mathbf{y}_n$. Weil K kompakt ist, besitzt die Folge (\mathbf{x}_n) eine gegen ein $\mathbf{x}_0 \in K$ konvergente Teilfolge $(\mathbf{x}_{n(k)})$. Wegen der Stetigkeit von f konvergiert dann $(\mathbf{y}_{n(k)})$ gegen $f(\mathbf{x}_0) \in f(K)$. Damit ist alles gezeigt. \square

Folgerung

Ist $K \subset \mathbb{R}^n$ kompakt und $f : K \rightarrow \mathbb{R}$ stetig, so nimmt f auf K Minimum und Maximum an.

BEWEIS: $f(K) \subset \mathbb{R}$ ist kompakt, also abgeschlossen und beschränkt. Damit existieren die Zahlen $y_- := \inf f(K)$ und $y_+ := \sup f(K)$, und sie müssen beide in $f(K)$ liegen. Dann gibt es aber Punkte $\mathbf{x}_- \in K$ und $\mathbf{x}_+ \in K$ mit $f(\mathbf{x}_-) = y_-$ und $f(\mathbf{x}_+) = y_+$. \square

Definition:

Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge. Man sagt, eine offene Teilmenge $V \subset \mathbb{R}^n$ liegt relativ-kompakt in U (und schreibt dafür: $V \subset\subset U$), falls gilt:

\bar{V} ist kompakt und in U enthalten.

Sehr nützlich ist oft die folgende Feststellung:

Existenz relativ-kompakter Umgebungen

Sei $U \subset \mathbb{R}^n$ offen und $K \subset U$ kompakt. Dann gibt es eine offene Menge $V \subset \mathbb{R}^n$ mit

$$K \subset V \subset\subset U.$$

BEWEIS: Zu jedem Punkt $\mathbf{x} \in K$ gibt es ein $\varepsilon(\mathbf{x}) > 0$, so daß $B_{\varepsilon(\mathbf{x})}(\mathbf{x}) \subset U$ ist. Nun sei $\delta(\mathbf{x}) := \frac{1}{2} \cdot \varepsilon(\mathbf{x})$ und $V_{\mathbf{x}} := B_{\delta(\mathbf{x})}(\mathbf{x})$. Dann liegt sogar $\bar{V}_{\mathbf{x}}$ in U .

Die offenen Mengen $V_{\mathbf{x}}$ überdecken K . Aber da K kompakt ist, kommt man mit endlich vielen solcher Mengen aus, die wir mit V_1, \dots, V_N bezeichnen wollen. Schließlich setzen wir $V := V_1 \cup \dots \cup V_N$. Dann gilt:

1. V ist offen, und K ist in V enthalten.
2. $\bar{V} = \bar{V}_1 \cup \dots \cup \bar{V}_N$ liegt noch in U .
3. Jede abgeschlossene Kugel ist eine abgeschlossene und beschränkte Menge, und eine endliche Vereinigung von abgeschlossenen und beschränkten Mengen ist wieder abgeschlossen und beschränkt. Also ist \bar{V} kompakt und damit $V \subset \subset U$.

□

Im Beweis wurde die folgende Aussage benutzt:

Hilfssatz

Sind $M_1, M_2 \subset \mathbb{R}^n$ beliebige Mengen, so ist

$$\overline{M_1 \cup M_2} = \overline{M_1} \cup \overline{M_2}.$$

BEWEIS: a) Mit $M_i \subset M_1 \cup M_2$ ist auch $\overline{M_i} \subset \overline{M_1 \cup M_2}$, für $i = 1, 2$, und damit auch $\overline{M_1} \cup \overline{M_2} \subset \overline{M_1 \cup M_2}$.

b) Da $M_i \subset \overline{M_i}$ ist, für $i = 1, 2$, ist $M_1 \cup M_2$ in der abgeschlossenen Menge $\overline{M_1} \cup \overline{M_2}$ enthalten. Dann muß aber auch die abgeschlossene Hülle $\overline{M_1 \cup M_2}$ in $\overline{M_1} \cup \overline{M_2}$ liegen. □

Definition:

Eine offene Menge $U \subset \mathbb{R}^n$ heißt *zusammenhängend* oder *ein Gebiet*, wenn gilt:

Sind $\mathbf{x}, \mathbf{y} \in U$ zwei beliebige Punkte, so gibt es eine stetige Kurve $\alpha : [0, 1] \rightarrow U$ mit $\alpha(0) = \mathbf{x}$ und $\alpha(1) = \mathbf{y}$.

Ein Gebiet kann nicht in zwei Teilgebiete zerlegt werden:

Unzerlegbarkeit von Gebieten

Sei $G \subset \mathbb{R}^n$ ein Gebiet und $U \subset G$ eine offene, nicht leere Teilmenge. Ist auch $G \setminus U$ offen, so muß schon $U = G$ sein.

BEWEIS: Sei $\mathbf{x}_0 \in U$ ein beliebiger Punkt. Wir wollen unter den gegebenen Voraussetzungen zeigen, daß $G \setminus U$ leer ist. Dazu benutzen wir das Widerspruchsprinzip und nehmen an, es gibt einen Punkt $\mathbf{y}_0 \in G \setminus U$. Weil G ein Gebiet ist, gibt es einen stetigen Weg $\alpha : [0, 1] \rightarrow G$, der \mathbf{x}_0 mit \mathbf{y}_0 verbindet. Und weil U und $G \setminus U$ beide offen sind, gibt es ein $\varepsilon > 0$, so daß $B_\varepsilon(\mathbf{x}_0)$ in U und $B_\varepsilon(\mathbf{y}_0)$ in $G \setminus U$ enthalten ist. Wegen der Stetigkeit von α gibt es dann ein $\delta > 0$, so daß $\alpha(t) \in B_\varepsilon(\mathbf{x}_0) \subset U$ für $|t| < \delta$ und $\alpha(t) \in B_\varepsilon(\mathbf{y}_0) \subset G \setminus U$ für $|t - 1| < \delta$ ist.

Nun sei $t_0 := \sup\{t \in [0, 1] : \alpha(t) \in U\}$. Offensichtlich ist

$$0 < \delta \leq t_0 \leq 1 - \delta < 1.$$

$\mathbf{z}_0 := \alpha(t_0)$ liegt in G . Wir unterscheiden zwei Fälle:

1. Fall: $\mathbf{z}_0 \in U$. Da U offen und α stetig ist, gibt es ein $r > 0$, so daß auch noch $\alpha(t) \in U$ für $|t - t_0| < r$ ist, also insbesondere für $t_0 \leq t < t_0 + r$. Aber das ist nicht möglich.

2. Fall: $\mathbf{z}_0 \in G \setminus U$. Mit dem gleichen Argument wie oben erhält man ein $r > 0$, so daß $\alpha(t) \in G \setminus U$ für $t_0 - r < t \leq t_0$ ist, und auch das ist unmöglich.

Also haben wir einen Widerspruch erzielt, die Annahme ist falsch und $U = G$. \square

Sind $\mathbf{x}_0, \mathbf{y}_0$ zwei Punkte im \mathbb{R}^n , so nennt man $\alpha : [0, 1] \rightarrow \mathbb{R}^n$ mit

$$\alpha(t) = \mathbf{x}_0 + t \cdot (\mathbf{y}_0 - \mathbf{x}_0) = (1 - t) \cdot \mathbf{x}_0 + t \cdot \mathbf{y}_0$$

die (*Verbindungs-*)*Strecke* von \mathbf{x}_0 und \mathbf{y}_0 . Manchmal versteht man unter der Verbindungsstrecke allerdings auch die Bildmenge

$$\alpha([0, 1]) = \{\mathbf{x} = (1 - t) \cdot \mathbf{x}_0 + t \cdot \mathbf{y}_0 \mid 0 \leq t \leq 1\}.$$

Aus dem Zusammenhang muß man entnehmen, was gerade gemeint ist.

Definition:

Eine Teilmenge $M \subset \mathbb{R}^n$ heißt *konvex*, falls mit je zwei Punkten \mathbf{x} und \mathbf{y} auch stets deren Verbindungsstrecke in M liegt.

Beispiele :

1. Jedes Intervall ist eine konvexe Teilmenge von \mathbb{R} .
2. Offene und abgeschlossene Kugeln im \mathbb{R}^n sind konvex:

Seien etwa $\mathbf{x}_0, \mathbf{y}_0 \in \overline{B_r(\mathbf{z}_0)}$, also

$$\|\mathbf{x}_0 - \mathbf{z}_0\| \leq r \text{ und } \|\mathbf{y}_0 - \mathbf{z}_0\| \leq r.$$

Dann hat jeder Punkt auf der Verbindungsstrecke die Gestalt $\mathbf{x} = (1 - t) \cdot \mathbf{x}_0 + t \cdot \mathbf{y}_0$, mit $0 \leq t \leq 1$. Da auch $\mathbf{z}_0 = (1 - t) \cdot \mathbf{z}_0 + t \cdot \mathbf{z}_0$ ist, folgt:

$$\begin{aligned} \|\mathbf{x} - \mathbf{z}_0\| &= \|(1 - t) \cdot (\mathbf{x}_0 - \mathbf{z}_0) + t \cdot (\mathbf{y}_0 - \mathbf{z}_0)\| \\ &\leq (1 - t) \cdot \|\mathbf{x}_0 - \mathbf{z}_0\| + t \cdot \|\mathbf{y}_0 - \mathbf{z}_0\| \\ &\leq (1 - t)r + tr = r. \end{aligned}$$

3. Jede konvexe offene Menge ist ein Gebiet.

§5 Partielle und totale Differenzierbarkeit

Sei nun $D \subset \mathbb{R}^n$ eine Teilmenge und $f : D \rightarrow \mathbb{R}^k$ eine Abbildung. Es gibt mehrere Stufen der Kompliziertheit:

- a) $n = 1$ und $k = 1$. Das ist die Situation der *reellen Funktionen von einer Veränderlichen*, die wir in Kapitel I betrachtet haben.
- b) $n = 1$ und k beliebig. Dann sprechen wir von einer *vektorwertigen Funktion* oder einer *parametrisierten Kurve* oder einem *parametrisierten Weg*. $f = (f_1, \dots, f_k)$ ist genau dann stetig bzw. differenzierbar, wenn alle Komponenten es sind. Solche Funktionen haben wir in Kapitel II behandelt. Das Schwergewicht lag allerdings auch da auf dem Fall $k = 1$.
- c) n beliebig und $k = 1$. Das ist der Fall einer *reellen Funktion von n Veränderlichen*. Die Stetigkeit einer solchen Funktion wird wie bei Abbildungen zwischen beliebigen metrischen Räumen definiert.

Die Differenzierbarkeit ist nicht so einfach zu beschreiben. Das wird im Wesentlichen der Inhalt dieses Paragraphen sein. Über die Veranschaulichung einer Funktion von mehreren Veränderlichen werden wir weiter unten sprechen. Allerdings darf man nicht zu sehr an geometrischen Vorstellungen kleben: das Bruttosozialprodukt ist z.B. eine Funktion von sehr vielen Variablen, und dieser funktionale Zusammenhang kann nicht mehr auf vernünftige Weise geometrisch dargestellt werden.

- d) n und k beliebig. Dann besteht $f = (f_1, \dots, f_k)$ aus k Funktionen von n Veränderlichen. Vieles kann man daher auf den Fall (c) zurückführen. Ist $k = n$, so spricht man auch von einem *Vektorfeld*.

Der Übergang von „skalaren“ Werten ($k = 1$) zu „vektoriellen“ Werten (k beliebig) bereitet i.a. nicht so große Schwierigkeiten wie der Übergang von einer Veränderlichen ($n = 1$) zu mehreren Veränderlichen (n beliebig).

Wir beschäftigen uns zunächst mit (skalaren) reellen Funktionen von n Veränderlichen:

Beispiel :

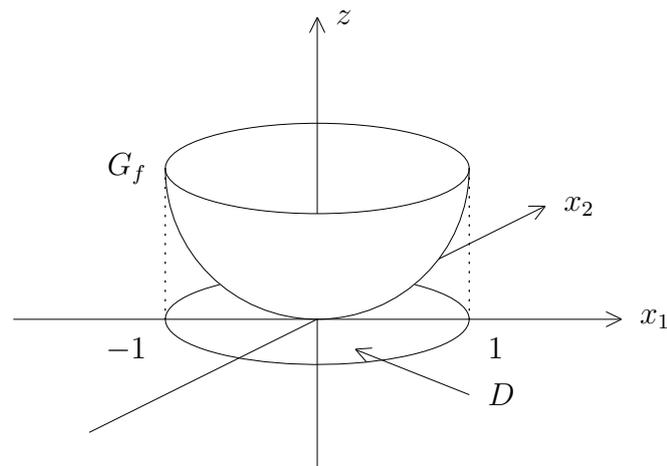
Sei $D := B_1(\mathbf{0}) \subset \mathbb{R}^2$ und $f : D \rightarrow \mathbb{R}$ definiert durch

$$f(x_1, x_2) := x_1^2 + x_2^2.$$

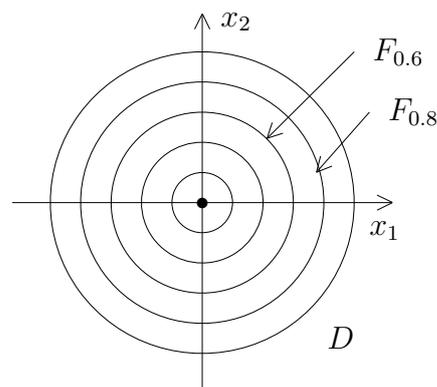
Wie kann man sich eine solche Funktion veranschaulichen? Der Graph

$$G_f := \{(x_1, x_2, z) \in D \times \mathbb{R} \mid z = f(x_1, x_2)\}$$

ist eine Fläche im \mathbb{R}^3 , die in folgendem Sinne über D liegt: Jede „vertikale Gerade“ $\{(x_1, x_2, z) \mid z \in \mathbb{R}\}$ durch einen festen Punkt $(x_1, x_2) \in D$ trifft den Graphen in genau einem Punkt.



Eine andere Möglichkeit der Darstellung ist die Benutzung von „Höhenlinien“. In D werden die *Niveaumengen* $F_c := \{\mathbf{x} \in D \mid f(\mathbf{x}) = c\}$ dargestellt, in unserem Beispiel sind das Kreislinien:



Stetigkeit von Polynomen

Polynome der Gestalt

$$\begin{aligned} p(\mathbf{x}) &= \sum_{i_1+i_2+\dots+i_n \leq m} a_{i_1 i_2 \dots i_n} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} \\ &= a_{00\dots 0} + a_{10\dots 0}x_1 + \dots + a_{0\dots 01}x_n + a_{20\dots 0}x_1^2 + \dots + a_{0\dots 0m}x_n^m \end{aligned}$$

sind stetige Funktionen auf \mathbb{R}^n .

BEWEIS: Die Funktionen $(x_1, \dots, x_n) \mapsto x_i$ und die konstanten Funktionen sind stetig, und Summen und Produkte stetiger Funktionen sind wieder stetig. \square

Die Schreibweise von Polynomen von mehreren Veränderlichen bereitet vielleicht am Anfang etwas Schwierigkeiten. Hier ist z.B. ein Polynom vom Grad 2 in 2 Veränderlichen:

$$\sum_{i_1+i_2 \leq 2} a_{i_1 i_2} x_1^{i_1} x_2^{i_2} = a_{00} + a_{10}x_1 + a_{01}x_2 + a_{20}x_1^2 + a_{11}x_1x_2 + a_{02}x_2^2.$$

Genauer nennt man die größte Zahl d , zu der es Indizes i_1, \dots, i_n mit $i_1 + \dots + i_n = d$ und $a_{i_1 \dots i_n} \neq 0$ gibt, den *Grad* des Polynoms. Im Beispiel des Polynoms vom Grad 2 müßte also wenigstens eine der Zahlen a_{20} , a_{11} oder a_{02} ungleich Null sein.

Definition:

Sei $B \subset \mathbb{R}^n$ offen, $\mathbf{a} \in B$ und $f : B \rightarrow \mathbb{R}$ eine Funktion. Für $\mathbf{v} \in \mathbb{R}^n$ bezeichnet man

$$D_{\mathbf{v}}f(\mathbf{a}) := \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a})}{t}$$

als *Richtungsableitung von f in \mathbf{a} in Richtung \mathbf{v}* (sofern der Grenzwert existiert).

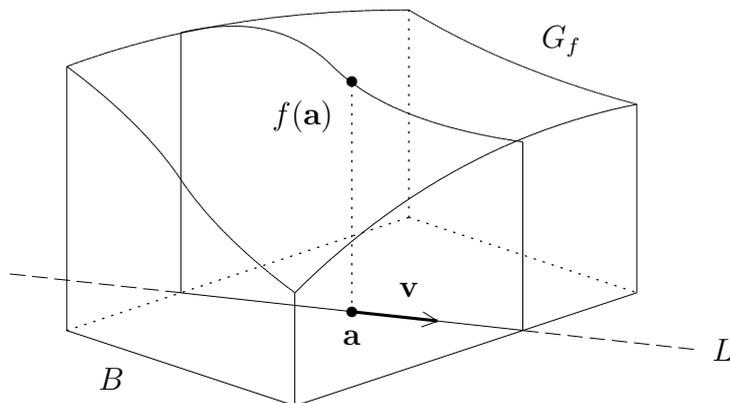
Was bedeutet das anschaulich?

$\alpha(t) := \mathbf{a} + t\mathbf{v}$ definiert eine parametrisierte Gerade $L \subset \mathbb{R}^n$. Sie geht durch den Punkt \mathbf{a} und hat den Richtungsvektor \mathbf{v} . Die Funktion

$$f_L(t) := f \circ \alpha(t) = f(\mathbf{a} + t\mathbf{v})$$

ist eine gewöhnliche Funktion einer Veränderlichen, und die Richtungsableitung von f in \mathbf{a} mit Richtung \mathbf{v} ist nichts anderes als die gewöhnliche Ableitung $(f_L)'(0)$.

Den Graphen von f_L erhält man, indem man den Graphen von f mit der über der Geraden L gelegenen Ebene $\{(\mathbf{x}, z) \in \mathbb{R}^n \times \mathbb{R} \mid \mathbf{x} \in L\}$ schneidet.



Beispiel:

Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert durch $f(x, y) := 1 - x^2 - y^2$. Ist $\mathbf{a} = (a_1, b_1)$ und $\mathbf{v} = (v_1, v_2)$, so ist

$$\begin{aligned} f_L(t) &= f(\mathbf{a} + t\mathbf{v}) = f(a_1 + tv_1, a_2 + tv_2) = 1 - (a_1 + tv_1)^2 - (a_2 + tv_2)^2 \\ &= (1 - a_1^2 - a_2^2) - 2t(a_1v_1 + a_2v_2) - t^2(v_1^2 + v_2^2). \end{aligned}$$

Also ist

$$D_{\mathbf{v}}f(\mathbf{a}) = (f_L)'(0) = (-2(a_1v_1 + a_2v_2) - 2t(v_1^2 + v_2^2)) \Big|_{t=0} = -2(a_1v_1 + a_2v_2).$$

Insbesondere ist $D_{\mathbf{v}}f(\mathbf{a}) = 0$ genau dann, wenn $\mathbf{a} \bullet \mathbf{v} = 0$ ist, wenn also diese beiden Vektoren aufeinander senkrecht stehen. In $\mathbf{a} = \mathbf{0}$ ist **jede** Richtungsableitung $= 0$.

Eigenschaften der Richtungsableitung

f und g seien in \mathbf{a} in Richtung \mathbf{v} differenzierbar, c sei eine Konstante. Dann sind auch $c \cdot f$, $f + g$ und $f \cdot g$ in \mathbf{a} in Richtung \mathbf{v} differenzierbar, und es gilt:

1. $D_{\mathbf{v}}(c \cdot f)(\mathbf{a}) = c \cdot D_{\mathbf{v}}f(\mathbf{a})$.
2. $D_{\mathbf{v}}(f + g)(\mathbf{a}) = D_{\mathbf{v}}f(\mathbf{a}) + D_{\mathbf{v}}g(\mathbf{a})$.
3. $D_{\mathbf{v}}(f \cdot g)(\mathbf{a}) = f(\mathbf{a}) \cdot D_{\mathbf{v}}g(\mathbf{a}) + D_{\mathbf{v}}f(\mathbf{a}) \cdot g(\mathbf{a})$.

BEWEIS: Es ist $(c \cdot f)_L = (c \cdot f) \circ \alpha = c \cdot (f \circ \alpha) = c \cdot f_L$, und analog $(f + g)_L = f_L + g_L$ und $(f \cdot g)_L = (f_L) \cdot (g_L)$. Mit der Definition der Richtungsableitung folgt nun ganz leicht die Behauptung.

Wir wollen das nur bei der Produktregel nachprüfen:

$$\begin{aligned} D_{\mathbf{v}}(f \cdot g)(\mathbf{a}) &= ((f \cdot g)_L)'(0) \\ &= (f_L \cdot g_L)'(0) \\ &= f_L(0) \cdot (g_L)'(0) + (f_L)'(0) \cdot g_L(0) \\ &= f(\mathbf{a}) \cdot D_{\mathbf{v}}g(\mathbf{a}) + D_{\mathbf{v}}f(\mathbf{a}) \cdot g(\mathbf{a}). \end{aligned}$$

□

Abhängigkeit vom Richtungsvektor

Sei f in \mathbf{a} in Richtung \mathbf{v} differenzierbar, $\mathbf{v} \neq \mathbf{0}$ und α eine Konstante. Dann ist f in \mathbf{a} auch in Richtung $\alpha\mathbf{v}$ differenzierbar, und es gilt:

$$D_{\alpha\mathbf{v}}f(\mathbf{a}) = \alpha \cdot D_{\mathbf{v}}f(\mathbf{a}).$$

BEWEIS: Die Ableitung in Richtung des Nullvektors existiert immer und ist $= 0$. Also können wir voraussetzen, daß auch $\alpha \neq 0$ ist. Dann gilt:

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t(\alpha\mathbf{v})) - f(\mathbf{a})}{t} &= \lim_{t \rightarrow 0} \left(\alpha \cdot \frac{f(\mathbf{a} + (t\alpha)\mathbf{v}) - f(\mathbf{a})}{t\alpha} \right) \\ &= \alpha \cdot \lim_{s \rightarrow 0} \frac{f(\mathbf{a} + s\mathbf{v}) - f(\mathbf{a})}{s} \\ &= \alpha \cdot D_{\mathbf{v}}f(\mathbf{a}). \end{aligned}$$

□

Es reicht daher, wenn man sich bei den Richtungsableitungen auf Einheitsvektoren beschränkt. Eine besondere Rolle spielen dabei die Standard-Einheitsvektoren $\mathbf{e}_1, \dots, \mathbf{e}_n$:

Definition:

Die Funktion f sei in \mathbf{a} in Richtung des i -ten Standard-Einheits-Vektors \mathbf{e}_i differenzierbar. Dann heißt

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) := D_{\mathbf{e}_i} f(\mathbf{a})$$

die i -te partielle Ableitung von f in \mathbf{a} . Man schreibt auch $f_{x_i}(\mathbf{a})$ dafür.

Wenn alle partiellen Ableitungen von f in \mathbf{a} existieren, dann heißt f in \mathbf{a} *partiell differenzierbar*.

Wie führt man die partielle Differentiation praktisch durch?

Sei $\mathbf{a} = (a_1, \dots, a_n)$. Dann gilt:

$$\begin{aligned} \frac{\partial f}{\partial x_i}(\mathbf{a}) &= D_{\mathbf{e}_i} f(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{e}_i) - f(\mathbf{a})}{t} \\ &= \lim_{t \rightarrow 0} \frac{1}{t} (f(a_1, \dots, a_i + t, \dots, a_n) - f(a_1, \dots, a_i, \dots, a_n)) \\ &= \lim_{s \rightarrow a_i} \frac{f(a_1, \dots, a_{i-1}, s, a_{i+1}, \dots, a_n) - f(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n)}{s - a_i} \\ &= \left. \frac{d}{ds} \right|_{s=a_i} f(a_1, \dots, a_{i-1}, s, a_{i+1}, \dots, a_n). \end{aligned}$$

Um also die i -te partielle Ableitung von f in \mathbf{a} auszurechnen, muß man in $f(x_1, \dots, x_n)$ die Variablen x_j , $j \neq i$, durch die entsprechenden Komponenten a_j von \mathbf{a} ersetzen. Danach hängt die Funktion nur noch von der einen verbliebenen Variablen x_i ab und kann im gewöhnlichen Sinne nach dieser Variablen an der Stelle a_i differenziert werden.

Beispiel:

Sei $f(x, y, z) := x^2 \cdot \cos(yz)$.

Um partiell nach x zu differenzieren, muß man die Variablen y und z festhalten und nur die Funktion $x \mapsto x^2 \cdot \cos(yz)$ betrachten. Also ist

$$\frac{\partial f}{\partial x}(x, y, z) = 2x \cdot \cos(yz).$$

Um partiell nach y zu differenzieren, muß man die Variablen x und z festhalten und nur die Funktion $y \mapsto x^2 \cdot \cos(yz)$ betrachten. So erhält man

$$\frac{\partial f}{\partial y}(x, y, z) = x^2 \cdot (-\sin(yz) \cdot z) = -x^2 z \sin(yz)$$

und analog

$$\frac{\partial f}{\partial z}(x, y, z) = -x^2 y \sin(yz).$$

Es sieht so aus, als hätte man die Verallgemeinerung der Differenzierbarkeit auf mehrere Veränderliche gefunden. Aber leider ist die partielle Differenzierbarkeit eine zu schwache Eigenschaft. Sie hat noch nicht einmal die Stetigkeit der Funktion selbst zur Folge:

Beispiel:

Wir betrachten die Funktion

$$f(x, y) := \begin{cases} \frac{xy^2}{x^2 + y^4} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0). \end{cases}$$

Die Funktionen $x \mapsto f(x, 0) \equiv 0$ und $y \mapsto f(0, y) \equiv 0$ sind sicherlich im Nullpunkt differenzierbar. Also ist f in $\mathbf{0} = (0, 0)$ partiell differenzierbar. Andererseits ist f dort nicht stetig:

Wenn man $\mathbf{y}_\nu := ((a_\nu)^2, a_\nu)$ setzt, mit einer Nullfolge (a_ν) , so konvergiert diese Folge gegen $(0, 0)$, aber es ist

$$\lim_{\nu \rightarrow \infty} f(\mathbf{y}_\nu) = \lim_{\nu \rightarrow \infty} \frac{(a_\nu)^4}{2(a_\nu)^4} = \frac{1}{2}.$$

Das dürfte nicht passieren, wenn f im Nullpunkt stetig wäre.

Eine weitere Schwäche der partiellen Differenzierbarkeit tritt auf, wenn man höhere Ableitungen betrachtet:

Ist $B \subset \mathbb{R}^n$ offen und $f : B \rightarrow \mathbb{R}$ in allen Punkten von B partiell differenzierbar, so bilden die partiellen Ableitungen $\frac{\partial f}{\partial x_i}(\mathbf{x})$ wieder reellwertige Funktionen auf B . Sind sie alle stetig, so nennt man f *stetig partiell differenzierbar*.

Definition:

Sei $B \subset \mathbb{R}^n$ offen, $f : B \rightarrow \mathbb{R}$ überall partiell differenzierbar und alle partiellen Ableitungen $\frac{\partial f}{\partial x_i}$ in einem Punkt $\mathbf{a} \in B$ wiederum partiell differenzierbar. Dann definiert man für $i, j = 1, \dots, n$:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}) := \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) (\mathbf{a}).$$

Man nennt diesen Ausdruck auch die *2-te partielle Ableitung von f nach x_i und x_j an der Stelle \mathbf{a}* , und schreibt dafür auch $f_{x_i x_j}(\mathbf{a})$.

Beispiel:

Sei $f(x_1, x_2) := e^{k \cdot x_1} \cdot \cos(x_2)$. Dann gilt:

$$\frac{\partial f}{\partial x_1}(\mathbf{x}) = k \cdot e^{k \cdot x_1} \cdot \cos(x_2) \quad \text{und} \quad \frac{\partial f}{\partial x_2}(\mathbf{x}) = -e^{k \cdot x_1} \cdot \sin(x_2),$$

sowie

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{a}) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{a}) = -k e^{k a_1} \sin(a_2).$$

Man kann sich nun fragen, ob man die 2-ten Ableitungen immer miteinander vertauschen kann, ob es also bei höheren partiellen Ableitungen nicht auf die Reihenfolge ankommt. Leider ist das nicht generell der Fall:

Beispiel:

$$\text{Sei } f(x, y) := \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0), \\ 0 & \text{für } (x, y) = (0, 0). \end{cases}$$

Dann gilt für $(x, y) \neq (0, 0)$:

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \frac{\partial}{\partial x} \left(\frac{x^3 y - y^3 x}{x^2 + y^2} \right) \\ &= \frac{(3x^2 y - y^3)(x^2 + y^2) - (x^3 y - y^3 x)2x}{(x^2 + y^2)^2} \\ &= \frac{x^4 y + 4x^2 y^3 - y^5}{(x^2 + y^2)^2}, \end{aligned}$$

also

$$\frac{\partial f}{\partial x}(0, y) = -y \quad (\text{für } y \neq 0).$$

Weiter ist

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{x \rightarrow 0} \frac{f(x, 0) - f(0, 0)}{x} = 0.$$

Also ist sogar $\frac{\partial f}{\partial x}(0, y) \equiv -y$ für alle y und $\frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1$.

Entsprechend erhalten wir für $(x, y) \neq (0, 0)$:

$$\begin{aligned} \frac{\partial f}{\partial y}(x, y) &= \frac{\partial}{\partial y} \left(\frac{x^3 y - y^3 x}{x^2 + y^2} \right) \\ &= \frac{(x^3 - 3y^2 x)(x^2 + y^2) - (x^3 y - y^3 x)2y}{(x^2 + y^2)^2} \\ &= \frac{x^5 - 4x^3 y^2 - xy^4}{(x^2 + y^2)^2}, \end{aligned}$$

also

$$\frac{\partial f}{\partial y}(x, 0) \equiv x \quad \text{für } x \neq 0,$$

und

$$\frac{\partial f}{\partial y}(0, 0) = \lim_{y \rightarrow 0} \frac{f(0, y) - f(0, 0)}{y} = 0.$$

Somit ist $\frac{\partial^2 f}{\partial x \partial y}(0, 0) = +1$.

Zum Glück gilt folgendes hinreichende Kriterium für die Gleichheit der gemischten zweiten Ableitungen:

Satz von Schwarz

Sei $B \subset \mathbb{R}^n$ offen und $f : B \rightarrow \mathbb{R}$ auf ganz B nach allen Variablen partiell differenzierbar, $\mathbf{a} \in B$.

Wenn die gemischten zweiten Ableitungen $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x})$ und $\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x})$ auf einer Umgebung von \mathbf{a} in B existieren und in \mathbf{a} stetig sind, so ist

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{a}).$$

Auf den etwas technischen Beweis verzichten wir hier. Es genügt übrigens schon, daß **eine** der beiden gemischten Ableitungen in der Nähe von \mathbf{a} existiert und in \mathbf{a} stetig ist. Dann folgt bereits die Existenz der anderen Ableitung und die Gleichheit.

Die Bildung einer partiellen Ableitung $\frac{\partial f}{\partial x_i}$ kann man auch als Anwendung eines „linearen Operators“ $\frac{\partial}{\partial x_i}$ auf die Funktion f auffassen.

Man faßt nun gerne die n partiellen Ableitungs-Operatoren zu einem vektoriellen Operator zusammen:

$$\nabla := \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right). \quad (\text{„Nabla“})$$

Dieser Operator kann auf verschiedene Weise wirken:

Sei $B \subset \mathbb{R}^n$ offen.

1. Ist $f : B \rightarrow \mathbb{R}$ eine stetig partiell differenzierbare Funktion, so heißt das **Vektorfeld**

$$\mathbf{grad}(f) := \nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$$

das *Gradientenfeld* von f . Der Wert $\mathbf{grad}(f)(\mathbf{a})$ in einem Punkt \mathbf{a} wird als *Gradient von f in \mathbf{a}* bezeichnet.

2. Sei $\mathbf{v} = (v_1, \dots, v_n) : B \rightarrow \mathbb{R}^n$ ein Vektorfeld, dessen sämtliche Komponenten v_i stetig partiell differenzierbar sind. Dann heißt die **Funktion**

$$\mathbf{div}(\mathbf{v}) := \nabla \bullet \mathbf{v} = \frac{\partial v_1}{\partial x_1} + \dots + \frac{\partial v_n}{\partial x_n}$$

die *Divergenz* von \mathbf{v} .

3. Sei jetzt speziell $n = 3$ und $\mathbf{v} : B \rightarrow \mathbb{R}^3$ ein stetig partiell differenzierbares Vektorfeld. Dann heißt das **Vektorfeld**

$$\mathbf{rot}(\mathbf{v}) := \nabla \times \mathbf{v} = \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right)$$

die *Rotation* von \mathbf{v} .

Man beachte, daß bei $\nabla \bullet \mathbf{v}$ und $\nabla \times \mathbf{v}$ nicht einfach nur Multiplikationen zwischen den Komponenten von ∇ und denen von \mathbf{v} durchgeführt werden, sondern daß die partiellen Ableitungen in ∇ als Operatoren auf den Komponenten von \mathbf{v} wirken! Die vereinfachte Schreibweise mit dem ∇ kann daher leicht zu Fehlern führen.

Divergenz und Rotation werden später ausführlicher in einem Kapitel über Vektoranalysis behandelt werden, mit dem Gradienten und seiner Bedeutung beschäftigen wir uns noch einmal weiter unten in diesem Paragraphen.

Zunächst wollen wir aber den Differenzierbarkeitsbegriff noch einmal überdenken. Bei der partiellen Differenzierbarkeit haben wir folgende Mängel festgestellt:

- Eine partiell differenzierbare Funktion braucht nicht stetig zu sein.
- Ist eine Funktion $2\times$ partiell differenzierbar, so hängen die Werte der zweiten Ableitungen von der Reihenfolge der Differentiation ab.

Vielleicht kann man ja die Situation in einer Veränderlichen auf geschicktere Art verallgemeinern:

Sei $I \subset \mathbb{R}$ ein offenes Intervall, $t_0 \in I$ und $f : I \rightarrow \mathbb{R}$ eine Funktion. Ist f in t_0 differenzierbar, so existiert der Grenzwert

$$f'(t_0) := \lim_{t \rightarrow t_0} \frac{f(t) - f(t_0)}{t - t_0}.$$

Führen wir die lineare Funktion

$$L(h) := f'(t_0) \cdot h$$

und die Restfunktion

$$r(h) := f(t_0 + h) - f(t_0) - L(h)$$

ein, so gilt:

1. $f(t) = f(t_0) + L(t - t_0) + r(t - t_0)$ für $t = t_0 + h$ nahe t_0 .
2. $\lim_{h \rightarrow 0} \frac{r(h)}{h} = 0$.
3. Der Graph der affin-linearen Funktion $T_f(t) := f(t_0) + L(t - t_0)$ ist die Tangente an den Graphen von f im Punkte t_0 .

Erfüllt f umgekehrt die Bedingungen (1) und (2), so ist f in t_0 differenzierbar, und $L(1)$ ist die Ableitung $f'(t_0)$. Die Bedingungen (1) und (2) lassen sich aber verhältnismäßig leicht auf die Situation mehrerer Veränderlicher übertragen:

Definition:

Sei $B \subset \mathbb{R}^n$ offen, $f : B \rightarrow \mathbb{R}$ eine Funktion und $\mathbf{a} \in B$ ein Punkt.

f heißt in \mathbf{a} (*total*) *differenzierbar*, wenn es eine lineare Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}$ und eine in der Nähe des Nullpunktes definierte „Restfunktion“ r gibt, so daß gilt:

1. $f(\mathbf{x}) = f(\mathbf{a}) + L(\mathbf{x} - \mathbf{a}) + r(\mathbf{x} - \mathbf{a})$ für \mathbf{x} nahe \mathbf{a} .
2. $\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = 0$.

Die (dadurch eindeutig bestimmte) lineare Abbildung L heißt die (*totale*) *Ableitung* oder das *totale Differential* von f in \mathbf{a} . Man bezeichnet sie auch mit

$$Df(\mathbf{a}) \quad , \quad df(\mathbf{a}) \quad \text{oder} \quad (df)_{\mathbf{a}}.$$

Bemerkung: Man kann die Bedingungen (1) und (2) für die totale Differenzierbarkeit in einem Punkt \mathbf{a} auch folgendermaßen zusammenfassen:

Es gibt eine lineare Abbildung $Df(\mathbf{a}) : \mathbb{R}^n \rightarrow \mathbb{R}$, so daß gilt:

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - Df(\mathbf{a})(\mathbf{h})}{\|\mathbf{h}\|} = 0.$$

Daß die Ableitung eindeutig bestimmt ist, folgt leicht, wenn auch nicht ganz so trivial wie in einer Veränderlichen. Wir werden diese Tatsache zunächst voraussetzen, um einige Ableitungen bestimmen zu können, werden den Beweis aber nachtragen.

Beispiele:

1. Sei $f(\mathbf{x}) \equiv c$ konstant. Da die Ableitung einer konstanten Funktion in einer Veränderlichen gleich Null ist, raten wir hier: $L = 0$ (also die Null-Abbildung). Tatsächlich ist dann

$$\frac{f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - L(\mathbf{h})}{\|\mathbf{h}\|} = \frac{c - c - 0}{\|\mathbf{h}\|} = 0$$

für jeden Vektor \mathbf{h} , und damit verschwindet auch der Grenzwert für $\mathbf{h} \rightarrow \mathbf{0}$. Also ist $Dc(\mathbf{a}) = 0$ in jedem Punkt \mathbf{a} .

2. Sei $f(\mathbf{x}) := \mathbf{u} \bullet \mathbf{x} = u_1 x_1 + \cdots + u_n x_n$ selbst schon eine Linearform. Dann ist

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = (f(\mathbf{a}) + f(\mathbf{h})) - f(\mathbf{a}) = f(\mathbf{h}).$$

Also gilt mit $L(\mathbf{h}) := f(\mathbf{h})$:

$$\frac{f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - L(\mathbf{h})}{\|\mathbf{h}\|} = 0 \quad \text{für jedes } \mathbf{h}.$$

Die Ableitung einer Linearform ist in jedem Punkt \mathbf{a} des \mathbb{R}^n wieder diese Linearform.

3. Nun sei $A = (a_{ij}) \in M_{n,n}(\mathbb{R})$ eine symmetrische Matrix und

$$f(\mathbf{x}) := \mathbf{x} \circ A \circ \mathbf{x}^\top = \sum_{i,j=1}^n a_{ij} x_i x_j$$

die zu A gehörige „quadratische Form“. Um die Ableitung zu bestimmen, bleiben wir besser bei der vektoriellen Schreibweise. Es ist

$$\begin{aligned} f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) &= (\mathbf{a} + \mathbf{h}) \circ A \circ (\mathbf{a} + \mathbf{h})^\top - \mathbf{a} \circ A \circ \mathbf{a}^\top \\ &= \mathbf{a} \circ A \circ \mathbf{a}^\top + \mathbf{h} \circ A \circ \mathbf{a}^\top + \mathbf{a} \circ A \circ \mathbf{h}^\top + \mathbf{h} \circ A \circ \mathbf{h}^\top \\ &\quad - \mathbf{a} \circ A \circ \mathbf{a}^\top \\ &= 2\mathbf{a} \circ A \circ \mathbf{h}^\top + \mathbf{h} \circ A \circ \mathbf{h}^\top. \end{aligned}$$

Jetzt sieht man schon etwas klarer: Wir versuchen es mit

$$L(\mathbf{h}) := 2\mathbf{a} \circ A \circ \mathbf{h}^\top \quad \text{und} \quad r(\mathbf{h}) := \mathbf{h} \circ A \circ \mathbf{h}^\top.$$

Offensichtlich ist L eine Linearform, und wir brauchen nur noch eine gute Abschätzung für den Restterm. Nach der Schwarzschen Ungleichung ist aber

$$|r(\mathbf{h})| = |(\mathbf{h} \circ A) \bullet \mathbf{h}| \leq \|\mathbf{h} \circ A\| \cdot \|\mathbf{h}\|,$$

also

$$\frac{|r(\mathbf{h})|}{\|\mathbf{h}\|} \leq \|\mathbf{h} \circ A\|.$$

Da $\mathbf{h} \mapsto \|\mathbf{h} \circ A\|$ als Zusammensetzung stetiger Abbildungen selbst stetig ist, folgt:

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = 0.$$

Somit ist $Df(\mathbf{a})(\mathbf{h}) = 2\mathbf{a} \circ A \circ \mathbf{h}^\top$.

Selbst bei relativ einfachen Funktionen ist die Suche nach der Ableitung recht mühsam. Wir brauchen einen einfachen Kalkül, und zum Glück gibt es den:

Berechnung der totalen Ableitung

Sei $f : B \rightarrow \mathbb{R}$ in $\mathbf{a} \in B$ differenzierbar. Dann existieren in \mathbf{a} auch sämtliche Richtungsableitungen von f , und es gilt:

$$Df(\mathbf{a})(\mathbf{v}) = D_{\mathbf{v}}f(\mathbf{a}).$$

Insbesondere ist f in \mathbf{a} nach allen Variablen partiell differenzierbar, und es gilt:

$$Df(\mathbf{a})(\mathbf{v}) = v_1 \frac{\partial f}{\partial x_1}(\mathbf{a}) + \cdots + v_n \frac{\partial f}{\partial x_n}(\mathbf{a}) = \mathbf{v} \bullet \nabla f(\mathbf{a}).$$

Aus diesem Satz folgt auch sofort die Eindeutigkeit der Ableitung!

BEWEIS: Sei $\mathbf{v} = (v_1, \dots, v_n)$ ein Richtungsvektor $\neq \mathbf{0}$ und $t \in \mathbb{R}$, $t \neq 0$. Dann gilt:

$$f(\mathbf{a} + t\mathbf{v}) = f(\mathbf{a}) + t \cdot Df(\mathbf{a})(\mathbf{v}) + r(t \cdot \mathbf{v}),$$

mit

$$\lim_{t \rightarrow 0} \frac{r(t \cdot \mathbf{v})}{t} = \pm \|\mathbf{v}\| \cdot \lim_{t \rightarrow 0} \frac{r(t \cdot \mathbf{v})}{\|t \cdot \mathbf{v}\|} = 0.$$

Also ist

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a}) - t \cdot Df(\mathbf{a})(\mathbf{v})}{t} = 0$$

und damit

$$D_{\mathbf{v}}f(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a})}{t} = Df(\mathbf{a})(\mathbf{v}).$$

Insbesondere existieren die partiellen Ableitungen $\frac{\partial f}{\partial x_i}(\mathbf{a}) = D_{\mathbf{e}_i}f(\mathbf{a})$ für $i = 1, \dots, n$, und es gilt:

$$Df(\mathbf{a})(\mathbf{v}) = Df(\mathbf{a})\left(\sum_{i=1}^n v_i \mathbf{e}_i\right) = \sum_{i=1}^n v_i Df(\mathbf{a})(\mathbf{e}_i) = \sum_{i=1}^n v_i D_{\mathbf{e}_i}f(\mathbf{a}) = \mathbf{v} \bullet \nabla f(\mathbf{a}).$$

□

Dieser Satz erlaubt es jetzt, totale Ableitungen mit Hilfe von partiellen Ableitungen auszurechnen, und für die letzteren brauchen wir ja nur den Kalkül aus der Theorie einer Veränderlichen zu übernehmen. Wir wollen das gleich einmal testen:

Beispiel:

Sei wieder

$$f(\mathbf{x}) = \mathbf{x} \circ A \circ \mathbf{x}^\top = \sum_{i,j=1}^n a_{ij} x_i x_j.$$

Dann ist

$$\begin{aligned} \frac{\partial}{\partial x_k} \sum_{i,j=1}^n a_{ij} x_i x_j &= \sum_{i,j=1}^n a_{ij} \frac{\partial}{\partial x_k} (x_i x_j) \\ &= \sum_{i,j=1}^n a_{ij} (\delta_{ik} x_j + \delta_{jk} x_i) \\ &= \sum_{j=1}^n a_{kj} x_j + \sum_{i=1}^n a_{ik} x_i \\ &= 2 \sum_{i=1}^n a_{ik} x_i \quad (\text{weil } A \text{ symmetrisch}). \end{aligned}$$

Also ist

$$\begin{aligned} Df(\mathbf{x})(\mathbf{h}) &= \sum_{k=1}^n h_k \frac{\partial f}{\partial x_k}(\mathbf{x}) \\ &= \sum_{k=1}^n h_k \cdot 2 \sum_{i=1}^n a_{ik} x_i \\ &= 2 \cdot \sum_{i,k=1}^n x_i a_{ik} h_k \\ &= 2\mathbf{x} \circ A \circ \mathbf{h}^\top. \end{aligned}$$

Das hatten wir schon auf anderem Wege herausbekommen.

Eine **Warnung** muß ausgesprochen werden! Der obige Satz ist nicht umkehrbar, es gibt Funktionen, die partiell, aber nicht total differenzierbar sind. Das ergibt sich aus folgender Feststellung:

Total differenzierbare Funktionen sind stetig

Ist f in \mathbf{a} total differenzierbar, so ist f dort auch stetig.

BEWEIS: Wir können schreiben:

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{h}) + r(\mathbf{h}),$$

mit $L(\mathbf{h}) \rightarrow 0$ und $r(\mathbf{h}) \rightarrow 0$ für $\mathbf{h} \rightarrow \mathbf{0}$.

Daher ist $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$. □

Wir haben schon ein Beispiel einer Funktion gesehen, die im Nullpunkt partiell differenzierbar, aber nicht stetig ist. Sie kann dann natürlich erst recht nicht total differenzierbar sein.

Wir stehen damit vor einem Dilemma: Berechnen wir die Ableitung einer Funktion f mit Hilfe der Definition der totalen Differenzierbarkeit, so haben wir damit automatisch auch die totale Differenzierbarkeit von f bewiesen. Aber dieser Weg ist meistens nicht durchführbar. Benutzen wir andererseits die besser handhabbaren partiellen Ableitungen, so müssen wir von Rechts wegen auch noch die totale Differenzierbarkeit beweisen. Also ist eigentlich nichts gewonnen. Zum Glück gibt es folgendes einfache Kriterium:

Differenzierbarkeitskriterium

Sei $B \subset \mathbb{R}^n$ offen, $f : B \rightarrow \mathbb{R}$ eine Funktion und $\mathbf{a} \in B$ ein Punkt.

Wenn es eine offene Umgebung U von \mathbf{a} in B gibt, so daß alle partiellen Ableitungen von f auf U existieren und in \mathbf{a} stetig sind, dann ist f in \mathbf{a} total differenzierbar.

Der Beweis ist nicht sehr schwer. Wie die totale Ableitung $L = Df(\mathbf{a})$ aussehen soll, wissen wir ja schon, wir müssen nur den Ausdruck

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - L(\mathbf{h})$$

abschätzen. Zu dem Zweck verbindet man \mathbf{a} und $\mathbf{a} + \mathbf{h}$ durch eine Kette von achsenparallelen Strecken. Die Differenzen der Funktionswerte von f an zwei aufeinanderfolgenden Endpunkten können jeweils mit Hilfe des Mittelwertsatzes durch partielle Ableitungen von f an geeigneten Zwischenpunkten ausgedrückt werden. Auf die Ausführung der technischen Einzelheiten verzichten wir hier.

Was ist die anschauliche Bedeutung der totalen Differenzierbarkeit?

In einer Veränderlichen gilt: Ist f in t_0 differenzierbar, so besitzt der Graph von f in $(t_0, f(t_0))$ eine Tangente. Parametrisiert wird die Tangente z.B. durch $t \mapsto (t, T_f(t))$, mit der affin-linearen Funktion $T_f(t) := f(t_0) + f'(t_0) \cdot (t - t_0)$. Dann gilt insbesondere:

1. $T_f(t_0) = f(t_0)$.
2. $\lim_{t \rightarrow t_0} \frac{f(t) - T_f(t)}{t - t_0} = 0$.

In mehreren Veränderlichen gilt etwas Analoges:

Existenz der Tangentenebene

Sei $f : B \rightarrow \mathbb{R}$ in $\mathbf{a} \in B \subset \mathbb{R}^n$ total differenzierbar.

Dann gibt es genau eine affin-lineare Funktion $T_f : \mathbb{R}^n \rightarrow \mathbb{R}$ mit folgenden Eigenschaften:

- a) $T_f(\mathbf{a}) = f(\mathbf{a})$.
- b) $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f(\mathbf{x}) - T_f(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|} = 0$.

Die Tangentenfunktion T_f ist gegeben durch

$$T_f(\mathbf{x}) := f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{x} - \mathbf{a}).$$

Den Graphen $\mathbf{T}_\mathbf{a}(f) := G_{T_f} \subset \mathbb{R}^{n+1}$ von T nennt man die (affine) Tangentenebene an G_f in $(\mathbf{a}, f(\mathbf{a}))$.

Diese Tangentenebene enthält insbesondere alle Tangenten an G_f in $(\mathbf{a}, f(\mathbf{a}))$, die sich aus den Richtungsableitungen ergeben.

BEWEIS: Sei F irgend eine affin-lineare Funktion, die die Bedingungen a) und b) erfüllt. Dann hat F die Gestalt

$$F(\mathbf{x}) = \alpha + L(\mathbf{x}),$$

mit einer Konstanten α und einer Linearform L . Wegen (a) ist $\alpha + L(\mathbf{a}) = f(\mathbf{a})$, also

$$F(\mathbf{x}) = f(\mathbf{a}) + L(\mathbf{x} - \mathbf{a}).$$

Wegen (b) folgt daraus:

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - L(\mathbf{h})}{\|\mathbf{h}\|} = 0.$$

Da f in \mathbf{a} differenzierbar und die Ableitung eindeutig bestimmt ist, muß $L = Df(\mathbf{a})$ sein. Das zeigt die Eindeutigkeit von T_f und Aussage (2).

Also ist

$$\mathbf{T}_\mathbf{a}(f) = \{(\mathbf{x}, z) \mid z = f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{x} - \mathbf{a})\}.$$

Sei $\mathbf{v} \in \mathbb{R}^n$ ein beliebiger Richtungsvektor, $\alpha : \mathbb{R} \rightarrow \mathbb{R}^{n+1}$ definiert durch

$$\alpha(t) := (\mathbf{a} + t\mathbf{v}, f(\mathbf{a} + t\mathbf{v})).$$

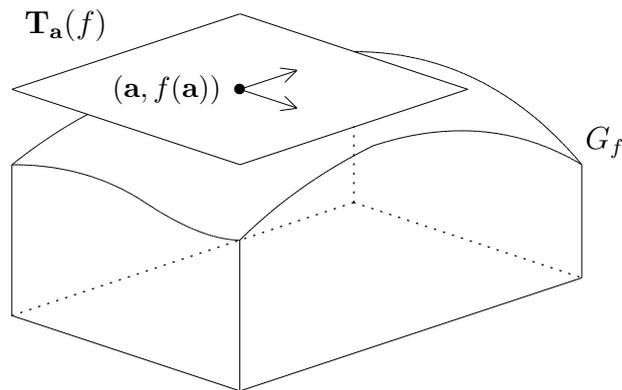
Dann ist α ein differenzierbarer Weg, der ganz innerhalb des Graphen G_f verläuft, mit $\alpha(0) = (\mathbf{a}, f(\mathbf{a}))$ und Tangentialvektor $\alpha'(0) = (\mathbf{v}, Df(\mathbf{a})(\mathbf{v}))$.

Behauptung: $(\mathbf{a}, f(\mathbf{a})) + \alpha'(0) \in \mathbf{T}_{\mathbf{a}}(f)$.

BEWEIS dafür:

$$(\mathbf{a}, f(\mathbf{a})) + \alpha'(0) = (\mathbf{a} + \mathbf{v}, f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{v})).$$

Setzt man $\mathbf{x} := \mathbf{a} + \mathbf{v}$ und $z := f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{v})$, so ist $z = f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{x} - \mathbf{a})$, also $(\mathbf{a}, f(\mathbf{a})) + \alpha'(0) \in \mathbf{T}_{\mathbf{a}}(f)$. \square



Beispiele :

1. Sei $f(x, y) := x^2 + y^2$.

Dann ist $f(0, 0) = 0$ und $Df(0, 0) = 0$. Also ist $T_f(x, y) \equiv 0$ und die Tangentenebene an den Graphen von f im Nullpunkt ist gegeben durch

$$\mathbf{T}_{(0,0)}(f) = \{(x, y, z) \in \mathbb{R}^3 \mid z = 0\}.$$

2. Sei $f(x, y) := e^x \cos(y)$ und $\mathbf{a} := (0, \frac{\pi}{4})$.

Dann ist $f(\mathbf{a}) = \frac{1}{2}\sqrt{2}$ und $Df(\mathbf{a})(v, w) = \frac{1}{2}\sqrt{2}(v - w)$. Also ist

$$T_f(x, y) = f(\mathbf{a}) + Df(\mathbf{a})(x, y - \frac{\pi}{4}) = \frac{1}{2}\sqrt{2}(1 + x - y + \frac{\pi}{4}).$$

Die Tangentenebene an den Graphen von f im Nullpunkt ist gegeben durch

$$\mathbf{T}_{(0,0)}(f) = \{(x, y, \frac{1}{2}\sqrt{2}(1 + x - y + \frac{\pi}{4})) \mid x, y \in \mathbb{R}^2\}.$$

3. Sei $f(x, y) := \begin{cases} \frac{xy^2}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0). \end{cases}$

Wir zeigen zunächst, daß f im Nullpunkt stetig ist: Sei (\mathbf{x}_ν) eine Nullfolge. Dann können wir schreiben:

$$\mathbf{x}_\nu = (r_\nu \cos(\varphi_\nu), r_\nu \sin(\varphi_\nu)), \text{ für } \nu \in \mathbb{N}.$$

Dabei konvergiert $r_\nu = \|\mathbf{x}_\nu\|$ gegen Null, und unabhängig von φ_ν ist

$$(\cos \varphi_\nu)^2 + (\sin \varphi_\nu)^2 = 1 \text{ und } 0 \leq |\cos \varphi_\nu|, |\sin \varphi_\nu| \leq 1.$$

Also konvergiert

$$|f(\mathbf{x}_\nu)| = \left| \frac{r_\nu^3 \cos \varphi_\nu (\sin \varphi_\nu)^2}{r_\nu^2} \right| \leq r_\nu$$

gegen Null.

Weiter ist $f(x, 0) \equiv 0$ und $f(0, y) \equiv 0$. Also ist f im Nullpunkt auch partiell differenzierbar, und es gilt:

$$\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0.$$

Es existieren sogar beliebige Richtungsableitungen:

Da $f(tx, ty) = t \cdot f(x, y)$ für alle t und beliebiges (x, y) gilt,² ist

$$D_{\mathbf{h}}f(\mathbf{0}) = \lim_{t \rightarrow 0} \frac{f(t\mathbf{h}) - f(\mathbf{0})}{t} = f(\mathbf{h}).$$

Man kann also an G_f im Nullpunkt in jeder beliebigen Richtung eine Tangente legen.

Wäre f in $\mathbf{0}$ total differenzierbar, so müßte $Df(\mathbf{0}) = 0$ sein. Für $\mathbf{h} := (r, r)$ ist aber

$$\frac{f(\mathbf{h}) - f(\mathbf{0}) - 0}{\|\mathbf{h}\|} = \frac{r^3}{2r^2 \cdot \sqrt{2}|r|} = \pm \frac{1}{2\sqrt{2}},$$

und das kann nicht gegen Null konvergieren.

Also ist f im Nullpunkt nicht total differenzierbar, und der Graph von f besitzt dort keine Tangentialebene. Wie soll man sich das vorstellen?

Da f homogen ist, gehört mit (\mathbf{x}, z) auch jeder Punkt $(t\mathbf{x}, tz)$ zum Graphen von f , also die ganze Gerade durch (\mathbf{x}, z) und den Nullpunkt. Diese Geraden sind dann natürlich auch Tangenten, und sie müßten daher auch in einer etwa existierenden Tangentialebene enthalten sein. Das ist nicht möglich, weil die Geraden gar nicht alle in einer Ebene liegen. Die Punkte $(1, 1, \frac{1}{2})$, $(1, -1, \frac{1}{2})$ und $(1, 0, 0)$ liegen z.B. auf G_f , sind aber linear unabhängig.

Tatsächlich hat G_f im Nullpunkt eine „Spitze“, und dieser Mangel an Glattheit verhindert die totale Differenzierbarkeit.

Wir wollen jetzt folgende Situation untersuchen:

Sei $B \subset \mathbb{R}^n$ offen, $\mathbf{a} \in B$ und $f : B \rightarrow \mathbb{R}$ in der Nähe von \mathbf{a} differenzierbar. Weiter sei $\alpha : I \rightarrow B$ ein differenzierbarer Weg, mit $\alpha(t_0) = \mathbf{a}$. Dann kann man f auf α einschränken und erhält

$$g := f \circ \alpha : I \rightarrow \mathbb{R},$$

²Man nennt f daher auch eine *homogene* Funktion.

eine reellwertige Funktion von **einer** Veränderlichen! Wir würden erwarten, daß g in t_0 differenzierbar ist, und daß man die Ableitung mit Hilfe der Kettenregel gewinnen kann. Aber wie?

Wir setzen

$$\Delta(t) = \begin{cases} \frac{\alpha(t) - \alpha(t_0)}{t - t_0} & \text{für } t \neq t_0, \\ \alpha'(t_0) & \text{in } t = t_0. \end{cases}$$

Dann ist Δ stetig in t_0 und $\alpha(t) = \alpha(t_0) + \Delta(t) \cdot (t - t_0)$. Weiter können wir wegen der Differenzierbarkeit von f schreiben:

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{h}) + \|\mathbf{h}\| \cdot e(\mathbf{h}),$$

mit $e(\mathbf{h}) := \frac{r(\mathbf{h})}{\|\mathbf{h}\|}$, also $\lim_{\mathbf{h} \rightarrow \mathbf{0}} e(\mathbf{h}) = 0$.

Dann ist

$$\begin{aligned} g(t) &= f \circ \alpha(t) \\ &= f(\mathbf{a}) + Df(\mathbf{a})(\alpha(t) - \alpha(t_0)) + \|\alpha(t) - \alpha(t_0)\| \cdot e(\alpha(t) - \alpha(t_0)), \end{aligned}$$

also

$$\frac{g(t) - g(t_0)}{t - t_0} = Df(\alpha(t_0))(\Delta(t)) \pm \|\Delta(t)\| \cdot e(\alpha(t) - \alpha(t_0)).$$

Läßt man t gegen t_0 gehen, so erhält man:

$$g'(t_0) = Df(\alpha(t_0))(\Delta(t_0)).$$

Damit haben wir bewiesen:

Spezielle Kettenregel

Ist $B \subset \mathbb{R}^n$ offen, $\alpha : I \rightarrow B$ in $t_0 \in I$ differenzierbar und $f : B \rightarrow \mathbb{R}$ in $\mathbf{a} := \alpha(t_0)$ differenzierbar, so ist auch $f \circ \alpha$ in t_0 differenzierbar, und es gilt:

$$(f \circ \alpha)'(t_0) = \nabla f(\mathbf{a}) \bullet \alpha'(t_0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\alpha(t_0)) \cdot \frac{d\alpha_i}{dt}(t_0).$$

Beispiele :

1. Sei $\alpha(t) := \mathbf{a} + t\mathbf{v}$ eine Gerade. Dann ist $\alpha'(t) \equiv \mathbf{v}$ und daher

$$(f \circ \alpha)'(0) = \nabla f(\mathbf{a}) \bullet \mathbf{v} = Df(\mathbf{a})(\mathbf{v}).$$

2. Sei $\alpha(t) := (\cos(t), \sin(t))$ und $f(x, y) = x + y$. Dann ist

$$(f \circ \alpha)'(t) = \frac{\partial f}{\partial x}(\alpha(t))\alpha_1'(t) + \frac{\partial f}{\partial y}(\alpha(t))\alpha_2'(t) = -\sin(t) + \cos(t).$$

Wir können jetzt das Wesen des Gradienten etwas besser ergründen:

Sei $B \subset \mathbb{R}^n$ offen und $f : B \rightarrow \mathbb{R}$ eine differenzierbare Funktion. Für $c \in \mathbb{R}$ sei

$$F_c := \{\mathbf{x} \in B \mid f(\mathbf{x}) = c\}$$

die entsprechende Niveau-Fläche von f .

Eigenschaften des Gradienten

Sei $\mathbf{a} \in B$, $f(\mathbf{a}) = c$ und $\nabla f(\mathbf{a}) \neq \mathbf{0}$.

1. $\nabla f(\mathbf{a})$ zeigt in die Richtung, in der f am schnellsten wächst.
2. Ist $\alpha : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^n$ ein differenzierbarer Weg mit $\alpha(0) = \mathbf{a}$, der ganz in F_c verläuft, so steht $\nabla f(\mathbf{a})$ auf $\alpha'(0)$ senkrecht.

BEWEIS: 1) Sei $\|\mathbf{v}\| = 1$. Dann gilt:

$$\begin{aligned} \left. \frac{d}{dt} \right|_0 f(\mathbf{a} + t\mathbf{v}) &= D_{\mathbf{v}}f(\mathbf{a}) \\ &= \nabla f(\mathbf{a}) \bullet \mathbf{v} \\ &= \|\nabla f(\mathbf{a})\| \cdot \|\mathbf{v}\| \cdot \cos \theta, \end{aligned}$$

wobei θ der Winkel zwischen \mathbf{v} und $\nabla f(\mathbf{a})$ ist.

Dieser Ausdruck wird maximal, wenn $\theta = 0$ ist, also $\mathbf{v} = \frac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|}$.

2) Verläuft α ganz in F_c , so ist $f \circ \alpha(t) \equiv c$, also

$$0 = (f \circ \alpha)'(t_0) = \nabla f(\mathbf{a}) \bullet \alpha'(t_0).$$

□

Definition:

Sei f wie oben gegeben, $f(\mathbf{a}) = c$ und $\nabla f(\mathbf{a}) \neq \mathbf{0}$. Dann nennt man

$$T_{\mathbf{a}}(F_c) := \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v} \bullet \nabla f(\mathbf{a}) = 0\}$$

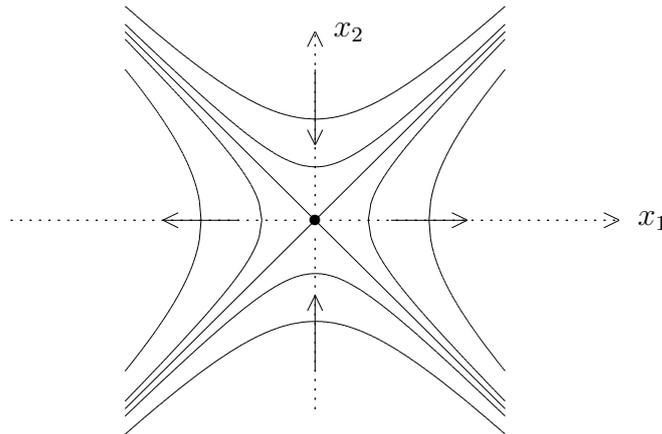
den *Tangententialraum* an F_c in \mathbf{a} (und $\mathbf{a} + T_{\mathbf{a}}(F_c)$ den *affinen Tangentialraum*).

Beispiele:

1. Sei $f(x_1, x_2) := x_1^2 + x_2^2$. Dann ist der Graph von f eine Schale in Form einer Halbkugel, und der Gradient $\nabla f(a_1, a_2) = 2(a_1, a_2)$ zeigt stets nach außen.
2. Bei der Funktion $f(x_1, x_2) := 1 - x_1^2 - x_2^2$ ist es genau umgekehrt. Der Graph ist eine umgestülpte Schale, und der Gradient $\nabla f(a_1, a_2) = -2(a_1, a_2)$ zeigt stets nach innen.

Wie im vorigen Beispiel verschwindet der Gradient im Nullpunkt. Dort hat er also auch keine Richtung, und das liegt daran, daß die Funktion im ersten Beispiel im Nullpunkt ein Minimum und im zweiten Beispiel ein Maximum besitzt.

3. Schließlich betrachten wir $f(x_1, x_2) := x_1^2 - x_2^2$. Hier ist $\nabla f(a_1, a_2) = 2(a_1, -a_2)$. Längs der x_1 -Achse zeigt der Gradient nach außen, längs der x_2 -Achse zeigt er nach innen, und im Nullpunkt verschwindet er. Dort liegt ein sogenannter „Sattelpunkt“ vor. Die Niveaulinien sehen folgendermaßen aus:



Setzt man $u := x_1 - x_2$ und $v := x_1 + x_2$, so sind die Niveaulinien Hyperbeln der Gestalt $u \cdot v = \text{const.}$.

4. Sei $g : B \rightarrow \mathbb{R}$ differenzierbar, $g(\mathbf{a}) = 0$ und $f : B \times \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$f(\mathbf{x}, z) := z - g(\mathbf{x}).$$

Dann stimmt die Niveaulfläche $F_0 = \{(\mathbf{x}, z) \mid f(\mathbf{x}, z) = 0\}$ mit dem Graphen $G_g = \{(\mathbf{x}, z) \mid z = g(\mathbf{x})\}$ überein.

Es ist $\nabla f(\mathbf{a}, 0) = (-\nabla g(\mathbf{a}), 1) \neq (\mathbf{0}, 0)$. Wir wollen die Tangentenhyperebene $\mathbf{T}_{\mathbf{a}}(g)$ mit dem affinen Tangentialraum $(\mathbf{a}, 0) + T_{(\mathbf{a}, 0)}(F_0)$ vergleichen.

Für $(\mathbf{h}, z) \in \mathbb{R}^{n+1}$ gilt:

$$\begin{aligned} \nabla f(\mathbf{a}, 0) \bullet (\mathbf{h}, z) = 0 &\iff (-\nabla g(\mathbf{a}), 1) \bullet (\mathbf{h}, z) = 0 \\ &\iff \nabla g(\mathbf{a}) \bullet \mathbf{h} = z \\ &\iff Dg(\mathbf{a})(\mathbf{h}) = z. \end{aligned}$$

Also liegt (\mathbf{x}, z) genau dann in $(\mathbf{a}, 0) + T_{(\mathbf{a}, 0)}(F_0)$, wenn gilt:

$$z = Dg(\mathbf{a})(\mathbf{x} - \mathbf{a}) = g(\mathbf{a}) + Dg(\mathbf{a})(\mathbf{x} - \mathbf{a}) = T_g(\mathbf{x}).$$

Aber das ist gleichbedeutend damit, daß (\mathbf{x}, z) in $\mathbf{T}_{\mathbf{a}}(g)$ liegt. Die beiden (auf ganz verschiedene Weise definierten) affinen Tangentialräume stimmen überein. Das hat natürlich auch damit zu tun, daß der Gradient einer Funktion der speziellen Gestalt $f(\mathbf{x}, z) = z - g(\mathbf{x})$ nie verschwindet.

§6 Extremwerte

Definition:

Sei $M \subset \mathbb{R}^n$ eine Teilmenge, $f : M \rightarrow \mathbb{R}$ stetig, $\mathbf{a} \in M$ ein Punkt.

f hat in \mathbf{a} auf M ein *relatives (oder lokales) Maximum* (bzw. ein *relatives (oder lokales) Minimum*), wenn es eine offene Umgebung $U(\mathbf{a}) \subset \mathbb{R}^n$ gibt, so daß

$$f(\mathbf{x}) \leq f(\mathbf{a}) \quad (\text{bzw.} \quad f(\mathbf{x}) \geq f(\mathbf{a}))$$

für alle $\mathbf{x} \in U \cap M$ ist. In beiden Fällen spricht man von einem *relativen (oder lokalen) Extremum*.

Gilt die Ungleichung sogar für alle $\mathbf{x} \in M$, so spricht man von einem *absoluten (oder globalen) Maximum oder Minimum*.

Notwendiges Kriterium für relative Extremwerte

Sei $B \subset \mathbb{R}^n$ offen und $f : B \rightarrow \mathbb{R}$ in $\mathbf{a} \in B$ differenzierbar.

Besitzt f in \mathbf{a} ein relatives Extremum, so ist $\nabla f(\mathbf{a}) = \mathbf{0}$.

BEWEIS: Für $i = 1, \dots, n$ besitzt auch $g_i(t) := f(\mathbf{a} + t\mathbf{e}_i)$ in $t = 0$ ein lokales Extremum. Nach dem notwendigen Kriterium aus der Differentialrechnung einer Veränderlichen muß dann $(g_i)'(0) = 0$ sein. Es ist aber

$$(g_i)'(0) = \frac{\partial f}{\partial x_i}(\mathbf{a}), \quad \text{für } i = 1, \dots, n.$$

Daraus folgt die Behauptung. □

Definition:

Ist f in \mathbf{a} differenzierbar und $\nabla f(\mathbf{a}) = \mathbf{0}$, so heißt \mathbf{a} ein *stationärer (oder kritischer) Punkt* von f .

Ein stationärer Punkt \mathbf{a} von f heißt *Sattelpunkt* von f , falls es in jeder Umgebung $U(\mathbf{a})$ Punkte \mathbf{b} und \mathbf{c} gibt, so daß

$$f(\mathbf{b}) < f(\mathbf{a}) < f(\mathbf{c})$$

ist.

Beispiele dazu werden wir später betrachten. Ein hinreichendes Kriterium für die Existenz eines Extremwertes erhält man in einer Veränderlichen durch Untersuchung der höheren

Ableitungen, insbesondere der zweiten Ableitung.

Wir kommen nun nicht umhin, die Taylorformel in n Veränderlichen zu beweisen, zumindest bis zur Ordnung 2.

Definition:

Eine Funktion $f : B \rightarrow \mathbb{R}$ heißt auf B k -mal stetig differenzierbar, wenn f partielle Ableitungen bis zur Ordnung k besitzt, also Ableitungen der Form

$$\frac{\partial^{i_1+\dots+i_n} f}{\partial x_1^{i_1} \partial x_2^{i_2} \dots \partial x_n^{i_n}} \quad \text{mit} \quad i_1 + \dots + i_n \leq k,$$

und wenn alle partiellen Ableitungen der Ordnung k auf B noch stetig sind.

Die Menge aller k -mal stetig differenzierbaren Funktionen auf B wird mit dem Symbol $C^k(B)$ bezeichnet.

Bemerkung: Ist $k \geq 1$ und $f \in C^k(B)$, so ist f insbesondere in jedem Punkt von B total differenzierbar. f ist dann sogar „ k -mal total differenzierbar“, aber auf diesen Begriff wollen wir hier nicht näher eingehen.

Wir betrachten nun eine Funktion $f \in C^2(B)$ und einen Punkt $\mathbf{a} \in B$. Für eine beliebige Richtung $\mathbf{h} \in \mathbb{R}^n$ sei $\alpha_{\mathbf{h}}(t) := \mathbf{a} + t\mathbf{h}$ und $g(t) := f \circ \alpha_{\mathbf{h}}(t) = f(\mathbf{a} + t\mathbf{h})$. Dann folgt aus der speziellen Kettenregel:

$$\begin{aligned} g'(t) &= \nabla f(\alpha_{\mathbf{h}}(t)) \circ \mathbf{h}^\top \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{a} + t\mathbf{h}) \cdot h_i. \end{aligned}$$

Da $\frac{\partial f}{\partial x_i}(\mathbf{x})$ nach Voraussetzung auf ganz B stetige partielle Ableitungen besitzt, also insbesondere total differenzierbar ist, ist auch $g'(t)$ ein weiteres Mal differenzierbar. Es gilt:

$$\begin{aligned} g''(t) &= \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \circ \alpha_{\mathbf{h}} \right)'(t) \cdot h_i \\ &= \sum_{i=1}^n \left(\sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\alpha_{\mathbf{h}}(t)) \cdot h_j \right) \cdot h_i \\ &= \sum_{i,j=1}^n h_i \cdot \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a} + t\mathbf{h}) \cdot h_j. \end{aligned}$$

Definition:

Sei f in der Nähe von $\mathbf{a} \in \mathbb{R}^n$ zweimal stetig differenzierbar. Dann heißt die symmetrische Matrix

$$H_f(\mathbf{x}) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) \mid i, j = 1, \dots, n \right)$$

die *Hesse-Matrix* von f in \mathbf{x} .

Wir haben gerade ausgerechnet, daß $g''(t) = \mathbf{h} \circ H_f(\mathbf{a} + t\mathbf{h}) \circ \mathbf{h}^\top$ ist.

Bemerkung: Die Symmetrie der Hesse-Matrix folgt aus der Vertauschbarkeit der 2. Ableitungen, und die ist nur gegeben, weil f in einer ganzen Umgebung von \mathbf{a} zweimal stetig differenzierbar ist. Diese Voraussetzung ist also wichtig!

Im Falle $n = 2$ ist $H_f = \begin{pmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{pmatrix}$.

Nun können wir die Taylorformel formulieren und beweisen:

Taylorformel 2.Ordnung

Sei $B = B_r(\mathbf{a})$ eine offene Kugel um \mathbf{a} , $f : B \rightarrow \mathbb{R}$ zweimal stetig differenzierbar. Dann gibt es eine auf $B_r(\mathbf{0})$ definierte Funktion R mit

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{R(\mathbf{h})}{\|\mathbf{h}\|^2} = 0,$$

so daß für $\|\mathbf{h}\| < r$ gilt:

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \circ \mathbf{h}^\top + \frac{1}{2} \mathbf{h} \circ H_f(\mathbf{a}) \circ \mathbf{h}^\top + R(\mathbf{h}).$$

BEWEIS: Ist $\mathbf{h} \in B_r(\mathbf{0})$, dann liegt $\alpha_{\mathbf{h}}(t) = \mathbf{a} + t\mathbf{h}$ für $t \in [-1, 1]$ in $B_r(\mathbf{a})$, und deshalb ist $g(t) := f \circ \alpha_{\mathbf{h}}(t)$ auf $[-1, 1]$ definiert und zweimal stetig differenzierbar. Wir wenden auf g in $t = 0$ den Satz von der Taylorentwicklung in einer Veränderlichen an:

Es gibt eine Funktion $\eta(t)$ mit $\lim_{t \rightarrow 0} \eta(t) = 0$, so daß gilt:

$$g(t) = g(0) + g'(0) \cdot t + \frac{1}{2} g''(0) \cdot t^2 + \eta(t) \cdot t^2.$$

Dabei ist

$$\eta(t) = \frac{1}{2} \cdot (g''(c) - g''(0)),$$

mit einer geeigneten (von t abhängigen) Zahl c mit $0 < c < t$. Setzen wir $t = 1$, so folgt die gewünschte Taylorformel, mit

$$\begin{aligned} R(\mathbf{h}) &:= \eta(1) \\ &= \frac{1}{2} \mathbf{h} \circ (H_f(\mathbf{a} + \mathbf{ch}) - H_f(\mathbf{a})) \circ \mathbf{h}^\top \end{aligned}$$

$$= \frac{1}{2} \sum_{i,j=1}^n \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a} + c\mathbf{h}) - \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}) \right) h_i h_j$$

und $0 < c < 1$. Diesen Ausdruck müssen wir noch abschätzen, um herauszubekommen, daß

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{R(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

ist.

Zunächst ist $|h_i| = |\mathbf{h} \bullet \mathbf{e}_i| \leq \|\mathbf{h}\| \cdot \|\mathbf{e}_i\| = \|\mathbf{h}\|$.

Die Summe enthält n^2 Summanden, und da f zweimal stetig differenzierbar ist, die zweiten partiellen Ableitungen also stetig sind, gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so daß

$$\left| \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a} + c\mathbf{h}) - \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}) \right| < \varepsilon$$

für $\mathbf{h} \in B_\delta(\mathbf{0})$ ist. Für solche \mathbf{h} ist dann

$$|R(\mathbf{h})| < \frac{\varepsilon}{2} \cdot n^2 \cdot \|\mathbf{h}\|^2.$$

Daraus ergibt sich die gewünschte Limesbeziehung. \square

Es gibt selbstverständlich auch Taylorformeln höherer Ordnung, aber mit denen werden wir uns hier nicht beschäftigen.

Ist nun f in \mathbf{a} stationär, also

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = \frac{1}{2} \mathbf{h} \circ H_f(\mathbf{a}) \circ \mathbf{h}^\top + R(\mathbf{h}),$$

so hängt das Verhalten von f in der Nähe von \mathbf{a} im Wesentlichen von der Hesse-Matrix ab, denn $R(\mathbf{h})$ verschwindet ja in \mathbf{a} von höherer Ordnung. Das führt uns zu einem ähnlichen hinreichenden Kriterium für Extremwerte, wie wir es aus der eindimensionalen Theorie kennen. Allerdings ist die Lage hier doch noch etwas komplizierter.

Ist $A \in M_{n,n}(\mathbb{R})$ eine symmetrische Matrix, so nennt man die Funktion

$$q(\mathbf{h}) := \mathbf{h} \circ A \circ \mathbf{h}^\top$$

bekanntlich eine *quadratische Form*. Es ist

$$q(t\mathbf{h}) = t^2 \cdot q(\mathbf{h}) \text{ für } t \in \mathbb{R} \text{ und } \mathbf{h} \in \mathbb{R}^n.$$

Insbesondere ist natürlich $q(\mathbf{0}) = 0$.

Definition:

Eine quadratische Form $q(\mathbf{h})$ heißt

- positiv semidefinit : $\iff q(\mathbf{h}) \geq 0$ für alle \mathbf{h} ,
- positiv definit : $\iff q(\mathbf{h}) > 0$ für alle $\mathbf{h} \neq \mathbf{0}$,
- negativ semidefinit : $\iff q(\mathbf{h}) \leq 0$ für alle \mathbf{h} ,
- negativ definit : $\iff q(\mathbf{h}) < 0$ für alle $\mathbf{h} \neq \mathbf{0}$,
- indefinit : $\iff \exists \mathbf{h}_1, \mathbf{h}_2$ mit $q(\mathbf{h}_1) < 0 < q(\mathbf{h}_2)$.

Erinnern wir uns an die Lineare Algebra! Da wurde gezeigt:

Ist $A \in M_{n,n}(\mathbb{R})$ eine symmetrische Matrix, so gibt es eine orthogonale Matrix $S \in \text{GL}(n, \mathbb{R})$, so daß $S^{-1} \circ A \circ S$ eine Diagonalmatrix ist. Die Einträge in der Diagonalmatrix sind die Eigenwerte von A . (Das war der Satz von der Hauptachsentransformation).

Daß S orthogonal ist, bedeutet, daß $S^T S = E_n$, also $S^{-1} = S^T$ ist. Setzen wir $\mathbf{v} := \mathbf{h} \circ S$, so ist

$$\begin{aligned} q_A(\mathbf{h}) &:= \mathbf{h} \circ A \circ \mathbf{h}^T = (\mathbf{v} \circ S^T) \circ A \circ (\mathbf{v} \circ S^T)^T \\ &= \mathbf{v} \circ (S^{-1} \circ A \circ S) \circ \mathbf{v}^T \\ &= \mathbf{v} \circ \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \circ \mathbf{v}^T \\ &= \sum_{i=1}^n \lambda_i (v_i)^2, \end{aligned}$$

wobei $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A sind. Also folgt:

$$\begin{aligned} q_A \text{ positiv definit} &\iff \mathbf{h} \circ A \circ \mathbf{h}^T > 0 \text{ für alle } \mathbf{h} \neq \mathbf{0} \\ &\iff \sum_{i=1}^n \lambda_i (v_i)^2 > 0 \text{ für alle } \mathbf{v} \neq \mathbf{0} \\ &\iff \lambda_1, \dots, \lambda_n > 0. \end{aligned}$$

Genauso ist q_A negativ definit, wenn alle Eigenwerte von A negativ sind. Und q_A ist positiv semidefinit (bzw. negativ semidefinit), wenn alle Eigenwerte von $A \geq 0$ (bzw. ≤ 0) sind. Gibt es wenigstens einen negativen und einen positiven Eigenwert, so ist q_A indefinit.

Im Falle $n = 2$ gibt es noch ein einfacheres Kriterium:

Definitheit von 2×2 -Matrizen

Sei $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix} \in M_{2,2}(\mathbb{R})$ eine symmetrische Matrix. Dann ist

$$q_A(h_1, h_2) = ah_1^2 + 2bh_1h_2 + dh_2^2,$$

und es gilt:

1. Ist $\det(A) < 0$, so ist q_A indefinit.
2. Ist $\det(A) > 0$ und $a > 0$, so ist q_A positiv definit.
3. Ist $\det(A) > 0$ und $a < 0$, so ist q_A negativ definit.

BEWEIS: Sei $\Delta := \det(A) = ad - b^2$. Zur Berechnung der Eigenwerte brauchen wir noch das charakteristische Polynom:

$$p_A(x) = \det \begin{pmatrix} a-x & b \\ b & d-x \end{pmatrix} = (a-x)(d-x) - b^2 = x^2 - (a+d)x + \Delta.$$

Die Eigenwerte λ_1, λ_2 von A sind die beiden Nullstellen dieses quadratischen Polynoms. Nach dem (hoffentlich aus der Schule bekannten) Satz von Vieta ist

$$\lambda_1 + \lambda_2 = a + d \quad \text{und} \quad \lambda_1 \cdot \lambda_2 = \Delta.$$

Ist $\Delta < 0$, so haben die beiden Eigenwerte verschiedenes Vorzeichen, und q_A ist indefinit. Ist $\Delta > 0$, so sind λ_1 und λ_2 beide $\neq 0$, und sie haben das gleiche Vorzeichen. Außerdem ist $ad = \Delta + b^2 > 0$. Ist nun $a > 0$, so ist auch $d > 0$ und damit $\lambda_1 + \lambda_2 > 0$. In diesem Fall ist q_A positiv definit. Genauso folgt aus $a < 0$, daß q_A negativ definit ist. \square

Nun haben wir endlich alles beisammen.

Hinreichendes Kriterium für Extremwerte

Sei $B \subset \mathbb{R}^n$ offen, $f \in \mathcal{C}^2(B)$. Weiter sei $\mathbf{a} \in B$ ein stationärer Punkt von f , also $\nabla f(\mathbf{a}) = \mathbf{0}$.

1. Ist $H_f(\mathbf{a})$ positiv definit, so besitzt f in \mathbf{a} ein relatives Minimum.
2. Ist $H_f(\mathbf{a})$ negativ definit, so besitzt f in \mathbf{a} ein relatives Maximum.
3. Ist $H_f(\mathbf{a})$ indefinit, so liegt in \mathbf{a} ein Sattelpunkt vor.

BEWEIS:

1) Sei $q(\mathbf{h}) := \mathbf{h} \circ H_f(\mathbf{a}) \circ \mathbf{h}^\top$. Da f in \mathbf{a} stationär ist, ergibt die Taylorformel:

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = \frac{1}{2}q(\mathbf{h}) + R(\mathbf{h}).$$

Die Funktion q ist stetig und nach Voraussetzung > 0 außerhalb des Nullpunktes. Insbesondere nimmt sie auf der abgeschlossenen und beschränkten und daher kompakten Menge

$$S^{n-1} := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}$$

ein Minimum $m > 0$ an. Daher gilt für beliebiges $\mathbf{h} \in \mathbb{R}^n$:

$$q(\mathbf{h}) = \|\mathbf{h}\|^2 \cdot q\left(\frac{\mathbf{h}}{\|\mathbf{h}\|}\right) \geq m \cdot \|\mathbf{h}\|^2.$$

Ist jetzt ein ε mit $0 < \varepsilon < \frac{m}{2}$ vorgegeben und dazu ein $r = r(\varepsilon)$ so gewählt, daß

$$|R(\mathbf{h})| \leq \varepsilon \cdot \|\mathbf{h}\|^2 \quad \text{für } \mathbf{h} \in B_r(\mathbf{0})$$

ist, so folgt:

$$\begin{aligned} f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) &= \frac{1}{2}q(\mathbf{h}) + R(\mathbf{h}) \\ &\geq \left(\frac{m}{2} - \varepsilon\right) \cdot \|\mathbf{h}\|^2 \geq 0 \end{aligned}$$

für alle $\mathbf{h} \in B_r(\mathbf{0})$.

Also ist $f(\mathbf{a} + \mathbf{h}) \geq f(\mathbf{a})$ für kleines \mathbf{h} , und es liegt ein relatives Minimum in \mathbf{a} vor.

2) Der Fall des Maximums kann durch Übergang von f zu $-f$ auf (1) zurückgeführt werden.

3) Ist q indefinit, so gibt es in jeder Umgebung von $\mathbf{0}$ Vektoren \mathbf{h}_1 und \mathbf{h}_2 mit $q(\mathbf{h}_1) < 0 < q(\mathbf{h}_2)$. Die Funktionen

$$\begin{aligned} f_1(t) &:= f(\mathbf{a} + t\mathbf{h}_1) \\ \text{und } f_2(t) &:= f(\mathbf{a} + t\mathbf{h}_2) \end{aligned}$$

sind dann definiert und zweimal differenzierbar, und es gilt:

$$(f_1)'(0) = (f_2)'(0) = 0, \quad (f_1)''(0) = q(\mathbf{h}_1) < 0 \quad \text{und} \quad (f_2)''(0) = q(\mathbf{h}_2) > 0.$$

Also besitzt f_1 in $t = 0$ ein isoliertes Maximum und f_2 in $t = 0$ ein isoliertes Minimum. Das bedeutet, daß f beliebig nahe bei \mathbf{a} sowohl Werte $< f(\mathbf{a})$ als auch Werte $> f(\mathbf{a})$ annimmt. Damit liegt ein Sattelpunkt vor. \square

Bemerkung: Ist $H_f(\mathbf{a})$ nur semidefinit, so kann man keine genaue Aussage machen!

Beispiele:

1. Sei $f(x, y) := x^2 + y^2$. Dann ist $\nabla f(x, y) = (2x, 2y)$, also $(0, 0)$ der einzige stationäre Punkt von f . Da $f(0, 0) = 0$ und allgemein $f(x, y) \geq 0$ ist, liegt ein absolutes Minimum vor. Tatsächlich ist

$$H_f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Da $2 > 0$ und $\det H_f(x, y) = 2 \cdot 2 - 0 \cdot 0 = 4 > 0$ ist, ist die Matrix positiv definit. (Das ist übrigens jede Diagonalmatrix mit nur positiven Einträgen). Das Hinreichende Kriterium sagt also auch, daß f im Nullpunkt ein lokales Minimum besitzt.

2. Sei $f(x, y) := 1 - x^2 - y^2$. Dann ist $\nabla f(x, y) = (-2x, -2y)$ und $H_f(x, y) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$ negativ definit. Hier liegt im Nullpunkt ein Maximum vor.

3. Sei $f(x, y) := x^2 - y^2$. Nun ist $\nabla f(x, y) = (2x, -2y)$ und $H_f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$. Da $\det H_f(x, y) < 0$ ist, liegt im Nullpunkt ein Sattelpunkt vor.

4. Sei $f(x, y) := e^{xy} + x^2 + \frac{1}{9}y^2$.

Dann ist $\nabla f(x, y) = (ye^{xy} + 2x, xe^{xy} + \frac{2}{9}y)$. Für die Hesse-Matrix ergibt sich:

$$H_f(x, y) = \begin{pmatrix} 2 + y^2 e^{xy} & (1 + xy)e^{xy} \\ (1 + xy)e^{xy} & \frac{2}{9} + x^2 e^{xy} \end{pmatrix}.$$

Der Nullpunkt ist sicher ein stationärer Punkt. Ist (x, y) irgendein anderer stationärer Punkt, so muß gelten:

$$xye^{xy} = -2x^2 \quad \text{und} \quad xye^{xy} = -\frac{2}{9}y^2,$$

also $x = \pm \frac{1}{3}y$.

Wäre $x = \frac{1}{3}y$, so wäre $0 = f_y(x, y) = \frac{y}{3}(e^{xy} + \frac{2}{3})$, also $y = 0$ (und damit auch $x = 0$) oder $e^{xy} = -\frac{2}{3}$, was nicht möglich ist. So bleibt nur die Gleichung $x = -\frac{1}{3}y$. Wegen der Bedingung $0 = f_x(x, y) = y(e^{xy} - \frac{2}{3})$ muß dann $e^{xy} = \frac{2}{3}$ sein, also $e^{-y^2/3} = \frac{2}{3}$.

Das führt auf die Gleichung $y^2 = -3 \ln(\frac{2}{3})$. Setzen wir $p := \sqrt{-3 \ln(\frac{2}{3})}$ (der Radius ist positiv, weil $\ln(\frac{2}{3}) < 0$ ist!), so sind die Punkte

$$\mathbf{x}_{1,2} := \pm(-\frac{1}{3}p, p)$$

weitere Kandidaten für stationäre Punkte, und mehr kann es nicht geben. Nun gilt:

$$\begin{aligned} f(\mathbf{0}) &= 1 \\ \text{und} \quad f(\mathbf{x}_{1,2}) &= e^{-p^2/3} + \frac{2}{9}p^2 \\ &= \frac{2}{3} \cdot (1 - \ln(\frac{2}{3})). \end{aligned}$$

Daß $\nabla f(\mathbf{0}) = (0, 0)$ ist, ist klar. Ist (x, y) einer der beiden Punkte \mathbf{x}_1 oder \mathbf{x}_2 , so ist $xy = -\frac{p^2}{3} = \ln(\frac{2}{3})$, also

$$\nabla f(x, y) = (\frac{2}{3}y + 2x, \frac{2}{3}x + \frac{2}{9}y) = \pm(\frac{2}{3}p - \frac{2}{3}p, -\frac{2}{9}p + \frac{2}{9}p) = (0, 0).$$

Da $H_f(0, 0) = \begin{pmatrix} 2 & 1 \\ 1 & \frac{2}{9} \end{pmatrix}$ ist, also $\det H_f(0, 0) = \frac{4}{9} - 1 < 0$, liegt im Nullpunkt ein Sattelpunkt vor! Und da $f(x, y) > \frac{1}{9}(x^2 + y^2)$ ist, gilt für $\|(x, y)\| \geq 3$, daß $f(x, y) > 1$ ist. Auf der kompakten Menge $\overline{B_3(\mathbf{0})}$ muß f als stetige Funktion ein globales Minimum ≤ 1 annehmen, und das muß sogar in der offenen Kugel $B_3(\mathbf{0})$ liegen, weil f auf dem Rand der Kugel schon Werte > 1 annimmt. In einem solchen Minimum muß f aber einen stationären Punkt besitzen. Dafür kommen nur die beiden Punkten \mathbf{x}_1 und \mathbf{x}_2 in Frage, und weil f in diesen Punkten den gleichen Wert annimmt, können wir schließen:

f besitzt in \mathbf{x}_1 und in \mathbf{x}_2 jeweils ein (globales) Minimum.

Manchmal interessiert man sich nur dafür, ob unter gewissen Nebenbedingungen ein Extremwert angenommen wird.

Beispiel:

Sei $f(x, y) := y$. Diese auf ganz \mathbb{R}^2 definierte Funktion mißt die Höhe über der x -Achse, und sie besitzt weder einen globalen noch einen lokalen Extremwert. Der Gradient $\nabla f(x, y) = (0, 1)$ verschwindet nirgends.

Wenn wir f allerdings auf die Parabel $P := \{(x, y) \mid y = x^2 + 1\}$ einschränken, so sehen wir:

$$f(x, y) = f(x, x^2 + 1) = x^2 + 1 \geq 1 \quad \text{und} \quad f(0, 1) = 1.$$

Also nimmt f auf P in $(0, 1)$ ein Minimum an. Setzt man $g(x, y) := y - x^2 - 1$, so kann man sagen:

$f(x, y)$ nimmt unter der Nebenbedingung $g(x, y) = 0$ in $(0, 1)$ ein Minimum an.

Für derartige Situationen wollen wir ein allgemeines Verfahren entwickeln. Zunächst beschäftigen wir uns allerdings mit einem einfachen Spezialfall.

Sei $B \subset \mathbb{R}^n$ offen und $g : B \rightarrow \mathbb{R}$ stetig differenzierbar. $N(g) := \{\mathbf{x} \in B \mid g(\mathbf{x}) = 0\}$ ist die „Nullstellenmenge“ von g .

Wir setzen voraus, daß $\nabla g(\mathbf{x}) \neq \mathbf{0}$ für alle $\mathbf{x} \in N(g)$ ist.

Weiter sei $\mathbf{a} \in N(g)$, $U(\mathbf{a}) \subset B$ eine offene Umgebung und $f : U \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion.

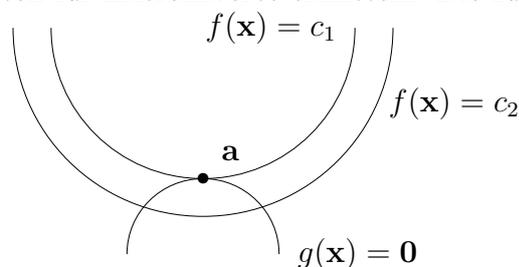
Definition:

f hat in \mathbf{a} ein *relatives Maximum* (bzw. *relatives Minimum*)

unter der Nebenbedingung $g(\mathbf{x}) = 0$,

falls $f(\mathbf{x}) \leq f(\mathbf{a})$ (bzw. $f(\mathbf{x}) \geq f(\mathbf{a})$) für alle $\mathbf{x} \in U \cap N(g)$ ist.

Wie kann man solche Extrema unter Nebenbedingungen bestimmen? Ein hinreichendes Kriterium ist schwer zu finden, aber zumindest können wir mit Hilfe eines notwendigen Kriteriums mögliche Kandidaten für Extremwerte ermitteln. Die Idee ist die folgende:



Damit die Werte von f in \mathbf{a} ein Maximum oder Minimum annehmen, müssen sich bei \mathbf{a} eine Niveauläche von f und die Menge M berühren. Also muß die Tangentialebene an M in \mathbf{a} mit der Tangentialebene an die Niveauläche von f übereinstimmen. Da die Gradienten auf den Tangentialebenen senkrecht stehen, muß dann der Gradient von f in die gleiche Richtung wie der Gradient der Funktionen g zeigen.

Das führt zu folgendem Kriterium:

Methode des Lagrangeschen Multiplikators

Hat f in \mathbf{a} ein relatives Extremum unter der Nebenbedingung

$$g(\mathbf{x}) = 0 \quad (\text{und } \nabla g(\mathbf{x}) \neq \mathbf{0}),$$

so gibt es eine Zahl $\lambda \in \mathbb{R}$, so daß gilt:

$$\nabla f(\mathbf{a}) = \lambda \cdot \nabla g(\mathbf{a}).$$

Die Zahl λ nennt man den Lagrangeschen Multiplikator. Man beachte, daß es sich hier wirklich nur um ein notwendiges Kriterium handelt! Die Punkte, die die angegebene Bedingung erfüllen, können Extremwerte sein. Ob sie es wirklich sind, muß man mit anderen Mitteln feststellen.

BEWEIS-Andeutung:

1) Wir benutzen den Begriff des Tangentialraums:

$$T_{\mathbf{a}}(N(g)) = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v} \bullet \nabla g(\mathbf{a}) = 0\}.$$

Sei $\alpha : (-\varepsilon, \varepsilon) \rightarrow N(g)$ ein stetig differenzierbarer Weg, $\alpha(0) = \mathbf{a}$. Dann ist $g \circ \alpha(t) \equiv 0$, also

$$0 = (g \circ \alpha)'(0) = \nabla g(\mathbf{a}) \bullet \alpha'(0).$$

Das bedeutet, daß $\alpha'(0)$ in $T_{\mathbf{a}}(N(g))$ liegt, wie es ja auch der Anschauung entspricht. Wir nehmen jetzt aber ohne Beweis zusätzlich an, daß es zu jedem $\mathbf{v} \in T_{\mathbf{a}}(N(g))$ einen Weg α in $N(g)$ mit $\alpha'(0) = \mathbf{v}$ gibt.

Da f in \mathbf{a} auf $N(g)$ ein relatives Extremum besitzen soll, hat auch $f \circ \alpha : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ in $t = 0$ ein relatives Extremum. Also ist

$$0 = (f \circ \alpha)'(0) = \nabla f(\mathbf{a}) \bullet \alpha'(0).$$

Das bedeutet, daß $\nabla f(\mathbf{a})$ auf allen Vektoren des Tangentialraums senkrecht steht, d.h.

$$\nabla f(\mathbf{a}) \in (T_{\mathbf{a}}(N(g)))^{\perp} = (\langle \nabla g(\mathbf{a}) \rangle^{\perp})^{\perp} = \langle \nabla g(\mathbf{a}) \rangle.$$

Dann muß es ein λ mit $\nabla f(\mathbf{a}) = \lambda \cdot \nabla g(\mathbf{a})$ geben.

2) In einem Spezialfall läßt sich die Zusatzannahme leicht verifizieren. Ist etwa

$$g(x_1, \dots, x_n) = x_n - h(x_1, \dots, x_{n-1}),$$

mit einer differenzierbaren Funktion h mit $h(a_1, \dots, a_{n-1}) = 0$, so ist $N(g)$ der Graph von h , und der Tangentialraum an $N(g)$ in $(a_1, \dots, a_{n-1}, 0)$ ist (bis auf die Verschiebung $(a_1, \dots, a_{n-1}, 0)$) die Tangentenebene an den Graphen G_h im Punkt (a_1, \dots, a_{n-1}) .

Ein Vektor $\mathbf{v} = (\mathbf{v}', v_n)$ (mit $\mathbf{v}' := (v_1, \dots, v_{n-1})$) liegt genau dann in der Tangentenebene, wenn $v_n = \nabla h(\mathbf{a}') \bullet \mathbf{v}'$ ist. Setzen wir

$$\alpha(t) := (\mathbf{a}' + t \cdot \mathbf{v}', h(\mathbf{a}' + t \cdot \mathbf{v}')),$$

so ist α ein differenzierbarer Weg in G_h mit $\alpha(0) = (\mathbf{a}', 0)$ und

$$\alpha'(0) = (\mathbf{v}', \nabla h(\mathbf{a}') \bullet \mathbf{v}') = \mathbf{v}.$$

Wir werden später mit dem Satz über implizite Funktionen zeigen, daß sich der Spezialfall immer durch eine einfache Vertauschung der Koordinaten herstellen läßt. \square

Beispiele:

1. Wir untersuchen die Funktion

$$f(x, y, z) := 3x^2 + 3y^2 + z^2$$

unter der Nebenbedingung $g(x, y, z) = 0$, mit $g(x, y, z) := x + y + z - 1$.

Hat f auf $N(g)$ in \mathbf{a} ein lokales Extremum, so muß es ein $\lambda \in \mathbb{R}$ geben, so daß gilt:

$$\begin{aligned} \nabla f(\mathbf{a}) &= \lambda \cdot \nabla g(\mathbf{a}) \\ \text{und } g(\mathbf{a}) &= 0. \end{aligned}$$

Das führt zu folgendem Gleichungssystem für $\mathbf{a} = (a, b, c)$:

$$\begin{aligned} 6a &= \lambda, \\ 6b &= \lambda, \\ 2c &= \lambda \\ \text{und } a + b + c &= 1. \end{aligned}$$

Es muß also $\frac{\lambda}{6} + \frac{\lambda}{6} + \frac{\lambda}{2} = 1$ sein. Damit ist $5\lambda = 6$ und $\lambda = \frac{6}{5}$. Einsetzen ergibt:

$$\mathbf{a} = (a, b, c) = \left(\frac{1}{5}, \frac{1}{5}, \frac{3}{5}\right).$$

Man überprüft sofort, daß \mathbf{a} tatsächlich auf $N(g)$ liegt. Außerdem ist

$$f(\mathbf{a}) = \frac{3}{25} + \frac{3}{25} + \frac{9}{25} = \frac{15}{25} = \frac{3}{5}.$$

Jetzt fängt der schwierige Teil an. Man muß irgendwie herausfinden, ob f in \mathbf{a} wirklich ein Extremum besitzt, und wenn ja, was für eins.

Es sollen hier drei verschiedene Methoden vorgestellt werden:

A) Berechnung von $f(\mathbf{a} + \mathbf{h})$ für kleines \mathbf{h} mit $g(\mathbf{a} + \mathbf{h}) = 0$.

Es ist

$$\begin{aligned} f(\mathbf{a} + \mathbf{h}) &= f\left(\frac{1}{5} + h_1, \frac{1}{5} + h_2, \frac{3}{5} + h_3\right) \\ &= 3\left(\frac{1}{5} + h_1\right)^2 + 3\left(\frac{1}{5} + h_2\right)^2 + \left(\frac{3}{5} + h_3\right)^2 \end{aligned}$$

$$\begin{aligned}
&= 3\left(\frac{1}{25} + \frac{2h_1}{5} + (h_1)^2\right) + 3\left(\frac{1}{25} + \frac{2h_2}{5} + (h_2)^2\right) + \left(\frac{9}{25} + \frac{6h_3}{5} + (h_3)^2\right) \\
&= \frac{3}{5} + \frac{6}{5}(h_1 + h_2 + h_3) + 3(h_1)^2 + 3(h_2)^2 + (h_3)^2 \\
&\geq \frac{3}{5} + \frac{6}{5}(h_1 + h_2 + h_3) \\
&= \frac{3}{5} = f(\mathbf{a}),
\end{aligned}$$

denn da $\mathbf{a} + \mathbf{h}$ auf $N(g)$ liegen soll, ist

$$1 = (a + h_1) + (b + h_2) + (c + h_3) = 1 + h_1 + h_2 + h_3,$$

also $h_1 + h_2 + h_3 = 0$.

Damit ist klar, daß f in \mathbf{a} ein Minimum unter der Nebenbedingung $g(\mathbf{x}) = 0$ besitzt.

Der Lagrangesche Multiplikator diene lediglich zum Auffinden des richtigen Punktes.

B) Abstrakte Argumentation:

Da $N(g)$ abgeschlossen ist, ist die Menge

$$M := \{\mathbf{x} \in \mathbb{R}^3 : g(\mathbf{x}) = 0 \text{ und } \|\mathbf{x}\| \leq 1\}$$

kompakt, und die stetige Funktion f nimmt auf M ein globales Minimum und ein globales Maximum an. Weil $\mathbf{a} \in M$, $(0, 0, 1) \in M$, $f(\mathbf{a}) = \frac{3}{5}$ und $f(0, 0, 1) = 1$ ist, kann \mathbf{a} nicht das Maximum sein, und im Minimum muß f einen Wert < 1 besitzen.

Also besitzt f sogar in der offenen Menge $B_1(\mathbf{0})$ unter der Nebenbedingung $g(\mathbf{x}) = 0$ ein globales (und damit erst recht lokales) Minimum. Dort kann man den Satz vom Lagrangeschen Multiplikator anwenden (die Offenheit des betrachteten Bereichs ist dabei wesentlich!). Es gibt aber nur einen Punkt, der dafür in Frage kommt, nämlich den Punkt \mathbf{a} .

C) Einsetzen der Nebenbedingung:

Diese Methode funktioniert nur dann, wenn man die Nebenbedingung nach einer Variablen auflösen kann. Das ist hier zum Glück der Fall:

$$g(x, y, z) = 0 \quad \Longleftrightarrow \quad z = 1 - x - y.$$

Nun suchen wir nach Extremwerten der Funktion

$$q(x, y) := f(x, y, 1 - x - y) = 3x^2 + 3y^2 + (1 - x - y)^2 = 4x^2 + 4y^2 + 2xy - 2x - 2y + 1.$$

Dabei können wir wie gewohnt mit dem Gradienten und der Hesseschen arbeiten. Es ist $\nabla q(x, y) = (8x + 2y - 2, 8y + 2x - 2)$. Man rechnet schnell nach, daß $\nabla q(x, y)$ genau dann verschwindet, wenn $x = \frac{1}{5}$ und $y = \frac{1}{5}$ ist, also $(x, y, z) = \mathbf{a}$.

Weiter ist $H_q(x, y) = \begin{pmatrix} 8 & 2 \\ 2 & 8 \end{pmatrix}$ positiv definit (unabhängig von (x, y)). Also muß q in $(\frac{1}{5}, \frac{1}{5})$ ein Minimum besitzen, und das bedeutet, daß f unter der Nebenbedingung $g(\mathbf{x}) = 0$ ein Minimum in \mathbf{a} besitzt.

2. Sei $A \in M_{n,n}(\mathbb{R})$ eine symmetrische Matrix und $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definiert durch $f(\mathbf{x}) := \mathbf{x} \circ A \circ \mathbf{x}^\top$. Da f stetig und

$$S^{n-1} = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$$

eine kompakte Menge ist, nimmt f auf S^{n-1} ein Maximum an. Dieses Maximum wollen wir suchen.

Die Nebenbedingung wird hier durch die Funktion $g(\mathbf{x}) := \mathbf{x} \circ \mathbf{x}^\top - 1$ definiert. Offensichtlich ist $\nabla g(\mathbf{x}) = 2\mathbf{x} \neq \mathbf{0}$ auf S^{n-1} , und es ist $\nabla f(\mathbf{x}) = 2\mathbf{x} \circ A$.

Wenn f in \mathbf{x} ein Maximum unter der Nebenbedingung $g(\mathbf{x}) = 0$ besitzt, so muß gelten:

$$\begin{aligned} \nabla f(\mathbf{x}) &= \lambda \cdot \nabla g(\mathbf{x}) \\ \text{und } g(\mathbf{x}) &= 0. \end{aligned}$$

Das führt zu dem Gleichungssystem:

$$\begin{aligned} 2\mathbf{x} \circ A &= 2\lambda\mathbf{x}, \\ \mathbf{x} \circ \mathbf{x}^\top &= 1. \end{aligned}$$

Insbesondere muß $(A - \lambda \cdot E_n) \circ \mathbf{x}^\top = 0$ sein, also \mathbf{x} Eigenvektor von A zum Eigenwert λ .

Weiter kann man sagen:

$$\lambda = \lambda \cdot (\mathbf{x} \circ \mathbf{x}^\top) = \mathbf{x} \circ (\lambda\mathbf{x})^\top = \mathbf{x} \circ (A \circ \mathbf{x}^\top) = f(\mathbf{x}).$$

Da die Existenz eines Maximums gesichert ist, haben wir gezeigt:

A besitzt wenigstens einen reellen Eigenwert.

Das ist ein neuer Beweis eines schon bekannten Resultats.

§7 Nichtlineare Probleme

Der Mittelwertsatz

Sei $B \subset \mathbb{R}^n$ offen und konvex, $f : B \rightarrow \mathbb{R}$ eine differenzierbare Funktion, $\mathbf{a}, \mathbf{b} \in B$ zwei Punkte.

Dann gibt es ein $t \in (0, 1)$ mit

$$f(\mathbf{b}) - f(\mathbf{a}) = \nabla f(\mathbf{a} + t(\mathbf{b} - \mathbf{a})) \bullet (\mathbf{b} - \mathbf{a}).$$

BEWEIS: Wir setzen $\alpha(t) := \mathbf{a} + t(\mathbf{b} - \mathbf{a})$ und $h(t) := f \circ \alpha(t)$. Dies ist eine auf $[0, 1]$ stetige und auf $(0, 1)$ differenzierbare Funktion. Nach dem Mittelwertsatz in einer Veränderlichen gibt es ein $t \in (0, 1)$, so daß

$$h(1) - h(0) = h'(t) \cdot (1 - 0) = h'(t)$$

ist. Es ist aber $h(1) - h(0) = f(\mathbf{b}) - f(\mathbf{a})$ und

$$h'(t) = \nabla f(\alpha(t)) \bullet \alpha'(t) = \nabla f(\mathbf{a} + t(\mathbf{b} - \mathbf{a})) \bullet (\mathbf{b} - \mathbf{a}).$$

□

Nun wenden wir uns vektorwertigen Funktionen von mehreren Veränderlichen zu.

Definition:

Sei $B \subset \mathbb{R}^n$ offen. Eine Abbildung $f = (f_1, \dots, f_m) : B \rightarrow \mathbb{R}^m$ heißt in $\mathbf{a} \in B$ differenzierbar, wenn alle Komponentenfunktionen f_1, \dots, f_m in \mathbf{a} differenzierbar sind.

Die durch $Df(\mathbf{a})(\mathbf{v}) := (Df_1(\mathbf{a})\mathbf{v}, \dots, Df_m(\mathbf{a})\mathbf{v})$ gegebene lineare Abbildung

$$Df(\mathbf{a}) : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

heißt die *Ableitung von f in \mathbf{a}* .

Definition:

Die Matrix $f'(\mathbf{a}) \in M_{m,n}(\mathbb{R})$, die $Df(\mathbf{a})$ beschreibt, nennt man *Funktionalmatrix* oder *Jacobische Matrix* von f in \mathbf{a} .

Gestalt der Funktionalmatrix

$$\text{Es ist } f'(\mathbf{a}) = \begin{pmatrix} \nabla f_1(\mathbf{a}) \\ \vdots \\ \nabla f_m(\mathbf{a}) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{pmatrix}.$$

BEWEIS: Die i -te Spalte der Funktionalmatrix $f'(\mathbf{a})$ ist gegeben durch

$$\vec{s}_i(f'(\mathbf{a})) = f'(\mathbf{a}) \circ \mathbf{e}_i^\top = (Df(\mathbf{a})(\mathbf{e}_i))^\top = \left(\frac{\partial f_1}{\partial x_i}(\mathbf{a}), \dots, \frac{\partial f_m}{\partial x_i}(\mathbf{a}) \right)^\top.$$

Das ergibt schon die Behauptung. □

Ist speziell $m = 1$, so ist $f'(\mathbf{a}) = \nabla f(\mathbf{a})$. Ist $n = 1$, so ist $f'(\mathbf{a})$ die gewöhnliche Ableitung, die wir korrekt eigentlich als Spaltenvektor schreiben müßten. Wir bleiben dann aber bei der alten Zeilenschreibweise und nehmen diese kleine Inkonsequenz bewußt in Kauf.

Definition:

Ist $n = m$, also $f'(\mathbf{x})$ eine quadratische Matrix, so heißt $J_f(\mathbf{x}) := \det f'(\mathbf{x})$ die *Funktionaldeterminante* oder *Jacobi-Determinante* von f in \mathbf{x} .

Beispiel:

Sei $f(x, y) := (e^{kx} \cos(y), e^{kx} \sin(y))$. Dann gilt:

$$f'(x, y) = \begin{pmatrix} ke^{kx} \cos(y) & -e^{kx} \sin(y) \\ ke^{kx} \sin(y) & e^{kx} \cos(y) \end{pmatrix}$$

und

$$J_f(x, y) = ke^{2kx} \cos^2(y) + ke^{2kx} \sin^2(y) = ke^{2kx}.$$

Wir können jetzt auch die Kettenregel verallgemeinern:

Erweiterte Kettenregel

Sei $B \subset \mathbb{R}^n$ offen, $f : B \rightarrow \mathbb{R}^m$ stetig differenzierbar, $G \subset \mathbb{R}^m$ offen mit $f(B) \subset G$ und $g : G \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion.

Dann ist auch $g \circ f : B \rightarrow \mathbb{R}$ stetig differenzierbar, und es gilt für $\mathbf{a} \in B$:

$$\nabla(g \circ f)(\mathbf{a}) = \nabla g(f(\mathbf{a})) \circ f'(\mathbf{a}).$$

BEWEIS: Sei $f = (f_1, \dots, f_m)$ und $\alpha_i : \mathbb{R} \rightarrow \mathbb{R}^n$ definiert durch

$$\alpha_i(t) := \mathbf{a} + t\mathbf{e}_i, \quad i = 1, \dots, n.$$

Daß f in \mathbf{a} partiell differenzierbar ist, bedeutet, daß alle Funktionen $f_j \circ \alpha_i$ in $t = 0$ differenzierbar sind. Sei $I = (-\varepsilon, \varepsilon) \subset \mathbb{R}$ so gewählt, daß $\alpha_i(I) \subset B$ für alle i gilt. Dann ist $\beta_i : I \rightarrow \mathbb{R}^m$ mit

$$\beta_i(t) := f \circ \alpha_i(t) = f(\mathbf{a} + t\mathbf{e}_i)$$

für jedes i ein in $t = 0$ differenzierbarer Weg in G , und es gilt:

$$\beta_i'(0) = ((f_1 \circ \alpha_i)'(0), \dots, (f_m \circ \alpha_i)'(0)) = \left(\frac{\partial f_1}{\partial x_i}(\mathbf{a}), \dots, \frac{\partial f_m}{\partial x_i}(\mathbf{a}) \right).$$

Aus der speziellen Kettenregel folgt nun: $(g \circ f) \circ \alpha_i = g \circ \beta_i$ ist ebenfalls in $t = 0$ differenzierbar, mit

$$\left. \frac{d}{dt} \right|_0 (g \circ f(\mathbf{a} + t\mathbf{e}_i)) = (g \circ \beta_i)'(0) = \nabla g(f(\mathbf{a})) \bullet \beta_i'(0).$$

Also ist $g \circ f$ in \mathbf{a} partiell differenzierbar, und es gilt:

$$\frac{\partial}{\partial x_i} (g \circ f)(\mathbf{a}) = \sum_{j=1}^m \frac{\partial g}{\partial y_j} (f(\mathbf{a})) \cdot \frac{\partial f_j}{\partial x_i}(\mathbf{a}).$$

Weil f und g nach Voraussetzung stetig differenzierbar sind, hängen auch die partiellen Ableitungen von $g \circ f$ stetig von \mathbf{a} ab.

In Matrixschreibweise lautet das Ergebnis: $\nabla(g \circ f)(\mathbf{a}) = \nabla g(f(\mathbf{a})) \circ f'(\mathbf{a})$. \square

Beispiel:

Sei $f(r, \varphi) = (r \cos(\varphi), r \sin(\varphi))$. Dann ist $f : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}^2$ eine differenzierbare Abbildung. Sie ordnet jedem Paar (r, φ) den Punkt (x, y) mit den Polarkoordinaten r und φ zu.

Sei $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ irgendeine differenzierbare Funktion. Dann gilt:

$$\begin{aligned} \frac{\partial(g \circ f)}{\partial r} &= \frac{\partial g}{\partial x} \cos(\varphi) + \frac{\partial g}{\partial y} \sin(\varphi) \\ \text{und} \\ \frac{\partial(g \circ f)}{\partial \varphi} &= -r \frac{\partial g}{\partial x} \sin(\varphi) + r \frac{\partial g}{\partial y} \cos(\varphi). \end{aligned}$$

Nun wollen wir die Kettenregel noch ein bißchen weiter verallgemeinern:

Ist $g : \mathbb{R}^m \rightarrow \mathbb{R}^k$ eine stetig differenzierbare Abbildung, so ist auch $g \circ f = (g_1 \circ f, \dots, g_k \circ f)$ stetig differenzierbar, und es gilt:

$$(g \circ f)'(\mathbf{a}) = \begin{pmatrix} \nabla(g_1 \circ f)(\mathbf{a}) \\ \vdots \\ \nabla(g_k \circ f)(\mathbf{a}) \end{pmatrix} = \begin{pmatrix} \nabla g_1(f(\mathbf{a})) \circ f'(\mathbf{a}) \\ \vdots \\ \nabla g_k(f(\mathbf{a})) \circ f'(\mathbf{a}) \end{pmatrix} = g'(f(\mathbf{a})) \circ f'(\mathbf{a}).$$

Zusammengefaßt:

Allgemeine Kettenregel

Sei $B \subset \mathbb{R}^n$ offen, $f : B \rightarrow \mathbb{R}^m$ stetig differenzierbar, $U \subset \mathbb{R}^m$ offen, $f(B) \subset U$ und $g : U \rightarrow \mathbb{R}^k$ stetig differenzierbar. Dann gilt in jedem $\mathbf{x} \in B$:

$$(g \circ f)'(\mathbf{x}) = g'(f(\mathbf{x})) \circ f'(\mathbf{x}).$$

Folgerung

Ist $n = m = k$, so ist $J_{g \circ f}(\mathbf{x}) = J_g(f(\mathbf{x})) \cdot J_f(\mathbf{x})$.

Beispiele:

1. Sei $f(\mathbf{x}) := \mathbf{x} + \mathbf{c}$ eine Translation. Dann ist $f'(\mathbf{x}) = E_n$ und $(g \circ f)'(\mathbf{x}) = g'(\mathbf{x} + \mathbf{c})$.
2. Sei $A \in M_{m,n}(\mathbb{R})$ und $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ die zugehörige lineare Abbildung, mit

$$f_A(\mathbf{x}) = (A \circ \mathbf{x}^\top)^\top = \mathbf{x} \circ A^\top.$$

Dann ist $f_A = (\lambda_1, \dots, \lambda_m)$, wobei die Komponenten $\lambda_j(\mathbf{x}) = \mathbf{z}_j(A) \circ \mathbf{x}^\top$ Linearformen sind. Also ist $D\lambda_j(\mathbf{x}) = \lambda_j$ und

$$\nabla \lambda_j(\mathbf{x}) = (\lambda_j(\mathbf{e}_1), \dots, \lambda_j(\mathbf{e}_n)) = \mathbf{z}_j(A).$$

Also ist $f'_A(\mathbf{x}) = A$ für jedes $\mathbf{x} \in \mathbb{R}^n$.

Ist $g : \mathbb{R}^m \rightarrow \mathbb{R}^k$ eine beliebige differenzierbare Abbildung, so ist

$$(g \circ f_A)'(\mathbf{x}) = g'(\mathbf{x} \circ A^\top) \circ A.$$

Ist umgekehrt $h : \mathbb{R}^q \rightarrow \mathbb{R}^n$ differenzierbar, so ist

$$(f_A \circ h)'(\mathbf{u}) = A \circ h'(\mathbf{u}).$$

Als nächstes wollen wir einen sehr allgemeinen Satz herleiten, der viele wichtige Anwendungen hat. Dazu noch ein paar Vorbemerkungen:

Sei X ein metrischer Raum, mit Metrik d . Wir wissen, wann eine Folge (x_n) in X konvergiert.

Wir können aber auch den Begriff der Cauchy-Folge auf metrische Räume übertragen: Eine Folge (x_n) in X heißt eine *Cauchy-Folge*, wenn gilt:

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N}, \text{ so daß für alle } n, m \geq n_0 \text{ gilt: } d(x_n, x_m) < \varepsilon.$$

Wie in \mathbb{R} folgt auch in metrischen Räumen, daß jede konvergente Folge eine Cauchy-Folge ist. Aber im Gegensatz zu der Situation in \mathbb{R} braucht in einem beliebigen metrischen Raum eine Cauchy-Folge nicht unbedingt zu konvergieren:

In $X := (0, 1)$ ist (x_n) mit $x_n := \frac{1}{n}$ eine Cauchy-Folge, aber sie hat keinen Grenzwert in X . In gewissem Sinne ist X unvollständig, es fehlt z.B. der Häufungspunkt 0.

Definition:

Ein metrischer Raum X heißt *vollständig*, wenn jede Cauchy-Folge in X auch einen Grenzwert in X besitzt.

Beispiele :

1. \mathbb{R} ist vollständig, das liegt an dem Vollständigkeitsaxiom.
2. Da man jede Vektorfolge in ihre Komponenten zerlegen kann, ist auch der \mathbb{R}^n vollständig.
3. Jede **kompakte** Teilmenge K des \mathbb{R}^n ist vollständig:

Eine Cauchy-Folge in K besitzt wegen der Kompaktheit eine in K konvergente Teilfolge. Aber aus der Cauchy-Eigenschaft folgert man leicht, daß auch die ursprüngliche Folge gegen den gleichen Grenzwert konvergiert.

4. Offene Teilmengen des \mathbb{R}^n sind – mit der Metrik des \mathbb{R}^n – nicht vollständig.

Banachscher Fixpunktsatz

Sei X ein vollständiger metrischer Raum, $f : X \rightarrow X$ eine stetige Abbildung. Wenn es ein λ mit $0 < \lambda < 1$ gibt, so daß

$$d(f(x), f(y)) \leq \lambda \cdot d(x, y)$$

für alle $x, y \in X$ ist, so besitzt f einen „Fixpunkt“, d.h. es gibt ein $x^* \in X$ mit $f(x^*) = x^*$. Dieser Fixpunkt ist eindeutig bestimmt.

Bemerkung: Man nennt die Abbildung f *kontrahierend*. Daß $\lambda < 1$ ist, ist entscheidend.

BEWEIS: Wir geben mit dem Beweis zugleich ein Konstruktionsverfahren an:

Sei $x_0 \in X$ beliebig gewählt. Die Punktfolge (x_n) sei induktiv durch

$$x_{n+1} := f(x_n)$$

definiert. Wir wollen zeigen, daß (x_n) gegen einen Fixpunkt konvergiert. Dazu schätzen wir ab:

$$\begin{aligned} d(x_n, x_{n+1}) &= d(f(x_{n-1}), f(x_n)) \\ &\leq \lambda \cdot d(x_{n-1}, x_n) \\ &\leq \dots \\ &\leq \lambda^n \cdot d(x_0, x_1). \end{aligned}$$

Da $0 < \lambda < 1$ ist, strebt der letzte Ausdruck gegen Null. Also kommen sich die Folgeglieder immer näher, die (x_n) bilden eine Cauchyfolge. Da X vollständig ist, konvergiert (x_n) gegen einen Punkt $x^* \in X$.

Weiter gilt:

$$\begin{aligned} d(f(x^*), x^*) &\leq d(f(x^*), f(x_n)) + d(f(x_n), x^*) \\ &\leq \lambda \cdot d(x^*, x_n) + d(x_{n+1}, x^*), \end{aligned}$$

und dieser Ausdruck wird beliebig klein. Das ist nur möglich, wenn $f(x^*) = x^*$ ist.

Zur Eindeutigkeit: Seien x und y zwei Fixpunkte. Ist $x \neq y$, so ist $d(x, y) > 0$ und daher

$$d(x, y) = d(f(x), f(y)) \leq \lambda \cdot d(x, y) < d(x, y).$$

Das kann aber nicht sein. □

Wir werden den Fixpunktsatz als erstes in folgender Form anwenden:

Sei $M \subset \mathbb{R}^n$, $g : M \rightarrow \mathbb{R}^n$ stetig, $\mathbf{x}_0 \in M$. Außerdem gebe es ein $r > 0$, so daß auch noch $X := \overline{B_r(\mathbf{x}_0)}$ ganz in M enthalten ist und folgende Bedingungen erfüllt sind:

1. (Kontraktionsbedingung)

Es gibt ein $\lambda \in \mathbb{R}$ mit $0 < \lambda < 1$, so daß die Ungleichung

$$\|g(\mathbf{x}) - g(\mathbf{y})\| \leq \lambda \cdot \|\mathbf{x} - \mathbf{y}\|$$

für alle $\mathbf{x}, \mathbf{y} \in M$ erfüllt ist.

2. (Startbedingung)

$$\|g(\mathbf{x}_0) - \mathbf{x}_0\| \leq (1 - \lambda) \cdot r.$$

Für $\|\mathbf{x} - \mathbf{x}_0\| \leq r$ ist dann

$$\begin{aligned} \|g(\mathbf{x}) - \mathbf{x}_0\| &\leq \|g(\mathbf{x}) - g(\mathbf{x}_0)\| + \|g(\mathbf{x}_0) - \mathbf{x}_0\| \\ &\leq \lambda \cdot \|\mathbf{x} - \mathbf{x}_0\| + (1 - \lambda) \cdot r \\ &\leq r. \end{aligned}$$

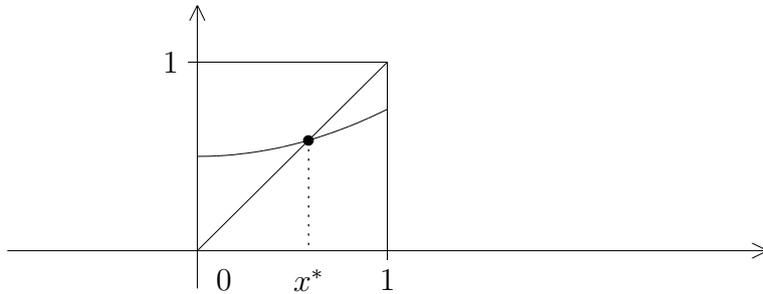
Also bildet g den vollständigen metrischen Raum X kontrahierend auf sich ab. Der Banachsche Fixpunktsatz besagt nun, daß es ein \mathbf{x}^* mit $\|\mathbf{x}^* - \mathbf{x}_0\| \leq r$ und $g(\mathbf{x}^*) = \mathbf{x}^*$ gibt.

Sei etwa $n = 1$ und $I := [x_0 - r, x_0 + r]$. Die Bedingung (2) ist für eine stetig differenzierbare Funktion $g : I \rightarrow \mathbb{R}$ besonders leicht herzustellen. Es gebe nämlich eine Konstante λ mit $0 < \lambda < 1$, so daß $|g'(x)| \leq \lambda$ für alle $x \in I$ ist. Anschaulich bedeutet das, daß g auf dem ganzen Intervall nicht zu stark wächst und nicht zu stark fällt.

Dann folgt aus dem Mittelwertsatz für beliebige Punkte $x, y \in I$: Es gibt ein ξ zwischen x und y , so daß $g(x) - g(y) = g'(\xi) \cdot (x - y)$ ist, also $|g(x) - g(y)| \leq \lambda \cdot |x - y|$.

Beispiel :

Sei $g(x) := \frac{1}{2} + \frac{1}{4}x^2$ auf $[0, 1]$. Wir wollen untersuchen, ob g einen Fixpunkt hat.



Es sei $x_0 = \frac{1}{2}$ und $r = 1$. Da $g'(x) = \frac{1}{2}x$ ist, können wir $\lambda = \frac{1}{2}$ setzen. Dann ist $|g'(x)| \leq \lambda$ und $|g(\frac{1}{2}) - \frac{1}{2}| = \frac{1}{16} < (1 - \lambda)r$. Also hat g genau einen Fixpunkt. Zu bestimmen ist dieser Fixpunkt ganz leicht, es gilt nämlich:

$$\begin{aligned} g(x) = x &\iff x^2 - 4x + 2 = 0 \\ &\iff x = 2 \pm \sqrt{2}. \end{aligned}$$

Da der Fixpunkt in $[0, 1]$ liegen soll, ist $x^* = 2 - \sqrt{2}$ der gesuchte Punkt.

Schiebt man den Startwert $g(x_0)$ zu weit nach oben oder läßt man $g'(x)$ zu groß werden, so bildet g nicht mehr das Intervall $[0, 1]$ auf sich ab, und es wird sinnlos, nach einem Fixpunkt zu fragen. Tatsächlich kann man sogar zeigen: Jede stetige Abbildung $g : [0, 1] \rightarrow [0, 1]$ besitzt einen Fixpunkt.

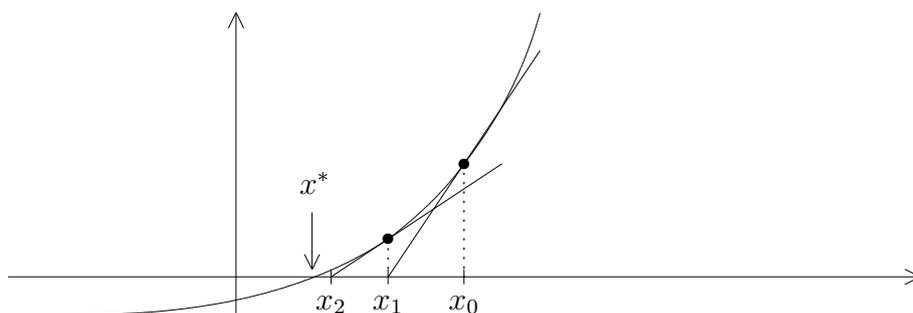
BEWEIS: Ist $g(0) = 0$ oder $g(1) = 1$, so ist man fertig. Andernfalls gilt für die stetige Funktion $h(x) := g(x) - x$, daß $h(0) > 0$ und $h(1) < 0$ ist. Nach dem Zwischenwertsatz muß es dann ein $x^* \in (0, 1)$ mit $h(x^*) = 0$ geben, also mit $g(x^*) = x^*$. \square

Der Vorteil des Banachschen Fixpunktsatzes liegt also nicht in seiner Existenzaussage, sondern in der Bereitstellung eines Approximationsverfahrens.

Als Anwendung betrachten wir das *Newton-Verfahren* zur Bestimmung einer Nullstelle einer differenzierbaren Funktion.

Wir beginnen mit dem Fall $n = 1$.

Sei I ein abgeschlossenes Intervall der Länge $2r$ und $f : I \rightarrow \mathbb{R}$ **zweimal stetig differenzierbar**. Außerdem setzen wir voraus, daß $f'(x) \neq 0$ für alle $x \in I$ ist. Wir suchen einen Punkt $x^* \in I$ mit $f(x^*) = 0$. Die Idee ist die folgende: Wir starten mit einem beliebigen Punkt $x_0 \in I$, und an Stelle von f betrachten wir die Tangente an f in x_0 , $T(x) := f(x_0) + f'(x_0) \cdot (x - x_0)$, und suchen die Nullstelle von T .



Es ist

$$\begin{aligned} T(x_1) = 0 &\iff 0 = f(x_0) + f'(x_0)(x_1 - x_0) \\ &\iff x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}. \end{aligned}$$

Die Anschauung sagt, daß x_1 näher an der gesuchten Nullstelle x^* ist als der Startpunkt x_0 . Also betrachten wir als nächstes die Tangente an f in x_1 . Deren Nullstelle x_2 erfüllt die Gleichung

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

Schrittweise konstruieren wir so eine Folge von Punkten x_1, x_2, x_3, \dots , von der wir hoffen, daß sie gegen die gesuchte Nullstelle konvergiert.

Definieren wir nun die (einmal stetig differenzierbare) Funktion

$$g(x) := x - \frac{f(x)}{f'(x)},$$

so ist $x_{n+1} = g(x_n)$, und es ist $f(x^*) = 0 \iff g(x^*) = x^*$. Unser Problem ist also ein Fixpunktproblem, und wir versuchen, den Fixpunktsatz auf die Funktion g anzuwenden. Zunächst ist

$$g'(x) = 1 - \frac{f'(x)^2 - f(x)f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)^2}.$$

Dieser Ausdruck sollte $\leq \lambda < 1$ auf dem ganzen Intervall sein. Weiter ist

$$g(x) - x = -\frac{f(x)}{f'(x)},$$

und um den Fixpunktsatz anwenden zu können, müssen wir noch die Startbedingung fordern:

$$|g(x_0) - x_0| \leq (1 - \lambda)r.$$

Das führt nun zu folgendem Satz:

Konvergenz des Newton-Verfahrens

Sei $f : I := [a, b] \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion und $f'(x) \neq 0$ für $x \in I$. Außerdem gebe es ein λ mit $0 < \lambda < 1$, so daß gilt:

$$\left| \frac{f(x)f''(x)}{f'(x)^2} \right| \leq \lambda \text{ für } x \in I \quad \text{und} \quad \left| \frac{f(x)}{f'(x)} \right| \leq (1 - \lambda) \frac{b - a}{2}.$$

Dann konvergiert die Folge x_0, x_1, x_2, \dots mit

$$x_0 \in I \text{ (beliebig), und } x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

gegen eine Nullstelle von f in I , und es gibt genau eine solche Nullstelle.

Außerdem hat man folgende Fehlerabschätzung:

$$\text{Ist } 0 < M \leq \min_I |f'(x)|, \text{ so ist } |x_n - x^*| \leq \frac{|f(x_n)|}{M}.$$

Ist f sogar dreimal stetig differenzierbar, so konvergiert die Folge „quadratisch“ gegen x^* , d.h. es ist

$$|x_{n+1} - x^*| \leq C \cdot (x_n - x^*)^2,$$

mit einer von n unabhängigen Konstanten $C > 0$.

BEWEIS: Unsere vorangegangenen Betrachtungen zeigen, daß die Funktion

$$g(x) := x - \frac{f(x)}{f'(x)}$$

die Voraussetzungen des Fixpunktsatzes erfüllt, also genau einen Fixpunkt x^* besitzt, und der ist dann natürlich die einzige Nullstelle von f .

Die Fehlerabschätzung erhält man aus dem Mittelwertsatz, angewandt auf f . Es ist nämlich

$$f(x_n) = f(x_n) - f(x^*) = f'(\xi) \cdot (x_n - x^*),$$

mit einem Punkt ξ zwischen x_n und x^* . Daraus folgt:

$$|x_n - x^*| = \left| \frac{f(x_n)}{f'(\xi)} \right| \leq \frac{|f(x_n)|}{M}.$$

Für die Abschätzung der Konvergenzgeschwindigkeit benutzen wir die Taylorformel. Ist f dreimal stetig differenzierbar, so ist g zweimal stetig differenzierbar, und es gilt:

$$g(x) = g(x^*) + g'(x^*)(x - x^*) + \frac{g''(c)}{2}(x - x^*)^2,$$

mit einem Punkt c zwischen x und x^* .

Es ist $g(x^*) = x^*$ und $g'(x^*) = 0$, weil $f(x^*) = 0$ ist. Setzen wir

$$C := \frac{1}{2} \cdot \max_I |g''(x)|,$$

so ist offensichtlich

$$|x_{n+1} - x^*| = |g(x_n) - x^*| \leq C \cdot |x_n - x^*|^2.$$

Damit ist alles gezeigt. □

Der Nachteil des obigen Satzes besteht darin, daß seine Voraussetzungen schwer nachzuprüfen sind. Wir geben daher noch eine einfachere Version an.

Einfache Voraussetzungen für das Newton-Verfahren

Sei $f : I := [a, b] \rightarrow \mathbb{R}$ dreimal stetig differenzierbar, $f(a) < 0$ und $f(b) > 0$. Außerdem gebe es Konstanten $C > 0$ und $\delta > 0$, so daß auf ganz I gilt:

$$f'(x) \geq \delta \quad \text{und} \quad 0 \leq f''(x) \leq C.$$

Wählt man dann den Startpunkt $x_0 \in I$ so, daß $f(x_0) > 0$ ist, so konvergiert die oben definierte Newton-Folge x_0, x_1, x_2, \dots gegen eine Nullstelle von f , und es gelten die gleichen Abschätzungen wie im vorigen Satz.

BEWEIS: Nach dem Zwischenwertsatz hat f auf $[a, b]$ eine Nullstelle, und da f streng monoton wachsend ist, kann es nur eine solche Nullstelle x^* in I geben. Ist $f(x_0) > 0$, so muß $x_0 > x^*$ sein.

Ist $x > x^*$, so ist $f(x) > 0$ und $f'(x) > 0$, also $g(x) = x - \frac{f(x)}{f'(x)} < x$. Das bedeutet, daß die Newton-Folge (mit $x_{n+1} = g(x_n)$) monoton fällt.

Behauptung: $x_n \geq x^*$ für alle n .

BEWEIS dafür durch Induktion nach n : Für $n = 0$ ist die Aussage wahr. Ist nun auch $x_n \geq x^*$, so betrachten wir die Funktion

$$\varphi(x) := f(x) - f(x_n) - f'(x_n)(x - x_n).$$

Es ist $\varphi'(x) = f'(x) - f'(x_n) \leq 0$ für $x \leq x_n$, denn nach Voraussetzung ist f' monoton wachsend. Das bedeutet aber, daß φ monoton fallend ist.

Da $\varphi(x_n) = 0$ ist, muß $\varphi(x) \geq 0$ für $x \leq x_n$ sein. Insbesondere ist auch $\varphi(x_{n+1}) \geq 0$. Aber da $x_{n+1} = g(x_n)$ ist, ist $\varphi(x_{n+1}) = f(x_{n+1})$. Daß $f(x_{n+1}) \geq 0$ ist, bedeutet aber, daß $x_{n+1} \geq x^*$ ist. Das war zu beweisen.

Nach dem Satz von der monotonen Konvergenz konvergiert (x_n) monoton fallend gegen ein x^{**} . Also konvergiert $g(x_n)$ gegen $g(x^{**})$. Andererseits konvergiert $g(x_n) = x_{n+1}$ auch gegen x^{**} . Das ist nur möglich, wenn $g(x^{**}) = x^{**}$ ist, also $f(x^{**}) = 0$. Da es nur eine Nullstelle gibt, muß $x^{**} = x^*$ sein. Das zeigt die Konvergenz der Newton-Folge gegen eine Nullstelle von f .

Da $f'(x)$ nach unten durch eine positive Konstante beschränkt ist und $f(x)$ in der Nähe von x^* beliebig klein wird, lassen sich die Voraussetzungen des Newton-Verfahrens in einer genügend kleinen Umgebung von x^* herstellen, und dort beweist man die Abschätzungen. □

Beispiele :

1. Sei $f(x) := x^2 - c$, $c > 1$. Wir suchen die Nullstelle \sqrt{c} von f mit Hilfe des Newton-Verfahrens.

Auf $[1, c]$ erfüllt f die Voraussetzungen von oben:

$f(1) < 0$, $f(c) > 0$, $f'(x) = 2x$, also $f'(x) \geq 2 > 0$ und $0 < f''(x) \leq 2$. Die Newton-Folge (x_n) mit dem Startwert $x_0 := c$ und

$$x_{n+1} = x_n - \frac{x_n^2 - c}{2x_n} = \frac{1}{2} \left(x_n + \frac{c}{x_n} \right)$$

muß also gegen $x^* = \sqrt{c}$ konvergieren.

Im Falle $c = 2$ erhält man z.B.:

n	x_n	$2/x_n$
0	2	1
1	1.5	1.33333
2	1.41666	1.41177
3	1.41421	1.41421

Das ergibt $\sqrt{2} \approx 1.41421$.

2. Sei $f(x) := x^3 + 2x - 5$, also $f'(x) = 3x^2 + 2$, $f''(x) = 6x$.

Hier ist $f(1) = -2 < 0$, $f(2) = 7 > 0$, $f'(x) \geq 5 > 0$ und $6 \leq f''(x) \leq 12$. Wir suchen die Nullstelle x^* von f in $[1, 2]$, deren Existenz schon durch den Zwischenwertsatz gesichert ist. Da die nötigen Voraussetzungen erfüllt sind, konvergiert die Newton-Folge mit $x_0 = 2$ gegen x^* . Außerdem haben wir die Fehlerabschätzung

$$|x_n - x^*| \leq \frac{1}{5} |f(x_n)|.$$

Wir erhalten:

n	x_n	$\frac{1}{5} f(x_n) $
0	2	1.4
1	1.5	0.275
2	1.342857	0.02144
3	1.32838	$< 10^{-3}$
4	1.32826	≈ 0

Damit ist $x^* \approx 1.3283$.

Bevor wir uns dem Fall $n > 1$ zuwenden können, müssen wir noch die Norm einer Matrix einführen.

Für eine Matrix

$$A = \left(a_{ij} \mid \begin{array}{l} i = 1, \dots, n \\ j = 1, \dots, m \end{array} \right) \in M_{n,m}(\mathbb{R})$$

setzen wir

$$\|A\| := \sqrt{\sum_{i,j} a_{ij}^2}.$$

Das ist nichts anderes, als wenn man A als ein Element des $\mathbb{R}^{n \cdot m}$ auffaßt und die gewöhnliche euklidische Norm von A bildet. Nun gilt:

1. $\|A + B\| \leq \|A\| + \|B\|$.
2. $\|\lambda \cdot A\| \leq |\lambda| \cdot \|A\|$ für $\lambda \in \mathbb{R}$.
3. $\|A\| = 0 \iff A = 0$.
4. Ist $A \in M_{n,m}(\mathbb{R})$ und $B \in M_{m,k}(\mathbb{R})$, so ist $\|A \circ B\| \leq \|A\| \cdot \|B\|$.

Diese Aussage muß noch bewiesen werden:

$A \circ B$ hat an der Stelle (i, j) den Eintrag

$$\sum_{k=1}^m a_{ik} b_{kj} = \mathbf{z}_i(A) \circ \vec{s}_j(B) = \mathbf{z}_i(A) \bullet \mathbf{s}_j(B).$$

Mit Hilfe der Schwarzschen Ungleichung folgt dann:

$$\begin{aligned} \|A \circ B\|^2 &= \sum_{i,j} (\mathbf{z}_i(A) \bullet \mathbf{s}_j(B))^2 \\ &\leq \sum_{i,j} \|\mathbf{z}_i(A)\|^2 \cdot \|\mathbf{s}_j(B)\|^2 \\ &= \left(\sum_i \|\mathbf{z}_i(A)\|^2 \right) \cdot \left(\sum_j \|\mathbf{s}_j(B)\|^2 \right) \\ &= \|A\|^2 \cdot \|B\|^2. \end{aligned}$$

Anwendung:

Verallgemeinerter Mittelwertsatz

Sei $B \subset \mathbb{R}^n$ offen und konvex, $f : B \rightarrow \mathbb{R}^m$ differenzierbar, $\mathbf{a}, \mathbf{b} \in B$. Dann gibt es einen Punkt \mathbf{z} auf der Verbindungsstrecke von \mathbf{a} und \mathbf{b} , so daß gilt:

$$\|f(\mathbf{b}) - f(\mathbf{a})\| \leq \|f'(\mathbf{z})\| \cdot \|\mathbf{b} - \mathbf{a}\|.$$

BEWEIS: Sei $\mathbf{u} := f(\mathbf{b}) - f(\mathbf{a}) \in \mathbb{R}^m$. Ist $\mathbf{u} = \mathbf{0}$, so ist die gewünschte Ungleichung trivialerweise erfüllt. Sei also $\mathbf{u} \neq \mathbf{0}$ und $h : [0, 1] \rightarrow \mathbb{R}$ definiert durch

$$h(t) := \mathbf{u} \bullet f(\mathbf{a} + t(\mathbf{b} - \mathbf{a})).$$

Dann ist $h'(t) = \mathbf{u} \bullet (f'(\mathbf{a} + t(\mathbf{b} - \mathbf{a})) \circ (\mathbf{b} - \mathbf{a})^\top)$. Nach dem Mittelwertsatz in einer Veränderlichen gibt es ein $\xi \in (0, 1)$ mit $h(1) - h(0) = h'(\xi)$. Also ist

$$\|\mathbf{u}\|^2 = |\mathbf{u} \bullet (f(\mathbf{b}) - f(\mathbf{a}))|$$

$$\begin{aligned}
&= |h(1) - h(0)| \\
&= |h'(\xi)| \\
&= |\mathbf{u} \bullet (f'(\mathbf{a} + \xi(\mathbf{b} - \mathbf{a})) \circ (\mathbf{b} - \mathbf{a})^\top)| \\
&\leq \|\mathbf{u}\| \cdot \|f'(\mathbf{a} + \xi(\mathbf{b} - \mathbf{a})) \circ (\mathbf{b} - \mathbf{a})^\top\| \\
&\leq \|\mathbf{u}\| \cdot \|f'(\mathbf{a} + \xi(\mathbf{b} - \mathbf{a}))\| \cdot \|\mathbf{b} - \mathbf{a}\|.
\end{aligned}$$

Setzt man $\mathbf{z} := \mathbf{a} + \xi(\mathbf{b} - \mathbf{a})$ und teilt die Ungleichung durch $\|\mathbf{u}\|$, so erhält man die Behauptung. \square

Folgerung

Sei $B \subset \mathbb{R}^n$ offen und zusammenhängend, $f : B \rightarrow \mathbb{R}^m$ differenzierbar und $f'(\mathbf{x}) \equiv \mathbf{0}$. Dann ist f konstant.

BEWEIS: Sei $\mathbf{x}_0 \in B$, $\mathbf{c} := f(\mathbf{x}_0) \in \mathbb{R}^m$ und

$$M := \{\mathbf{x} \in B \mid f(\mathbf{x}) = \mathbf{c}\}.$$

Dann ist M nicht leer. Ist $\mathbf{y} \in M$, so ist auch $f(\mathbf{y}) = \mathbf{c}$, und außerdem gibt es eine offene und konvexe Umgebung U von \mathbf{y} , die noch ganz in B liegt. Nach dem verallgemeinerten Mittelwertsatz gilt dann für beliebiges $\mathbf{x} \in U$:

$$\|f(\mathbf{x}) - f(\mathbf{y})\| = 0, \quad \text{also } f(\mathbf{x}) = f(\mathbf{y}) = \mathbf{c}.$$

Das bedeutet, daß U in M enthalten ist. Weil $\mathbf{y} \in M$ beliebig gewählt wurde, folgt, daß M offen ist.

Andererseits ist – weil f stetig ist – auch die Menge

$$B \setminus M = \{\mathbf{x} \in B \mid f(\mathbf{x}) \neq \mathbf{c}\}$$

offen. Nach einem von uns bewiesenen Satz über zusammenhängende Mengen muß nun $M = B$ sein, also $f(\mathbf{x}) \equiv \mathbf{c}$ auf ganz B . \square

Nun kommen wir zum Newton-Verfahren im \mathbb{R}^n .

Sei $B \subset \mathbb{R}^n$ offen, $f : B \rightarrow \mathbb{R}^m$ zweimal stetig differenzierbar. Wir wollen versuchen, die Gleichung

$$f(\mathbf{x}) = \mathbf{0}$$

zu lösen. Dabei gehen wir wie im Falle $n = 1$ vor. Allerdings lassen wir zunächst die Konvergenzfrage außer Acht. Wir nehmen an, es gebe eine (eindeutig bestimmte) Lösung \mathbf{x}^* , und wir wählen einen Startpunkt \mathbf{x}_0 , von dem wir glauben, daß er hinreichend nahe bei \mathbf{x}^* liegt. Dann suchen wir eine Nullstelle der „Tangenten-Abbildung“

$$T(\mathbf{x}) := f(\mathbf{x}_0) + Df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = f(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0) \circ f'(\mathbf{x}_0)^\top.$$

Offensichtlich gilt:

$$T(\mathbf{x}_1) = \mathbf{0} \iff Df(\mathbf{x}_0)(\mathbf{x}_1) = Df(\mathbf{x}_0)(\mathbf{x}_0) - f(\mathbf{x}_0).$$

Man muß also ein lineares Gleichungssystem lösen, um \mathbf{x}_1 zu bestimmen. Man kann dieses System auch umschreiben:

1. $Df(\mathbf{x}_0)(\mathbf{z}_1) = -f(\mathbf{x}_0)$.
2. $\mathbf{x}_1 := \mathbf{x}_0 + \mathbf{z}_1$.

Wie im Falle $n = 1$ fährt man nun fort und konstruiert eine Folge $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots$, wobei die Rekursion in zwei Schritten erfolgt:

1. Bestimme \mathbf{z}_{k+1} aus der Gleichung

$$Df(\mathbf{x}_k)(\mathbf{z}_{k+1}) = -f(\mathbf{x}_k).$$

2. Setze $\mathbf{x}_{k+1} := \mathbf{x}_k + \mathbf{z}_{k+1}$.

Ist $Df(\mathbf{x}_k)$ ein Isomorphismus (also $f'(\mathbf{x}_k)$ invertierbar), so ist $\mathbf{z}_{k+1} = -Df(\mathbf{x}_k)^{-1}(f(\mathbf{x}_k))$. Auf jeden Fall muß bei jedem Schritt entweder ein lineares Gleichungssystem gelöst oder eine Matrix invertiert werden. Das ist kaum praktikabel, und Bedingungen für die Konvergenz des Verfahrens sind schwer zu formulieren und noch schwerer zu verifizieren.

Man hat sich daher ein einfacheres Verfahren ausgedacht.

Vereinfachtes Newton-Verfahren

Sei $B \subset \mathbb{R}^n$ offen, $f : B \rightarrow \mathbb{R}^n$ zweimal stetig differenzierbar und $f(\mathbf{x}^*) = \mathbf{0}$ für ein $\mathbf{x}^* \in B$. (Die Existenz einer Nullstelle und ihre ungefähre Lage ist meist schon vorher bekannt). Außerdem gebe es eine Umgebung $U(\mathbf{x}^*) \subset B$, so daß $f'(\mathbf{x})$ für alle $\mathbf{x} \in U$ invertierbar ist.

$\mathbf{x}_0 \in U$ sei genügend nahe bei \mathbf{x}^* gewählt, $A := f'(\mathbf{x}_0)$ gesetzt und die Folgen \mathbf{z}_k und \mathbf{x}_k rekursiv definiert durch

1. $\mathbf{z}_{k+1}^\top := -A^{-1} \circ f(\mathbf{x}_k)^\top$.
2. $\mathbf{x}_{k+1} := \mathbf{x}_k + \mathbf{z}_{k+1}$.

Dann konvergiert (\mathbf{x}_k) gegen \mathbf{x}^* .

BEWEIS: Wir können ein ε mit $0 < \varepsilon < 1$ finden, so daß $f'(\mathbf{x})$ für $\|\mathbf{x} - \mathbf{x}^*\| \leq \varepsilon$ invertierbar ist. Dann gibt es eine Konstante $c > 0$, so daß dort $\|f'(\mathbf{x})\| > c$ ist.

Als nächstes sei eine Zahl r mit $0 < r < \frac{\varepsilon}{2}$ so gewählt, daß gilt:

$$\|f'(\mathbf{y})^{-1} \circ (f'(\mathbf{y}) - f'(\mathbf{x}))\| < \frac{1}{2} \text{ für } \|\mathbf{x} - \mathbf{y}\| \leq r \text{ und } \mathbf{x}, \mathbf{y} \in \overline{B_\varepsilon(\mathbf{x}^*)}.$$

Und schließlich sei noch ein δ mit $0 < \delta < r$ und folgender Eigenschaft gewählt:

$$\|f(\mathbf{x})\| < \frac{r \cdot c}{2} \text{ für } \|\mathbf{x} - \mathbf{x}^*\| < \delta.$$

Es ist offensichtlich, daß man geeignete Zahlen ε , r und δ finden kann.

Nun dürfen wir den Startwert \mathbf{x}_0 beliebig in $B_\delta(\mathbf{x}^*)$ wählen. Das ist das, was in der Formulierung des Satzes unter „genügend nahe“ verstanden wird. Natürlich ist δ nicht explizit bekannt, aber mit etwas Glück und hilfreichen Hinweisen des Experimentators wird man schon einen guten Startpunkt finden. Mit $A := f'(\mathbf{x}_0)$ gilt dann:

1. $\|\mathbf{x}^* - \mathbf{x}_0\| < r$.
2. Für $\mathbf{x} \in B_r(\mathbf{x}_0)$ ist $\|\mathbf{x} - \mathbf{x}_0\| < r$ und $\|\mathbf{x} - \mathbf{x}^*\| < \varepsilon$, also

$$\|A^{-1} \circ (A - f'(\mathbf{x}))\| \leq \frac{1}{2}.$$

3. $\|f(\mathbf{x}_0) \circ (A^{-1})^\top\| \leq \|f(\mathbf{x}_0)\| \cdot \|A^{-1}\| < \frac{r \cdot c}{2} \cdot \frac{1}{c} = \frac{r}{2}$.

Sei $g : B \rightarrow \mathbb{R}^n$ definiert durch

$$g(\mathbf{x}) := \mathbf{x} - (A^{-1} \circ f(\mathbf{x}))^\top = \mathbf{x} - f(\mathbf{x}) \circ (A^{-1})^\top.$$

Dann ist

$$g'(\mathbf{x}) = E_n - A^{-1} \circ f'(\mathbf{x}) = A^{-1} \circ (A - f'(\mathbf{x})),$$

also $\|g'(\mathbf{x})\| < \frac{1}{2}$ und $\|g(\mathbf{x}_0) - \mathbf{x}_0\| < \frac{r}{2}$. Setzen wir $\lambda := \frac{1}{2}$, so ist

$$\|g(\mathbf{x}_0) - \mathbf{x}_0\| < (1 - \lambda)r,$$

und nach dem verallgemeinerten Mittelwertsatz ist auch

$$\|g(\mathbf{y}) - g(\mathbf{x})\| \leq \lambda \cdot \|\mathbf{y} - \mathbf{x}\|, \text{ für } \mathbf{x}, \mathbf{y} \in \overline{B_r(\mathbf{x}_0)}.$$

Also erfüllt g die Voraussetzungen des Fixpunktsatzes, und die Newton-Folge konvergiert gegen \mathbf{x}^* . □

Als nächstes betrachten wir allgemeinere nichtlineare Gleichungen:

Sei $B \subset \mathbb{R}^n$ eine offene Menge und $f : B \rightarrow \mathbb{R}^m$ eine differenzierbare Abbildung, $\mathbf{y}_0 \in \mathbb{R}^m$ vorgegeben. Wir suchen nach Lösungen der Gleichung

$$f(\mathbf{x}) = \mathbf{y}_0.$$

Wir untersuchen erst einmal den Fall $n = m$ und beginnen mit $n = 1$.

Sei $f : I \rightarrow \mathbb{R}$ stetig differenzierbar. Wann ist im Falle einer Veränderlichen die Gleichung $f(x) = y$ überhaupt lösbar, und wann ist das eindeutig möglich? Ist f surjektiv, so gibt es immer eine Lösung. Ist f injektiv, so ist die Lösung stets eindeutig bestimmt. Beides ist z.B. dann erfüllt, wenn f streng monoton wachsend oder fallend ist, und das gilt wiederum, wenn $f'(x) \neq 0$ für alle $x \in I$ ist. Ist nur bekannt, daß $f'(t_0) \neq 0$ für ein $t_0 \in I$ ist, so gilt das wegen der Stetigkeit von f noch in einer ganzen Umgebung von t_0 , und dort existiert auch die Umkehrfunktion f^{-1} , durch die man die Gleichung lösen kann.

Sinngemäß läßt sich dieser Sachverhalt auf höhere Dimensionen übertragen:

Satz von der Umkehrabbildung

Sei $B \subset \mathbb{R}^n$ offen, $f : B \rightarrow \mathbb{R}^n$ stetig differenzierbar. Ist $\mathbf{x}_0 \in B$, $f(\mathbf{x}_0) = \mathbf{y}_0$ und $J_f(\mathbf{x}_0) \neq 0$, so gibt es offene Umgebungen $U(\mathbf{x}_0) \subset B$ und $V(\mathbf{y}_0) \subset \mathbb{R}^n$, so daß gilt:

1. $J_f(\mathbf{x}) \neq 0$ für alle $\mathbf{x} \in U$.
2. $f : U \rightarrow V$ ist bijektiv.
3. $f^{-1} : V \rightarrow U$ ist wieder differenzierbar.
4. $D(f^{-1})(\mathbf{y}) = (Df(f^{-1}(\mathbf{y})))^{-1}$ für alle $\mathbf{y} \in V$.

Bemerkung: Ist $J_f(\mathbf{x}) \neq 0$, so sagt man auch, f ist *regulär* in \mathbf{x} . Das hat zur Folge, daß $f'(\mathbf{x})$ eine invertierbare Matrix und die zugehörige lineare Abbildung $Df(\mathbf{x})$ ein Isomorphismus ist. Also kann man die Umkehrabbildung $(Df(\mathbf{x}))^{-1}$ bilden.

Der Satz von der Umkehrabbildung nimmt einen zentralen Platz in der Theorie der Funktionen von mehreren Veränderlichen ein! Sein Beweis ist nicht ganz einfach, wir wollen aber doch die entscheidenden Teile dieses Beweises besprechen.

BEWEIS: 1) Ohne Beweis geben wir an, daß es reicht, den Satz unter folgenden speziellen Annahmen zu beweisen:

Sei $\mathbf{x}_0 = \mathbf{0}$, $f(\mathbf{0}) = \mathbf{0}$ und $f'(\mathbf{0}) = E_n$, also $J_f(\mathbf{0}) = 1$.

Wir schreiben f in der Form $f = (f_1, \dots, f_n)$ und setzen

$$F(\mathbf{u}_1, \dots, \mathbf{u}_n) := \begin{pmatrix} \nabla f_1(\mathbf{u}_1) \\ \vdots \\ \nabla f_n(\mathbf{u}_n) \end{pmatrix}.$$

Dann ist $\det \circ F : B \times \dots \times B \rightarrow \mathbb{R}$ eine stetige Funktion und $\det(F(\mathbf{u}, \dots, \mathbf{u})) = J_f(\mathbf{u})$, und man kann eine Umgebung $U = U(\mathbf{0}) \subset B$ finden, so daß $\det F(\mathbf{u}_1, \dots, \mathbf{u}_n) \neq 0$ für $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$ ist. Insbesondere ist dann $J_f(\mathbf{u}) \neq 0$ für $\mathbf{u} \in U$.

2) Beweis der lokalen Injektivität:

Sei $f(\mathbf{a}) = f(\mathbf{b})$ für $\mathbf{a}, \mathbf{b} \in U$. Nach dem Mittelwertsatz gibt es Punkte \mathbf{z}_ν zwischen \mathbf{a} und \mathbf{b} , so daß gilt:

$$0 = f_\nu(\mathbf{b}) - f_\nu(\mathbf{a}) = \nabla f(\mathbf{z}_\nu) \circ (\mathbf{b} - \mathbf{a})^\top,$$

also $F(\mathbf{z}_1, \dots, \mathbf{z}_n) \circ (\mathbf{b} - \mathbf{a})^\top = \mathbf{0}^\top$. Da die Matrix $F(\mathbf{z}_1, \dots, \mathbf{z}_n)$ invertierbar ist, folgt: $\mathbf{a} = \mathbf{b}$.

3) Beweis der lokalen Surjektivität:

Wir müssen versuchen, die Gleichung $f(\mathbf{x}) = \mathbf{y}_0$ für \mathbf{y}_0 in der Nähe des Nullpunktes zu lösen. Dazu benutzen wir die Abbildung

$$\tilde{f}(\mathbf{x}) := f(\mathbf{x}) - \mathbf{y}_0.$$

Offensichtlich ist $f(\mathbf{x}) = \mathbf{y}_0 \iff \tilde{f}(\mathbf{x}) = \mathbf{0}$. Für die neue Gleichung machen wir den Newton-Ansatz, mit $\mathbf{x}_0 := \mathbf{0}$:

$$g(\mathbf{x}) := \mathbf{x} - (\tilde{f}'(\mathbf{0})^{-1} \circ \tilde{f}(\mathbf{x})^\top)^\top = \mathbf{x} - f(\mathbf{x}) + \mathbf{y}_0.$$

Offensichtlich ist dann $g(\mathbf{x}) = \mathbf{x} \iff f(\mathbf{x}) = \mathbf{y}_0$.

Es ist $g'(\mathbf{x}) = E_n - f'(\mathbf{x}) = f'(\mathbf{0}) - f'(\mathbf{x})$ und $\|g(\mathbf{0}) - \mathbf{0}\| = \|\mathbf{y}_0\|$. Wählt man \mathbf{x} und \mathbf{y}_0 jeweils nahe genug am Nullpunkt, so lassen sich die Bedingungen für den Fixpunktsatz sehr leicht herstellen. Also gibt es eine Lösung.

4) Die Differenzierbarkeit der Umkehrabbildung läßt sich nun nicht allzu schwer nachweisen, weil wir schon einen Kandidaten für die Ableitung $D(f^{-1})(\mathbf{y})$ haben, nämlich $(Df(f^{-1}(\mathbf{y})))^{-1}$. Auf das Vorführen der technischen Details verzichten wir hier aber. \square

Beispiele:

1. Die Polarkoordinaten $(x, y) = f(r, \varphi) = (r \cos(\varphi), r \sin(\varphi))$ haben wir schon an früherer Stelle betrachtet. Es ist

$$f'(r, \varphi) = \begin{pmatrix} \cos(\varphi) & -r \sin(\varphi) \\ \sin(\varphi) & r \cos(\varphi) \end{pmatrix},$$

also $J_f(r, \varphi) = r \cos^2(\varphi) + r \sin^2(\varphi) = r$. In jedem Punkt (r, φ) mit $r > 0$ und $\varphi \in \mathbb{R}$ ist f damit lokal umkehrbar.

$$f : \mathbb{R}_+ \times [0, 2\pi) \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}$$

ist sogar global umkehrbar.

2. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definiert durch

$$f(x, y) := (x^2 - y^2, 2xy).$$

Dann gilt:

$$f'(x, y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}, \text{ also } J_f(x, y) = 4(x^2 + y^2).$$

Damit ist $J_f(x, y) \neq 0$ für $(x, y) \neq (0, 0)$ und f überall außerhalb des Nullpunktes lokal umkehrbar.

f ist aber nicht global umkehrbar, denn es ist ja $f(-x, -y) = f(x, y)$

Nun wollen wir uns noch der Lösung einer Gleichung

$$f(x_1, \dots, x_n) = (y_1, \dots, y_m)$$

zuwenden, wobei f auf einer offenen Menge B des \mathbb{R}^n definiert und stetig differenzierbar ist und Werte in einem \mathbb{R}^m mit $m < n$ hat.

Im linearen Fall ergibt sich aus der Dimensionsformel sofort, daß die Gleichung nicht eindeutig lösbar ist, vielmehr erhält man entweder gar keine Lösung oder einen affinen

Lösungsraum. Im nichtlinearen Fall kennen wir die Situation auch schon, wenn $m = 1$ ist. Dann ist die Lösungsmenge M nämlich eine „Niveaufläche“:

$$M = F_{y_0}(f) := \{\mathbf{x} \in B \mid f(\mathbf{x}) = y_0\}.$$

Beispiel:

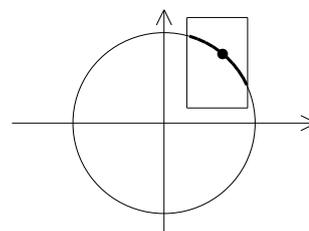
Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert durch $f(x, y) := x^2 + y^2 - 1$. Dann ist

$$\{(x, y) \in \mathbb{R}^2 \mid f(x, y) = 0\} = \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| = 1\} =: S^1$$

der Einheitskreis.

Ist $(a, b) \in S^1$ und etwa $a > 0$ und $b > 0$, so gilt in der Nähe von (a, b) :

$$\begin{aligned} f(x, y) = 0 &\iff y^2 = 1 - x^2 \\ &\iff y = g(x) := \sqrt{1 - x^2}. \end{aligned}$$



Die Lösungsmenge sieht also in der Nähe von (a, b) wie der Graph einer Funktion g aus. Es gibt eine Umgebung $U = V \times W$ von (a, b) , so daß gilt:

$$\{(x, y) \in U \mid f(x, y) = 0\} = \{(x, g(x)) \mid x \in V\}.$$

Man sagt auch, die Funktion g ist „implizit“ durch die Gleichung $f = 0$ gegeben.

Die Beziehung $f(x, g(x)) \equiv 0$ kann man übrigens benutzen, um die Ableitung von g besonders einfach zu berechnen. Mit der Kettenregel folgt nämlich:

$$0 = \frac{\partial f}{\partial x}(x, g(x)) \cdot 1 + \frac{\partial f}{\partial y}(x, g(x)) \cdot g'(x),$$

also

$$g'(x) = -\frac{f_x(x, g(x))}{f_y(x, g(x))} = -\frac{2x}{2g(x)} = -\frac{x}{g(x)}.$$

Bei dieser Umformung hätten wir natürlich erst einmal überprüfen müssen, ob $f_y(x, g(x))$ in der Nähe von $x = a$ nicht verschwindet. Tatsächlich ist $f_y(a, b) = 2b > 0$, weil wir das so gefordert haben. Wir werden sehen, daß das gerade die Bedingung für die Auflösbarkeit nach y ist.

Der Kreis S^1 ist eine so symmetrische Figur, daß nicht einzusehen ist, warum die Gleichung $f(x, y) = 0$ nicht überall eine implizite Funktion definieren sollte. Wenn wir etwa im Punkt $(1, 0)$ das Koordinatensystem um 90° drehen, dann sieht der Kreis auch dort lokal wie ein Graph aus, allerdings wie der Graph einer Funktion $x = h(y)$. Tatsächlich ist dort $h(y) = \sqrt{1 - y^2}$.

Wie kann man erkennen, nach welcher Variablen aufgelöst werden kann? Der Kreis kann überall dort als Graph einer Funktion $y = g(x)$ aufgefaßt werden, wo er keine vertikale Tangente besitzt, und er kann überall dort als Graph einer Funktion $x =$

$h(y)$ aufgefaßt werden, wo er keine horizontale Tangente besitzt. Nun ist S^1 eine Niveaulinie von f , und der Gradient von f steht jeweils auf der Tangenten senkrecht. Damit haben wir:

$$\begin{aligned}
 S^1 \text{ Graph von } y = g(x) &\iff \text{Tangente nicht vertikal} \\
 &\iff \nabla f(x, y) \text{ nicht horizontal} \\
 &\iff \frac{\partial f}{\partial y}(x, y) \neq 0 \\
 \text{und } S^1 \text{ Graph von } x = h(y) &\iff \text{Tangente nicht horizontal} \\
 &\iff \nabla f(x, y) \text{ nicht vertikal} \\
 &\iff \frac{\partial f}{\partial x}(x, y) \neq 0.
 \end{aligned}$$

Wir betrachten nun eine offene Menge $B \subset \mathbb{R}^k \times \mathbb{R}^m$ und eine stetig differenzierbare Abbildung $f : B \rightarrow \mathbb{R}^m$, also ein System von m nichtlinearen Gleichungen für $k + m$ Variable. Den Satz der ersten k Variablen x_1, \dots, x_k fassen wir zu einem Vektor \mathbf{x} , den der folgenden m Variablen x_{k+1}, \dots, x_{k+m} zu einem Vektor \mathbf{y} zusammen. Dann definiert man:

$$\frac{\partial f}{\partial \mathbf{x}} := \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_k} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_k} \end{pmatrix}$$

und

$$\frac{\partial f}{\partial \mathbf{y}} := \begin{pmatrix} \frac{\partial f_1}{\partial x_{k+1}} & \cdots & \frac{\partial f_1}{\partial x_{k+m}} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_{k+1}} & \cdots & \frac{\partial f_m}{\partial x_{k+m}} \end{pmatrix}.$$

Damit ist

$$f'(\mathbf{x}, \mathbf{y}) = \left(\frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{y}) \mid \frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}, \mathbf{y}) \right).$$

Satz über implizite Funktionen

Ist $f(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{0}$ und die Matrix $\frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0) \in M_{m,m}(\mathbb{R})$ regulär, so gibt es Umgebungen $U(\mathbf{x}_0)$, $V(\mathbf{y}_0)$ mit $U \times V \subset B$ und genau eine stetig differenzierbare Abbildung $g : U \rightarrow V$, so daß gilt:

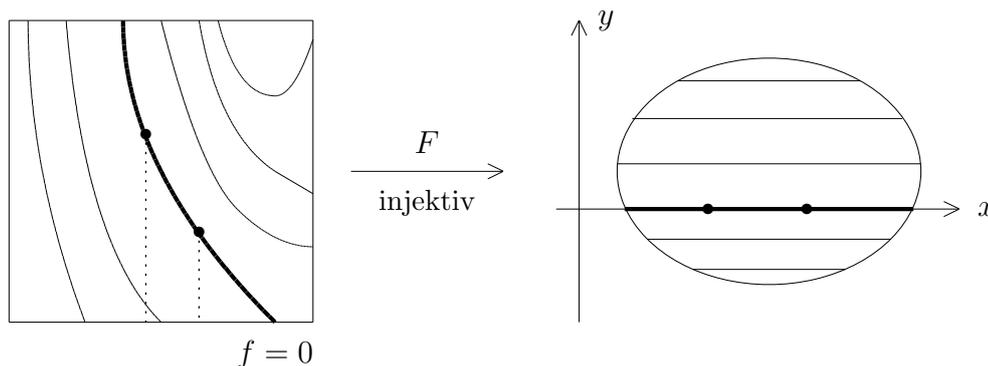
1. $g(\mathbf{x}_0) = \mathbf{y}_0$.
2. $f(\mathbf{x}, g(\mathbf{x})) \equiv \mathbf{0}$ auf U .
3. $g'(\mathbf{x}) = - \left(\frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}, g(\mathbf{x})) \right)^{-1} \circ \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}, g(\mathbf{x}))$ auf U .

BEWEIS: Sei $n := k + m$. Der Trick besteht darin, den Raum um $(\mathbf{x}_0, \mathbf{y}_0)$ herum so differenzierbar zu verbiegen, daß aus der Nullstellenmenge

$$N := \{(x_1, \dots, x_n) \mid f_1(x_1, \dots, x_n) = \dots = f_m(x_1, \dots, x_n) = 0\}$$

ein Ebenenstück wird. Zu diesem Zweck definieren wir $F : B \rightarrow \mathbb{R}^n$ durch

$$F(\mathbf{x}, \mathbf{y}) := (\mathbf{x}, f(\mathbf{x}, \mathbf{y})).$$



Dann ist

$$F'(\mathbf{x}_0, \mathbf{y}_0) = \left(\begin{array}{c|c} E_k & \mathbf{0} \\ \hline \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}_0, \mathbf{y}_0) & \frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0) \end{array} \right),$$

und daher

$$\det F'(\mathbf{x}_0, \mathbf{y}_0) = \det \frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0) \neq 0.$$

Das bedeutet, daß F in $(\mathbf{x}_0, \mathbf{y}_0)$ lokal umkehrbar ist. Wir setzen $H := F^{-1}$ (in der Nähe von $(\mathbf{x}_0, \mathbf{y}_0)$).

Offensichtlich ist $H(\mathbf{u}, \mathbf{v}) = (\mathbf{u}, h(\mathbf{u}, \mathbf{v}))$.

Weil $H(\mathbf{u}, \mathbf{0}) = (\mathbf{u}, h(\mathbf{u}, \mathbf{0}))$ ist, setzen wir

$$g(\mathbf{x}) := h(\mathbf{x}, \mathbf{0}).$$

Dann ist einerseits

$$F(\mathbf{x}, g(\mathbf{x})) = (\mathbf{x}, f(\mathbf{x}, g(\mathbf{x})))$$

und andererseits

$$F(\mathbf{x}, g(\mathbf{x})) = F(\mathbf{x}, h(\mathbf{x}, \mathbf{0})) = F \circ H(\mathbf{x}, \mathbf{0}) = (\mathbf{x}, \mathbf{0}).$$

Daraus folgt: $f(\mathbf{x}, g(\mathbf{x})) \equiv \mathbf{0}$.

Weil $F(\mathbf{x}_0, \mathbf{y}_0) = (\mathbf{x}_0, f(\mathbf{x}_0, \mathbf{y}_0)) = (\mathbf{x}_0, \mathbf{0})$ ist, ist

$$(\mathbf{x}_0, \mathbf{y}_0) = H(\mathbf{x}_0, \mathbf{0}) = (\mathbf{x}_0, h(\mathbf{x}_0, \mathbf{0})) = (\mathbf{x}_0, g(\mathbf{x}_0)),$$

also $g(\mathbf{x}_0) = \mathbf{y}_0$.

Schließlich ist g differenzierbar, und weil $f(\mathbf{x}, g(\mathbf{x})) \equiv \mathbf{0}$ ist, folgt mit der Kettenregel:

$$\mathbf{0} = \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}, g(\mathbf{x})) \circ E_k + \frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}, g(\mathbf{x})) \circ g'(\mathbf{x}),$$

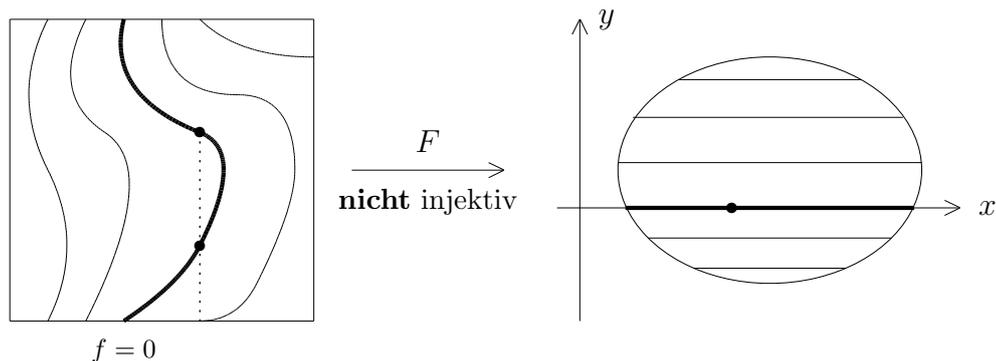
also

$$g'(\mathbf{x}) = - \left(\frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}, g(\mathbf{x})) \right)^{-1} \circ \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}, g(\mathbf{x})).$$

Man beachte hier die Reihenfolge bei der Matrizenmultiplikation!

Das Ganze gilt in der Nähe von $(\mathbf{x}_0, \mathbf{y}_0)$. Wählt man die Umgebungen U und V klein genug, so ist alles gezeigt. \square

Bemerkung: Das Verfahren kann nicht funktionieren, wenn man die Regularität von $\frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)$ nicht fordert. Man kann anschaulich sehen, daß die Abbildung F dann nicht mehr injektiv ist.



Häufig lassen sich dann aber die Koordinaten so vertauschen, daß anschließend doch die Voraussetzungen erfüllt sind.

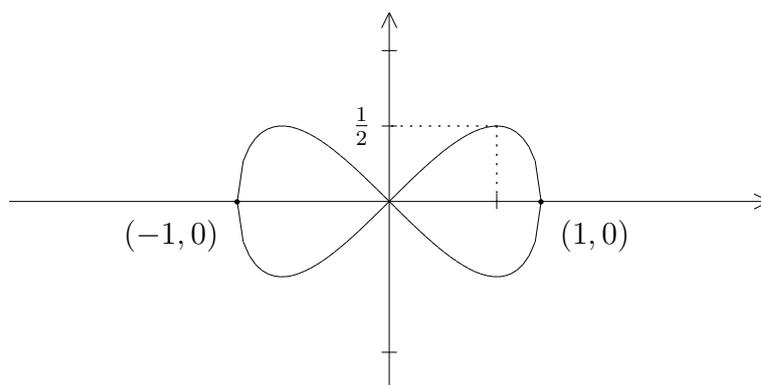
Beispiele :

1. Betrachten wir noch einmal den Kreis $S^1 = \{(x, y) \mid f(x, y) := x^2 + y^2 - 1 = 0\}$. Für $y \neq 0$ (also $x \neq \pm 1$) ist $\frac{\partial f}{\partial y}(x, y) = 2y \neq 0$. Also kann man den Satz über implizite Funktionen anwenden und die Gleichung $f(x, y) = 0$ lokal nach y auflösen: $y = g(x)$. Die Formel für die Ableitung von g ergibt hier:

$$g'(x) = - \frac{f_x(x, g(x))}{f_y(x, g(x))} = - \frac{x}{g(x)} = - \frac{x}{\sqrt{1-x^2}}.$$

Leider ist die Auflösung nicht immer so schön konkret durchführbar!

2. Sei $f(x, y) := x^2(1-x^2) - y^2$. Die Kurve $C := \{(x, y) \mid f(x, y) = 0\}$ ist eine sogenannte „Lemniskate“:



Wir berechnen die partiellen Ableitungen:

$$\begin{aligned} f_x(x, y) &= 2x - 4x^3 = 2x(1 - 2x^2), \\ f_y(x, y) &= -2y. \end{aligned}$$

Im Nullpunkt ist die Gleichung überhaupt nicht auflösbar. Das liegt anschaulich daran, daß der dort auftretende Kreuzungspunkt aus keiner Richtung wie ein Graph aussieht.

In den Punkten $(1, 0)$ und $(-1, 0)$ ist jeweils $f_y(x, y) = 0$, also keine Auflösung nach y möglich. Allerdings ist dort $f_x(x, y) \neq 0$, wir können also lokal nach x auflösen. Das ist hier sogar konkret möglich, die Gleichung $x^4 - x^2 + y^2 = 0$ führt auf

$$x = \pm \frac{1}{2} \sqrt{2 \pm 2\sqrt{1 - 4y^2}}.$$

Läßt man y gegen Null gehen, so muß x^2 gegen 1 streben. Das schließt unter der ersten Wurzel das Minus-Zeichen aus, und man bekommt:

$$\begin{aligned} x &= +\frac{1}{2} \sqrt{2 + 2\sqrt{1 - 4y^2}} \quad \text{bei } (1, 0) \\ \text{und } x &= -\frac{1}{2} \sqrt{2 + 2\sqrt{1 - 4y^2}} \quad \text{bei } (-1, 0). \end{aligned}$$

In allen anderen Punkten ist $f_y(x, y) \neq 0$, denn wenn $y = 0$ und $f(x, y) = 0$ ist, dann kann nur $x = 0$ oder $x = \pm 1$ sein. Dann ist

$$y = \pm \sqrt{x^2(1 - x^2)},$$

wobei das Vorzeichen davon abhängt, ob man sich gerade in der oberen oder in der unteren Halbebene befindet.

Rechnen wir noch im Falle der oberen Halbebene die Ableitung von $y = g(x)$ aus:

$$\begin{aligned} g'(x) &= -\frac{f_x(x, g(x))}{f_y(x, g(x))} \\ &= -\frac{2x - 4x^3}{-2g(x)} \\ &= \frac{x(1 - 2x^2)}{\sqrt{x^2(1 - x^2)}}. \end{aligned}$$

Diese Beziehung gilt natürlich nicht bei $x = 0$. Für $0 < x < 1$ ist $g'(x) = 0$ genau dann erfüllt, wenn $1 - 2x^2 = 0$ ist, also $x = \frac{1}{2}\sqrt{2}$. Dort ist $y = \frac{1}{2}$. Offensichtlich liegt ein Maximum vor, und mit dieser Information kann man schon eine recht gute Skizze der Lemniskate erstellen.

Nun können wir auch das Verfahren mit dem Lagrangeschen Multiplikator verallgemeinern.

Sei $B \subset \mathbb{R}^n$ offen, $g : B \rightarrow \mathbb{R}^m$ stetig differenzierbar und $\text{rg } g'(\mathbf{x}) = m$ für alle $x \in B$.

Weiter sei $M := \{\mathbf{x} \in B \mid g(\mathbf{x}) = \mathbf{0}\}$, $\mathbf{a} \in M$, $U(\mathbf{a}) \subset B$ eine offene Umgebung und $f : U \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion.

Definition:

f hat in \mathbf{a} ein *relatives Maximum* (bzw. *relatives Minimum*) unter den Nebenbedingungen

$$g_1(\mathbf{x}) = \dots = g_m(\mathbf{x}) = 0,$$

falls $f(\mathbf{x}) \leq f(\mathbf{a})$ (bzw. $f(\mathbf{x}) \geq f(\mathbf{a})$) für alle $\mathbf{x} \in U \cap M$.

Hilfssatz

Seien $U, V \subset \mathbb{R}^n$ zwei Untervektorräume.

Ist $U \subset V$, so ist $V^\perp \subset U^\perp$.

BEWEIS: Sei $\mathbf{v} \in V^\perp$ festgehalten, $\mathbf{u} \in U$ beliebig. Dann ist $\mathbf{u} \in V$, nach Voraussetzung, und daher $\mathbf{v} \bullet \mathbf{u} = 0$. Also ist auch $\mathbf{v} \in U^\perp$. \square

Das führt zu folgendem Kriterium:

Methode der Lagrangeschen Multiplikatoren

Hat f in \mathbf{a} ein relatives Extremum unter den Nebenbedingungen

$$g_1(\mathbf{x}) = \dots = g_m(\mathbf{x}) = 0,$$

so gibt es Zahlen $\lambda_1, \dots, \lambda_m \in \mathbb{R}$, so daß gilt:

$$\nabla f(\mathbf{a}) = \lambda_1 \cdot \nabla g_1(\mathbf{a}) + \dots + \lambda_m \cdot \nabla g_m(\mathbf{a}).$$

Die Zahlen $\lambda_1, \dots, \lambda_m$ nennt man Lagrangesche Multiplikatoren. Man beachte, daß es sich hier wieder nur um ein notwendiges Kriterium handelt! Die Punkte, die die angegebene Bedingung erfüllen, können Extremwerte sein. Ob sie es wirklich sind, muß man mit anderen Mitteln feststellen.

BEWEIS: Es sei $d := n - m$ und $g := (g_1, \dots, g_m)$. Dann ist $\mathbf{a} \in M$ und $\text{rg } g'(\mathbf{a}) = n - d$. O.B.d.A. können wir annehmen, daß die letzten m Spalten der Funktionalmatrix linear unabhängig sind. Damit sind die Voraussetzungen des Satzes über implizite Funktionen erfüllt. Schreibt man \mathbf{a} in der Form

$$\mathbf{a} = (\mathbf{a}^*, \mathbf{a}^{**}) = (a_1, \dots, a_d, a_{d+1}, \dots, a_n),$$

so gibt es offene Umgebungen $U(\mathbf{a}^*)$ und $V(\mathbf{a}^{**})$, sowie eine stetig differenzierbare Abbildung $\varphi : U \rightarrow V$, so daß gilt:

$$\{(\mathbf{x}, \mathbf{y}) \in U \times V \mid g(\mathbf{x}, \mathbf{y}) = \mathbf{0}\} = \{(\mathbf{x}, \varphi(\mathbf{x})) \mid \mathbf{x} \in U\}.$$

Außerdem ist $\varphi'(\mathbf{a}^*) = -\frac{\partial g}{\partial \mathbf{y}}(\mathbf{a})^{-1} \circ \frac{\partial g}{\partial \mathbf{x}}(\mathbf{a})$.

Der Satz ist bewiesen, wenn wir gezeigt haben, daß

$$\nabla f(\mathbf{a}) \in \langle \nabla g_1(\mathbf{a}), \dots, \nabla g_m(\mathbf{a}) \rangle$$

ist. Da für einen Unterraum $U \subset \mathbb{R}^n$ stets die Beziehung $(U^\perp)^\perp = U$ gilt, reicht es wegen des Hilfssatzes zu zeigen:

$$\langle \nabla g_1(\mathbf{a}), \dots, \nabla g_m(\mathbf{a}) \rangle^\perp \subset \langle \nabla f(\mathbf{a}) \rangle^\perp.$$

Das wollen wir nun tun!

Sei $\mathbf{v} \in \langle \nabla g_1(\mathbf{a}), \dots, \nabla g_m(\mathbf{a}) \rangle^\perp$. Das bedeutet:

$$Dg(\mathbf{a})(\mathbf{v}) = (\nabla g_1(\mathbf{a}) \bullet \mathbf{v}, \dots, \nabla g_m(\mathbf{a}) \bullet \mathbf{v}) = (0, \dots, 0) = \mathbf{0},$$

oder mit Hilfe der Funktionalmatrix $g'(\mathbf{a}) = \left(\frac{\partial g}{\partial \mathbf{x}}(\mathbf{a}), \frac{\partial g}{\partial \mathbf{y}}(\mathbf{a}) \right)$

und $\mathbf{v} = (\mathbf{v}^*, \mathbf{v}^{**}) \in \mathbb{R}^d \times \mathbb{R}^{n-d}$ ausgedrückt:

$$\begin{aligned} \mathbf{0} = g'(\mathbf{a}) \circ \mathbf{v}^\top &= \left(\frac{\partial g}{\partial \mathbf{x}}(\mathbf{a}) \mid \frac{\partial g}{\partial \mathbf{y}}(\mathbf{a}) \right) \circ \begin{pmatrix} (\mathbf{v}^*)^\top \\ (\mathbf{v}^{**})^\top \end{pmatrix} \\ &= \frac{\partial g}{\partial \mathbf{x}}(\mathbf{a}) \circ (\mathbf{v}^*)^\top + \frac{\partial g}{\partial \mathbf{y}}(\mathbf{a}) \circ (\mathbf{v}^{**})^\top, \end{aligned}$$

also

$$\frac{\partial g}{\partial \mathbf{x}}(\mathbf{a}) \circ (\mathbf{v}^*)^\top = -\frac{\partial g}{\partial \mathbf{y}}(\mathbf{a}) \circ (\mathbf{v}^{**})^\top.$$

Bis jetzt haben wir nur die Voraussetzungen interpretiert und umgeformt. Nun kommt der entscheidende Schritt:

Durch $\alpha(t) := (\mathbf{a}^* + t\mathbf{v}^*, \varphi(\mathbf{a}^* + t\mathbf{v}^*))$ wird ein Weg definiert, der ganz in der Fläche M verläuft. Insbesondere ist

$$\begin{aligned} \alpha(0) &= (\mathbf{a}^*, \varphi(\mathbf{a}^*)) = \mathbf{a} \\ \text{und } \dot{\alpha}(0) &= (\mathbf{v}^*, (\varphi'(\mathbf{a}^*) \circ (\mathbf{v}^*)^\top)^\top) \\ &= (\mathbf{v}^*, (-\frac{\partial g}{\partial \mathbf{y}}(\mathbf{a})^{-1} \circ \frac{\partial g}{\partial \mathbf{x}}(\mathbf{a}) \circ (\mathbf{v}^*)^\top)^\top) \\ &= (\mathbf{v}^*, \mathbf{v}^{**}) = \mathbf{v}. \end{aligned}$$

Jetzt haben wir nahezu alle Voraussetzungen ausgeschlachtet. Wir müssen aber noch benutzen, daß f auf M in \mathbf{a} ein lokales Extremum hat. Dann hat auch $f \circ \alpha$ in 0 ein lokales Extremum, und es muß gelten:

$$0 = (f \circ \alpha)'(0) = \nabla f(\mathbf{a}) \bullet \mathbf{v}.$$

Also ist $\mathbf{v} \in \langle \nabla f(\mathbf{a}) \rangle^\perp$, und wir sind fertig!

□