

# Mathematik für Physiker II

Ausarbeitung einer Vorlesung  
vom Sommersemester 2006

JOACHIM WEIDMANN

Fachbereich Informatik und Mathematik  
der Universität Frankfurt

Stand 10. Januar 2007

## 1 Lineare Gleichungssysteme und Determinanten

Wie löst man (rechnerisch) ein u.U. großes lineares Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m &= b_1 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m &= b_n. \end{aligned}$$

Schauen wir uns das zunächst für 2 Gleichungen mit 2 Unbekannten an ( $n = m = 2$ ):

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2. \end{aligned}$$

Sind *alle*  $a_{ij} = 0$ , so ist offenbar nichts zu tun: Ist wenigstens ein  $b_i \neq 0$ , so gibt es keine Lösung des Systems (linke Seite = 0, rechte Seite  $\neq 0$ ). Sind auch alle  $b_i = 0$ , so ist jedes Paar  $(x_1, x_2)$  eine Lösung.

Sei also mindestens ein  $a_{ij} \neq 0$ . Durch Vertauschung der Zeilen (der Gleichungen) oder /und der Spalten (der Variablen) kann erreicht werden, daß  $a_{11} \neq 0$  gilt. (Aus Gründen der numerischen Genauigkeit wird man  $|a_{11}|$  außerdem möglichst groß wählen. Falls Spalten vertauscht werden, muß man sich dies natürlich merken, damit man am Ende die Werte der einzelnen Variablen richtig zuordnen kann.)

Addiert man nun das  $-\frac{a_{21}}{a_{11}}$ -Fache der 1. Zeile zur 2. Zeile, so erhält man das neue System

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ \left(a_{22} - \frac{a_{21}}{a_{11}}a_{12}\right)x_2 &= b_2 - \frac{a_{21}}{a_{11}}b_1. \end{aligned}$$

Ist  $a_{11}a_{22} - a_{12}a_{21} = 0$ , also wegen  $a_{11} \neq 0$  auch  $a_{22} - \frac{a_{21}}{a_{11}}a_{12} = 0$ , so ist das System nur dann lösbar, wenn auch  $b_2 - \frac{a_{21}}{a_{11}}b_1 = 0$  ist, d. h.  $a_{11}b_2 - a_{21}b_1 = 0$ . Dann ist  $x_2$  beliebig wählbar, und  $x_1$  kann (in Abhängigkeit von  $x_2$ ) berechnet werden.

Ist  $a_{11}a_{22} - a_{12}a_{21} \neq 0$ , also  $a_{22} - \frac{a_{21}}{a_{11}}a_{12} \neq 0$ , so kann das System (für beliebige  $b_1, b_2$ ) von unten her aufgelöst werden:

$$\begin{aligned} x_2 &= \frac{b_2 - \frac{a_{21}}{a_{11}}b_1}{a_{22} - \frac{a_{21}}{a_{11}}a_{12}} = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}}, \\ x_1 &= \frac{1}{a_{11}} \left( b_1 - a_{12} \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}} \right) = \frac{1}{a_{11}} \frac{a_{11}a_{22}b_1 - a_{21}a_{12}b_1 - a_{11}a_{12}b_2 + a_{12}a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}} \\ &= \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{21}a_{12}}. \end{aligned}$$

Später werden wir die *Determinante* quadratischer Matrizen kennen lernen, für  $2 \times 2$ -Matrizen

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := ad - bc.$$

Damit ist also

$$x_1 = \frac{\det \begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{21} \end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}, \quad x_2 = \frac{\det \begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}.$$

Dies ist ein Spezialfall der später zu beweisenden *Cramerschen Regel*.

Das obige Verfahren läßt sich analog auf Systeme beliebig vieler ( $n$ ) Gleichungen mit beliebig vielen ( $m$ ) Unbekannten anwenden. Dies liefert das sogenannte *GAUSSsche* Eliminationsverfahren für lineare Gleichungssysteme

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m &= b_2 \\ \vdots & \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m &= b_n. \end{aligned}$$

Sind alle  $a_{ij} = 0$ , so ist das Eliminationsverfahren beendet; das System ist genau dann lösbar, wenn alle  $b_i$  verschwinden, und alle  $m$ -Tupel  $(x_1, \dots, x_m)$  sind Lösungen.

Andernfalls können wir (ggf. wieder nach Vertauschung von Zeilen oder/und Spalten) annehmen, daß  $a_{11} \neq 0$  ist (oder sogar möglichst groß). Addiert man nun das  $-\frac{a_{k1}}{a_{11}}$  der 1. Gleichung zur  $k$ -ten Gleichung ( $k \geq 2$ ), so erhält man ein System der Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m &= b_1 \\ \tilde{a}_{22}x_2 + \dots + \tilde{a}_{2m}x_m &= \tilde{b}_2 \\ \vdots & \\ \tilde{a}_{n2}x_2 + \dots + \tilde{a}_{nm}x_m &= \tilde{b}_n. \end{aligned}$$

Nun verfährt man mit den neu entstandenen Gleichungen 2 bis  $n$  (das sind  $n - 1$  Gleichungen für  $m - 1$  Unbekannte) entsprechend. (Man muß beachten, daß man bei einer Vertauschung der Spalten die entsprechende Vertauschung auch in der ersten Zeile vornehmen muß.)

Das Verfahren ist spätestens nach  $n - 1$  Schritten beendet, oder früher, wenn in dem neu entstandenen System links vom Gleichheitszeichen nur noch Nullen stehen. Es entsteht also ein System der folgenden Form, wobei wir nur noch die Koeffizienten angeben und die Unbekannten  $x_1, \dots, x_m$  sowie die  $+$ -Zeichen weglassen

$$\begin{array}{cccccccc}
\alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1k} & \alpha_{1,k+1} & \dots & \alpha_{1m} & \beta_1 \\
& \alpha_{22} & \alpha_{23} & \dots & \alpha_{2k} & \alpha_{2,k+1} & \dots & \alpha_{2m} & \beta_2 \\
& & \alpha_{33} & \dots & \alpha_{3k} & \alpha_{3,k+1} & \dots & \alpha_{3m} & \beta_3 \\
& & & \ddots & \vdots & \vdots & & \vdots & \vdots \\
& & & & \alpha_{kk} & \alpha_{k,k+1} & \dots & \alpha_{km} & \beta_k \\
& & & & & & & & 0 & \beta_{k+1} \\
& & & & & & & & \vdots & \\
& & & & & & & & 0 & \beta_n.
\end{array}$$

Die  $\alpha_{ii}$  sind alle von 0 verschieden; die Zeilen mit den Nullen links vom Gleichheitszeichen müssen i. allg. nicht auftreten. Jedenfalls ist immer  $k \leq \min\{n, m\}$ .

Sind nicht alle  $\beta_{k+1}, \dots, \beta_n$  gleich 0, so ist das System nicht lösbar. Sind diese alle gleich 0 (bzw. treten sie gar nicht auf,  $k = n$ ), so können  $x_{k+1}, \dots, x_m$  beliebig gewählt werden und die  $x_k, \dots, x_1$  können daraus rekursiv berechnet werden.

Das (ursprüngliche) Gleichungssystem heißt *homogen*, wenn  $b_1 = b_2 = \dots = b_n = 0$  gilt. In diesem Fall ist natürlich auch  $\beta_1 = \beta_2 = \dots = \beta_n = 0$ . Ein homogenes Gleichungssystem hat also immer mindestens die Triviale (Null-)Lösung  $(x_1, \dots, x_m) = (0, \dots, 0)$ . Zusammen mit der vorherigen Überlegung ist damit klar:

**Satz 1.1** *Ein homogenes Gleichungssystem mit weniger Gleichungen als Unbekannten ( $n < m$ ) hat nichttriviale Lösungen, d. h.  $(x_1, \dots, x_m) \neq (0, \dots, 0)$ .*

*Beweis.* Die obige Überlegung zeigt, daß mindestens ein  $x_j$  beliebig (also insbesondere  $\neq 0$ ) gewählt werden kann ■

**Satz 1.2** *Ist  $u_0$  eine beliebige Lösung des inhomogenen Systems und  $U$  die Menge der Lösungen des zugehörigen homogenen Systems (offenbar ein linearer Raum), so erhält man alle Lösungen des inhomogenen Systems in der Form*

$$u_0 + u \text{ mit } u \in U$$

*Beweis.* a) Zwei Lösungen des inhomogenen Systems unterscheiden sich um eine Lösung des homogenen Systems. Ist also  $w$  eine beliebige Lösung des inhomogenen Systems, so gilt

$$w = u_0 + u \text{ mit } u := w - u_0 \in U$$

b) Für beliebiges  $u \in U$  ist  $u_0 + u$  Lösung des inhomogenen Systems. ■

Bezeichnen wir die Spaltenvektoren der Matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ a_{21} & \dots & a_{2m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \quad \text{mit} \quad a_j := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix},$$

so kann das Gleichungssystem auch in der Form

$$\sum_{j=1}^m x_j a_j = b \quad \text{mit} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

geschrieben werden. Daraus ergibt sich:

**Satz 1.3** a) *Das System ist genau dann lösbar, wenn  $b$  Linearkombination der  $a_1, \dots, a_m$  ist:*

$$b \in L\{a_1, \dots, a_m\}.$$

- b) *Das System hat höchstens eine Lösung, wenn die Vektoren  $a_1, \dots, a_m$  linear unabhängig sind.*  
 c) *Das System ist genau dann eindeutig lösbar, wenn die  $a_1, \dots, a_m$  linear unabhängig sind mit  $b \in L\{a_1, \dots, a_m\}$ .*  
 d) *Das System ist genau dann für alle  $b$  eindeutig lösbar, wenn  $n = m$  gilt und  $\{a_1, \dots, a_n\}$  eine Basis in  $\mathbb{K}^n$  bildet.*

*Beweis.* a) Ist offensichtlich auf Grund der obigen Schreibweise des Systems.

b) Gibt es höchstens eine Lösung, so ist jede nichttriviale Linearkombination der  $a_1, \dots, a_m$  ungleich 0, d. h.  $a_1, \dots, a_m$  sind linear unabhängig.

c) Folgt unmittelbar aus a) und b).

d) Aus c) folgt, daß  $a_1, \dots, a_m$  linear unabhängig sind, und daß  $\mathbb{K}^n = L\{a_1, \dots, a_m\}$  gilt, d. h.  $\{a_1, \dots, a_m\}$  ist eine Basis von  $\mathbb{K}^n$ , also ist die Anzahl  $m$  der Elemente von  $\{a_1, \dots, a_m\}$  gleich  $\dim \mathbb{K}^n = n$ . ■

## 1: 20. April 2006

Beim Eliminationsverfahren sind die „Null“-Zeilen dadurch entstanden, daß Linearkombinationen der ersten  $k$  Zeilen addiert (bzw. subtrahiert) wurden. Also waren die Zeilen  $k+1, \dots, n$  in der linearen Hülle der ersten  $k$  Zeilen enthalten. Andererseits sind die (neuen) ersten  $k$  Zeilen offensichtlich linear unabhängig (warum?). Es ist also

$$k = \dim L\{\text{Zeilenvektoren von } A\} =: \text{Zeilenrang von } A.$$

**Satz 1.4** Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$ ,  $\dim V =: m < \infty$ ,  $A$  eine lineare Abbildung von  $V$  nach  $W$ ,

$$\text{Kern}(A) = \text{Nullraum von } A := \{v \in V : Av = 0\},$$

$$\text{Bild}(A) = \text{Wertebereich von } A := \{Av : v \in V\}.$$

$\text{Kern}(A)$  und  $\text{Bild}(A)$  sind Untervektorräume von  $V$  bzw.  $W$  und es gilt

$$\dim \text{Kern}(A) + \dim \text{Bild}(A) = \dim V.$$

*Beweis.* Sei  $\{v_1, \dots, v_k\}$  eine Basis von  $\text{Kern}(A)$ ,  $k := \dim \text{Kern}(A)$ . Dann gibt es  $v_{k+1}, \dots, v_m$  so, daß  $\{v_1, \dots, v_m\}$  eine Basis von  $V$  bildet. Jedes Element von  $\text{Bild}(A)$  hat die Gestalt

$$Ax = A \sum_{j=1}^m c_j v_j = \sum_{j=k+1}^m c_j Av_j,$$

d. h.  $\dim \text{Bild}(A) \leq m - k$ . Andererseits ist keine nichttriviale Linearkombination der  $\{Av_j : j = k+1, \dots, m\}$  gleich 0, sonst wäre die entsprechende Linearkombination der  $\{v_j : j = k+1, \dots, m\}$  im  $\text{Kern}(A)$ . Also ist  $\dim \text{Bild}(A) = m - k$ , womit die Behauptung beweisen ist. ■

**Satz 1.5** Sei  $V_0$  ein  $k$ -dimensionaler Teilraum von  $\mathbb{K}^m$ ,  $V_0^\perp$  das (bezüglich des üblichen Skalarprodukts) orthogonale Komplement von  $V_0$ ,

$$V_0^\perp := \left\{ x \in \mathbb{K}^m : \langle x, v \rangle = 0 \quad \forall v \in V_0 \right\}.$$

Dann gilt  $\dim V_0^\perp = m - k$ .

*Beweis.* Zunächst gilt  $V_0 \cap V_0^\perp = \{0\}$ , da für jedes  $x \in V_0 \cap V_0^\perp$  gilt  $\|x\|^2 = \langle x, x \rangle = 0$ . Also gilt

$$\dim V_0 + \dim V_0^\perp = \dim(V_0 + V_0^\perp) \leq m$$

und somit  $\dim V_0^\perp \leq m - k$ .

Die Elemente von  $V_0^\perp$  sind die Lösungen des homogenen Gleichungssystems

$$\langle x, v_j \rangle = 0 \quad \text{für } j = 1, \dots, k, \quad \{v_1, \dots, v_k\} \text{ Basis von } V_0$$

mit  $k$  Gleichungen für  $m$  Unbekannte. Nach den obigen Überlegungen ist dieser Lösungsraum mindestens  $(m - k)$ -dimensional, d. h. es gilt  $\dim V_0^\perp \geq m - k$ , womit die Gleichheit folgt. ■

**Satz 1.6** Für jede Matrix  $A$  gilt

$$\text{Zeilenrang von } A = \text{Spaltenrang von } A =: \text{rang}(A).$$

*Beweis.* Der Spaltenrang ist gleich der Dimension des Bildraums der durch  $A$  erzeugten linearen Abbildung.

Der Kern ist gleich dem orthogonalen Komplement des durch die Zeilenvektoren aufgespannten Teilraums, also

$$\begin{aligned}\dim \text{Kern}(A) &= m - \text{Zeilenrang}(A) \\ m &= \dim \text{Kern}(A) + \dim \text{Bild}(A) \\ &= m - \text{Zeilenrang}(A) + \text{Spaltenrang}(A).\end{aligned}$$

Daraus folgt die behauptete Gleichheit. ■

Ein wichtiger Begriff im Zusammenhang mit der Lösung linearer Gleichungssysteme, der linearen Un-/Abhängigkeit von Vektoren, oder allgemeiner der Untersuchung linearer Abbildungen ist die *Determinante*. Den Spezialfall der Determinante von  $2 \times 2$ -Matrizen haben wir bereits oben kennen gelernt.

Sei  $\mathbb{K}^{n,n}$  (bzw.  $\mathbb{R}^{n,n}$  oder  $\mathbb{C}^{n,n}$ ) die Menge der  $n \times n$ -Matrizen mit Einträgen aus  $\mathbb{K}$  (bzw.  $\mathbb{R}$  oder  $\mathbb{C}$ ). Eine Abbildung (Funktion)  $\det : \mathbb{K}^{n,n} \rightarrow \mathbb{K}$  heißt eine *Determinantenfunktion* (oder kurz *Determinante*), wenn gilt:

$$(a) \det \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = 1,$$

(b)  $\det(\cdot)$  ist linear in jeder Spalte, d. h. (wenn  $a^j$  bzw.  $b^j$  jeweils für eine entsprechende Spalte [Spaltenvektor] steht)

$$\begin{aligned}\det(a^1, \dots, a^{j-1}, \lambda a^j + \mu b^j, a^{j+1}, \dots, a^n) \\ = \lambda \det(a^1, \dots, a^j) + \mu \det(a^1, \dots, a^{j-1}, b^j, a^{j+1}, \dots, a^n) \quad (j = 1, \dots, n)\end{aligned}$$

(es wird sich zeigen, daß auch die Linearität in jeder Zeile gilt),

(c) bei Vertauschung zweier Spalten geht die Determinante in ihr Negatives über.

Die Eigenschaften (b) und (c) bedeuten, daß  $\det$  eine *alternierende Multilinearform* ist.

Es wird weiter unten zu zeigen sein, daß eine solche Funktion existiert. Zunächst wollen wir aus der Definition einige wichtige Eigenschaften ableiten:

**Satz 1.7** *Eine Determinantenfunktion hat die folgenden Eigenschaften:*

a) *Es gilt  $\det A = 0$ , falls*

(i) *eine Spalte gleich 0 ist, oder*

(ii) *zwei Spalten gleich sind.*

b) Es gilt  $\det A = 0$  genau dann, wenn die Spalten linear abhängig sind. Oder:  $\det A \neq 0$  genau dann, wenn die Spalten linear unabhängig sind.

c) Addition eines Vielfachen einer Spalte zu einer anderen Spalte ändert die Determinante nicht:

$$\det(a^1, \dots, a^{j-1}, a^j + \lambda a^k, a^{j+1}, \dots, a^n) = \det(a^1, \dots, a^n) \text{ für } \lambda \in \mathbb{K}, k \neq j.$$

*Beweis.* a) (i) folgt aus der Linearität bezüglich der jeweiligen Spalte: Ist z.B. die  $j$ -te Spalte gleich 0, so gilt

$$\begin{aligned} \det(a^1, \dots, a^n) &= \det(a^1, \dots, a^{j-1}, 0 \cdot a^j, a^{j+1}, \dots, a^n) \\ &= 0 \det(a^1, \dots, a^n) = 0. \end{aligned}$$

(ii) Vertauschung der beiden gleichen Spalten ändert einerseits nichts, führt aber andererseits zum negativen Wert.

b)  $\Leftarrow$ : Gilt z.B.  $a^k = \sum_{j \neq k} \alpha_j a^j$ , so folgt aus der Linearität bezüglich der  $k$ -ten Spalte

$$\det A = \sum_{j \neq k} \alpha_j \det(a^1, \dots, a^{k-1}, a^j, a^{k+1}, \dots, a^n) = 0,$$

da in den letzten Determinanten jeweils zwei Spalten gleich sind (die  $k$ -te und die  $j$ -te).

$\Rightarrow$ : Seien jetzt  $a^1, \dots, a^n$  linear unabhängig, also eine Basis von  $\mathbb{K}^n$ . Dann läßt sich insbesondere jedes  $e^k$  der kanonischen Basis ( $e^k = (0, \dots, 0, 1, 0, \dots, 0)$  mit 1 an der  $k$ -ten Stelle) durch die  $a^j$  darstellen:

$$e^k = \sum_{j=1}^n x_{kj} a^j \quad \text{mit} \quad x_{kj} \in \mathbb{K}.$$

Damit folgt

$$\begin{aligned} 1 &= \det(e^1, \dots, e^n) = \det\left(\sum_{j_1=1}^n x_{1j_1} a^{j_1}, \sum_{j_2=1}^n x_{2j_2} a^{j_2}, \dots, \sum_{j_n=1}^n x_{nj_n} a^{j_n}\right) \\ &= \sum_{j_1=1}^n \dots \sum_{j_n=1}^n x_{1j_1} x_{2j_2} \dots x_{nj_n} \det(a^{j_1}, \dots, a^{j_n}) \\ &\quad \text{(jede dieser Determinanten ist } \pm \det(a^1, \dots, a^n), \text{ da sie aus} \\ &\quad \det(a^1, \dots, a^n) \text{ durch endlich viele Vertauschungen hervorgeht oder 0,} \\ &\quad \text{da mindestens 2 Spalten gleich sind)} \\ &= \alpha \det(a^1, \dots, a^n) \quad \text{mit } \alpha \in \mathbb{K}. \end{aligned}$$

Also ist  $\det(a^1, \dots, a^n) \neq 0$ .

c) Mit Teil a(i) folgt

$$\begin{aligned} &\det(a^1, \dots, a^{j-1}, a^j + \lambda a^k, a^{j+1}, \dots, a^n) \\ &= \det(a^1, \dots, a^n) + \lambda \det(a^1, \dots, a^{j-1}, a^k, a^{j+1}, \dots, a^n) \\ &= \det(a^1, \dots, a^n) + \lambda 0 = \det(a^1, \dots, a^n). \end{aligned}$$





Im folgenden sei  $S_n$  die *Symmetrische Gruppe* der Permutationen der Elemente  $1, \dots, n$ . Dies ist *offensichtlich* eine Gruppe (das *neutrale Element* ist die Identität [jedes  $j$  wird  $j$  zugeordnet]; das *inverse Element* ist die Umkehrung der Zuordnung.  $S_n$  enthält  $n!$  Elemente (dem ersten können  $n$  Elemente zugeordnet werden, dem zweiten  $n - 1$ , dem dritten  $n - 2 \dots$ ). Die Elemente  $\sigma$  von  $S_n$  werden in der Form

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \dots & \sigma(n) \end{pmatrix}$$

notiert. Jedes  $\sigma$  läßt sich als Zusammensetzung von endlich vielen Transpositionen (=Vertauschung zweier Elemente) darstellen.

Jedes  $\sigma$  läßt sich entweder als Produkt einer geraden oder einer ungeraden Anzahl von Transpositionen schreiben (den Beweis geben wir hier nicht an, vgl. z. B. G. Fischer: Lineare Algebra) Je nachdem heißt  $\sigma$  eine gerade oder ungerade Permutation. Man definiert das Signum einer Permutation durch

$$\text{sign } \sigma := \begin{cases} 1 & \text{falls } \sigma \text{ gerade,} \\ -1 & \text{falls } \sigma \text{ ungerade.} \end{cases}$$

Für zwei Permutationen  $\sigma$  und  $\tau$  gilt offenbar

$$\text{sign } \sigma\tau = \text{sign } \sigma \text{ sign } \tau.$$

## 2:25.04.06

### Satz 1.8 (Existenz und Eindeutigkeit der Determinante)

$$\det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \sum_{\sigma \in S_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \text{sign } \sigma = \sum_{\sigma \in S_n} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \text{sign } \sigma$$

ist die eindeutig bestimmte Determinantenfunktion.

*Beweis.* Auf Grund der obigen Rechenregeln gilt

$$\begin{aligned} \det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} &= \det \left( \sum_{k_1=1}^n a_{k_1 1} e^{k_1}, \dots, \sum_{k_n=1}^n a_{k_n n} e^{k_n} \right) \\ &= \sum_{k_1=1}^n \cdots \sum_{k_n=1}^n a_{k_1 1} \cdots a_{k_n n} \det(e^{k_1}, \dots, e^{k_n}) \\ &\quad (\text{alle anderen Determinanten verschwinden,} \\ &\quad \text{da mindestens zwei Spalten gleich sind}) \\ &= \sum_{\sigma \in S_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \det(e^{\sigma(1)}, \dots, e^{\sigma(n)}) \\ &= \sum_{\sigma \in S_1} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \text{sign } \sigma. \end{aligned}$$

Diese Summe ist offenbar gleich

$$\begin{aligned} &= \sum_{\sigma \in S_n} a_{1,\sigma^{-1}(1)} \cdots a_{n,\sigma^{-1}(n)} \operatorname{sign} \sigma \\ &= \sum_{\sigma \in S_n} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \operatorname{sign} \sigma. \end{aligned}$$

■

**Folgerung** Aus dieser Formel für  $\det$  ergibt sich sofort

- $\det A^t = \det A$ ,  $\det(A^*) = \overline{\det A}$ ,
- alle bisherigen Aussagen bezüglich der Spalten gelten auch bezüglich der Zeilen.

Eine weitere wichtige Konsequenz ist der

**Satz 1.9 (Entwicklungssatz für Determinanten)** Für alle  $ij \in \{1, \dots, n\} \times \{1, \dots, n\}$  sei  $A_{ij}$  die  $(n-1) \times (n-1)$ -Matrix, die  $i, j$ -te Minorante, die durch Streichung der  $i$ -ten Zeile und  $j$ -ten Spalte entsteht. Dann gilt für jedes  $i \in \{1, \dots, n\}$  bzw.  $j \in \{1, \dots, n\}$

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij})$$

die Entwicklung nach der  $i$ -ten Zeile bzw. nach der  $j$ -ten Spalte.

*Beweis.* Wir beweisen die Entwicklung nach der  $i$ -ten Zeile:

Durch  $i-1$  Zeilenvertauschungen und später  $j-1$  Spaltenvertauschungen erhält man

$$\begin{aligned} \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} &= (-1)^{i-1} \det \begin{pmatrix} a_{i1} & \cdots & a_{in} \\ a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \\ &= (-1)^{i-1} \sum_{j=1}^n \det \begin{pmatrix} 0 & \cdots & a_{ij} & \cdots & 0 \\ a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j} & \cdots & a_{i+1,n} \\ a_{n,1} & \cdots & a_{n,j} & \cdots & a_{n,n} \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
&= (-1)^{i-1} \sum_{j=1}^n (-1)^{j-1} \begin{pmatrix} a_{ij} & 0 & \dots & 0 & 0 & \dots & 0 \\ a_{1j} & a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,j} & a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i+1,j} & a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n,j} & a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{n,n} \end{pmatrix} \\
&\quad (\text{da aus der ersten Spalte nur } a_{ij} \text{ einen Beitrag liefert}) \\
&= \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij}.
\end{aligned}$$

■

Als **Spezialfall** erhält man für  $n = 3$  die **Sarrussche Regel**

$$\begin{aligned}
\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\
&\quad - (a_{13}a_{22}a_{31} + a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33}),
\end{aligned}$$

die man sich so merken kann: Man schreibt rechts von der Matrix die 1. und die 2. Spalte (in dieser Reihenfolge) nochmals. Die Produkte aus den absteigenden Diagonalen werden positiv, die aus den aufsteigenden Diagonalen negativ gerechnet.

Jede  $n \times m$ -Matrix  $A$  definiert in natürlicher Weise eine lineare Abbildung  $A : \mathbb{R}^m \rightarrow \mathbb{R}^n$  durch

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m a_{1j}x_j \\ \vdots \\ \sum_{j=1}^m a_{nj}x_j \end{pmatrix}$$

Umgekehrt wird jeder linearen Abbildung  $\mathbb{R}^m \rightarrow \mathbb{R}^n$  durch

$$Ae_j =: \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}, \quad A = \begin{pmatrix} a_{1j} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{nj} & & a_{nm} \end{pmatrix}$$

eine  $n \times m$ -Matrix zugeordnet.

Ist  $A$  eine  $n \times m$ -Matrix und  $B$  eine  $\ell \times n$ -Matrix, so wird die entsprechende Abbildung  $BA : \mathbb{R}^m \rightarrow \mathbb{R}^1$  durch das *Matrixprodukt*

$$BA = \begin{pmatrix} \sum_{j=1}^n b_{1j}a_{j1} & \dots & \sum_{j=1}^n b_{1j}a_{jn} \\ \vdots & & \vdots \\ \sum_{j=1}^n b_{\ell j}a_{j1} & \dots & \sum_{j=1}^n b_{\ell j}a_{jn} \end{pmatrix}.$$

Dies folgt aus

$$\begin{aligned} (BA)x = B(Ax) &= \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{\ell 1} & \dots & b_{\ell n} \end{pmatrix} \begin{pmatrix} \sum_{j=1}^m a_{1j}x_j \\ \vdots \\ \sum_{j=1}^m a_{nj}x_j \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=1}^n b_{1i} \sum_{j=1}^m a_{ij}x_j \\ \vdots \\ \sum_{i=1}^n b_{\ell i} \sum_{j=1}^m a_{ij}x_j \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m \left( \sum_{i=1}^n b_{1i}a_{ij} \right) x_j \\ \vdots \\ \sum_{j=1}^m \left( \sum_{i=1}^n b_{\ell i}a_{ij} \right) x_j \end{pmatrix}. \end{aligned}$$

Speziell für quadratische Matrizen ergibt sich damit der

**Satz 1.10 (Determinantenmultiplikationssatz)** Für  $A, B \in \mathbb{K}^{n,n}$  gilt

$$\det(AB) = \det A \det B = \det(BA).$$

*Beweis.* 1. Fall:  $\det B = 0$ . Dann sind die Spaltenvektoren linear abhängig. Das Bild der entsprechenden linearen Abbildung  $B$  (dies ist die lineare Hülle der Spaltenvektoren) ist also  $\neq \mathbb{K}^n$ . Dann ist aber auch das Bild des Produktes  $AB$  ungleich  $\mathbb{K}^n$ , d. h. die Spaltenvektoren von  $AB$  sind auch linear abhängig. Also ist  $0 = \det(AB) = \det A \cdot 0 = \det A \det B$ .

2. Fall:  $\det B \neq 0$ . Dann gilt jedenfalls

$$\det AB = \frac{\det AB}{\det B} \det B.$$

Für den Ausdruck  $\frac{\det AB}{\det B}$  gilt auf Grund der Eigenschaften einer Determinante: Er ist

- linear in jeder Spalte von  $A$  (man betrachte das Matrixprodukt  $AB$  und stelle fest, daß diese linear von jeder Spalte von  $A$  abhängt),
- 0, falls die Spalten von  $A$  linear abhängig sind,
- 1, falls  $A = E = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$  ist.

Also ist  $A \mapsto \frac{\det AB}{\det B}$  eine/die Determinante, d. h.

$$\frac{\det AB}{\det B} = \det A,$$

woraus die Behauptung folgt.

Die zweite Aussage ist nun offensichtlich. ■

**Folgerung** Ist  $A \in \mathbb{K}^{n,n}$  invertierbar, d. h. es gibt eine Matrix  $A^{-1}$  mit  $A^{-1}A = AA^{-1} = E$ , so gilt  $\det(A^{-1}) = (\det A)^{-1}$ . (Da  $A$  genau dann invertierbar ist, wenn die Zeilen oder Spalten linear unabhängig sind, ist  $A$  genau dann invertierbar, wenn  $\det A \neq 0$  gilt).

Wir können nun die Lösung von linearen Gleichungssystemen mit  $n$  Gleichungen und  $n$  Unbekannten im Fall der eindeutigen Lösbarkeit geschlossen angeben:

**Satz 1.11 (Cramersche Regel)** Sei  $A \in \mathbb{K}^{n,n}$  mit  $\det A \neq 0$ . Die eindeutig bestimmte Lösung

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ von } Ax = b \text{ mit } b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

ist gegeben durch

$$x_i = \frac{1}{\det A} \det \begin{pmatrix} a_{11} & \dots & a_{1,i-1} & b_1 & a_{1,i+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & & a_{n,i-1} & b_n & a_{n,i+1} & \dots & a_{nn} \end{pmatrix} \text{ für } i = 1, \dots, n.$$

*Beweis.* Das System  $Ax = b$  können wir auch in der Form

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

schreiben, bzw. für jedes  $i$

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \dots + x_{i-1} \begin{pmatrix} a_{1,i-1} \\ \vdots \\ a_{n,i-1} \end{pmatrix} + \begin{pmatrix} x_i a_{1i} - b_1 \\ \vdots \\ x_i a_{ni} - b_n \end{pmatrix} + x_{i+1} \begin{pmatrix} a_{1,i+1} \\ \vdots \\ a_{n,i+1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} = 0.$$

Die Spalten der Matrix

$$\begin{pmatrix} a_{11} & \dots & a_{1,i-1} & x_i a_{1i} - b_1 & a_{1,i+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,i-1} & x_i a_{ni} - b_n & a_{n,i+1} & \dots & a_{nn} \end{pmatrix}$$

sind also linear abhängig. Für jedes  $i$  folgt also wegen der Linearität in der  $i$ -ten Spalte

$$x_i \det A - \det \begin{pmatrix} a_{11} & \dots & a_{1,i-1} & b_1 & a_{1,i+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,i-1} & b_n & a_{n,i+1} & \dots & a_{nn} \end{pmatrix} = 0$$

Auflösung nach  $x_i$  liefert die Cramersche Regel. ■

**Satz 1.12 (Matrixinversion)** Sei  $A \in \mathbb{K}^{n,n}$  invertierbar (d. h.  $\det A \neq 0$ ). Dann ist

$$A^{-1} = \frac{1}{\det A} B$$

mit

$$B = (b_{ij}) = \left( (-1)^{i+j} \det A_{ji} \right) = \begin{pmatrix} \det A_{11} & -\det A_{21} & \det A_{31} & \dots \\ -\det A_{12} & \det A_{22} & -\det A_{32} & \dots \\ \det A_{13} & -\det A_{23} & \det A_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Speziell ist im Fall  $n = 2$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

*Beweis.* Aus  $AA^{-1} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$  folgt, daß die  $j$ -te Spalte von  $A^{-1}$  (mit  $(a_{ij}^-)$  bezeichnet) die Lösung des Gleichungssystems

$$Ax = e^j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \text{mit 1 an } j\text{-ter Stelle}$$

ist. Also folgt mit der Cramerschen Regel

$$a_{ij}^- = \frac{1}{\det A} (-1)^{i+j} \det A_{ji}.$$

Das ist die Behauptung. ■

**Beispiel 1.13** Eine wichtige Determinante ist die *Vandermondesche Determinante*

$$\Delta(x_1, \dots, x_n) := \det \begin{pmatrix} 1 & \dots & 1 \\ x_1 & & x_n \\ \vdots & & \vdots \\ x_1^{n-1} & \dots & x_n^{n-1} \end{pmatrix} \quad \text{für } x_1, \dots, x_n \in \mathbb{C}.$$

Indem man sukzessive in der ersten Spalte Nullen erzeugt (der Leser wird erkennen, wie das gemacht wird), erhält man

$$\Delta(x_1, \dots, x_n) = \det \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ x_1 & x_2 & x_3 & \dots & x_n \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-2} & x_2^{n-2} & x_3^{n-2} & \dots & x_n^{n-2} \\ 0 & x_2^{n-2}(x_2 - x_1) & x_3^{n-2}(x_3 - x_1) & \dots & x_n^{n-2}(x_n - x_1) \end{pmatrix}$$

$$\begin{aligned}
&= \det \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ x_1 & x_2 & x_3 & \dots & x_n \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-3} & x_2^{n-3} & x_3^{n-3} & \dots & x_n^{n-3} \\ 0 & x_2^{n-3}(x_2 - x_1) & x_3^{n-3}(x_3 - x_1) & \dots & x_n^{n-3}(x_n - x_1) \\ 0 & x_2^{n-2}(x_2 - x_1) & x_3^{n-2}(x_3 - x_1) & \dots & x_n^{n-2}(x_n - x_1) \end{pmatrix} \\
&\quad \vdots \\
&= \det \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & (x_2 - x_1) & (x_3 - x_1) & \dots & (x_n - x_1) \\ 0 & x_2(x_2 - x_1) & x_3(x_3 - x_1) & \dots & x_n(x_n - x_1) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_2^{n-2}(x_2 - x_1) & x_3^{n-2}(x_3 - x_1) & \dots & x_n^{n-2}(x_n - x_1) \end{pmatrix} \\
&= \prod_{j=2}^n (x_j - x_1) \det \begin{pmatrix} 1 & \dots & 1 \\ x_2 & \dots & x_n \\ \vdots & & \vdots \\ x_2^{n-2} & \dots & x_n^{n-2} \end{pmatrix}
\end{aligned}$$

Indem man mit dieser Determinante entsprechend verfährt, usw., erhält man

$$\begin{aligned}
\Delta(x_1, \dots, x_n) &= \prod_{j_1=2}^n (x_{j_1} - x_1) \prod_{j_2=3}^n (x_{j_2} - x_2) \dots \prod_{j_{n-1}=n}^n (x_{j_{n-1}} - x_{n-1}) \\
&\quad (\text{das letzte Produkt besteht nur aus } x_n - x_{n-1}) \\
&= \prod_{i < j} (x_j - x_i).
\end{aligned}$$

Bemerkenswert ist, daß diese Determinante genau dann von 0 verschieden ist, wenn die Zahlen  $x_1, \dots, x_n$  paarweise verschieden sind.  $\square$

**3: 27. April 2006**

## 2 Eigenwerte und Eigenvektoren

Sei  $V$  ein  $\mathbb{K}$ -Vektorraum,  $A : V \rightarrow V$  eine lineare Abbildung. Ein  $\lambda \in \mathbb{K}$  heißt ein *Eigenwert* (EW) von  $A$ , wenn es ein  $x \in V \setminus \{0\}$  gibt mit  $Ax = \lambda x$ ; ein solches  $x$  heißt *Eigenvektor* (EV) oder auch *Eigenelement* (EE) von  $A$  zum Eigenwert  $\lambda$ . (Mit  $x$  ist natürlich auch  $cx$  Eigenvektor für jedes  $c \in \mathbb{K} \setminus \{0\}$ .)

Ist  $V$  ein  $n$ -dimensionaler  $\mathbb{K}$ -Vektorraum und  $\{x^1, \dots, x^n\}$  eine Basis von  $V$ , so ist

$$U : V \rightarrow \mathbb{K}^n, \quad \sum_{j=1}^n c_j e^j \mapsto (c_1, \dots, c_n)$$

eine bijektive lineare Abbildung,

$$U^{-1} : \mathbb{K}^n \rightarrow V, \quad (c_1, \dots, c_n) \mapsto \sum_{j=1}^n c_j e^j.$$

Ist  $A$  eine lineare Abbildung in  $V$  mit

$$Ax^j = \sum_{i=1}^n a_{ij}x^i \quad \text{für } j = 1, \dots, n,$$

so ist

$$\begin{aligned} UAU^{-1}(c_1, \dots, c_n) &= UA \sum_{j=1}^n c_j x^j = U \sum_{i=1}^n \sum_{j=1}^n a_{ij} c_j x^i \\ &= \left( \sum_{j=1}^n a_{1j} c_j, \dots, \sum_{j=1}^n a_{nj} c_j \right), \end{aligned}$$

d. h.  $UAU^{-1}$  ist die durch die Matrix  $(a_{ij})$  in  $\mathbb{K}^n$  erzeugte lineare Abbildung. Man sagt,  $A$  wird bezüglich der Basis  $\{x^1, \dots, x^n\}$  durch die Matrix  $(a_{ij})$  erzeugt.

Offenbar ist  $\lambda$  genau dann ein Eigenwert von  $A$  mit einem Eigenvektor  $x$ , wenn  $\lambda$  Eigenvektor von  $UAU^{-1}$  ist mit Eigenvektor  $Ux$  (nachrechnen). Wir können uns deshalb im endlichdimensionalen Fall auf die Untersuchung von linearen Abbildungen in  $\mathbb{K}^n$ , die bekanntlich durch  $n \times n$ -Matrizen erzeugt werden, beschränken.

Man nennt eine lineare Abbildung  $A$  *diagonalisierbar*, wenn es eine Basis gibt, bezüglich der  $A$  durch eine Diagonalmatrix (alle nicht-Diagonalelemente verschwinden) dargestellt wird. Offenbar gilt

Eine lineare Abbildung  $A$  in einem endlich-dimensionalen Vektorraum  $V$  ist genau dann diagonalisierbar, wenn  $V$  eine Basis aus Eigenvektoren von  $A$  besitzt. ( $A$  ist dann bezüglich dieser Basis durch die Diagonalmatrix erzeugt, auf der Diagonalen die Eigenwerte von  $A$  stehen.

Offenbar sind diagonalisierbare Abbildungen besonders einfach zu untersuchen. Es ist deshalb eine wichtige Frage, woran man die Diagonalisierbarkeit erkennt.

**Beispiel 2.1** Die Spiegelung  $S$  an der  $e^1$ -Achse in  $\mathbb{R}^2$  (gilt entsprechend in  $\mathbb{C}^2$ , aber ist es da noch eine Spiegelung?) wird durch die Matrix  $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  dargestellt, hat also die Eigenwerte 1 und  $-1$  mit (den) zugehörigen Eigenvektoren  $e^1$  und  $e^2$ .  $\square$

**Beispiel 2.2** Die Drehung  $D$  um  $90^\circ \sim \pi/2$  in  $\mathbb{R}^2$  (in mathematisch positiver Richtung) wird bekanntlich (vgl. Teil I der Vorlesung, Satz 7.2) durch die Matrix

$$D = \begin{pmatrix} \cos \pi/2 & -\sin \pi/2 \\ \sin \pi/2 & \cos \pi/2 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

dargestellt. Sie hat keinen reellen Eigenwert. Dies erscheint geometrisch offensichtlich. Rechnerisch sieht man das so:

$$D \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Leftrightarrow \begin{matrix} -x_2 = \lambda x_1 \\ x_1 = \lambda x_2 \end{matrix} \Leftrightarrow \begin{matrix} -x_2 = \lambda^2 x_1 \\ x_1 = -\lambda^2 x_1 \end{matrix}.$$



Das ist offenbar für kein  $\lambda \in \mathbb{R}$  möglich (man beachte, daß  $(x_1, x_2) \neq (0, 0)$  sein muß).

In  $\mathbb{C}^2$  hat die gleiche Matrix offenbar die Eigenwerte  $\lambda = \pm i$  mit den Eigenvektoren  $u_+ = \begin{pmatrix} i \\ 1 \end{pmatrix}$  für  $\lambda = i$  und  $u_- = \begin{pmatrix} 1 \\ i \end{pmatrix}$  für  $\lambda = -i$ .  $\square$

Wie findet man Eigenwerte und Eigenvektoren einer Matrix  $A = (a_{ij})$ ? Damit  $\lambda$  Eigenwert ist, muß offenbar die homogene Gleichung

$$\begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = 0$$

eine nicht-triviale Lösung haben, d. h. der Kern von  $A - \lambda E$  muß  $\neq \{0\}$  sein, bzw.  $\det(A - \lambda E) = 0$ .  $\det(A - \lambda E)$  ist ein Polynom vom Grad  $n$ , das *charakteristische Polynom* von  $A$ . Seine Nullstellen sind die Eigenwerte von  $A$ . Der Fundamentalsatz der Algebra liefert:

**Satz 2.3** *Jede Matrix  $A \in \mathbb{C}^{n,n}$  hat mindestens einen Eigenwert und höchstens  $n$  verschiedene aus  $\mathbb{C}$ .*

Das Gaußsche Eliminationsverfahren liefert dann eine Möglichkeit, zugehörige Eigenvektoren zu bestimmen.

#### Beispiel 2.4

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad A - \lambda E = \begin{pmatrix} 2 - \lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & -1 & -\lambda \end{pmatrix}.$$

Das charakteristische Polynom ist

$$p_A(\lambda) = (2 - \lambda)(\lambda^2 + 1),$$

mit den Nullstellen  $\lambda = 2, i, -i$ . Zugehörige Eigenvektoren sind

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ i \end{pmatrix}, \begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}.$$

$\square$

**Satz 2.5** *Sei  $A \in \mathbb{K}^{n,n}$  und  $\lambda_1, \dots, \lambda_r$  verschiedene Eigenwerte von  $A$  mit zugehörigen Eigenvektoren  $x^1, \dots, x^r$ .*

*Dann ist  $\{x^1, \dots, x^r\}$  linear unabhängig. (Gilt entsprechend für beliebig viele verschiedene Eigenwerte einer linearen Abbildung in einem beliebigen Vektorraum.)*

*Beweis.* Induktion nach  $r$ .

$r = 1$ : Richtig, da  $x^1 \neq 0$ .

$r - 1 \Rightarrow r$ : Seien also  $x^1, \dots, x^{r-1}$  linear unabhängig. Aus

$$\sum_{i=1}^r c_i x^i = 0 \quad \text{folgt} \quad \sum_{i=1}^r \lambda_i c_i x^i = A \sum_{i=1}^r c_i x^i = 0,$$

also

$$0 = \lambda_r \sum_{i=1}^r c_i x^i - \sum_{i=1}^r \lambda_i c_i x^i = \sum_{i=1}^r (\lambda_r - \lambda_i) c_i x^i = \sum_{i=1}^{r-1} (\lambda_r - \lambda_i) c_i x^i$$

und somit (wegen der linearen Unabhängigkeit von  $x^1, \dots, x^{r-1}$ ),  $(\lambda_r - \lambda_i) c_i = 0$  für  $i = 1, \dots, r-1$ ,  $c_i = 0$  für  $i = 1, \dots, r-1$ . Wegen  $\sum_{i=1}^r c_i x^i = 0$  ist dann auch  $c_r = 0$ , d. h.  $\{x^1, \dots, x^r\}$  ist linear unabhängig. ■

Besonders einfach und übersichtlich ist die Situation bei symmetrischen bzw. hermiteschen Matrizen. Dazu betrachten wir Vektorräume mit Skalarprodukt (vgl. Teil I, Abschnitt 7). Diese Räume nennt man

- euklidisch im reellen Fall (speziell  $\mathbb{R}^n$ )
- unitär im komplexen Fall (speziell  $\mathbb{C}^n$ ).

Das übliche Skalarprodukt ist

$$\langle x, y \rangle = \begin{cases} \sum_{j=1}^n x_j y_j & \text{in } \mathbb{R}^n, \\ \sum_{j=1}^n \bar{x}_j y_j & \text{in } \mathbb{C}^n, \end{cases} \quad \text{für } x = (x_1, \dots, x_n), y = (y_1, \dots, y_n).$$

In Vektorräumen mit Skalarprodukt spielen Orthonormalbasen eine wichtige Rolle. Sei  $(V, \langle \cdot, \cdot \rangle)$  ein Vektorraum mit Skalarprodukt. Eine Basis  $\{e^1, \dots, e^n\}$  heißt eine *Orthonormalbasis* (ONB) von  $V$ , wenn gilt

$$\langle e_i, e_j \rangle = \delta_{ij} \quad \text{für } i, j = 1, \dots, n$$

(d. h. die Elemente der Basis sind paarweise orthogonal und normiert). Gibt es solche Basen?

**Satz 2.6** *Jeder endlichdimensionale Vektorraum mit Skalarprodukt besitzt eine Orthonormalbasis.*

*Beweis.* Sei  $\{f^1, \dots, f^n\}$  eine Basis von  $V$ . Das folgende *Gram-Schmidtsche Orthonormalisierungsverfahren* liefert eine Orthonormalbasis:

$$\begin{aligned} e^1 &:= \frac{1}{\|f^1\|} f^1 \\ e^2 &:= \frac{1}{\|f^2 - \langle e^1, f^2 \rangle e^1\|} (f^2 - \langle e^1, f^2 \rangle e^1) \\ &\vdots \\ e^{j+1} &:= \frac{1}{\|f^{j+1} - \sum_{i=1}^j \langle e^i, f^{j+1} \rangle e^i\|} \left( f^{j+1} - \sum_{i=1}^j \langle e^i, f^{j+1} \rangle e^i \right) \end{aligned}$$

Offenbar sind die  $e^j$  normiert und paarweise orthogonal und damit linear unabhängig (Beweis!). Da sie außerdem die gleiche lineare Hülle haben wie  $\{f^1, \dots, f^n\}$ , bilden sie eine Basis. ■

In  $\mathbb{R}^n$  bzw.  $\mathbb{C}^n$  bilden natürlich die kanonischen Einheitsvektoren  $\{e^1, \dots, e^n\}$ ,  $e^j = (0, \dots, 1, \dots, 0)$  mit 1 an der  $j$ -ten Stelle eine ONB.

Eine Abbildung  $U : V_1 \rightarrow V_2$  heißt *orthogonal* (im reellen Fall), *unitär* (im komplexen Fall), wenn gilt

$$\langle Ux, Uy \rangle_{V_2} = \langle x, y \rangle_{V_1} \quad \text{für alle } x, y \in V_1.^1$$

#### 4: 2. Mai 2006

**Satz 2.7** Seien  $V_1, V_2$   $n$ -dimensionale Vektorräume mit Skalarprodukt (beide über  $\mathbb{R}$  oder beide über  $\mathbb{C}$ ). Ist  $\{e^1, \dots, e^n\}$  eine ONB von  $V_1$  und  $U$  eine lineare Abbildung von  $V_1$  nach  $V_2$ , so ist  $U$  genau dann orthogonal/unitär, wenn  $\{f^1, \dots, f^n\}$  mit  $f^j := Ue^j$  eine ONB in  $V_2$  ist.

*Beweis.*  $\Rightarrow$ : Wegen der Orthonormalität von  $\{e^1, \dots, e^n\}$  und der Orthogonalität/Unitarität von  $U$  folgt

$$\langle f^i, f^j \rangle = \langle Ue^i, Ue^j \rangle = \langle e^i, e^j \rangle = \delta_{ij} \quad \text{für } i, j = 1, \dots, n.$$

Da dann  $\{f^1, \dots, f^n\}$  auch linear unabhängig ist (aus  $\sum_j c_j f^j = 0$  folgt  $0 = \|\sum_j c_j f^j\|^2 = \sum_j |c_j|^2$ , also  $c_j = 0$  für alle  $j$ ), ist  $\{f^1, \dots, f^n\}$  eine ONB.

$\Leftarrow$ : Beliebige Elemente  $x, y \in V_1$  haben die Form

$$x = \sum_{j=1}^n c_j e^j \quad \text{mit} \quad \sum_{j=1}^n |c_j|^2 = \|x\|^2, \quad y = \sum_{j=1}^n d_j e^j \quad \text{mit} \quad \sum_{j=1}^n |d_j|^2 = \|y\|^2.$$

Damit folgt, da  $\{f^1, \dots, f^n\}$  eine ONB in  $V_2$  ist,

$$\begin{aligned} \langle Ux, Uy \rangle_{V_2} &= \left\langle U \sum_{j=1}^n c_j e^j, U \sum_{j=1}^n d_j e^j \right\rangle_{V_2} = \left\langle \sum_{j=1}^n c_j Ue^j, \sum_{j=1}^n d_j Ue^j \right\rangle_{V_2} \\ &= \left\langle \sum_{j=1}^n c_j f^j, \sum_{j=1}^n d_j f^j \right\rangle_{V_2} = \sum_{j=1}^n \bar{c}_j d_j \\ &= \left\langle \sum_{j=1}^n c_j e^j, \sum_{j=1}^n d_j e^j \right\rangle_{V_1} = \langle x, y \rangle_{V_1}. \end{aligned}$$

■

**Satz 2.8** Zu jeder linearen Abbildung  $A$  in einem endlichdimensionalen Vektorraum mit Skalarprodukt  $(V, \langle \cdot, \cdot \rangle)$  gibt es genau eine lineare Abbildung  $A^*$ , die zu  $A$  adjungierte Abbildung, mit

$$\langle x, Ay \rangle = \langle A^*x, y \rangle \quad \text{für alle } x, y \in V.$$

<sup>1</sup>Insbesondere ist  $U$  isometrisch ( $\|Ux\|_{V_2} = \|x\|_{V_1}$  für alle  $x \in V_1$ ).

(Das entsprechende gilt für eine lineare Abbildung  $A$  von einem Vektorraum  $(V_1, \langle \cdot, \cdot \rangle_{V_1})$  nach  $(V_2, \langle \cdot, \cdot \rangle_{V_2})$ ,  $A^*$  ist dann eine Abbildung von  $V_2$  nach  $V_1$ .)

*Beweis. Eindeutigkeit:* Seien  $B_1$  und  $B_2$  lineare Abbildungen mit der für  $A^*$  geforderten Eigenschaft. Dann gilt

$$\langle (B_1 - B_2)x, y \rangle = \langle x, Ay \rangle - \langle x, Ay \rangle = 0 \quad \text{für alle } x, y \in V,$$

also

$$(B_1 - B_2)x = 0 \quad \text{für alle } x \in V, \quad \text{d. h. } B_1 = B_2.$$

*Existenz: Spezialfall*  $V = \mathbb{K}^n$  mit dem üblichen Skalarprodukt: Wird  $A$  durch die Matrix  $(a_{ij})$  erzeugt, so wird  $A^*$  durch die adjungierte Matrix  $(\overline{a_{ij}})^t = (\overline{a_{ji}})$  erzeugt:

$$\langle x, Ay \rangle = \sum_{i=1}^n \overline{x_i} \sum_{j=1}^n a_{ij} y_j = \sum_{j=1}^n y_j \sum_{i=1}^n a_{ij} \overline{x_i} = \sum_{j=1}^n y_j \sum_{i=1}^n \overline{a_{ij}} x_i,$$

also ist  $A^*x = \left( \sum_{i=1}^n \overline{a_{i1}} x_i, \dots, \sum_{i=1}^n \overline{a_{in}} x_i \right)$ .

*Allgemeiner Fall*  $(V, \langle \cdot, \cdot \rangle)$ : Es gibt eine orthonormale/unitäre Abbildung  $U : \mathbb{K}^n \rightarrow V$ . Dann ist  $B := U^{-1}AU$  eine lineare Abbildung in  $\mathbb{K}^n$  mit einer Adjungierten  $B^*$ . Für alle  $v, w \in V$  gilt

$$\begin{aligned} \langle v, Aw \rangle_V &= \langle U^{-1}v, U^{-1}AUU^{-1}w \rangle = \langle U^{-1}v, BU^{-1}w \rangle \\ &= \langle B^*U^{-1}v, U^{-1}w \rangle = \langle UB^*U^{-1}v, w \rangle, \end{aligned}$$

d. h.  $UB^*U^{-1}$  ist die Adjungierte von  $A$ . ■

Eine lineare Abbildung  $A$  in einem Vektorraum mit Skalarprodukt  $(V, \langle \cdot, \cdot \rangle)$  heißt *selbstadjungiert* (d. h. zu sich selbst adjungiert), wenn  $A = A^*$  gilt, d. h.

$$\langle x, Ay \rangle = \langle Ax, y \rangle \quad \text{für alle } x, y \in V.$$

**Bemerkung 2.9** Die durch eine Matrix  $(a_{ij})$  in  $\mathbb{K}^n$  mit dem üblichen Skalarprodukt erzeugte lineare Abbildung ist genau dann selbstadjungiert, wenn  $a_{ij} = \overline{a_{ji}}$  gilt. Eine solche Matrix heißt hermitesch, im reellen Fall auch symmetrisch.

**Satz 2.10** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein endlichdimensionaler Vektorraum (reell oder komplex) mit Skalarprodukt,  $A$  eine selbstadjungierte lineare Abbildung in  $V$ . Dann gilt

- a) Alle Eigenwerte von  $A$  sind reell.
- b)  $A$  hat mindestens einen Eigenwert.
- c) Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.

*Beweis.* a) Hier ist natürlich nur im komplexen Fall etwas zu beweisen. Ist  $\lambda$  Eigenwert von  $A$  mit Eigenvektor  $x$ , so gilt

$$\begin{aligned}\overline{\lambda}\|x\|^2 &= \overline{\lambda}\langle x, x \rangle = \langle \lambda x, x \rangle = \langle Ax, x \rangle = \langle x, Ax \rangle = \langle x, \lambda x \rangle \\ &= \lambda \langle x, x \rangle = \lambda \|x\|^2,\end{aligned}$$

d. h.  $\lambda$  ist reell.

b) Ist  $U : \mathbb{K}^m \rightarrow V$  eine unitäre/orthogonale Abbildung, so ist  $B := U^{-1}AU$  selbstadjungiert in  $\mathbb{K}^m$ , denn es gilt

$$\begin{aligned}\langle Bx, y \rangle_{\mathbb{K}^m} &= \langle U^{-1}AUx, y \rangle_{\mathbb{K}^m} = \langle UU^{-1}AUx, Uy \rangle_V = \langle AUx, Uy \rangle_V \\ &= \langle Ux, AUy \rangle_V = \langle U^{-1}Ux, U^{-1}AUy \rangle_{\mathbb{K}^m} = \langle x, By \rangle_{\mathbb{K}^m}.\end{aligned}$$

$B$  wird also durch eine hermitesche (im reellen Fall symmetrische) Matrix erzeugt. Im komplexen Fall hat  $B$  mindestens einen Eigenwert, der nach Teil a reell ist. Im reellen Fall erzeugt die reelle Matrix  $B$  in  $\mathbb{C}^m$  eine selbstadjungierte Abbildung, die (wie eben gezeigt) mindestens einen reellen Eigenwert  $\lambda$  hat, d. h. es gilt  $p_B(\lambda) = 0$  für das charakteristische Polynom von  $B$ . Also ist  $B - \lambda E$  nicht injektiv in  $\mathbb{R}^m$  und somit ist  $\lambda$  auch im reellen Fall Eigenwert von  $B$ .

Ist  $v$  der zu  $\lambda$  gehörige Eigenvektor von  $B$ , so gilt

$$A(Uv) = U(U^{-1}AU)v = UBv = U\lambda v = \lambda(Uv),$$

d. h.  $Uv$  ist Eigenvektor von  $A$  zum Eigenwert  $\lambda$ .

c) Sind  $\lambda_1 \neq \lambda_2$  Eigenwerte mit zugehörigen Eigenvektoren  $v_1, v_2$ , so gilt (da  $\lambda_j$  reell sind)

$$\begin{aligned}(\lambda_1 - \lambda_2)\langle v_1, v_2 \rangle &= \langle v_1, \lambda_2 v_2 \rangle - \langle \lambda_1 v_1, v_2 \rangle \\ &= \langle v_1, Av_2 \rangle - \langle Av_1, v_2 \rangle = 0,\end{aligned}$$

also  $\langle v_1, v_2 \rangle = 0$ . ■

**Satz 2.11** Sei  $A$  eine selbstadjungierte lineare Abbildung im  $m$ -dimensionalen Vektorraum  $V$  mit Skalarprodukt  $\langle \cdot, \cdot \rangle$ . Dann existieren reelle Zahlen  $\lambda_1, \dots, \lambda_m$  und eine ONB  $\{e_1, \dots, e_m\}$  in  $V$  mit

$$\begin{aligned}Ae_i &= \lambda_i e_i \quad \text{für } i = 1, \dots, m, \\ A \sum_{i=1}^m c_i e^i &= \sum_{i=1}^m \lambda_i c_i e^i \quad \text{für alle } c_i \in \mathbb{K}.\end{aligned}$$

*Beweis.* Nach obigem Satz existiert ein reelles  $\lambda_1$  und ein  $f_1 \in V \setminus \{0\}$  mit  $Af_1 = \lambda_1 f_1$ . Wir setzen

$$e_1 := \frac{1}{\|f_1\|} f_1 \quad \text{und} \quad V_1 = \{e_1\}^\perp = \{x \in V : \langle x, e_1 \rangle = 0\}.$$

Dann ist  $V = V_1 + L\{e_1\}$ ,  $\dim V_1 = m - 1$  und  $A|_{V_1}$  ist selbstadjungiert in  $V_1$ . Dazu ist nur  $AV_1 \subset V_1$  zu zeigen: Für  $x \in V_1$  ist  $\langle e_1, Vx \rangle = \langle Ae_1, x \rangle = \lambda_1 \langle e_1, x \rangle = 0$ , also  $Vx \in \{e_1\}^\perp = V_1$ .

Ist  $m - 1 = 0$ , so sind wir fertig.

Ist  $m - 1 > 0$ , so existieren  $\lambda_2 \in \mathbb{R}$  und  $f_2 \in V_1 \setminus \{0\}$  mit  $Af_2 = \lambda_2 f_2$ . Setze  $e_2 := \frac{1}{\|f_2\|} f_2$ . Dann ist  $e_2 \perp e_1$ . Setze  $V_2 := \{e_1, e_2\}^\perp \dots$ . Bricht nach  $m$  Schritten ab. ■

**Satz 2.12 (Hauptachsentransformation)** Sei  $A$  eine selbstadjungierte lineare Abbildung in einem  $m$ -dimensionalen Vektorraum  $V$  mit Skalarprodukt  $\langle \cdot, \cdot \rangle$ . Dann existiert eine unitäre / orthogonale Abbildung  $U : \mathbb{K}^m \rightarrow V$  so, daß

$$U^{-1}AU = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}.$$

Ist  $V = \mathbb{K}^m$  mit dem üblichen Skalarprodukt, so ist

$$U = (u^1 : u^2 : \dots : u^m) \quad \text{und} \quad U^{-1} = U^* = \begin{pmatrix} \overline{u^1} \\ \dots \\ \vdots \\ \dots \\ \overline{u^m} \end{pmatrix}$$

(Spaltenvektoren von  $U =$  orthonormierte Eigenvektoren von  $A$ ; Zeilenvektoren von  $U^{-1} = U^* =$  konjugierte der Eigenvektoren von  $A$ ).

*Beweis.* Sei  $e^i$  der  $i$ -te kanonische Einheitsvektor in  $\mathbb{K}^m$ . Dann ist  $U : \mathbb{K}^m \rightarrow V$  mit

$$Ue^i = u^i, \quad U^{-1}u^i = e^i \quad \text{für } i = 1, \dots, m$$

eine unitäre/orthogonale Abbildung. Also gilt für  $x = (x_1, \dots, x_m)^t \in \mathbb{K}^m$

$$U^{-1}AUx = U^{-1}A \sum_{j=1}^m x_j e^j = U^{-1} \sum_{j=1}^m \lambda_j x_j e^j = (\lambda_1 x_1, \dots, \lambda_m x_m)^t,$$

d. h.  $U^{-1}AU$  wird durch die im Satz angegebene Diagonalmatrix erzeugt. Ist  $V = \mathbb{R}^m$ , so haben  $U$  und  $U^{-1}$  die oben angegebene Gestalt. ■

#### 5: 4. Mai 2006

Eine selbstadjungierte lineare Abbildung  $A$  in einem Vektorraum mit Skalarprodukt heißt

- *positiv semidefinit* (bzw. *nichtnegativ*,  $A \geq 0$ ), wenn gilt  $\langle x, Ax \rangle \geq 0$  für alle  $x \in V$ ,
- *positiv definit* (strikt positiv,  $A > 0$ ), wenn gilt  $\langle x, Ax \rangle > 0$  für alle  $x \in V \setminus \{0\}$ .

**Satz 2.13** Sei  $A$  selbstadjungiert im Vektorraum  $(V, \langle \cdot, \cdot \rangle)$  mit Skalarprodukt,  $\lambda_j$  die Eigenwerte von  $A$ . Dann gilt:

- a)  $A \geq 0 \Leftrightarrow \lambda_j \geq 0$  für alle  $j$ ,  
 b)  $A > 0 \Leftrightarrow \lambda_j > 0$  für alle  $j$ .

*Beweis.* Sind  $u^j$  die zu den  $\lambda^j$  gehörigen orthonormierten Eigenvektoren (eine Basis), so gilt

$$\left\langle A \sum_{j=1}^m c_j u^j, \sum_{j=1}^m c_j u^j \right\rangle = \sum_{j=1}^m \lambda_j |c_j|^2,$$

also

- a)  $\left\langle A \sum c_j u^j, \sum c_j u^j \right\rangle \geq 0$  für alle  $c_j \Leftrightarrow \lambda_j \geq 0$  für alle  $j$ ,  
 b)  $\left\langle A \sum c_j u^j, \sum c_j u^j \right\rangle > 0$  für alle  $c_j$  mit  $(c_1, \dots, c_m) \neq (0, \dots, 0) \Leftrightarrow \lambda_j > 0$  für alle  $j$ .

■

**Beispiel 2.14** Die Matrix  $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$  erzeugt in  $\mathbb{K}^2$  eine nichtnegative selbstadjungierte Abbildung: Das charakteristische Polynom ist  $p_A(\lambda) = (1 - \lambda)^2 - 1$ ; die Eigenwerte sind also  $\lambda_{1/2} = 0/2$ , beide  $\geq 0$ . Die zugehörigen normierten Eigenelemente (eindeutig bis Faktoren vom Betrag 1) sind

$$u^1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad u^2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Die Nichtnegativität kann man aber auch so sehen:

$$\begin{aligned} \left\langle A \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle &= \left\langle \begin{pmatrix} x+y \\ x+y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = |x|^2 + \bar{y}x + \bar{x}y + |y|^2 \\ &= |x+y|^2 \geq 0 \quad \text{für alle } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{K}^2. \end{aligned}$$

(Beachte, für  $x = -y$  ist dieser Ausdruck gleich 0.)

□

**Beispiel 2.15** Die Matrix

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

hat das charakteristische Polynom

$$\begin{aligned} p_A(\lambda) &= \det \begin{pmatrix} 1-\lambda & 1 & 0 \\ 1 & 1-\lambda & 1 \\ 0 & 1 & 1-\lambda \end{pmatrix} \\ &\quad \text{(Entwicklung nach 1. Zeile oder 1. Spalte)} \\ &= (1-\lambda)[(1-\lambda)^2 - 1] - (1-\lambda) = (1-\lambda)[\lambda^2 - 2\lambda - 1]. \end{aligned}$$

Also hat  $A$  die Eigenwerte  $\lambda_1 = 1, \lambda_{2/3} = 1 \pm \sqrt{1+1} = 1 \pm \sqrt{2}$ . Da  $A$  positive und negative Eigenwerte hat, ist  $A$  nicht definiert. Wir bestimmen die Eigenvektoren:

Zu  $\lambda_1 = 1$ :

$$(A - 1E) \begin{pmatrix} x_1^1 \\ x_2^1 \\ x_3^1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1^1 \\ x_2^1 \\ x_3^1 \end{pmatrix} = 0.$$

Daraus folgt  $x_2^1 = 0$  und  $x_3^1 = -x_1^1$ , also für den normierten Eigenvektor (z. B. )

$$x^1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

Zu  $\lambda_2 = 1 + \sqrt{2}$ :

$$(A - (1 + \sqrt{2})E) \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \end{pmatrix} = \begin{pmatrix} -\sqrt{2} & 1 & 0 \\ 1 & -\sqrt{2} & 1 \\ 0 & 1 & -\sqrt{2} \end{pmatrix} \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \end{pmatrix} = 0,$$

woraus folgt

$$x^2 = \frac{1}{2} \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix},$$

Zu  $\lambda_3 = 1 - \sqrt{2}$  entsprechend

$$x^3 = \frac{1}{2} \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}.$$

Damit kann auch die Hauptachsentransformation

$$U = \begin{pmatrix} -1/\sqrt{2} & 1/2 & 1/2 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/2 & 1/2 \end{pmatrix}$$

gewonnen werden. □

Wir wissen jetzt: Eine lineare Abbildung  $A$ , die eine Basis aus Eigenvektoren besitzt, ist diagonalisierbar. Dies gilt insbesondere dann, wenn  $A$  selbstadjungiert ist (auch wenn  $A$  normal ist,  $A^*A = AA^*$ , was wir hier nicht zeigen). Ein anderer Spezialfall ist:

**Satz 2.16** *Ist  $A$  eine lineare Abbildung in einem  $m$ -dimensionalen Vektorraum mit  $m$  verschiedenen Eigenwerten, so ist  $A$  diagonalisierbar.*

*Beweis.* Nach Satz ?? sind die  $m$  Eigenvektoren zu den  $m$  verschiedenen Eigenwerten linear unabhängig. Sie bilden also eine Basis  $V$ . ■



**Beispiel 2.17** Im allgemeinen hat eine lineare Abbildung keine Basis aus Eigenvektoren, wie man schon an der durch die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

in  $\mathbb{K}^2$  erzeugten linearen Abbildung erkennt. Sie hat den Eigenvektor  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  zum Eigenwert 0. Das charakteristische Polynom  $p_A(\lambda) = \lambda^2$  liefert nur den Eigenwert 0, zu dem es aber keinen zweiten (linear unabhängigen) Eigenvektor gibt.  $\square$

Das vorausgehende Beispiel ist ein Spezialfall für die *JORDANsche Normalform* einer Matrix.

**Satz 2.18 (Jordansche Normalform)** *Ist  $A$  eine beliebige lineare Abbildung in einem endlichdimensionalen Vektorraum, so gibt es eine Basis in  $V$ , bezüglich der  $A$  dargestellt wird durch eine Matrix der Form*

$$\begin{pmatrix} C_1 & & 0 \\ & C_2 & \\ & & \ddots \\ 0 & & & C_r \end{pmatrix}$$

mit sogenannten *JORDAN-Kästen*

$$C_k = \begin{pmatrix} & 1 & & & 0 \\ \lambda_k & & & & \\ & \lambda_k & 1 & & \\ & & \lambda_k & & \\ & & & \ddots & 1 \\ & & & & \lambda_k \end{pmatrix} \in \mathbb{K}^{\nu_k, \nu_k}, \lambda_k \in \mathbb{K}.$$

Dabei ist natürlich  $\sum \nu_k = \dim V$ . Die  $\lambda_k$  für verschiedene  $k$  können gleich sein.

Man beachte, daß  $C_k$  in  $\mathbb{K}^{\nu_k}$  den einzigen Eigenwert  $\lambda_k$  hat mit dem einzigen Eigenelement  $e^1$ .  $\lambda_k$  ist  $\nu_k$ -fache Nullstelle des charakteristischen Polynoms von  $C_k$ .  $\lambda_k$  ist Eigenwert von  $C_k$  mit geometrischer Vielfachheit 1 und algebraischer Vielfachheit  $\nu_k$ . (Geometrische Vielfachheit eines Eigenwerts = Zahl der linear unabhängigen Eigenvektoren zu diesem Eigenwert = Dimension des zugehörigen Eigenraumes; algebraische Vielfachheit = Vielfachheit des Eigenwerts als Nullstelle des charakteristischen Polynoms = Gesamtdimension der zugehörigen Jordan-Kästen.)

**Beispiel 2.19** Die Matrix  $A = \begin{pmatrix} 1 & -1 \\ 4 & -3 \end{pmatrix}$  hat das charakteristische Polynom

$$p_A(\lambda) := \det(A - \lambda E) = \det \begin{pmatrix} 1 - \lambda & -1 \\ 4 & -3 - \lambda \end{pmatrix} = (\lambda + 1)^2,$$

d. h.  $\lambda = -1$  ist Eigenwert mit algebraischer Vielfachheit 2. Da

$$A - \lambda E = A + E = \begin{pmatrix} 2 & -1 \\ 4 & -2 \end{pmatrix} \neq 0$$

ist, gibt es (bis auf Vielfache) nur einen Eigenvektor, z. B.

$$u_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Mit dem zweiten Vektor  $u_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  gilt:

$$Au_1 = -u_1 \quad \text{und} \quad Au_2 = u_1 - u_2,$$

d. h. bezüglich der Basis  $\{u_1, u_2\}$  wird  $A$  durch seine Jordansche Normalform

$$\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$$

dargestellt. □

### 3 Stetige Funktionen mehrerer Variablen

Bei diesem Abschnitt handelt es sich teilweise um eine Wiederholung von Stoff aus dem 1. Semester.

Auf  $\mathbb{R}^m$  wird durch

$$d(x, y) := \left\{ \sum_{j=1}^m (\xi_j - \eta_j)^2 \right\}^{1/2} \quad \text{für} \quad \begin{array}{l} x = (\xi_1, \dots, \xi_m) \\ y = (\eta_1, \dots, \eta_m) \end{array}$$

ein *Abstand* definiert (analog in  $\mathbb{C}^m$ , wobei dann  $(\xi_j - \eta_j)^2$  durch  $(\xi_j - \eta_j)\overline{(\xi_j - \eta_j)}$  zu ersetzen ist).

Genauer hat dieses  $d$  die Eigenschaften einer *Metrik*:

(M1)  $d(x, y) \geq 0$  für alle  $x, y$ ;  $d(x, y) = 0 \Leftrightarrow x = y$ .

(M2)  $d(x, y) = d(y, x)$ , *Symmetrie*.

(M3)  $d(x, z) \leq d(x, y) + d(y, z)$ , *Dreiecksungleichung*.

Eine Menge  $X$  ausgestattet mit einer Metrik  $d$  nennt man einen *metrischen Raum*  $(X, d)$ . Eine Teilmenge  $Y$  von  $X$  ausgestattet mit der Metrik  $d$  aus  $(X, d)$  ist wieder ein metrischer Raum (ein Teilraum von  $(X, d)$  mit der *induzierten Metrik*).

Für eine Folge  $(x_n)$  aus  $X$  sagt man  $(x_n)$  *konvergiert gegen*  $x$ ,  $x_n \rightarrow x$ ,  $\lim_{n \rightarrow \infty} x_n = x$ , wenn gilt  $d(x_n, x) \rightarrow 0$ . Offenbar gilt  $x_n \rightarrow x$  in  $\mathbb{R}^m$  genau dann, wenn  $\xi_{n,j} \rightarrow \xi_j$  für  $j = 1, \dots, m$  gilt ( $x = (\xi_1, \dots, \xi_m)$ ,  $x_n = (\xi_{n,1}, \dots, \xi_{n,m})$ ).

Sind  $(X, d)$  und  $(Y, e)$  metrische Räume, so heißt eine Abbildung (im Fall  $Y = \mathbb{R}$  oder  $\mathbb{C}$ , oder auch  $\mathbb{R}^m$  oder  $\mathbb{C}^m$  redet man meist von einer Funktion)  $f : X \rightarrow Y$  *stetig* im Punkt  $x \in X$ , wenn gilt

- $\forall \varepsilon > 0 \exists \delta > 0$  so dass  $\forall y \in X$  mit  $d(x, y) < \delta$  gilt  $e(f(x), f(y)) < \varepsilon$ , oder
- für jede Folge  $(x_n)$  aus  $X$  mit  $x_n \rightarrow x \in X$  gilt  $f(x_n) \rightarrow f(x)$ .  $f$  heißt *stetig*, wenn es in jedem  $x \in X$  stetig ist.

Die  $d$ -Kugel um  $x$  mit dem Radius  $\varepsilon \geq 0$  ist die Menge

$$K_\varepsilon(x) := \{y \in X : d(x, y) < \varepsilon\}$$

(insbesondere ist  $K_0(x) = \emptyset$ ).

Eine Teilmenge  $O$  von  $X$  heißt *offen*, wenn zu jedem  $x \in O$  ein  $\varepsilon > 0$  existiert mit  $K_\varepsilon(x) \subset O$ .

Insbesondere ist  $K_\varepsilon(x)$  offen: Ist  $\varepsilon = 0$ , so ist  $K_\varepsilon(x) = \emptyset$  und die Bedingung ist offensichtlich erfüllt. Ist  $\varepsilon > 0$ , so ist für jedes  $y \in K_\varepsilon(x)$  die Kugel  $K_{\varepsilon-d(x,y)}(y)$  in  $K_\varepsilon(x)$  enthalten, denn nach der Dreiecksungleichung gilt für  $z \in K_{\varepsilon-d(x,y)}(y)$

$$d(x, z) = d(x, y) + d(y, z) < d(x, y) + \varepsilon - d(x, y) = \varepsilon.$$

**6: 9. Mai 2006** Eine Teilmenge  $A$  von  $X$  heißt *abgeschlossen*, wenn ihr Komplement  $X \setminus A$  offen ist. Insbesondere ist jede (sogenannte abgeschlossene) Kugel

$$\overline{K}_\varepsilon(x) := \{y \in X : d(x, y) \leq \varepsilon\} \quad (\varepsilon \geq 0)$$

abgeschlossen: Ist  $y \in X \setminus \overline{K}_\varepsilon(x)$ , so ist  $d(x, y) > \varepsilon$ , also nach der Dreiecksungleichung  $K_{d(x,y)-\varepsilon}(y) \subset X \setminus \overline{K}_\varepsilon(x)$ .

Endliche Vereinigungen/Durchschnitte offener/abgeschlossener Mengen sind offen/abgeschlossen. Es ist leicht zu sehen, daß beliebige Vereinigungen offener Mengen offen sind, während beliebige Durchschnitte abgeschlossener Mengen abgeschlossen sind (Beispiele!?).

Oben wurde die Stetigkeit punktweise definiert. Auf Grund des folgenden Satzes ist die Stetigkeit (in allen Punkten) äquivalent zu einer schönen globalen Eigenschaft:

**Satz 3.1** Seien  $(X, d)$  und  $(Y, e)$  metrische Räume.  $f : X \rightarrow Y$  ist genau dann stetig, wenn das Urbild jeder offenen Menge offen ist, d. h.

$$O \text{ offen in } Y \Rightarrow f^{-1}(O) := \{x \in X : f(x) \in O\} \text{ ist offen.}$$

(Die Schreibweise  $f^{-1}$  impliziert nicht, daß  $f$  injektiv ist.)

*Beweis.*  $\Leftarrow$ : Für beliebiges  $x \in X$  ist die Stetigkeit im Punkt  $x$  zu beweisen: Zu jedem  $\varepsilon > 0$  ist  $K_\varepsilon(f(x))$  offen in  $Y$  und  $x \in f^{-1}(K_\varepsilon(f(x)))$ . Also existiert ein  $\delta > 0$  mit  $K_\delta(x) \subset f^{-1}(K_\varepsilon(f(x)))$ .

Damit gilt für jedes  $y \in K_\delta(x)$

$$f(y) \in K_\varepsilon(f(x)),$$

d. h.  $f$  ist stetig in  $x$ .

$\Rightarrow$ : Sei  $O$  offen in  $Y$ ,  $x \in f^{-1}(O)$ . Da  $O$  offen ist, existiert ein  $\varepsilon > 0$  mit  $K_\varepsilon(f(x)) \subset O$ . Da  $f$  in  $x$  stetig ist, existiert ein  $\delta > 0$  mit

$$f(y) \in K_\varepsilon(f(x)) \text{ für alle } y \in K_\delta(x).$$

Also ist  $K_\delta(x) \subset f^{-1}(K_\varepsilon(f(x))) \subset f^{-1}(O)$ , d. h.  $f^{-1}(O)$  ist offen. ■

In  $\mathbb{R}^m$  bzw.  $\mathbb{C}^m$  kann man außer der Stetigkeit noch definieren:

- *partiell stetig* (d. h. stetig in den einzelnen Komponenten von  $x$ ), und
- *richtungsstetig*.

Wir haben gesehen:

stetig  $\implies$  richtungsstetig  $\implies$  partiell stetig,  
 aber (vgl. Teil I, Beispiel 8.1)  
 partiell stetig  $\not\Rightarrow$  richtungsstetig  $\not\Rightarrow$  stetig.

Für eine Teilmenge  $A$  eines metrischen Raumes definiert man:

die *abgeschlossene Hülle* von  $A$

$\bar{A} := \bigcap$  aller abgeschlossenen Mengen, die  $A$  enthalten  
 = kleinste abgeschlossene Menge, die  $A$  enthält =  $\{y \in X : \exists (x_n) \text{ aus } A \text{ mit } x_n \rightarrow y\}$

und den *offenen Kern* von  $A$

$\overset{\circ}{A} := \bigcup$  aller offenen Teilmengen von  $A$  = größte offene Teilmenge von  $A$   
 =  $\{x \in A : \exists \varepsilon > 0 \text{ mit } K_\varepsilon(x) \subset A\}$ .

$(X, d)$  heißt *vollständig*, wenn jede Cauchyfolge in  $X$  konvergiert. Insbesondere sind  $\mathbb{R}^m$  und  $\mathbb{C}^m$  mit den oben angegebenen Metriken vollständig.

**Satz 3.2** *Seien  $(X, d)$ ,  $(Y, e)$  metrische Räume,  $Y$  vollständig,  $A \subset X$  und  $f : A \rightarrow Y$  gleichmäßig stetig<sup>2</sup>. Dann läßt sich  $f$  eindeutig stetig auf  $\bar{A}$  fortsetzen.*

*Beweis.* Sei  $x \in \bar{A}$  (beliebig, es darf auch aus  $A$  sein). Es existiert eine Folge  $(x_n)$  aus  $A$  mit  $x_n \rightarrow x$ . Wegen der gleichmäßigen Stetigkeit ist dann  $(f(x_n))$  eine Cauchyfolge, also konvergent. Man kann also definieren

$$\tilde{f}(x) := \lim_{n \rightarrow \infty} f(x_n).$$

Diese Definition ist unabhängig von der Wahl der Folge  $(x_n)$  (mit  $x_n \rightarrow x$ ) (Beweis!); für  $x \in A$  ist  $\tilde{f}(x) = f(x)$ .

Bleibt zu zeigen: Für jedes  $x \in \bar{A}$  ist  $\tilde{f}$  stetig in  $x$ . Dazu sei  $(x_n)$  eine Folge aus  $\bar{A}$  mit  $x_n \rightarrow x$ . Dann existieren  $y_n \in A$  mit  $d(y_n, x_n) < \frac{1}{n}$  und  $e(f(y_n), f(x_n)) < \frac{1}{n}$ , also

$$e(f(x), f(x_n)) \leq e(f(x), f(y_n)) + e(f(y_n), f(x_n)) \rightarrow 0$$

wegen  $e(f(x), f(y_n)) \rightarrow 0$ , da  $y_n \rightarrow x$  gilt, und  $e(f(y_n), f(x_n)) < \frac{1}{n}$ . ■

Ein metrischer Raum  $(X, d)$  heißt *kompakt*, wenn jede Folge  $(x_n)$  aus  $X$  eine in  $X$  konvergente Teilfolge enthält. Dies ist äquivalent zu der folgenden Überdeckungseigenschaft (die in allgemeineren Situationen als Definition benutzt wird; in  $\mathbb{R}^m$  ist dies der Satz von Heine–Borel, vgl. z. B. H. Heuser, Analysis1, S. 157):

<sup>2</sup>D. h.  $\forall \varepsilon > 0 \exists \delta > 0 \forall x, z \in X$  mit  $d(x, z) < \delta$  gilt  $e(f(x), f(z)) < \varepsilon$ .

Jede offene Überdeckung  $\{O_\alpha, \alpha \in A\}$  von  $X$  enthält eine endliche Überdeckung (d. h. es gibt  $n \in \mathbb{N}$  und  $\alpha_1, \dots, \alpha_n \in A$  mit  $X \subset \bigcup_{j=1}^n A_{\alpha_j}$ ).

Jetzt noch drei wichtige Sätze über Funktionen auf metrischen Räumen.

**Satz 3.3 (Satz vom Maximum)** Sei  $(X, d)$  ein kompakter Raum,  $f : X \rightarrow \mathbb{R}$  stetig. Dann existieren  $x_m$  und  $x_M$  in  $X$  mit

$$f(x_m) = \min_{x \in X} f(x) \quad \text{und} \quad f(x_M) = \max_{x \in X} f(x)$$

*Beweis.* Dieser geht genauso wie für Funktionen auf  $\mathbb{R}$ . Zunächst wird die Beschränktheit von  $f$  gezeigt: Wäre  $f$  unbeschränkt, so gäbe es eine Folge  $(x_n)$  mit  $|f(x_n)| \geq n$ . Da  $X$  kompakt ist, gibt es eine Teilfolge  $(x_{n_k})$  mit  $x_{n_k} \rightarrow x_0$  für  $k \rightarrow \infty$ . Wegen der Stetigkeit von  $f$  folgt daraus  $f(x_{n_k}) \rightarrow f(x_0)$  im Widerspruch zu  $|f(x_{n_k})| \geq n_k$ .

Also ist die Menge  $\{f(x) : x \in X\}$  beschränkt, besitzt also ein Supremum (Infimum), d. h. es gibt eine Folge  $(x_n)$  mit  $f(x_n) \rightarrow \sup f$  ( $f(x_n) \rightarrow \inf f$ ). Wieder gibt es eine Teilfolge  $(x_{n_k})$  mit  $x_{n_k} \rightarrow x_M$  ( $x_{n_k} \rightarrow x_m$ ) und somit wegen der Stetigkeit von  $f$

$$f(x_M) = \lim_{k \rightarrow \infty} f(x_{n_k}) = \sup f \quad (f(x_m) = \lim_{k \rightarrow \infty} f(x_{n_k}) = \inf f).$$

$f$  nimmt also in  $x_M$  bzw.  $x_m$  sein Maximum bzw. Minimum an. ■

**Satz 3.4 (Banachscher Fixpunktsatz)** Sei  $(X, d)$  ein vollständiger metrischer Raum. Ist  $f : X \rightarrow X$  eine Kontraktion, d. h. es gilt

$$d(f(x), f(y)) \leq Ld(x, y) \quad \forall x, y \in X$$

mit einem  $L < 1$ , so besitzt  $f$  genau einen Fixpunkt  $\bar{x}$ ,  $f(\bar{x}) = \bar{x}$ . Genauer gilt: Für jedes  $x \in X$  konvergiert die Folge  $(f_n(x))$  mit  $f_0(x) = x$ ,  $f_{n+1}(x) = f(f_n(x))$  gegen  $\bar{x}$ , und es gilt

$$d(f_n(x), \bar{x}) \leq \frac{L^n}{1-L} d(f(x), x) \quad \forall n \in \mathbb{N}.$$

*Beweis.* Sei  $x \in X$ . Dann gilt

$$d(f_{n+1}(x), f_n(x)) \leq Ld(f_n(x), f_{n-1}(x)) \leq \dots \leq L^n d(f(x), x), \quad (*)$$

und entsprechend

$$d(f_m(x), f_n(x)) \leq L^n d(f_{m-n}(x), x) \quad \text{für } m \geq n.$$

Durch Induktion zeigt man

$$d(f_n(x), x) \leq \frac{1-L^n}{1-L} d(f(x), x) :$$

$n = 1$ : Offensichtlich richtig.

$n \Rightarrow n + 1$ :

$$\begin{aligned} d(f_{n+1}(x), x) &\leq d(f_{n+1}(x)f_n(x)) + d(f_n(x), x) \\ &\quad (\text{Ungleichung } (*) \text{ u. Induktionsannahme}) \\ &\leq L^n d(f(x), x) + \frac{1-L^n}{1-L} d(f(x), x) \\ &= \frac{1-L^{n+1}}{1-L} d(f(x), x). \end{aligned}$$

Damit folgt dann

$$d(f_m(x), f_n(x)) \leq L^n d(f_{m-n}(x), x) \leq L^n \frac{1-L^{m-n}}{1-L} d(f(x), x).$$

Insbesondere ist  $(f_n(x))$  eine Cauchyfolge, also konvergent,  $\bar{x} := \lim_{n \rightarrow \infty} f_n(x)$ .

Da  $f$  stetig ist, gilt (beachte  $f_{n-1}(x) \rightarrow \bar{x}$ )

$$\bar{x} = \lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} f(f_{n-1}(x)) = f(\bar{x}).$$

Also ist  $\bar{x}$  ein Fixpunkt von  $f$ . Für  $m \rightarrow \infty$  folgt aus obiger Ungleichung

$$d(\bar{x}, f_n(x)) \leq \frac{L^n}{1-L} d(f(x), x).$$

Es bleibt die *Eindeutigkeit* zu zeigen: Sind  $\bar{x}$  und  $\hat{x}$  Fixpunkte, so folgt

$$d(\bar{x}, \hat{x}) = d(f(\bar{x}), f(\hat{x})) \leq L d(\bar{x}, \hat{x}).$$

Das ist nur möglich, wenn  $d(\bar{x}, \hat{x}) = 0$ , also  $\bar{x} = \hat{x}$ , gilt. ■

**Beispiel 3.5** In  $x = [0, \infty)$  mit der üblichen Metrik sei

$$f(x) = x + \frac{1}{1+x}.$$

Dann gilt  $|f(x) - f(y)| < |x - y|$ : Für  $x > y$  folgt das aus

$$f(x) - f(y) = x - y + \frac{1}{1+x} - \frac{1}{1+y} = x - y + \frac{1+y-1-x}{(1+x)(1+y)} \begin{cases} > 0, \\ < x - y, \end{cases}$$

entsprechend für  $x < y$ .  $f$  hat aber keinen Fixpunkt, denn

$$\text{aus } \bar{x} = \bar{x} + \frac{1}{1+\bar{x}} \text{ folgt } 0 = \frac{1}{1+\bar{x}},$$

ein Widerspruch. (In obigem Satz ist also  $d(f(x), f(y)) \leq Ld(x, y)$  mit  $L < 1$  nicht durch  $d(f(x), f(y)) < d(x, y)$  ersetzbar.) □

**Beispiel 3.6** In  $X = (0, \infty)$  mit der üblichen Metrik (beachte  $X$  ist nicht vollständig) sei  $f(x) = Lx$  mit  $0 < L < 1$ . Offenbar ist  $f$  eine Kontraktion, aber  $f$  hat keinen Fixpunkt. In obigem Satz kann also auf die Vollständigkeit nicht verzichtet werden.  $\square$

**Satz 3.7** Seien  $(X, d)$  und  $(Y, e)$  metrische Räume,  $X$  kompakt und  $f : X \rightarrow Y$  stetig, so ist  $f$  gleichmäßig stetig.

*Beweis.* Annahme:  $f$  ist *nicht* gleichmäßig stetig. Dann gilt

$$\exists \varepsilon > 0 \quad \forall \delta > 0 \quad \exists x, z \in X \quad \text{mit } d(x, z) < \delta \quad \text{und } e(f(x), f(z)) > \varepsilon$$

oder

$$\exists \varepsilon > 0 \quad \forall n \in \mathbb{N} \quad \exists x_n, z_n \in X \quad \text{mit } d(x_n, z_n) < \frac{1}{n} \quad \text{und } e(f(x_n), f(z_n)) > \varepsilon.$$

Da  $X$  kompakt ist, existiert eine Teilfolge  $(n_k)$  von  $\mathbb{N}$  mit  $(x_{n_k}) \rightarrow x$ ,  $(z_{n_k}) \rightarrow z$ . Wegen  $d(x_{n_k}, z_{n_k}) < \frac{1}{n_k}$  ist  $x = z$ , also  $\lim f(x_{n_k}) = \lim f(z_{n_k})$  im Widerspruch zu  $e(f(x_{n_k}), f(z_{n_k})) > \varepsilon$ .  $\blacksquare$

## 4 Differenzierbare Funktionen mehrerer Variablen

Aus der Analysis der Funktionen einer Variablen kennen wir zwei äquivalente Definitionen der Differenzierbarkeit: Eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  ist *differenzierbar* im Punkt  $x_0$ , wenn gilt

– der Limes

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \rightarrow x_0} \frac{f(x_0 + h) - f(x_0)}{h}$$

existiert; dieser Limes ist die *Ableitung* von  $f$  im Punkt  $x_0$  (= *Steigerung* der Tangente an die Kurve im Punkt  $(x_0, f(x_0))$  = Tangens des Steigungswinkels), oder

–  $f$  ist in der Umgebung von  $x_0$  *linear approximierbar*, d. h. es gilt

$$f(x) = f(x_0) + m(x - x_0) + \varphi_{x_0}(x)$$

mit

$$\frac{\varphi_{x_0}(x)}{|x - x_0|} \rightarrow 0 \quad \text{für } x \rightarrow x_0;$$

hier ist  $m$  die *Ableitung* von  $f$  im Punkt  $x_0$ .

Für Funktionen, die auf  $\mathbb{R}^m$  definiert sind (gleichgültig, ob sie nach  $\mathbb{R}$  oder  $\mathbb{R}^n$  abbilden) ist die erste Definition nicht brauchbar, da nicht durch einen Vektor  $x - x_0$  dividiert werden kann. Die zweite Definition läßt sich dagegen unmittelbar verallgemeinern.

Sei  $A \subset \mathbb{R}^m$  offen (damit jeder Punkt in  $A$  eine Umgebung hat, die ganz in  $A$  liegt),  $f : A \rightarrow \mathbb{R}^n$ ,  $x_0 \in A$ .  $f$  heißt bei  $x_0$  (*total*) *differenzierbar*, wenn eine (natürlich i. allg. von  $x_0$  abhängige) lineare Abbildung  $L : \mathbb{R}^m \rightarrow \mathbb{R}^n$  existiert mit

$$f(x) = f(x_0) + L(x - x_0) + \varphi_{x_0}(x) \quad \text{mit} \quad \frac{\varphi_{x_0}(x)}{|x - x_0|} \rightarrow 0 \quad \text{für } x \rightarrow x_0$$

bzw.

$$\frac{1}{|x - x_0|} (f(x) - f(x_0) - L(x - x_0)) \rightarrow 0 \quad \text{für } x \rightarrow x_0.$$

(Die letzte Zeile macht auch deutlich, daß  $\varphi_{x_0}(x) = f(x) - f(x_0) - L(x - x_0)$  ist, und daß die Forderung an  $\varphi_{x_0}(\cdot)$  sich eigentlich auf diesen Ausdruck bezieht.)

**Satz 4.1** *Ist  $f$  in  $x_0$  differenzierbar, so ist die lineare Abbildung  $L$  eindeutig bestimmt.*

*Beweis.* Haben  $L_1$  und  $L_2$  die gewünschte Eigenschaft, so gilt für  $M := L_1 - L_2$

$$\begin{aligned} \lim_{x \rightarrow x_0} \frac{Mx}{|x|} &= \lim_{x \rightarrow x_0} \frac{M(x - x_0)}{|x - x_0|} = \lim_{x \rightarrow x_0} \frac{L_1(x - x_0) - L_2(x - x_0)}{|x - x_0|} \\ &= \lim_{x \rightarrow x_0} \frac{\varphi_{2,x_0}(x) - \varphi_{1,x_0}(x)}{|x - x_0|} = 0 \end{aligned}$$

(wobei  $\varphi_{j,x_0}(\cdot)$  die Funktion ist, die in obiger Definition von  $L_j$  auftritt). Also gilt für alle  $x \in \mathbb{R}^m$

$$Mx = \frac{M\left(\frac{x}{n}\right)}{\left|\frac{x}{n}\right|} |x| \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$



Da die linke Seite nicht von  $n$  abhängt, gilt also  $Mx = 0$  für alle  $x \in \mathbb{R}^m$ , d. h.  $M = 0$ ,  $L_1 = L_2$ . ■

Diese eindeutig bestimmte lineare Abbildung  $L : \mathbb{R}^m \rightarrow \mathbb{R}^n$  heißt die *Ableitung* von  $f$  im Punkt  $x_0$ . Man schreibt dafür  $Df(x_0)$ ,  $f'(x_0)$ .

$f : A \rightarrow \mathbb{R}^n$  heißt *differenzierbar*, wenn es in jedem Punkt von  $A$  differenzierbar ist. Dann ist  $Df : A \rightarrow L(\mathbb{R}^m, \mathbb{R}^n)$ ,  $x \mapsto Df(x)$  wieder eine Funktion, jetzt von  $A \subset \mathbb{R}^m$  in den Raum  $L(\mathbb{R}^m, \mathbb{R}^n)$  der linearen Abbildungen von  $\mathbb{R}^m$  nach  $\mathbb{R}^n$ .

**Satz 4.2**  $A \subset \mathbb{R}^m$  offen,  $f : A \rightarrow \mathbb{R}^n$  differenzierbar (im Punkt  $x_0 \in A$ ). Dann ist  $f$  stetig (im Punkt  $x_0$ ).

*Beweis.* Da jede lineare Abbildung von  $\mathbb{R}^m \rightarrow \mathbb{R}^n$  stetig ist (Beweis?), folgt dies direkt mit der obigen Definition der Differenzierbarkeit. ■

**Beispiel 4.3**  $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ ,  $f(x) = Ax$  mit einer linearen Abbildung  $A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ . Offenbar ist  $f(x) = f(x_0) + A(x - x_0) + 0$ , d. h.  $Df(x) = A$  für alle  $x$ . □

**Beispiel 4.4**  $f : \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $f(x) = |x|^2 = \sum_{j=1}^m \xi_j^2$  ist in 0 differenzierbar mit Ableitung 0, denn es gilt

$$f(x) - f(0) - 0(x - 0) = f(x) =: \varphi_0(x), \quad \varphi_0(x) = |x|^2,$$

also

$$\frac{\varphi_0(x)}{|x - 0|} = \frac{|x|^2}{|x|} \rightarrow 0 \quad \text{für } |x| \rightarrow 0.$$

□

Für allgemeinere Fälle (und auch in obigem Beispiel für  $x \neq 0$ ) ist es mit unseren bisherigen Kenntnissen nicht leicht, die Ableitung explizit auszurechnen bzw. die Differenzierbarkeit zu zeigen.

Deshalb schauen wir uns zunächst die einzelnen Komponenten von  $x$  und  $f$  an: Sei also  $f$  in  $x_0$  differenzierbar,

$$L := Df(x_0) = (\ell_{jk})_{\substack{j=1, \dots, n \\ k=1, \dots, m}},$$

also (mit  $x_0 + x$  statt  $x$ )

$$f(x_0 + x) - f(x_0) - \varphi_{x_0}(x_0 + x) = Lx = \left( \sum_{p=1}^m \ell_{jp} \xi_p \right).$$

Speziell erhalten wir für die  $j$ -te Komponente von  $f$

$$\frac{1}{|x|} \left( f_j(x_0 + x) - f_j(x_0) - \sum_{p=1}^m \ell_{jp} \xi_p \right) = \frac{\varphi_{x_0, j}(x_0 + x)}{|x|}.$$

Spezialisieren wir nun weiter

$$x = he^k = (0, \dots, 0, h, 0, \dots, 0) \quad (h \text{ an der } k\text{-ten Stelle}),$$

so folgt

$$\text{für } h > 0 \text{ (also } |x| = h\text{): } \frac{1}{h} \{f_j(x_0 + he^k) - f_j(x_0)\} - \ell_{jk} \rightarrow 0 \text{ für } h \rightarrow 0,$$

$$\text{für } h < 0 \text{ (also } |x| = -h\text{): } \frac{1}{-h} \{f_j(x_0 + he^k) - f_j(x_0)\} + \ell_{jk} \rightarrow 0 \text{ für } h \rightarrow 0,$$

d. h. es gilt

$$\ell_{jk} = \lim_{h \rightarrow 0} \frac{1}{h} (f_j(x_0 + he^k) - f_j(x_0)) =: \frac{\partial f_j}{\partial \xi_k}(x_0).$$

Dieser Grenzwert heißt die *partielle Ableitung* von  $f_j$  nach der  $k$ -ten Komponente im Punkt  $x_0$ . (Für  $\frac{\partial f_j}{\partial \xi_k}$  schreibt man auch  $D_k f_j, \partial_k f_j, f_{j,k}$ .) Wir haben also gezeigt:

**8: 16. Mai 2006**

**Satz 4.5**  $A \subset \mathbb{R}^m$  offen,  $f : A \rightarrow \mathbb{R}^n$  differenzierbar in  $x_0$ . Dann existieren alle partiellen Ableitungen  $\frac{\partial f_j}{\partial \xi_k}(x_0)$  ( $j = 1, \dots, n, k = 1, \dots, m$ ) und es gilt

$$Df(x_0) = \left( \frac{\partial f_j}{\partial \xi_k}(x_0) \right)_{\substack{j=1, \dots, n \\ k=1, \dots, m}} \quad \text{JACOBI-Matrix im Punkt } x_0.$$

Diese Matrix wird auch als *Differentialmatrix* oder *Funktionalmatrix* im Punkt  $x_0$  bezeichnet; man schreibt auch  $\left( \frac{\partial (f_1, \dots, f_n)}{\partial (\xi_1, \dots, \xi_m)} \right)(x_0)$ .

Man definiert nun für  $f : A \rightarrow \mathbb{R}^n$ :

- $f$  heißt *im Punkt  $x_0$  partiell differenzierbar*, wenn dort alle partiellen Ableitungen existieren,
- $f$  heißt *partiell differenzierbar*, wenn es in allen Punkten von  $A$  partiell differenzierbar ist,
- $f$  heißt *stetig partiell differenzierbar*, wenn es partiell differenzierbar ist und alle partiellen Ableitungen stetig sind.

Der obige Satz besagt: *Aus der Differenzierbarkeit folgt die partielle Differenzierbarkeit.*

Gilt auch die Umkehrung? **Nein**, wie folgendes Beispiel zeigt:

**Beispiel 4.6** Die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x) = f(\xi_1, \xi_2) = \begin{cases} \frac{\xi_1 \xi_2}{\xi_1^2 + \xi_2^2} & \text{für } (\xi_1, \xi_2) \neq (0, 0), \\ 0 & \text{für } (\xi_1, \xi_2) = (0, 0) \end{cases}$$

ist jedenfalls außerhalb 0 (sogar stetig) partiell differenzierbar. Da  $f$  auf beiden Achsen identisch verschwindet, existieren die partiellen Ableitungen auch in 0 und sind beide gleich 0.  $f$  ist aber im Nullpunkt nicht stetig, und somit sicher nicht differenzierbar: Für alle  $t \neq 0$  ist  $f(t, t) = \frac{1}{2}$ , während  $f(0, 0) = 0$  ist.  $\square$

In den Punkten  $x \neq 0$  ist die Funktion dieses Beispiels total differenzierbar, wie aus folgendem Satz folgt:

**Satz 4.7**  $A \subset \mathbb{R}^m$  offen,  $f : A \rightarrow \mathbb{R}^n$  stetig partiell differenzierbar. Dann ist  $f$  differenzierbar. (Tatsächlich wird gezeigt: Ist  $f$  partiell differenzierbar nahe  $x_0$  und sind die partiellen Ableitungen bei  $x_0$  stetig, so ist  $f$  bei  $x_0$  total differenzierbar. Wenn alle partiellen Ableitungen stetig sind, sind alle Einträge in der Jacobi-Matrix stetig; man sagt dann, daß  $f$  stetig differenzierbar ist.)

*Beweis.* Sei  $x_0 \in A$ ,  $L := (D_j f_k(x_0))_{\substack{j=1, \dots, m \\ k=1, \dots, n}}$ . Wir haben zu zeigen: zu jedem  $\varepsilon > 0$  existiert ein  $\delta > 0$  so, dass für alle  $x \in A$  mit  $|x - x_0| < \delta$  gilt

$$\frac{\varphi_{x_0}(x)}{|x - x_0|} := \frac{|f(x) - f(x_0) - L(x - x_0)|}{|x - x_0|} < \varepsilon.$$

Sei also  $\varepsilon > 0$ . Da alle  $D_j f_k(\cdot)$  stetig sind, existiert ein  $\delta > 0$  mit

$$|D_j f_k(x) - D_j f_k(x_0)| < \frac{\varepsilon}{2\sqrt{nm}} \text{ für } x \in A \text{ mit } |x - x_0| < \delta;$$

für  $x, y \in K_\delta(x_0)$  ist dann  $|D_j f_k(x) - D_j f_k(y)| < \varepsilon/\sqrt{nm}$ .

Sei  $x \in K_\delta(x_0)$ ,  $t = (\tau_1, \dots, \tau_m)^t := x - x_0$ ,

$$t_j := (\tau_1, \dots, \tau_j, 0, \dots, 0) = \sum_{\ell=1}^j \tau_\ell e^\ell \text{ für } j = 0, \dots, m.$$

( $t_0 = 0$ ). Dann gilt für  $k = 1, \dots, n$

$$f_k(x) - f_k(x_0) = \sum_{j=1}^m \left( f_k(x_0 + t_j) - f_k(x_0 + t_{j-1}) \right).$$

Nach dem 1-dimensionalen Mittelwertsatz der Differentialrechnung (vgl. Teil I, Satz 10.3) existieren  $\sigma_{j,k}$  zwischen 0 und  $\tau_j$  mit

$$\begin{aligned} f_k(x_0 + t_j) - f_k(x_0 + t_{j-1}) &= f_k(x_0 + t_{j-1} + \tau_j e^j) - f_k(x_0 + t_{j-1}) \\ &= \tau_j D_j f_k(x_0 + t_{j-1} + \sigma_{j,k} e^j), \\ \left| f_k(x_0 + t_j) - f_k(x_0 + t_{j-1}) - \tau_j D_j f_k(x_0) \right| &\leq |\tau_j| \left| D_j f_k(x_0 + t_{j-1} + \sigma_{j,k} e^j) - D_j f_k(x_0) \right| \\ &\leq |x - x_0| \frac{\varepsilon}{\sqrt{nm}}, \\ \left| f_k(x) - f_k(x_0) - \sum_{j=1}^m D_j f_k(x_0) \tau_j \right| &\leq \sum_{j=1}^m \left| f_k(x_0 + t_j) - f_k(x_0 + t_{j-1}) - \tau_j D_j f_k(x_0) \right| \\ &\leq |x - x_0| \frac{\varepsilon}{\sqrt{n}}, \\ \left| f(x) - f(x_0) - L(x - x_0) \right|^2 &= \sum_{k=1}^n \left| f_k(x) - f_k(x_0) - \sum_{j=1}^m D_j f_k(x_0) \tau_j \right|^2 \\ &\leq |x - x_0|^2 \frac{\varepsilon^2}{n} = |x - x_0|^2 \varepsilon^2, \end{aligned}$$

und damit schließlich

$$\frac{|f(x) - f(x_0) - L(x - x_0)|}{|x - x_0|} < \varepsilon \quad \text{für } |x - x_0| < \delta,$$

womit die Behauptung bewiesen ist. ■

Im Fall  $\mathbf{m} = \mathbf{1}$  ist die Ableitung nichts anderes als der aus den Ableitungen der einzelnen Komponenten von  $f$  gebildete Vektor:

$$f'(x) = Df(x) = \begin{pmatrix} f'_1(x) \\ \vdots \\ f'_n(x) \end{pmatrix}$$

Wie bei der Untersuchung der Stetigkeit passiert hier nichts wesentlich neues.

Interessanter ist der Fall  $\mathbf{m} > \mathbf{1}$  und  $\mathbf{n} = \mathbf{1}$ , eine reellwertige Funktion von  $m$  Variablen

$$f'(x) = Df(x) = \left( \frac{\partial}{\partial \xi_1} f(x), \dots, \frac{\partial}{\partial \xi_m} f(x) \right) =: \text{grad } f(x),$$

der *Gradient* von  $f$ . Als Matrix ist dies eine  $1 \times m$ -matrix; als Vektor (s. unten) schreibt man es wohl besser als Spalte, wie die Vektoren in  $\mathbb{R}$ . Für  $y = (\eta_1, \dots, \eta_m) \in \mathbb{R}^m$  ist

$$Df(x_0)y = \sum_{k=1}^m \frac{\partial}{\partial \xi_k} f(x_0) \eta_k = \langle Df(x_0), y \rangle.$$

Um die (geometrische) Bedeutung des Gradienten zu erkennen, betrachten wir zunächst *richtungsableitungen* von  $f$ . Die Richtungsableitung von  $f$  im Punkt  $x_0$  in Richtung  $e$  (mit  $\|e\| = 1$ ) ist definiert durch (falls der Limes existiert)

$$\left. \frac{\partial}{\partial t} f(x_0 + te) \right|_{t=0} := \lim_{t \rightarrow 0} \frac{1}{t} (f(x_0 + te) - f(x_0)).$$

Die partielle Ableitung  $\frac{\partial}{\partial \xi_k} f(x_0)$  ist also nichts anderes als die Richtungsableitung in Richtung des  $k$ -ten Einheitsvektors  $e^k$ .

**Satz 4.8 (Geometrische Deutung des Gradienten)** Sei  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  differenzierbar im Punkt  $x_0$ .

- Ist  $\text{grad } f(x_0) = 0$ , so sind alle Richtungsableitungen gleich 0.
- Ist  $\text{grad } f(x_0) \neq 0$ , so zeigt der Vektor  $\text{grad } f(x_0)$  in die Richtung  $e$  mit der größten Richtungsableitung (genauer: die Richtung, in der  $f$  am schnellsten wächst); die Norm des Gradienten ist gleich dieser Richtungsableitung.

*Beweis.* Ist  $e \in \mathbb{R}^m$  mit  $\|e\| = \left\{ \sum_{j=1}^m \varepsilon_j^2 \right\}^{1/2} = 1$ , so ist die Richtungsableitung von  $f$  in Richtung  $e$  im Punkt  $x_0$  (aufgrund der linearen Approximierbarkeit)

$$\begin{aligned}
 \frac{\partial}{\partial t} f(x_0 + te) \Big|_{t=0} &= \lim_{t \rightarrow 0} \frac{1}{t} \left( f(x_0 + te) - f(x_0) \right) \\
 &= \lim_{t \rightarrow 0} \frac{1}{t} \left\{ f(x_0) + \langle \text{grad } f(x_0), te \rangle + \varphi_{x_0}(x_0 + te) - f(x_0) \right\} \\
 &\quad \left( \text{da } \frac{1}{t} \varphi_{x_0}(x_0 + te) \rightarrow 0 \text{ gilt für } t \rightarrow 0 \right) \\
 &= \langle \text{grad } f(x_0), e \rangle \leq |\text{grad } f(x_0)| \\
 &\quad \left( \text{ab hier gilt es nur für } \text{grad } f(x_0) \neq 0 \right) \\
 &= \left\langle \text{grad } f(x_0), \frac{1}{|\text{grad } f(x_0)|} \text{grad } f(x_0) \right\rangle \\
 &\quad \left( \text{erneute Anwendung des dritten „=“-Zeichens, jetzt mit} \right. \\
 &\quad \left. e := \frac{1}{|\text{grad } f(x_0)} \text{grad } f(x_0) \right) \\
 &= \frac{\partial}{\partial t} f \left( x_0 + \frac{t}{|\text{grad } f(x_0)} \text{grad } f(x_0) \right) \Big|_{t=0}.
 \end{aligned}$$

Hieraus folgen beide Aussagen. ■

Geht man ein (nicht zu großes) Stück in Richtung des Gradienten, so vergrößert sich der Funktionswert. Diese Tatsache macht man sich beim *Gradientenverfahren* zur Bestimmung von Extremwerten zu Nutze.

#### Beispiel 4.9 Die Funktion

$$f : \mathbb{R}^m \rightarrow \mathbb{R}, f(x) = |x| = (\xi_1^2 + \dots + \xi_m^2)^{1/2}$$

ist für  $x \neq 0$  differenzierbar, nicht für  $x = 0$ . Geometrisch ist offensichtlich, daß gilt

$$\text{grad } f(x) = \frac{1}{|x|} (\xi_1, \dots, \xi_m) = \frac{1}{|x|} x \text{ für } x \neq 0.$$

Man rechne das explizit nach! □

Sind  $f, g : \mathbb{R}^m \rightarrow \mathbb{R}$  stetig differenzierbar, so folgt mit der Produktregel für Funktionen einer Variablen (hier angewandt auf partielle Ableitungen)

$$\text{grad}(f, g)(x) = f(x) \text{grad } g(x) + g(x) \text{grad } f(x).$$

Sei  $A \subset \mathbb{R}^m$ . Eine Abbildung  $f : A \rightarrow \mathbb{R}^m$  (gleiches  $m$ ) heißt ein *Vektorfeld* auf  $A$ . (Man stellt sich das so vor, dass in jedem Punkt von  $A$  ein Vektor angeheftet wird.) Ist  $f$  differenzierbar, so ist die *Divergenz* des Vektorfeldes  $f$  definiert durch

$$\text{div } f(x) := \sum_{j=1}^m \frac{\partial f_j}{\partial \xi_j}(x).$$

diese (reellwertige) Funktion wird auch als *Quellendichte* des Vektorfeldes  $f$  bezeichnet, was wir erst später verstehen werden.

Gelegentlich ist es nützlich, den „formalen Vektor“

$$\text{Nabla} = \nabla = \left( \frac{\partial}{\partial \xi_1}, \dots, \frac{\partial}{\partial \xi_m} \right)$$

zu betrachten (Nabla=antikes Saiteninstrument mit der form  $\nabla$ ). Damit kann man (formal) schreiben:

$$\text{grad } f = \nabla f \quad \text{für } f : \mathbb{R}^m \rightarrow \mathbb{R},$$

$$\text{div } f = \langle \nabla, f \rangle \quad \text{für } f : \mathbb{R}^m \rightarrow \mathbb{R}^m.$$

Ist  $A \subset \mathbb{R}^m$ ,  $f : A \rightarrow \mathbb{R}^m$ ,  $g : A \rightarrow \mathbb{R}$  stetig differenzierbar, so gilt

$$\begin{aligned} \text{div}(g, f)(x) &= \sum_{j=1}^m \frac{\partial}{\partial \xi_j} (fg_j)(x) = \sum_{j=1}^m \left\{ \left( \frac{\partial}{\partial \xi_j} g(x) \right) f_j(x) + g(x) \frac{\partial}{\partial \xi_j} f_j(x) \right\} \\ &= \langle \text{grad } g(x), f(x) \rangle + g(x) \text{div } f(x) = \langle \nabla g(x), f(x) \rangle + g(x) \langle \nabla, f(x) \rangle. \end{aligned}$$

**Beispiel 4.10** Das Vektorfeld  $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$ ,  $f(x) = x$  hat die Divergenz

$$\text{div } f(x) = \sum_{j=1}^m \frac{\partial \xi_j}{\partial \xi_j} = \sum_{j=1}^m 1 = m.$$

Für  $f : \mathbb{R}^m \setminus \{0\} \rightarrow \mathbb{R}^m$ ,  $f(x) = \frac{1}{|x|}x$  (= Einheitsvektor in Richtung  $x$ ) gilt

$$\begin{aligned} \text{div } f(x) &= \langle \text{grad } \frac{1}{|x|}, x \rangle + \frac{1}{|x|} \text{div } x \\ &= -\langle \frac{1}{|x|^3}x, x \rangle + \frac{m}{|x|} = \frac{m-1}{|x|} \end{aligned}$$

für  $x \neq 0$ . □

**9: 18. Mai 2006** Ähnlich wie im 1-dimensionalen gelten auch hier einige nützliche Differentiationsregeln:

**Satz 4.11 (Differentiationsregeln)** Sei  $A \subset \mathbb{R}^m$  offen.

- a) Sind  $f, g : A \rightarrow \mathbb{R}^n$  differenzierbar in  $x_0 \in A$ ,  $a, b \in \mathbb{R}$ , so ist auch  $af + bg$  differenzierbar in  $x_0$  mit

$$D(af + bg)(x_0) = aDf(x_0) + bDg(x_0).$$

- b) Produktregel: Sind  $f : A \rightarrow \mathbb{R}^n$  und  $g : A \rightarrow \mathbb{R}$  differenzierbar in  $x_0 \in A$ , so ist auch  $gf : A \rightarrow \mathbb{R}^n$  differenzierbar in  $x_0$  mit

$$D(g \cdot f)(x_0) : \mathbb{R}^m \rightarrow \mathbb{R}^n, \quad y \mapsto g(x_0)Df(x_0)y + (Dg(x_0)y)f(x_0),$$

speziell ist also

$$D_k(gf)(x_0) = g(x_0)D_k f(x_0) + D_k g(x_0)f(x_0).$$

- c) Kettenregel: Ist  $f : A \rightarrow \mathbb{R}^n$  differenzierbar in  $x_0$ ,  $B \subset \mathbb{R}^n$  offen mit  $f(A) \subset B$  und  $g : B \rightarrow \mathbb{R}^p$  differenzierbar im Punkt  $f(x_0)$ , so ist  $h := g \circ f$  differenzierbar in  $x_0$  mit

$$Dh(x_0) = Dg(f(x_0)) \cdot Df(x_0).$$

*Beweis.* a) Da  $f$  und  $g$  in  $x_0$  differenzierbar sind gilt (lineare Approximierbarkeit)

$$\begin{aligned} f(x) &= f(x_0) + Df(x_0)(x - x_0) + \varphi_{1,x_0}(x) \\ g(x) &= g(x_0) + Dg(x_0)(x - x_0) + \varphi_{2,x_0}(x) \end{aligned}$$

mit

$$\frac{\varphi_{j,x_0}(x)}{|x - x_0|} \rightarrow 0 \text{ für } x \rightarrow x_0, \quad j = 1, 2.$$

Damit folgt

$$\begin{aligned} (af + bg)(x) &= (af + bg)(x_0) + (aDf(x_0) + bDg(x_0))(x - x_0) + \varphi_{x_0}(x) \\ \text{mit } \varphi_{x_0}(x) &= a\varphi_{1,x_0}(x) + b\varphi_{2,x_0}(x), \text{ also } \frac{\varphi_{x_0}(x)}{|x - x_0|} \rightarrow 0 \text{ für } x \rightarrow x_0. \end{aligned}$$

- b) Für  $f$  und  $g$  gelten Darstellung wie oben. Damit folgt

$$g(x)f(x) = g(x_0)f(x_0) + g(x_0)Df(x_0)(x - x_0) + (Dg(x_0)(x - x_0)f(x_0) + \omega_{x_0}(x))$$

mit

$$\omega_{x_0}(x) = \text{Terme mit mindestens einem } \varphi_{j,x_0}(x),$$

also  $\frac{1}{|x - x_0|}\omega_{x_0}(x) \rightarrow 0$  für  $x \rightarrow x_0$ . Also ist

$$D(gf)(x_0)y = g(x_0)Df(x_0)y + (Dg(x_0)y)f(x_0).$$

- c) Wieder gelten für  $f$  und  $g$  Darstellungen wie oben; für  $g$  bedeutet das

$$g(y) = g(f(x_0)) + Dg(f(x_0))(y - f(x_0) + \varphi_{2,f(x_0)}(y))$$

mit

$$\frac{\varphi_{2,f(x_0)}(y)}{|y - f(x_0)|} \rightarrow 0 \text{ für } y \rightarrow f(x_0).$$

Damit folgt

$$\begin{aligned} h(x) &= g(f(x)) = g\left(f(x_0) + Df(x_0)(x - x_0) + \varphi_{1,x_0}(x)\right) \\ &= g(f(x_0)) + Dg(f(x_0))\left(Df(x_0)(x - x_0) + \varphi_{1,x_0}(x)\right) \\ &\quad + \varphi_{2,f(x_0)}\left(f(x_0) + Df(x_0)(x - x_0) + \varphi_{1,x_0}(x)\right) \\ &= h(x_0) + Dg(f(x_0))Df(x_0)(x - x_0) \\ &\quad + Dg(f(x_0))\varphi_{1,x_0}(x) + \varphi_{2,f(x_0)}(f(x)) \end{aligned}$$

mit

$$\frac{Dg(f(x_0))\varphi_{1,x_0}(x)}{|x - x_0|} = Dg(f(x_0))\frac{\varphi_{1,x_0}(x)}{|x - x_0|} \rightarrow 0 \text{ für } x \rightarrow x_0$$

und

$$\frac{\varphi_{2,f(x_0)}f(x)}{|x-x_0|} = \begin{cases} 0 & \text{für } f(x) = f(x_0), \\ \frac{\varphi_{2,f(x_0)}(f(x))}{|f(x)-f(x_0)|} \frac{|f(x)-f(x_0)|}{|x-x_0|} \rightarrow 0 & \text{für } f(x) \neq f(x_0), \end{cases} \rightarrow 0 \text{ für } x \rightarrow 0.$$

da für den letzten Term gilt

$$\frac{|f(x)-f(x_0)|}{|x-x_0|} = \left| \frac{Df(x_0)(x-x_0)}{|x-x_0|} + \frac{\varphi_{1,x_0}(x)}{|x-x_0|} \right| \leq C \text{ für } x \text{ nahe } x_0.$$

■



**Folgerung 4.12** *Speziell folgt aus der Kettenregel:*

$$D_k h(x_0) = Dh(x_0)e_k = Dg(f(x_0))Df(x_0)e_k = Dg(f(x_0))D_k f(x_0),$$

$$D_k h_j(x_0) = Dg_j(f(x_0))D_k f(x_0) = \sum_{\ell=1}^n D_\ell f_j(f(x_0))D_k f_\ell(x_0).$$

Wie im eindimensionalen Fall hat man auch hier verschiedene Versionen des Mittelwertsatzes:

**Satz 4.13 (Mittelwertsatz)** *Sei  $A \subset \mathbb{R}^m$  offen,  $f : A \rightarrow \mathbb{R} = \mathbb{R}^1$  differenzierbar,  $x, y$  und die Verbindungsstrecke von  $x$  und  $y$  liegen in  $A$ . Dann gibt es ein  $\vartheta \in (0, 1)$  mit*

$$f(y) - f(x) = Df(x + \vartheta(y - x))(y - x).$$

*Beweis.* Sei

$$g(t) := x + t(y - x) \quad \text{für } t \in [0, 1].$$

Dann ist

$$h := f \circ g : [0, 1] \rightarrow \mathbb{R}$$

differenzierbar mit

$$h'(t) = Dh(t) = Df(g(t))Dg(t) = Df(x + t(y - x))(y - x).$$

Mit dem 1-dimensionalen Mittelwertsatz folgt: Es gibt ein  $\vartheta \in (0, 1)$  mit

$$f(y) - f(x) = h(1) - h(0) = h'(\vartheta)(1 - 0) = Df(x + \vartheta(y - x))(y - x).$$

Das ist die Behauptung. ■

Daraus ergeben sich einige für Anwendungen nützliche Abschätzungen:

**Folgerung 4.14** *Sei  $A \subset \mathbb{R}^m$  offen,  $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$  differenzierbar.*

- a) *Gilt  $(D_k f_j(x)) \leq M$  für alle  $x \in A$  und  $k = 1, \dots, m, j = 1, \dots, n$  und liegt mit  $x$  und  $y$  auch die Verbindungsstrecke in  $A$ , so gilt*

$$|f(y) - f(x)| \leq M\sqrt{nm}|y - x|.$$

- b) *Ist  $A$  zusammenhängend (d. h. zu beliebigen  $x, y \in A$  existieren  $x_1, \dots, x_\ell \in A$  so, daß alle Strecken  $\overline{x_1 x_2}, \overline{x_2 x_3}, \dots, \overline{x_{\ell-1} x_\ell}$  in  $A$  liegen) und gilt  $Df(x) = 0$  für alle  $x \in A$ , so ist  $f$  konstant.*

- c) *Liegen  $x, y$  und die Verbindungsstrecke in  $A$ , so existiert ein  $z$  zwischen  $x$  und  $y$  mit*

$$|f(x) - f(y)| \leq |Df(z)| |y - x|,$$

wobei  $|C|$  die Operatornorm der Matrix  $C$  ist, d. h.  $|C| = \max\{|Cx| : |x| \leq 1\}$ .

*Beweis.* a) Zunächst gilt für alle  $j = 1, \dots, n$  nach obigem Mittelwertsatz

$$\begin{aligned} |f_j(y) - f_j(x)| &= |Df_j(x + \vartheta(y-x))(y-x)| \\ &= \left| \sum_{k=1}^m D_k f_j(x + \vartheta(y-x))(\eta_k - \xi_k) \right| \leq M \sum_{k=1}^m |\eta_k - \xi_k| \\ &\leq M m^{1/2} \left( \sum_{k=1}^m |\eta_k - \xi_k|^2 \right)^{1/2} = M m^{1/2} |y-x|, \end{aligned}$$

also

$$|f(y) - f(x)| = \left( \sum_{j=1}^n |f_j(y) - f_j(x)|^2 \right)^{1/2} = M m^{1/2} n^{1/2} |y-x|.$$

b) Mit Teil a folgt für  $x, y, x_1, \dots, x_\ell$  wie oben

$$f(x) = f(x_1) = \dots = f(x_\ell) = f(y).$$

c) Betrachten wir zunächst den Fall, daß ein  $j$  existiert mit  $f_k(x) = f_k(y)$  für  $k \neq j$ . Nach dem Mittelwertsatz existiert dann ein  $z$  zwischen  $x$  und  $y$  mit

$$f_j(y) - f_j(x) = Df_j(z)(y-x),$$

also

$$\begin{aligned} |f(y) - f(x)| &= |f_j(y) - f_j(x)| \leq |Df_j(z)| |y-x| \\ &\leq |Df(z)| |y-x| \end{aligned}$$

(da die Operatornorm von  $Df_j(z)$  sicher  $\leq$  der Operatornorm von  $Df(z)$  ist).

Im allgemeinen Fall existiert eine orthogonale Abbildung  $U : \mathbb{R}^n \rightarrow \mathbb{R}^n$  und eine Translation  $T$  in  $\mathbb{R}^n$  so, daß

$$UT(f(y) - f(x)) = |f(x) - f(y)| e^1$$

gilt. Wegen

$$|D(UTf)(z)| = |UD(Tf)(z)| = |Df(z)|$$

folgt die Behauptung aus dem ersten Teil des Beweises. ■

## 5 Höhere Ableitungen, der Satz von Taylor, Extrema

Ist  $f$  partiell differenzierbar und ist die partielle Ableitung  $D_j f$  wieder partiell differenzierbar, so bildet man z. B.  $D_k D_j f = D_k(D_j f)$  usw., die *zweiten partiellen Ableitungen* von  $f$ .

Entsprechend definiert man, falls  $f$   $p$  mal partiell differenzierbar ist, die  *$p$ -ten partiellen Ableitungen*  $D_{j_1} D_{j_2} \dots D_{j_p} f$ .

Bei allen einfachen Beispielen (man lasse sich etwas mehr oder weniger beliebiges einfallen) gilt  $D_i D_j f = D_j D_i f$  und entsprechend, falls  $f$  hinreichend oft partiell differenzierbar ist, für die höheren partiellen Ableitungen. Gilt das immer? Bzw. unter welchen zusätzlichen Bedingungen gilt es? Das folgende Beispiel zeigt, daß es nicht immer gilt. Der dann folgende Satz liefert dann eine einfache Bedingung unter der die Ableitungen vertauscht werden dürfen.

**Beispiel 5.1** Die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $f(\xi_1, \xi_2) = \begin{cases} \frac{\xi_1 \xi_2^3}{\xi_1^2 + \xi_2^2} & \text{für } x \neq 0, \\ 0 & \text{für } x = 0 \end{cases}$  (die außerhalb 0 beliebig oft differenzierbar ist) ist in 0 zwei mal partiell differenzierbar (allerdings nicht stetig partiell differenzierbar) und es gilt  $D_1 D_2 f(0, 0) \neq D_2 D_1 f(0, 0)$ :

Zunächst gilt für  $x \neq 0$

$$\text{grad } f(\xi_1, \xi_2) = (D_1 f(x), D_2 f(x)) = \left( \frac{\xi_2^3 (\xi_2^2 - \xi_1^2)}{(\xi_1^2 + \xi_2^2)^2}, \frac{\xi_1 \xi_2^2 (3\xi_1^2 + \xi_2^2)}{(\xi_1^2 + \xi_2^2)^2} \right).$$

Da  $f$  auf beiden Achsen und im Nullpunkt verschwindet, gilt  $D_1 f(0, 0) = D_2 f(0, 0) = (0, 0)$ . Für die übrigen Punkte auf den Achsen folgt aus obiger Formel

$$D_1 f(0, \xi_2) = \xi_2 \quad \text{und} \quad D_2 f(\xi_1, 0) = 0 \quad \text{für alle } \xi_1, \xi_2.$$

Damit folgt

$$D_2 D_1 f(0, \xi_2) = 1 \quad \text{und} \quad D_1 D_2 f(\xi_1, 0) = 0 \quad \text{für alle } \xi_1, \xi_2.$$

Also gilt  $D_2 D_1 f(0, 0) \neq D_1 D_2 f(0, 0)$ . □

**10: 23. Mai 2006**

**Satz 5.2 (Vertauschbarkeit der Diff.-Reihenfolge)** Sei  $A \subset \mathbb{R}^m$  offen,  $f : A \rightarrow \mathbb{R}^n$  zwei mal stetig differenzierbar. Dann gilt für alle  $i, j = 1, 2, \dots, m$

$$D_i D_j f = D_j D_i f.$$

(Unter der obigen Voraussetzung sind also zwei beliebige partielle Ableitungen vertauschbar; ist  $f$   $p$  mal stetig differenzierbar, so sind in  $D_{j_1} \dots D_{j_p} f$  die Ableitungen beliebig vertauschbar.)

*Beweis.* Zunächst wird das Problem in 3 Schritten vereinfacht.

**O. E.  $n = 1$ :** Es genügt  $D_i D_j f_k = D_j D_i f_k$  zu zeigen.

**O. E.  $m = 2$  und  $i, j = 1, 2$ :** Für  $i = j$  ist nichts zu zeigen.

Da die Koordinaten  $x_k$  mit  $k \neq i$  und  $k \neq j$  keine Rolle spielen, können sie als fest betrachtet werden.

**O. E. z. z.  $D_1 D_2 f(0, 0) = D_2 D_1 f(0, 0)$ :** Das ist immer durch eine Translation zu erreichen, die am Problem nichts ändert.

Sei also  $A \subset \mathbb{R}^2$  offen,  $f : A \rightarrow \mathbb{R}$  zwei mal stetig differenzierbar,  $(0, 0) \in A$ . Für hinreichend kleine  $s, t \neq 0$  gilt nach dem 1-dim Mittelwertsatz für  $x \mapsto f(x, t) - f(x, 0)$  bzw.  $x \mapsto D_1(s_1, x)$

$$\begin{aligned} (f(s, t) - f(s, 0)) - (f(0, t) - f(0, 0)) &= s \left( D_1 f(s_1, t) - D_1 f(s_1, 0) \right) \\ &= st D_2 D_1(s_1, t_1), \end{aligned}$$

und entsprechend bei anderer Klammerung mit dem Mittelwertsatz für  $x \mapsto f(s, x) - f(0, x)$  bzw.  $x \mapsto D_2 f(x, t_2)$

$$\begin{aligned} (f(s, t) - f(0, t)) - (f(s, 0) - f(0, 0)) &= t \left( D_2 f(s, t_2) - D_2 f(0, t_2) \right) \\ &= ts D_1 D_2 f(s_2, t_2). \end{aligned}$$

Da die linken Seiten identisch sind und  $s, t \neq 0$ , folgt

$$D_2 D_1 f(s_1, t_1) = D_1 D_2 f(s_2, t_2)$$

(man beachte, daß in der Regel  $s_1 \neq s_2$  und  $t_1 \neq t_2$  sein wird). Die  $s_j$  und  $t_j$  liegen zwischen 0 und  $s$  bzw. zwischen 0 und  $t$ . Mit  $s, t \rightarrow 0$  gilt also auch  $s_j t_j \rightarrow 0$  für  $j = 1, 2$ . Da die zweiten Ableitungen stetig sind folgt

$$D_2 D_1 f(0, 0) = D_1 D_2 f(0, 0).$$

■

Mit  $C^p(A, \mathbb{R}^n)$  oder  $C^p(A, \mathbb{C}^n)$  bezeichnet man den Vektorraum (reell oder komplex) der  $p$  mal stetig differenzierbaren Funktionen  $f : A \rightarrow \mathbb{R}^n$  (oder  $\mathbb{C}^n$ ).

Für  $f \in C^p(A, \mathbb{R}^n)$  und  $q \leq p$ ,  $j_1, j_2, \dots, j_q \subset \{1, \dots, m\}$  gilt dann

$$D_{j_1} \dots D_{j_q} f = \prod_{j=1}^q D_j^{\alpha_j} f =: D^\alpha f,$$

wobei  $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}_0^m$  ein *Multiindex* ist;  $\alpha_j$  gibt an, wie oft  $j$  unter den  $j_1, \dots, j_q$  vorkommt.

Mit

$$|\alpha| = \sum \alpha_j$$

bezeichnet man die *Länge* oder *Ordnung* von  $\alpha$ . Schließlich definiert man noch

$$x^\alpha := \xi_1^{\alpha_1} \xi_2^{\alpha_2} \dots \xi_m^{\alpha_m}, \quad \alpha! := \prod_{j=1}^m \alpha_j!.$$

Damit können wir nun den Taylorsche Satz auf Funktionen von  $m$  Variablen übertragen:

**Satz 5.3 (Taylor)** Sei  $A \subset \mathbb{R}^m$  offen,  $f \in C^p(A, \mathbb{R})$ ,  $p > 0$ ,  $x_0 \in A$ . Dann gilt für alle  $x \in A$ , für die die Verbindungsstrecke von  $x_0$  nach  $x$  in  $A$  liegt,

$$f(x) = \sum_{|\alpha| \leq p-1} \frac{1}{\alpha!} D^\alpha f(x_0) (x - x_0)^\alpha + R_p(x)$$

mit

$$R_p(x) = \sum_{|\alpha|=p} \frac{1}{\alpha!} D^\alpha f(x + \vartheta(x - x_0))(x - x_0)^\alpha \quad \text{mit } 0 < \vartheta < 1.$$

Die wichtigsten Spezialfälle sind:

$p = 1$ : Mittelwertsatz,

$$p = 2: f(x) = f(x_0) + \sum_{k=1}^m D_k f(x_0)(\xi_k - \xi_{0,k}) + R_2(x),$$

$$p = 3: f(x) = f(x_0) + \sum_{k=1}^m D_k f(x_0)(\xi_k - \xi_{0,k}) \\ + \frac{1}{2} \sum_{j,k=1}^m D_j D_k f(x_0)(\xi_j - \xi_{0,j})(\xi_k - \xi_{0,k}) + R_3(x).$$

*Beweis.* Der Beweis ergibt sich – mit einiger Rechnung – direkt aus dem 1-dimensionalen Satz. Sei

$$y := (x - x_0) = (\eta_1, \dots, \eta_m),$$

$$g : [-\varepsilon, 1 + \varepsilon] \rightarrow \mathbb{R}, \quad g(t) := f(x_0 + ty) = f \circ h(t)$$

mit  $h(t) = x_0 + ty$ . Für kleines  $\varepsilon > 0$  ist  $g$  auf  $[-\varepsilon, 1 + \varepsilon]$  definiert mindestens  $p$  mal stetig differenzierbar.

Mit der Kettenregel folgt (beachte  $h_j(t) = \xi_{0,j} + t\eta_j$ ,  $\frac{d}{dt}h_j(t) = \eta_j$ )

$$g'(t) = Dg(t) = Df(h(t))Dh(t) \\ = \sum_{j=1}^m D_j f(x_0 + ty) \frac{d}{dt}h_j(t) = \sum_{j=1}^m \eta_j D_j f(x_0 + ty).$$

Ist  $p \geq 2$ , so ist  $g'$  auch stetig differenzierbar mit

$$g''(t) = \sum_{j=1}^m \eta_j \frac{\partial}{\partial t} D_j f(x_0 + ty) = \sum_{j=1}^m \eta_j \sum_{i=1}^m \eta_i D_i D_j f(x_0 + ty) \\ = \left( \sum_{j=1}^m \eta_j D_j \right)^2 f(x_0 + ty).$$

wobei hier  $(\dots)^2$  formal auszurechnen und der entstehende Ausdruck auf  $f(x_0 + ty)$  anzuwenden ist.

Durch Induktion nach  $k$  folgt

$$g^{(k)}(t) = \left( \sum_{j=1}^m \eta_j D_j \right)^k f(x_0 + ty).$$

Mit dem *Polynomischen Satz* (oder *Polynomialformel*, vgl. Binominalformel)<sup>3</sup>

$$\left( \sum_{j=1}^m \xi_j \right)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha \quad (\text{mit } x = (\xi_1, \dots, \xi_m))$$

<sup>3</sup> *Beweis* durch Induktion nach  $m$ :  $m = 1$  ist offensichtlich,  $m = 2$  ist die Binominalformel.  $m \Rightarrow m + 1$ : Sei die Formel nicht für  $m$ . Dann folgt mit der Binominalformel

$$(\xi_1 + \dots + \xi_m + \xi_{m+1})^k = ((\xi_1 + \dots + \xi_m) + \xi_{m+1})^k$$

folgt

$$\frac{1}{k!}g^{(k)}(t) = \sum_{|\alpha|=k} \frac{y^\alpha}{\alpha!} D^\alpha f(x_0 + ty).$$

Einsetzen in den 1-dimensionalen Satz von Taylor,

$$g(t) = \sum_{k=0}^{p-1} \frac{1}{k!} g^{(k)}(0) t^k + \tilde{R}_p(t), \quad \tilde{R}_p(t) = \frac{1}{p!} t^p g^{(p)}(\vartheta t)$$

mit  $0 < \vartheta < 1$ , liefert

$$\begin{aligned} f(x) &= g(1) = \sum_{k=0}^{p-1} \frac{1}{k!} g^{(k)}(0) \tilde{R}_p(1) \\ &= \sum_{k=0}^{p-1} \sum_{|\alpha|=k} \frac{1}{\alpha!} y^\alpha D^\alpha f(x_0) + \tilde{R}_p(1) \\ &= \sum_{|\alpha| \leq p-1} \frac{1}{\alpha!} D^\alpha f(x_0) (x - x_0)^\alpha + \tilde{R}_p(1) \end{aligned}$$

mit

$$R_p(x) = \tilde{R}_p(1) = \frac{1}{p!} g^{(p)}(\vartheta) = \sum_{|\alpha|=p} \frac{1}{\alpha!} D^\alpha f(x_0 + \vartheta(x - x_0)) (x - x_0)^\alpha.$$

■

Eine wichtige Anwendung des Satzes von Taylor ist die Bestimmung lokaler Extrema bei hinreichend differenzierbaren Funktionen.

Sei  $A \subset \mathbb{R}^m$  offen,  $f : A \rightarrow \mathbb{R}$ . Wir erinnern:

- *lokales Maximum (lokales Minimum)* in  $x_0 \in A$ :  
 $\exists \delta > 0$  mit  $f(x) \leq f(x_0)$  ( $f(x) \geq f(x_0)$ ) für  $x \in A$  mit  $|x - x_0| < \delta$ ,
- *strenges lokales Maximum (strenges lokales Minimum)* in  $x_0 \in A$ :  $\dots f(x) < f(x_0)$  ( $f(x) > f(x_0)$ ) für alle  $x \in A$  mit  $0 < |x - x_0| < \delta$ .

11: 30. Mai 2006

**Satz 5.4** Sei  $A \subset \mathbb{R}^m$  offen,  $f : A \rightarrow \mathbb{R}$ ,  $x_0 \in A$ .

$$\begin{aligned} &= \sum_{p_m=0}^k \binom{k}{p_m} (\xi_1 \dots + \xi_m)^{k-p_m} \xi_{m+1}^{p_m} \\ &= \sum_{p_m=0}^k \frac{k!}{p_m!(k-p_m)!} \sum_{|\beta|=k-p_m} \frac{(k-p_m)!}{\beta!} (\xi_1, \dots, \xi_m)^\beta \xi_{m+1}^{p_m} \\ &= \sum_{|\alpha|=k} \frac{k!}{\alpha!} (\xi_1, \dots, \xi_m, \xi_{m+1})^\alpha = \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha. \end{aligned}$$

a) **Notwendige Bedingungen für Extrema:** Sei  $f$  differenzierbar und bei  $x_0$  liege ein Extremum (Maximum oder Minimum) vor.

(i) Dann gilt  $\text{grad } f(x_0) = 0$  (d. h. in  $x_0$  verschwinden alle partiellen Ableitungen).

(ii) Ist außerdem  $f \in C^2(A, \mathbb{R})$ , so ist die Hesse-Matrix  $(D_i D_j f(x_0))_{i,j=1,\dots,m}$  negativ semidefinit (im Fall eines Maximums), positiv semidefinit (im Fall eines Minimums).

b) **Hinreichende Bedingung für Extrema:** Ist  $f \in C^2(A, \mathbb{R})$ ,  $\text{grad } f(x_0) = 0$  und ist die Hesse-Matrix  $(D_i D_j f(x_0))_{i,j=1,\dots,m}$  negativ definit (bzw. positiv definit), so liegt bei  $x_0$  ein strenges lokales Maximum (bzw. ein strenges lokales Minimum) vor.

*Beweis.* a) (i) Da  $f$  differenzierbar ist, sind auch die „partiellen Funktionen“  $g_j(t) := f(x + te^j)$  bei 0 differenzierbar und haben dort ein lokales Extremum. Nach dem entsprechenden 1-dimensionalen Satz gilt also  $g'_j(0) = 0$  und somit

$$D_j f(x_0) = g'_j(0) = 0 \quad \text{für } j = 1, \dots, m.$$

(ii) Bei  $x_0$  liege ein lokales Maximum vor.  $\text{grad } f(x_0) = 0$  folgt wie in (i). Nehmen wir an, daß  $(D_i D_j f(x_0))$  nicht negativ semidefinit ist, d. h.

$$\exists y \in \mathbb{R}^m \quad \text{mit} \quad \langle (D_i D_j f(x_0)) y, y \rangle > 0.$$

Da die Funktionen  $D_i D_j f(x_0)$  stetig sind, existiert ein  $\delta > 0$  mit  $\langle (D_i D_j f(x)) y, y \rangle > 0$  für alle  $x$  mit  $|x - x_0| < \delta$ . Mit dem Taylorschen Satz ( $p = 2$ ) folgt für  $0 \leq t \leq 1$  ( $0 < \vartheta_t < 1$ )

$$\begin{aligned} f(x_0 + ty) &= f(x_0) + 0 + \sum_{|\alpha|=2} \frac{1}{\alpha!} D^\alpha f(x_0 + \vartheta_t ty) t^{|\alpha|} y^\alpha \\ &= f(x_0) + \frac{1}{2} \sum_{i,j=1}^m t^2 D_i D_j f(x_0 + \vartheta_t ty) \eta_i \eta_j > f(x_0), \end{aligned}$$

im Widerspruch dazu, daß bei  $x_0$  ein lokales Maximum vorliegt. Entsprechend für Minimum.

b) Sei  $\text{grad } f(x_0) = 0$ ,  $(D_i D_j f(x_0))$  positiv definit. Es gibt also ein  $\lambda > 0$  (der kleinste Eigenwert der Matrix  $(D_i D_j f(x_0))$ ) mit

$$\sum_{i,j=1}^m D_i D_j f(x_0) \eta_i \eta_j \geq \lambda \quad \text{für alle } y \in \mathbb{R}^m \quad \text{mit } |y| = 1.$$

Da die  $D_i D_j f(\cdot)$  stetig sind, existiert ein  $\delta > 0$  mit

$$\sum_{i,j=1}^m D_i D_j f(x) \eta_i \eta_j \geq \frac{\lambda}{2} \quad \text{für } y \in \mathbb{R}^m \quad \text{mit } |y| = 1, \quad x \in A \quad \text{mit } |x - x_0| < \delta.$$

Also gilt

$$\langle (D_i D_j f(x)) y, y \rangle \geq \frac{\lambda}{2} |y|^2 \quad \text{für } y \in \mathbb{R}^m, \quad x \in A \quad \text{mit } |x - x_0| < \delta,$$

und somit für alle  $x \in A$  mit  $|x - x_0| < \delta$

$$\begin{aligned} f(x) - f(x_0) &= \sum_{|\alpha|=2} \frac{1}{\alpha!} D^\alpha f(x_0 + \vartheta_x(x - x_0))(x - x_0)^\alpha \\ &= \frac{1}{2} \sum_{i,j=1}^m D_i D_j f(x_0 + \vartheta_x(x - x_0)) (\xi_i - \xi_{0,i}) (\xi_j - \xi_{0,j}) > 0. \end{aligned}$$

Bei  $x_0$  liegt also ein strenges lokales Minimum vor. entsprechend verfährt man im anderen Fall. ■

**Beispiel 5.5**  $f_\lambda : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $f_\lambda(x, y) = e^{xy} + x^2 + \lambda y^2$  ( $\lambda > 0$ )

Zunächst einige Vorüberlegungen zum Verhalten von  $f_\lambda$ :

$$\begin{aligned} - f_\lambda(x, 0) &= 1 + x^2, \quad f_\lambda(0, y) = 1 + \lambda y^2, \\ - f_\lambda(t, t) &= e^{t^2} + (1 + \lambda)t^2, \quad f_\lambda(t, -t) = e^{-t^2} + (1 + \lambda)^2 t^2. \end{aligned}$$

Die partiellen Ableitungen von  $f$ :

$$\begin{aligned} D_1 f_\lambda(x, y) &= y e^{xy} + 2x, & D_2 f_\lambda(x, y) &= x e^{xy} + 2\lambda y, \\ D_1^2 f_\lambda(x, y) &= y^2 e^{xy} + 2, & D_2^2 f_\lambda(x, y) &= x^2 e^{xy} + 2\lambda, \\ D_1 D_2 f_\lambda(x, y) &= D_2 D_1 f_\lambda(x, y) = (1 + xy) e^{xy}. \end{aligned}$$

Im Nullpunkt gilt  $f_\lambda(0, 0) = 1$ ,  $\text{grad } f_\lambda(0, 0) = (0, 0)$ . Die Hessematrix ist dort

$$\begin{pmatrix} D_1^2 f_\lambda(0, 0) & D_1 D_2 f_\lambda(0, 0) \\ D_1 D_2 f_\lambda(0, 0) & D_2^2 f_\lambda(0, 0) \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2\lambda \end{pmatrix}.$$

Mit dem charakteristischen Polynom

$$\det \begin{pmatrix} 2 - \mu & 1 \\ 1 & 2\lambda - \mu \end{pmatrix} = (2 - \mu)(2\lambda - \mu) - 1 = \mu^2 - 2(1 + \lambda)\mu + 4\lambda - 1,$$

also mit den Eigenwerten

$$\mu = 1 + \lambda \pm \sqrt{(1 + \lambda)^2 - 4\lambda + 1} = 1 + \lambda \pm \sqrt{\lambda^2 - 2\lambda + 2}.$$

**Für  $\lambda = \frac{1}{4}$ :**  $\mu = \frac{5}{4} \pm \sqrt{\frac{25}{16}} = \frac{5}{2}$  bzw. 0.

**Für  $\lambda > \frac{1}{4}$ :** Beide Eigenwerte  $> 0$ , denn dann gilt  $\mu_1 \mu_2 = \det \begin{pmatrix} 2 & 1 \\ 1 & 2\lambda \end{pmatrix} = 4\lambda - 1 > 0$ .

**Für  $\lambda < \frac{1}{4}$ :** Ein Eigenwert  $> 0$ , ein Eigenwert  $< 0$ , da dann  $\mu_1, \mu_2 = \det \begin{pmatrix} 2 & 1 \\ 1 & 2\lambda \end{pmatrix} < 0$ .

Damit folgt:

**Für  $\lambda > \frac{1}{4}$**  liegt (nach obigem Satz) bei  $(0, 0)$  ein *strenges lokales Minimum* für  $f_\lambda$  vor.

**Für  $\lambda = \frac{1}{4}$**  ist  $f_{1/4}(x, y) = 1 + \left(x + \frac{1}{2}y\right)^2 + (e^{xy} - 1 - xy) > 1 + \left(x + \frac{1}{2}y\right)^2$  für  $xy \neq 0$ .



Also ist

$$f_{1/4}(x, y) \begin{cases} > 1 + \left(x + \frac{1}{2}y\right)^2 & \text{für } xy \neq 0, \\ = x^2 + \frac{1}{4}y^2 + 1 & \text{für } xy = 0, \\ = 1 & \text{für } (x, y) = (0, 0). \end{cases}$$

also hat  $f_{1/4}$  in  $(0, 0)$  ein strenges globales Minimum.

Für  $\lambda > \frac{1}{4}$  gilt  $f_\lambda(x, y) \geq f_{1/4}(x, y)$  und  $f_\lambda(0, 0) = 1$ ; also ist das für  $\lambda > \frac{1}{4}$  strenge lokale Minimum tatsächlich ein strenges globales Minimum in  $(0, 0)$ .

Für  $\lambda < \frac{1}{4}$  liegt bei  $(0, 0)$  ein Sattelpunkt vor. Gibt es weitere Punkte mit  $\text{grad } f_\lambda = 0$ ? Aus  $D_1 f_\lambda = D_2 f_\lambda = 0$  folgt  $xye^{xy} = -2x^2 = -2\lambda y^2$ , also  $x = \pm\sqrt{\lambda}y$ .  $x = \sqrt{\lambda}y$  kommt nicht in Frage, da dort  $D_1 f_\lambda = ye^{xy} + 2x \neq 0$  gilt. Bleibt  $x = -\sqrt{\lambda}y$ , also  $e^{-\sqrt{\lambda}y^2} = -\frac{2x^2}{xy} = -\frac{2x}{y} = 2\sqrt{\lambda}$ . Wegen  $e^{-\sqrt{\lambda}y^2} < 1$  ist eine Lösung  $y \neq 0$  nur möglich, wenn  $2\sqrt{\lambda} < 1$ , also  $\lambda < \frac{1}{4}$ , gilt. Die *positive* Lösung ist dann (im 2. Quadranten)

$$y = \sqrt{\frac{-\ln(2\sqrt{\lambda})}{\sqrt{\lambda}}}, \quad x = -\sqrt{\lambda}y = -\sqrt{-\sqrt{\lambda}\ln(2\sqrt{\lambda})}.$$

Entsprechend (im 4. Quadranten)

$$y = -\sqrt{\frac{-\ln(2\sqrt{\lambda})}{\sqrt{\lambda}}}, \quad x = -\sqrt{\lambda}y = \sqrt{-\sqrt{\lambda}\ln(2\sqrt{\lambda})}.$$

Wegen  $f(x, y) \rightarrow \infty$  für  $|(x, y)| \rightarrow \infty$  und der Punktsymmetrie von  $f$  ( $f(x, y) = f(-x), -y$ ) sind dies notwendige globale Minima.

*Zusammengefaßt:* für  $\lambda \geq \frac{1}{4}$  hat  $f_\lambda$  in  $(0, 0)$  ein globales Minimum.

Für  $\lambda < \frac{1}{4}$  hat  $f_\lambda$  in

$$\pm \left( -\sqrt{-\sqrt{\lambda}\ln(2\sqrt{\lambda})}, \quad \sqrt{\frac{-\ln 2(\sqrt{\lambda})}{\sqrt{\lambda}}} \right)$$

globale Maxima (mit gleichem Funktionswert  $\frac{1}{2\sqrt{\lambda}} - 2\sqrt{\lambda}\ln(2\sqrt{\lambda})$  und in  $(0, 0)$  einen Sattelpunkt.

□

12: 01. Juni 2006

## 6 Kurven, Kurvenlänge und Kurvenintegrale

### 6.1 Parameterdarstellung

eine Kurve  $\Gamma$  in  $\mathbb{R}^m$  denken wir uns in der Regel gegeben durch eine Parameterdarstellung

$$\gamma : [a, b] \rightarrow \mathbb{R}^m, \quad t \mapsto \gamma(t),$$

wobei z. B.  $\gamma(t)$  den Ort eines Teilchens (hier also in  $\mathbb{R}^1, \mathbb{R}^2$  oder  $\mathbb{R}^3$ ) zur Zeit  $t$  beschreibt. Meist werden wir annehmen, daß  $\gamma$  zumindest stückweise stetig differenzierbar ist.

$\gamma'(t)$  (in der Physik meist  $\dot{\gamma}(t)$ ) ist dann die Geschwindigkeit, mit der der Punkt  $\gamma(t)$  durchlaufen wird (bzw. die Geschwindigkeit zur Zeit  $t$ ).

Einfache Beispiele sind:

**Beispiel 6.1** a) *Strecke* von  $x$  nach  $y$ :

$$\gamma : [0, 1] \rightarrow \mathbb{R}^m, \quad \gamma(t) := x + t(y - x).$$

Es kann aber auch eine andere Parameterdarstellung interessant sein:

$$\begin{aligned} \tilde{\gamma} : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] &\rightarrow \mathbb{R}^m, & \tilde{\gamma}(t) &:= \frac{1}{2}(x + y + (\sin t)(y - x)), \\ \hat{\gamma} : [-1, 1] &\rightarrow \mathbb{R}^m, & \hat{\gamma}(t) &:= \frac{x + y}{2} + t\frac{y - x}{2}. \end{aligned}$$

b) *Kreislinie* mit Radius  $r$  um den Nullpunkt in  $\mathbb{R}^2$  (im mathematisch positiven Sinn, d. h. entgegen dem Uhrzeigersinn, durchlaufen):

$$\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \gamma(t) := (r \cos t, r \sin t),$$

oder

$$\tilde{\gamma} : [0, 1] \rightarrow \mathbb{R}^2, \quad \tilde{\gamma}(t) := (r \cos 2\pi t, r \sin 2\pi t),$$

oder, mehrfach durchlaufen ( $k$ -fach):

$$\gamma_k : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \gamma_k(t) := (r \cos kt, r \sin kt).$$

c) *Schraubenlinie* mit Radius  $r$  und Ganghöhe  $c$ :

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^3, \quad \gamma(t) := (r \cos t, r \sin t, \frac{c}{2\pi}t),$$

oder, nur ein *Umlauf*:

$$\gamma_1 : [0, 1] \rightarrow \mathbb{R}^3, \quad \gamma_1(t) := (r \cos 2\pi t, r \sin 2\pi t, ct).$$

c) *NEILSche Parabel*:  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2, \gamma(t) = (t^2, t^3)$ .

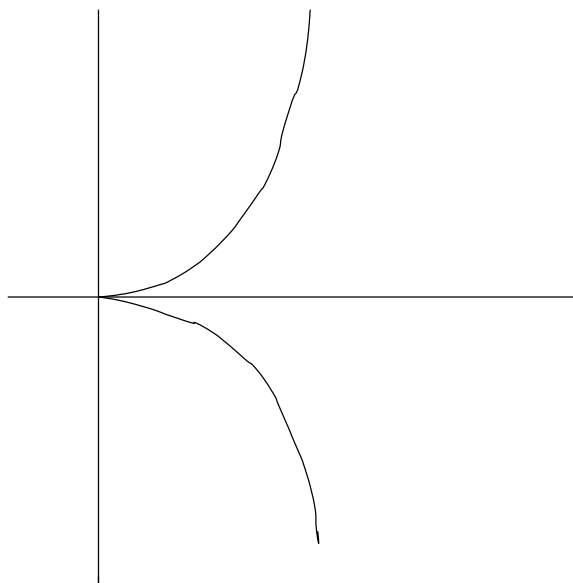


Abbildung 1: Neilsche Parabel

Hier fällt auf, daß die Kurve, obwohl die Parameterdarstellung beliebig differenzierbar ist, im Nullpunkt nicht „glatt“ ist. Woran liegt das? Zunächst fällt auf, daß im Nullpunkt die „Geschwindigkeit“  $\gamma'(t)$  verschwindet.  $\square$

Die meisten Kurven lassen sich, zumindest stückweise, als Graphen darstellen:  $\gamma(t) = (t, f(t))$ .

*Kreis:*  $\gamma : [-r, r] \rightarrow \mathbb{R}^2$ ,  $\gamma(t) = (t, \pm\sqrt{r^2 - t^2})$ .

*Neilsche Parabel:* Offenbar ist  $y(x) = \pm x^{3/2}$  ( $x > 0$ ) oder  $x(y) = |y|^{2/3}$ . Dementsprechend erhält man

$$\gamma : [0, \infty) \rightarrow \mathbb{R}^2, \quad \gamma(t) := (t, \pm t^{3/2})$$

oder

$$\tilde{\gamma} : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \tilde{\gamma}(t) := (|t|^{2/3}, t).$$

Eine Kurve  $\Gamma$  mit Parameterdarstellung  $\gamma$  heißt *regulär*, bzw. *regulär im Punkt*  $\gamma(t)$ , wenn  $\gamma'$  nicht verschwindet, bzw.  $\gamma'(t) \neq 0$  ist.

Die Neilsche Parabel ist nicht regulär im Nullpunkt. Das könnte zunächst an der Parameterdarstellung liegen. Tatsächlich kann aber die Neilsche Parabel keine überall reguläre Parameterdarstellung haben, da sie im Nullpunkt keine wohldefinierte Richtung hat.

Andererseits ist die Strecke von  $x$  nach  $y$  (vgl. obiges Beispiel a) mit der Darstellung  $\gamma$  regulär, mit der Darstellung  $\tilde{\gamma}$  nicht regulär in den beiden Randpunkten. Mit der Darstellung

$$\hat{\gamma} : [-1, 1] \rightarrow \mathbb{R}^m, \quad \hat{\gamma}(t) := \frac{x+y}{2} - t^3 \frac{y-x}{2}$$

ist sie nicht regulär im Mittelpunkt.

Wenn sich zwei Kurven  $(\Gamma_1, \gamma_1)$  und  $(\Gamma_2, \gamma_2)$  in einem Punkt  $\gamma_1(t_1) = \gamma_2(t_2)$  schneiden und beide Kurven in diesem Punkt regulär sind, gilt

$$\langle \gamma'_1(t_1), \gamma'_2(t_2) \rangle = \cos \text{Winkel}(\gamma'_1(t_1), \gamma'_2(t_2)) \|\gamma'_1(t_1)\| \|\gamma'_2(t_2)\|,$$

d. h. der *Schnittwinkel* ist gegeben durch

$$\text{Winkel } (\gamma'_1(t_1), \gamma'_2(t_2)) = \arccos \frac{\langle \gamma'_1(t_1), \gamma'_2(t_2) \rangle}{\|\gamma'_1(t_1)\| \|\gamma'_2(t_2)\|}.$$

## 6.2 Rektifizierbare Kurven, Kurvenlänge

Eine Kurve  $(\Gamma, \gamma)$  heißt *rektifizierbar*, wenn ihr auf folgendem Weg eine Länge zugeschrieben werden kann: Ist  $\gamma : [a, b] \rightarrow \mathbb{R}^m$  eine Parameterdarstellung von  $\Gamma$  und gibt es ein  $C$  so, daß für alle Zerlegungen  $\alpha = t_0 < t_1 < \dots < t_n = b$  des Intervalls  $[a, b]$  gilt

$$L_\gamma(t_0, t_1, \dots, t_n) := \sum_{i=1}^n |\gamma(t_i) - \gamma(t_{i-1})| \leq C,$$

so existiert das Supremum aller  $L_\gamma(t_0, \dots, t_n)$  ( $n \in \mathbb{N}$ ). Dieses wird mit  $L_\Gamma =$  *Länge der Kurve*  $\Gamma$  bezeichnet.

Nun ist zwar ziemlich klar, daß z. B. die Schraubenlinie oder die Neilsche Parabel keine endliche Länge haben. Ist aber das Parameterintervall beschränkt (und abgeschlossen) und  $\gamma$  stetig, so wird man das schon erwarten. Tatsächlich gilt das nicht:

**Beispiel 6.2** Die Kurve  $\Gamma$  mit  $\gamma : [0, 1] \rightarrow \mathbb{R}^2, t \mapsto \left(t, t \sin \frac{1}{t}\right)$  ( $\gamma(0) = 0$ ) ist nicht rektifizierbar. Das sieht man, wenn man für die Zerlegung von  $[0, 1]$  immer mehr Punkte nimmt in denen  $\sin \frac{1}{t} = \pm 1$  gilt.  $\square$

**Satz 6.3** Eine Kurve  $\Gamma$  mit stetig differenzierbarer Parameterdarstellung  $\gamma : [a, b] \rightarrow \mathbb{R}^m$  ist rektifizierbar und hat die Länge

$$L_\gamma = \int_a^b |\gamma'(t)| dt = \int_a^b \left( \gamma'_1(t)^2 + \gamma'_2(t)^2 + \dots + \gamma'_m(t)^2 \right)^{1/2} dt.$$

*Beweis.* Für jedes  $i$  gilt

$$|\gamma(t_i) - \gamma(t_{i-1})| = \left\{ \sum_{k=1}^m |\gamma_k(t_i) - \gamma_k(t_{i-1})|^2 \right\}^{1/2} = (t_i - t_{i-1}) \left\{ \sum_{k=1}^m \gamma'_k(\tau_{i,k})^2 \right\}^{1/2}$$

mit  $\tau_{i,k}$  zwischen  $t_{i-1}$  und  $t_i$ . Also unterscheidet sich  $\sum_{i=1}^n |\gamma(t_i) - \gamma(t_{i-1})|$  bei zunehmender Verfeinerung immer weniger von der Riemann-Summe<sup>4</sup>

$$\sum_{i=1}^n \left\{ \sum_{k=1}^m \gamma'_k(t_i)^2 \right\}^{1/2} (t_i - t_{i-1}) = \sum_{i=1}^n |\gamma'(t_i)| (t_i - t_{i-1})$$

<sup>4</sup>Ist man  $\{ |t_i - t_{i-1}| : i = 1, \dots, n \} < \delta$  mit  $|\gamma'_k(t) - \gamma'_k(s)| < \varepsilon$  für  $|t - s| < \delta$  und  $k = 1, \dots, m$ , so ist

$$\left| |\gamma'(t_i)| - \left\{ \sum_{k=1}^m \gamma'_k(\tau_{i,k})^2 \right\}^{1/2} \right| \leq \left\{ \sum_{k=1}^m \left( \gamma'_k(t_i) - \gamma'_k(\tau_{i,k}) \right)^2 \right\}^{1/2} \leq \sqrt{m\varepsilon}.$$

Also unterscheiden sich die beiden Summen für eine solche Zerlegung höchstens um  $(b-a)\sqrt{m\varepsilon}$ .

für das Integral  $\int_a^b |\gamma'(t)| dt$ . ■

Bis auf die Neilsche Parabel sind die Lösungen der obigen Kurve leicht zu berechnen (und elementargeometrisch bekannt). Hier wollen wir die Länge einer weiteren Kurve berechnen:

**Beispiel 6.4** Die *Zykloide* ist die Bahn eines Punktes auf dem Rand des Einheitskreises, der auf der  $x$ -Achse abrollt:

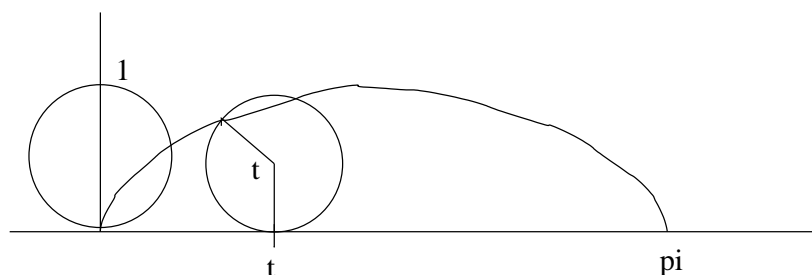


Abbildung 2: Zykloide

$$\begin{aligned} \gamma : \mathbb{R} &\rightarrow \mathbb{R}^2, \quad \gamma(t) = (t - \sin t, 1 - \cos t), \\ \gamma'(t) &= (1 - \cos t, \sin t), \\ |\gamma'(t)|^2 &= (1 - \cos t)^2 + \sin^2 t = 1 - 2 \cos t + \cos^2 t + \sin^2 t \\ &= 2 - 2 \cos t = 2 \left( \sin^2 \frac{t}{2} + \cos^2 \frac{t}{2} \right) - 2 \cos^2 \frac{t}{2} + 2 \sin^2 \frac{t}{2} = 4 \sin^2 \frac{t}{2}, \\ |\gamma'(t)| &= 2 \left| \sin \frac{t}{2} \right| \end{aligned}$$

und damit die Länge (einer Periode) der Zykloide:

$$L = \int_0^{2\pi} 2 \left| \sin \frac{t}{2} \right| dt = 2 \int_0^{2\pi} \sin \frac{t}{2} dt = -4 \cos \frac{t}{2} \Big|_0^{2\pi} = 8.$$

(Es ist überraschend, daß hier kein Ausdruck entsteht, der  $\pi$  enthält, sondern eine ganze Zahl.)  $\square$

Es ist noch interessant zu sehen, wie die Kurvenlänge in zwei speziellen Formen der Parameterdarstellung berechnet werden kann:

**Satz 6.5** a) Ist die Kurve als Graph der stetig differenzierbaren Funktion  $g : [a, b] \rightarrow \mathbb{R}^{m-1}$  gegeben, so ist die Kurvenlänge

$$L = \int_a^b \left( 1 + |g'(t)|^2 \right)^{1/2} dt.$$

b) Ist eine ebene Kurve in Polarkoordinaten  $r : [a, b] \rightarrow \mathbb{R}^2$ ,  $\varphi \mapsto r(\varphi)$  mit stetig differenzierbarem  $r(\cdot)$  gegeben, so ist die Kurvenlänge

$$L = \int_a^b \left( r'(\varphi)^2 + r(\varphi)^2 \right)^{1/2} d\varphi.$$

*Beweis.* a) Hier ist  $\gamma(t) = (t, g(t))$ , also  $|\gamma'(t)| = (1 + |g'(t)|^2)^{1/2}$ , woraus die Behauptung folgt.

b) Hier ist  $\gamma(\varphi) = (r(\varphi) \cos \varphi, r(\varphi) \sin \varphi)$ , also

$$\begin{aligned} |\gamma'(\varphi)|^2 &= (r'(\varphi) \cos \varphi - r(\varphi) \sin \varphi)^2 + (r'(\varphi) \sin \varphi + r(\varphi) \cos \varphi)^2 \\ &= r'(\varphi)^2 + r(\varphi)^2, \end{aligned}$$

woraus auch in diesem Fall die Behauptung folgt. ■

**13: 6. Juni 2006**

### 6.3 Kurvenintegrale

Die bei einer Bewegung eines Körpers im Kraftfeld  $k(\cdot)$  geleistete Arbeit ist gleich der Weglänge mal der Kraftkomponente in Richtung des Weges. Durchläuft der Körper eine Kurve  $(\Gamma, \gamma)$ , so zerlegt man zunächst das Parameterintervall

$$a = t_0 < t_1 < \dots < t_n = b.$$

Die im (Zeit-)Intervall  $[t_{k-1}, t_k]$  geleistete Arbeit ist näherungsweise

$$\left\langle k(\gamma(t_{k-1})), \gamma(t_k) - \gamma(t_{k-1}) \right\rangle \sim \left\langle k(\gamma(t_{k-1})), \gamma'(t_{k-1}) \right\rangle (t_k - t_{k-1}).$$

Die über das ganze Zeitintervall  $[a, b]$  geleistete Arbeit ist also

$$\sum_{k=1}^n \left\langle k(\gamma(t_{k-1})), \gamma'(t_{k-1}) \right\rangle (t_k - t_{k-1}).$$

Dies ist die/eine Riemannsumme für das Integral

$$\int_{\Gamma} k(x) dx := \int_a^b \left\langle k(\gamma(t)), \gamma'(t) \right\rangle dt.$$

Aus der Vorüberlegung zu dieser Definition ist offensichtlich, daß dieses Integral nicht von der Wahl der Parameterdarstellung abhängt.

Haben wir nun allgemein in einer offenen Menge  $A \subset \mathbb{R}^m$  ein stetiges Vektorfeld  $k : A \rightarrow \mathbb{R}^m$  und eine Kurve  $\Gamma$  in  $A$  mit stetig differenzierbarer Parameterdarstellung  $\gamma : [a, b] \rightarrow \mathbb{R}^m$ , so definieren wir das *Integral des Vektorfeldes  $k$  längs der Kurve  $\Gamma$*  durch

$$\int_{\Gamma} k(x) dx := \int_a^b \left\langle k(\gamma(t)), \gamma'(t) \right\rangle dt.$$

Gibt es eine stetig differenzierbare Funktion  $V : A \rightarrow \mathbb{R}$  mit  $k(x) = \text{grad } V(x)$ , so heißt  $k$  ein *Gradientenfeld* und  $V$  ein *Potential* von  $k$ . Der folgende Satz charakterisiert die Gradientenfelder durch ihre Kurvenintegrale:

**Satz 6.6** Sei  $A \subset \mathbb{R}^m$  offen,  $k : A \rightarrow \mathbb{R}^m$  ein stetiges Vektorfeld auf  $A$ . Dann ist  $k$  genau dann ein Gradientenfeld, wenn die Kurvenintegrale über  $k$  wegunabhängig sind (d. h. der Wert des Integrals hängt nur von Anfangs- und Endpunkt der Kurve ab). Es gilt dann

$$\int_{\Gamma} k(x) dx = \int_a^b \langle k(\gamma(t)), \gamma'(t) \rangle dt = k(\gamma(b)) - k(\gamma(a)).$$

*Beweis.*  $\Rightarrow$ : Sei  $k = \text{grad } V$ . Dann gilt

$$\int_a^b \langle k(\gamma(t)), \gamma'(t) \rangle dt = \int_a^b \sum_{j=1}^m D_j V(\gamma(t)) \gamma'_j(t) dt = \int_a^b \frac{d}{dt} V(\gamma(t)) dt = V(\gamma(b)) - V(\gamma(a)).$$

$\Leftarrow$ : Sei das Kurvenintegral wegunabhängig. Wir konstruieren ein Potential von  $k$ . Dazu sei  $x_0 \in A$  fest gewählt,  $\Gamma_x$  ein beliebiger Weg von  $x_0$  nach  $x$  in  $A$ ,

$$V(x) := \int_{\Gamma_x} k(z) dz$$

(damit ist  $V(\cdot)$  wohldefiniert, da das Integral wegunabhängig ist). Dann gilt

$$\begin{aligned} D_j V(x) &= \lim_{h \rightarrow 0} \frac{1}{h} (V(x + h e^j) - V(x)) = \lim_{h \rightarrow 0} \frac{1}{h} \left\{ \int_{x_0}^{x+h e^j} \dots - \int_{x_0}^x k(z) dz \right\} \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+h e^j} k(z) dz = \lim_{h \rightarrow 0} \frac{1}{h} \int_0^h \langle k(x + t e^j), e^j \rangle dt \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \int_0^h k_j(x + t e_j) dt = k_j(x), \end{aligned}$$

d. h.  $V$  ist ein Potential von  $k$ . ■

**Beispiel 6.7** Das Vektorfeld

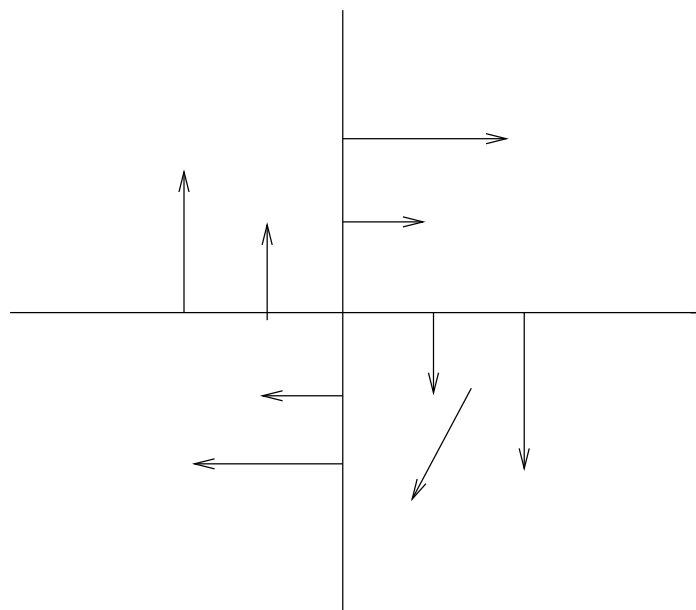
$$k : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad k(x) = (\xi_2, -\xi_1)$$

hat kein Potential. Auf Grund der Skizze

sieht man sofort ein, daß das Kurvenintegral über Kreise um Null nicht verschwindet.

Außerdem: Wäre  $k = \text{grad } V = (D_1 V, D_2 V)$ , so würde folgen (beachte, daß dieses  $k$  beliebig oft stetig differenzierbar ist)

$$\text{rot } k(x) := D_1 k_2 - D_2 k_1 = D_1 D_2 V - D_2 D_1 V = 0.$$

Abbildung 3: Vektorfeld  $k(x) = (\xi_2, -\xi_1)$ 

Tatsächlich gilt aber  $\text{rot } k = 2$ . (Die Bedeutung der *Rotation*  $\text{rot}$  werden wir später kennen lernen.)

□

Ist  $k$  ein stetig differenzierbares Gradientenfeld  $k = \text{grad } V$ , so gilt (da dann  $V$  zwei mal stetig differenzierbar ist) für alle  $i, j \in \{1, \dots, m\}$

$$D_i k_j - D_j k_i = D_i D_j V - D_j D_i V = 0.$$

Man kann sich also fragen, ob das ausreicht um zu garantieren, daß  $k$  ein Gradientenfeld ist. Im allgemeinen Nein!

**Beispiel 6.8** Sei  $k : \mathbb{R}^2 \setminus \{0\}$  das Vektorfeld

$$k(x) = \left( \frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right).$$

Dann gilt zwar  $D_1 k_2 = D_2 k_1 = 0$ , aber die Integrale über Kreise um 0 verschwinden nicht, d. h.  $k$  ist kein Gradientenfeld. Der folgende Satz macht deutlich, woran das liegt. □

**Satz 6.9 (Integrabilitätsbedingung)** Sei  $G \in \mathbb{R}^m$  offen und einfach zusammenhängend (d. h. jede geschlossene Kurve in  $G$  läßt sich in  $G$  [d. h. ohne das man dabei  $G$  verläßt] stetig auf einen Punkt zusammenziehen<sup>5</sup>),  $k : G \rightarrow \mathbb{R}^m$  ein stetig differenzierbares Vektorfeld mit  $D_i k_j - D_j k_i = 0$  für alle  $i, j \in \{1, \dots, m\}$ . Dann ist  $k$  ein Gradientenfeld.

<sup>5</sup>In  $\mathbb{R}^2$  bedeutet das, dass im Innern der Kurve keine Punkte liegen, die nicht zu  $G$  gehören;  $\mathbb{R}^2 \setminus \{0\}$  ist nicht einfach zusammenhängend; dagegen ist  $\mathbb{R}^m \setminus \{0\}$  für  $m \geq 3$  einfach zusammenhängend



*Beweis.* Der Einfachheit halber betrachten wir nur sternförmige Gebiete: Ein Gebiet  $G \subset \mathbb{R}^m$  heißt *sternförmig* bezüglich  $x_0$ , wenn mit jedem  $x \in G$  auch die Verbindungsstrecke mit  $x_0$  in  $G$  liegt.

Ohne Einschränkung sei  $G$  sternförmig bezüglich  $0$ . Wir definieren  $(\gamma : [0, 1] \rightarrow \mathbb{R}^m, \gamma(t) = tx)$

$$V(x) := \int_0^x k(z) dz = \int_0^1 \langle k(tx), x \rangle dt.$$

$V$  ist damit wohldefiniert und offensichtlich (?) stetig in  $G$ . Weiter gilt

$$\begin{aligned} \frac{\partial}{\partial x_j} V(x) &= \int_0^1 \frac{\partial}{\partial x_j} \langle k(tx), x \rangle dt = \int_0^1 \frac{\partial}{\partial x_j} \sum_{\ell=1}^m k_\ell(tx) x_\ell dt \\ &= \int_0^1 \left\{ k_j(tx) + \sum_{\ell=1}^m x_\ell \frac{\partial}{\partial x_j} k_\ell(tx) \right\} dt \\ &\quad \text{( hier wird die Voraussetzung benutzt )} \\ &= \int_0^1 \left\{ k_j(tx) + \sum_{\ell=1}^m x_\ell \frac{\partial}{\partial x_\ell} k_j(tx) \right\} dt \\ &= \int_0^1 \left\{ k_j(tx) + t \langle \text{grad } k_j(tx), x \rangle \right\} dt \\ &\quad \left( \text{Produktregel für } t \cdot k_j \text{ und Kettenregel für } \frac{\partial}{\partial t} k_j(tx) \right) \\ &= \int_0^1 \frac{\partial}{\partial t} (tk_j(tx)) dt = 1 \cdot k_j(1x) - 0 \cdot k_j(0x) = k_j(x) \end{aligned}$$

Also ist  $V$  ein Potential von  $k$ . ■

## 7 Umkehrfunktion und implizit definierte Funktion

### 7.1 Umkehrfunktion

Aus dem 1-dimensionalen Fall wissen wir: Ist  $f : I \rightarrow \mathbb{R}$  stetig und streng monoton, so ist auch das Bild  $f(I)$  ein Intervall und  $f$  besitzt eine *Umkehrfunktion (Inverse)*  $f^{-1} : f(I) \rightarrow I$ . Ist zusätzlich  $f$  differenzierbar mit  $f'(x_0) \neq 0$ , so ist  $f^{-1}$  an der Stelle  $y_0 = f(x_0)$  differenzierbar mit

$$(f^{-1})(f(x_0)) = \frac{1}{f'(x_0)} \quad \text{bzw.} \quad (f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))}.$$

Ist  $f$  stetig differenzierbar mit  $f'(x) \neq 0$ , so ist auch  $f^{-1}$  stetig differenzierbar (die strenge Monotonie folgt dann bereits aus der Forderung  $f'(x) \neq 0$ ).

Hier soll das Resultat auf Funktionen mehrerer Variablen verallgemeinert werden. Die Monotonie läßt sich nicht übertragen. Dagegen erweist es sich als sinnvoll, die Forderung  $f'(x) \neq 0$  durch  $\det Df(x) \neq 0$  zu ersetzen.

Wir sehen dies zunächst an folgendem einfachen Beispiel:

**Beispiel 7.1** Sei  $L$  eine  $m \times m$ -Matrix,  $a \in \mathbb{R}^m$ ,  $f$  die *affine Funktion*

$$f : \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad f(x) = a + Lx.$$

Offensichtlich ist  $f$  in jedem  $x \in \mathbb{R}^m$  differenzierbar mit Ableitung  $Df(x) = L$  für alle  $x \in \mathbb{R}^m$ , denn es gilt

$$\begin{aligned} f(z) &= a + Lz = a + Lx + L(z - x) \\ &= f(x) + L(z - x) + \varphi_x(z) \quad \text{mit } \varphi_x(z) \equiv 0. \end{aligned}$$

Da  $y = a + Lx$  genau dann für alle  $Y$  eindeutig nach  $x$  auflösbar ist, wenn  $L$  invertierbar ist,  $x = L^{-1}(y - a)$ , ist  $f$  genau dann invertierbar, wenn  $L$  invertierbar ist, und es gilt

$$f^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad f^{-1}(y) = L^{-1}(y - a) = b + L^{-1}y \quad \text{mit } b = -L^{-1}a.$$

Insbesondere ist auch  $f^{-1}$  differenzierbar mit Ableitung  $(Df)^{-1}(y) = L^{-1}$  für alle  $y \in \mathbb{R}^m$ .  $\square$

Da eine differenzierbare Funktion  $f : A \rightarrow \mathbb{R}^m$  (mit  $A \subset \mathbb{R}^m$  offen) in der Nähe von  $x_0$  durch die affine Funktion

$$F_0 : \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad F_{x_0}(x) = f(x_0) + Df(x_0)(x - x_0)$$

*sehr gut approximiert* wird, liegt es nahe, daß  $f$  in der Nähe von  $x_0$  umkehrbar ist, wenn die Matrix  $Df(x_0)$  invertierbar ist. Genauer: annähernd sollte für  $y$  nahe  $f(x_0)$  gelten

$$f^{-1}(y) \sim F_{x_0}^{-1}(y) = Df(x_0)^{-1}(y - f(x_0)) + x_0.$$

Die Ableitung von  $f^{-1}$  bei  $x_0$  sollte also gerade  $(Df(x_0))^{-1}$  sein. Da die Approximation nur lokal gut ist, ist aber zu erwarten, daß auch diese Aussage nur lokal gilt (vgl. dazu den folgenden Satz und das darauffolgende Beispiel).

Im folgenden Satz benutzen wir eine neue Sprechweise: ist  $x_0 \in A \subset \mathbb{R}^m$ , so heißt  $U$  eine *offene Umgebung von  $x_0$  in  $A$* , wenn  $U$  offen ist mit  $x_0 \in U \subset A$ . Im Fall  $A = \mathbb{R}^m$  nennen wir  $U$  einfach eine *offene Umgebung von  $x_0$* .

**Satz 7.2 (Umkehrfunktion)** Sei  $A \subset \mathbb{R}^m$  offen,  $f : A \rightarrow \mathbb{R}^m$  stetig differenzierbar,  $x_0 \in A$ ,  $y_0 = f(x_0)$ ,  $Df(x_0)$  invertierbar. Dann gibt es eine offene Umgebung  $U$  von  $x_0$  in  $A$  und eine offene Umgebung  $V$  von  $y_0$  so, daß

$$f : U \rightarrow V \quad \text{bijektiv}$$

ist und  $Df(x)$  für alle  $x \in U$  invertierbar ist. Die Abbildung

$$g := (f|_U)^{-1} : V \rightarrow U$$

ist stetig differenzierbar mit

$$Dg(f(x)) = (Df(x))^{-1} \quad \text{für } x \in U \quad \text{bzw.} \quad Dg(y) = (Df(g(y)))^{-1} \quad \text{für } y \in V.$$

**Beispiel 7.3** Wie oben schon vermutet, gilt die Aussage des Satzes (insbesondere die Invertierbarkeit) nicht global. Dies wird durch folgendes Beispiel belegt:

$$A = \mathbb{R}^2 \setminus \{0\} \subset \mathbb{R}^2, \quad f(\xi_1, \xi_2) = (\xi_1^2 - \xi_2^2, 2\xi_1 \xi_2).$$

(Setzt man  $z = \xi_1 + i\xi_2$ , so ist offenbar  $f(\xi_1, \xi_2) = (\operatorname{Re} z^2, \operatorname{Im} z^2)$ , d. h. bei  $f$  handelt es sich um die Darstellung der komplexen Quadrierung in  $\mathbb{R}^2$ .)

Es gilt

$$Df(\xi_1, \xi_2) = \begin{pmatrix} 2\xi_1 & -2\xi_2 \\ 2\xi_2 & 2\xi_1 \end{pmatrix}, \quad \text{also } \det D(f(\xi_1, \xi_2)) = 4(\xi_1^2 + \xi_2^2) \neq 0 \quad \text{für alle } x \in A.$$

Trotzdem ist  $f$  nicht injektiv, denn es gilt für jedes  $x = (\xi_1, \xi_2) \in A$

$$f(-\xi_1, -\xi_2) = f(\xi_1, \xi_2).$$

□

*Beweis von Satz ??.* Da  $A$  offen ist, und  $f$  stetig differenzierbar, existiert ein  $\delta > 0$  mit  $U := K_\delta(x_0) \subset A$  und <sup>6</sup>

$$|Df(x) - Df(x_0)| \leq \frac{1}{2} \left| Df(x_0)^{-1} \right|^{-1} =: \lambda \quad \text{für } x \in A \text{ mit } |x - x_0| < \delta.$$

Wir werden zeigen: Mit diesem  $U$  und  $V := f(U)$  gilt die Behauptung des Satzes.

**1. Schritt:**  $Df(x)$  ist invertierbar für  $x \in U$ : Für alle  $x \in U$  und  $y \in \mathbb{R}^m \setminus \{0\}$  gilt (dabei benutzen wir, daß für eine invertierbare Matrix  $L$  gilt  $|y| \leq |L^{-1}| |Ly|$  bzw.  $|Ly| \geq |L^{-1}|^{-1} |y|$ )

$$\begin{aligned} |Df(x)y| &\geq |Df(x_0)y| - |Df(x)y - Df(x_0)y| \\ &\geq \left| Df(x_0)^{-1} \right|^{-1} |y| - \frac{1}{2} \left| Df(x_0)^{-1} \right|^{-1} |y| \\ &= \frac{1}{2} \left| Df(x_0)^{-1} \right|^{-1} |y| > 0, \end{aligned}$$

<sup>6</sup>dabei ist für eine  $m \times m$ -Matrix  $L$  mit  $|L|$  die „Operatorennorm“ von  $L$  gemeint,  $|L| := \sup\{|Lx| : x \in \mathbb{R}^m, |x| \leq 1\}$ . Es gilt z. B.  $|L| \leq \sum_{i,j=1}^m |l_{ij}|$  und  $|L| \leq \left\{ \sum_{i,j=1}^m |l_{ij}|^2 \right\}^{1/2}$ ; insbesondere ist  $|L|$  klein, wenn alle  $a_{ij}$  klein sind.

d. h. für  $x \in U$  ist  $Df(x)$  invertierbar.

**2. Schritt:**  $f$  ist auf  $U$  injektiv: Für  $x \in U$  schreiben wir  $f$  in der Form

$$f(x) = Df(x_0)x + F(x) \text{ mit } F(x) := f(x) - Df(x_0)x.$$

Dann gilt für alle  $x, x' \in U$  mit  $x \neq x'$  (wieder benutzen wir  $|Ly| \geq |L^{-1}|^{-1}|y|$ )

$$\begin{aligned} |f(x) - f(x')| &= |Df(x_0)(x - x') + F(x) - F(x')| \\ &\geq |Df(x_0)(x - x')| - |F(x) - F(x')| \\ &\geq |Df(x_0)^{-1}|^{-1}|x - x'| - |DF(\xi)||x - x'| \\ &= 2\lambda|x - x'| - |Df(\xi) - Df(x_0)||x - x'| \\ &\geq 2\lambda|x - x'| - \lambda|x - x'| = \lambda|x - x'| > 0. \end{aligned}$$

Also ist  $f|_U$  injektiv.

Damit ist jedenfalls  $f : U \rightarrow V$  bijektiv.

**3. Schritt:**  $V$  ist offen: Dazu zeigen wir: Ist

$$y' = f(x') \in V = f(U), \quad \delta' := \delta - |x' - x_0|, \quad \varepsilon = \varepsilon(x') := \frac{\lambda}{2}\delta',$$

so ist  $K_\varepsilon(y') \in V$ , d. h.: zu jedem  $z \in K_\varepsilon(y')$  gibt es ein  $x \in U = K_\delta(x_0)$  mit  $f(x) = z$ .

Wegen  $|z - y'| < \varepsilon = \frac{\lambda}{2}\delta'$  gibt es ein  $r < \delta'$  mit  $|z - y'| < \frac{\lambda}{2}r$ .

Wir betrachten die Hilfsfunktion

$$\psi : \overline{K}_r(x') \rightarrow \mathbb{R}_+, \quad x \mapsto |f(x) - z|^2.$$

Gesucht ist also ein  $x \in K_r(x') \subset U$  mit  $f(x) = z$ , d. h. mit  $\psi(x) = 0$ ; da  $\psi$  nicht negativ ist, ist dies ein Minimum von  $\psi$ . – Da  $\psi$  auf  $\overline{K}_r(x')$  stetig ist, nimmt es dort sein Minimum an; es ist zu zeigen, daß dieser minimale Wert 0 ist. Für  $x$  auf dem Rand von  $\overline{K}_r(x')$ , d. h.  $|x - x'| = r$ , gilt

$$\begin{aligned} \psi(x) &= |f(x) - z|^2 \geq \left\{ |f(x) - f(x')| - |y' - z| \right\}^2 \\ &\geq \left\{ \lambda|x - x'| - \frac{\lambda}{2}r \right\}^2 = \left\{ \lambda r - \frac{\lambda}{2}r \right\}^2 = \frac{1}{4}\lambda^2 r^2. \end{aligned}$$

Andererseits gilt im Mittelpunkt  $x'$  von  $\overline{K}_r(x')$

$$\psi(x') = |f(x') - z|^2 = |y' - z|^2 < \left(\frac{1}{2}\lambda r\right)^2 = \frac{1}{4}\lambda^2 r^2.$$

Also nimmt  $\psi$  sein Minimum im Innern der Kugel  $K_r(x')$  an, etwa in  $\tilde{x} \in K_r(x')$ . Dann muß aber gelten

$$\begin{aligned} 0 &= \text{grad } \psi(\tilde{x}) = \left( \dots, D_j \sum_{k=1}^m |f_k(\tilde{x}) - z_k|^2, \dots \right) \\ &= \left( \dots, 2 \sum_{k=1}^m D_j f_k(\tilde{x})(f_k(\tilde{x}) - z_k), \dots \right) = 2(Df(\tilde{x}))^t(f(\tilde{x}) - z). \end{aligned}$$

Da  $Df(\tilde{x})$  invertierbar ist, gilt dies auch für  $(Df(\tilde{x}))^t$ , und somit ist  $f(\tilde{x}) = z$ , d. h.  $\tilde{x}$  ist das gesuchte  $x$  mit  $z = f(x)$ . Damit ist bewiesen, daß  $V$  offen ist.

**4. Schritt:** Es bleibt schließlich nur noch zu zeigen  $\mathbf{g} := (\mathbf{f}|_U)^{-1}$  ist in  $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$  differenzierbar mit  $D\mathbf{g}(\mathbf{y}_0) = D\mathbf{f}(\mathbf{x}_0)^{-1}$ : (O. E. können wir uns auf  $x_0$  beschränken, da jedes  $x \in A$  als Ausgangspunkt gewählt werden kann.) Für  $y \in \mathbb{R}^m$  mit  $|y| < \varepsilon$  ( $\varepsilon$  wie oben) gilt

$$g(y_0 + y) = x_0 + x \in U = K_\delta(x_0).$$

Wegen

$$y = f(x_0 + x) - y_0 = f(x_0 + x) - f(x_0)$$

folgt aus den obigen Überlegungen

$$|y| \geq \lambda|x| \text{ bzw. } |x| \leq \frac{1}{\lambda}|y|.$$

Außerdem gilt (da  $f$  in  $x_0$  die Ableitung  $Df(x_0)$  hat)

$$y - Df(x_0)x = f(x_0 + x) - f(x_0) - Df(x_0)x = \varphi_{f,x_0}(x_0 + x),$$

$$Df(x_0)^{-1}y - x = Df(x_0)^{-1}\varphi_{f,x_0}(x_0 + x)$$

mit

$$\frac{1}{|x|}\varphi_{f,x_0}(x_0 + x) \rightarrow 0 \text{ für } x \rightarrow 0.$$

Damit erhalten wir

$$\begin{aligned} |g(y_0 + y) - g(y_0) - Df(x_0)^{-1}y| &= |(x_0 + x) - x_0 - Df(x_0)^{-1}y| \\ &= |x - Df(x_0)^{-1}y| = |Df(x_0)^{-1}\varphi_{f,x_0}(x_0 + x)| \\ &\leq |Df(x_0)^{-1}||\varphi_{f,x_0}(x_0 + x)| \end{aligned}$$

mit

$$\frac{1}{|y|}|Df(x_0)^{-1}||\varphi_{f,x_0}(x_0 + x)| \leq \frac{1}{\lambda}|Df(x_0)^{-1}|\frac{|\varphi_{f,x_0}(x_0 + x)|}{|x|} \rightarrow 0$$

für  $y \rightarrow 0$  (was mit  $x \rightarrow 0$  gleichbedeutend ist). Also gilt

$$g(y_0 + y) = g(y_0) + Df(x_0)^{-1}y + \varphi_{g,y_0}(y_0 + y)$$

mit

$$\frac{1}{|y|}\varphi_{g,y_0}(y_0 + y) \rightarrow 0 \text{ für } y \rightarrow 0,$$

d. h.  $g$  ist in  $y_0 = f(x_0)$  differenzierbar mit  $Dg(y_0) = Df(x_0)^{-1}$ . ■

## 7.2 Implizit definierte Funktionen

Wir betrachten erst die wesentlich überschaubarere Situation in  $\mathbb{R}^2$ . Ist  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  eine „vernünftige“ (was das genau heißt sehen wir später; z. B. stetig differenzierbare) Funktion, so wird durch die Nullstellenmenge

$$M := \left\{ (x, y) \in \mathbb{R}^2 : g(x, y) = 0 \right\}$$

in vielen Fällen eine Kurve in  $\mathbb{R}^2$  beschrieben. So ist z. B. für  $g(x, y) = x^2 + y^2 - 1$  die Menge  $M$  gerade die Einheitskreislinie. Da sich Kurven in der Regel (zumindest stückweise) als Graph darstellen lassen, kommen wir zu der

**Frage:** Gibt es eine Menge  $U \subset \mathbb{R}$  und ein  $f : U \rightarrow \mathbb{R}$  so, daß der Graph von  $f$  ein Stück der Nullstellenmenge  $M$  von  $g$  beschreibt. In anderen Worten, kann die Gleichung  $g(x, y) = 0$  (wenigstens lokal) nach  $y$  aufgelöst werden,  $y = f(x)$  mit  $g(x, f(x)) = 0$ ? Eine solche Funktion  $f$  heißt durch  $g$  *implizit definiert*.

Offensichtlich ist  $M$  sicher dann ein Graph einer (allerdings i. Allg. nicht notwendig stetigen und überall definierten) Funktion  $f$ , wenn die Funktion

$$g_x : y \mapsto g(x, y)$$

für jedes  $x$  streng monoton ist, denn dann gibt es für jedes  $x$  höchstens ein  $y$  mit  $g_x(y) = g(x, y) = 0$ .

Dies gilt aber schon für so einfache Funktionen wie die oben genannte

$$g(x, y) = x^2 + y^2 - 1$$

nicht global. Allerdings läßt sich die Einheitskreislinie stückweise als Graph von stetig differenzierbaren Kurven darstellen, nämlich

$$y = \pm\sqrt{1-x^2} \quad \text{für } -1 < x < 1 \quad \text{bzw.} \quad x = \pm\sqrt{1-y^2} \quad \text{für } -1 < y < 1.$$

Bevor wir den allgemeinen Fall betrachten, formulieren und beweisen wir den entsprechenden Satz für den bisher betrachteten einfachen Fall, in dem die Situation noch geometrisch einfach erfaßbar ist.

**Satz 7.4** Sei  $A \subset \mathbb{R}^2$  offen,  $g \in C^1(A) = C^1(A, \mathbb{R})$ ,

$$(x_0, y_0) \in M := \left\{ (x, y) \in A : g(x, y) = 0 \right\} \quad \text{mit } D_2g(x_0, y_0) \neq 0.$$

Dann existieren offene Intervalle  $U, V \subset \mathbb{R}$  mit

$$x_0 \in U, \quad y_0 \in V, \quad U \times V \subset A$$

und eine stetige Funktion

$$f : U \rightarrow V \quad \text{mit } M \cap (U \times V) = G_f := \{(x, f(x)) : x \in U\}$$

und

$$Df(x) = f'(x) = -\frac{D_1g(x, f(x))}{D_2g(x, f(x))} \quad \text{für } x \in U.$$

(Insbesondere kann die Ableitung von  $f$  angegeben werden, ohne daß man die Funktion  $f$  genau kennt.)

*Beweis.* Wegen  $D_2g(x_0, y_0) \neq 0$  und der Stetigkeit von  $D_2g$  existiert ein Rechteck  $U' \times V$  um  $(x_0, y_0)$  mit

$$D_2g(x, y) \neq 0 \quad \text{für } (x, y) \in U' \times V,$$

d. h.  $g(x, \cdot)$  ist streng monoton in  $V$  für jedes  $x \in U'$ . Also gibt es zu jedem  $x \in U'$  höchstens ein  $y \in V$  mit  $g(x, y) = 0$ ; es gibt also ein Teilintervall  $U \subset U'$  mit  $x_0 \in U$  so, daß für jedes  $x \in U$  genau ein  $y =: f(x) \in V$  existiert mit  $g(x, y) = 0$ .

Wenn wir hier die Differenzierbarkeit von  $f$  als gegeben betrachten (im Beweis des folgenden Satzes ergibt sich diese für den allgemeinen Fall), dann folgt aus  $h(x) := g(x, f(x)) \equiv 0$  mit  $k(x) := (x, f(x))$

$$\begin{aligned} 0 &= h'(x) = Dh(x) = Dg(k(x))Dk(x) \\ &= \left\langle \left( D_1g(x, f(x)), D_2g(x, f(x)) \right), (1, f'(x)) \right\rangle \\ &= D_1g(x, f(x)) + f'(x)D_2g(x, f(x)). \end{aligned}$$

woraus die behauptete Formel für  $f'$  folgt. ■

**Beispiel 7.5** Für  $g(x, y) = x - y^2$  ist das Nullstellengebilde

$$\left\{ (x, y) \in \mathbb{R}^2 : x = y^2 \right\}, \quad \left\{ (x, y) \in \mathbb{R}^2 : x > 0, y = \pm\sqrt{x} \right\}.$$

Die durch  $g$  implizit definierte Funktion ist also

$$(0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto \sqrt{x} \quad \text{oder} \quad x \mapsto -\sqrt{x}.$$

Im Nullpunkt ist der Satz nicht anwendbar, da dort  $D_2g(0, 0) = 0$  ist (dort ist aber auch die Wurzel nicht differenzierbar).

Für andere Punkte gilt

$$\frac{d}{dx} \left( \pm\sqrt{x} \right) = -\frac{D_1g(x, y(x))}{D_2g(x, y(x))} = -\frac{1}{-2y(x)} = \frac{1}{2(\pm\sqrt{x})}.$$

Wir haben damit einen anderen Weg gefunden, die Ableitung der Wurzelfunktion zu bestimmen. Dies kann natürlich entsprechend auf andere Umkehrfunktionen angewandt werden. □

**15: 13. Juni 2006**

**Korollar 7.6** Sei  $A \subset \mathbb{R}^2$  offen,  $g : A \rightarrow \mathbb{R}$  stetig differenzierbar mit

$$\text{grad } g(x, y) \neq 0 \quad \text{für alle } (x, y) \in M := \{(x, y) \in A : g(x, y) = 0\}.$$

Dann ist  $M$  endliche Vereinigung von Graphen stetig differenzierbarer Funktionen  $f_j$  (zum Teil  $y$  als Funktion von  $x$ , zum Teil  $x$  als Funktion von  $y$ ).

*Beweis.* In jedem Punkt  $(x_0, y_0) \in M$  ist  $D_1g(x_0, y_0) \neq 0$  oder  $D_2g(x_0, y_0) \neq 0$ . Deshalb gibt es ein ganzes Rechteck  $U \times V$  um  $(x_0, y_0)$  mit  $D_1g(x, y) \neq 0$  für alle  $(x, y) \in U \times V$  oder  $D_2g(x, y) \neq 0$  für alle  $(x, y) \in U \times V$ , also eine Funktion

$$f : U \rightarrow V \quad \text{oder} \quad \tilde{f} : V \rightarrow U$$

mit

$$M \cap (U \times V) = G_f \quad \text{oder} \quad M \cap (U \times V) = G_{\tilde{f}}^{-1},$$

(wobei  $G_h^{-1}$  aus  $G_h$  durch Vertauschen der Komponenten hervorgeht). In konkreten Fällen ist klar, daß sich  $M$  durch endlich viele solcher Rechtecke überdecken läßt; im allgemeinen Fall wird hier der Satz von Heine–Borel benötigt (vgl. z. B. H. Heuser, Lehrbuch der Analysis 2, §157). ■

**Satz 7.7 (Implizite Funktionen)** Sei  $m, n \in \mathbb{N}$ ,  $A \subset \mathbb{R}^{m+n}$  offen,  $g \in C^1(A, \mathbb{R}^n)$ . Die Punkte aus  $\mathbb{R}^{m+n}$  schreiben wir in der Form

$$(x, y) = (\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n).$$

Es sei

$$L_1(x, y) := (D_i g_j(x, y))_{\substack{i=1, \dots, m \\ j=1, \dots, n}} = \begin{pmatrix} D_1 g_1(x, y) & \dots & D_m g_1(x, y) \\ \vdots & \ddots & \vdots \\ D_1 g_n(x, y) & \dots & D_m g_n(x, y) \end{pmatrix},$$

$$L_2(x, y) := (D_{m+j} g_i(x, y))_{i,j=1, \dots, n} = \begin{pmatrix} D_{m+1} g_1(x, y) & \dots & D_{m+n} g_1(x, y) \\ \vdots & \ddots & \vdots \\ D_{m+1} g_n(x, y) & \dots & D_{m+n} g_n(x, y) \end{pmatrix}$$

und somit

$$Dg(x, y) = (L_1(x, y) : L_2(x, y)).$$

Für ein

$$(x_0, y_0) \in M := \{(x, y) \in A : g(x, y) = 0\}$$

sei die Matrix  $L_2(x_0, y_0)$  invertierbar.

Dann existieren offene Mengen  $U \in \mathbb{R}^m$  und  $V \subset \mathbb{R}^n$  mit

$$(x_0, y_0) \in U \times V \subset A$$

und eine stetig differenzierbare Funktion

$$f : U \rightarrow V \quad \text{mit} \quad M \cap (U \times V) = G_f,$$

$$Df(x) = -L_2(x, f(x))^{-1} L_1(x, f(x)).$$



*Beweisidee:* In der Nähe von  $(x_0, y_0)$  gilt mit „großer Genauigkeit“

$$g(x, y) \sim g(x_0, y_0) + L_1(x_0, y_0)(x - x_0) + L_2(x_0, y_0)(y - y_0).$$

Auflösung der Gleichung  $g(x, y) = 0$  nach  $y$  in der Nähe von  $(x_0, y_0)$  mit  $g(x_0, y_0) = 0$  sollte also möglich sein, wenn die Gleichung

$$L_1(x_0, y_0)(x - x_0) + L_2(x_0, y_0)(y - y_0) = 0$$

nach  $y$  auflösbar ist. Das ist genau dann möglich, wenn  $L_2(x_0, y_0)$  invertierbar ist. Die Lösung ist

$$f(x) = y = y_0 + L_2(x_0, y_0)^{-1} L_1(x_0, y_0)(x - x_0).$$

Als Ableitung erhalten wir, insbesondere im Punkt  $x = x_0$

$$Df(x_0) = L_2(x_0, y_0)^{-1} L_1(x_0, y_0).$$

*Beweis.* Wir definieren die folgenden Hilfsfunktionen

$$\begin{aligned} G &: A \rightarrow \mathbb{R}^{m+n} & (x, y) &\mapsto (x, g(x, y)), \\ j_1 &: \mathbb{R}^m \rightarrow \mathbb{R}^{m+n} & x &\mapsto (x, 0) \quad (\text{Injektion}), \\ p_2 &: \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n, & (x, y) &\mapsto y \quad (\text{Projektion}). \end{aligned}$$

Im folgenden sei  $E_m$  die  $m$ -dimensionale Einheitsmatrix,  $0_{m,n}$  die  $m$ -zeilige und  $n$ -spaltige Nullmatrix. Damit gilt offensichtlich

$$DG(x_0, y_0) = \begin{pmatrix} E_m & 0_{m,n} \\ L_1(x_0, y_0) & L_2(x_0, y_0) \end{pmatrix},$$

$$\det DG(x_0, y_0) = \det L_2(x_0, y_0) \neq 0$$

da  $L_2(x_0, y_0)$  invertierbar ist. Also gibt es eine Umgebung von  $(x_0, y_0)$  in  $A$ , ohne Einschränkung in der Form  $U_1 \times V$  wählbar mit offenen Umgebungen  $U_1$  und  $V$  von  $x_0$  bzw.  $y_0$ , so, daß  $G : U_1 \times V \rightarrow \mathbb{R}^{m+n}$  eine stetig differenzierbare Inverse  $(G|_{U_1 \times V})^{-1}$  hat; wegen  $G(x_0, y_0) = (x_0, g(x_0, y_0)) = (x_0, 0)$  ist  $(G|_{U \times V})^{-1}$  jedenfalls auf  $U \times \{0\}$  definiert für eine geeignete Umgebung  $U \subset U_1$  von  $x_0$  in  $\mathbb{R}^m$ .

Wir zeigen, daß

$$f := p_2 \circ G^{-1} \circ j_1 : U \rightarrow V$$

die gewünschte Funktion ist:

- $j_1$  bildet  $U$  auf  $U \times \{0\}$  ab,  $x \mapsto (x, 0)$ ,
- $G^{-1}$  bildet  $U \times \{0\}$  in  $U \times V$  ab,  $(x, 0) \mapsto (x, y)$  mit  $g(x, y) = 0$ ,
- $p_2$  bildet  $U \times V$  nach  $V$  ab,  $(x, y) \mapsto y$ , wobei für die hier in Frage stehenden  $(x, y)$  das  $y$  die Eigenschaft hat, daß  $g(x, y) = 0$  ist.

Insgesamt bildet also  $f$  jedes  $x \in U$  auf das  $y \in V$  ab, für das  $g(x, y) = 0$  gilt; wir haben also die gesuchte Funktion von  $f$  auf einer Umgebung  $U$  von  $x_0$  „explizit“ dargestellt.

Nun bleibt nur noch die Ableitung zu berechnen:

$$Df(x) = Dp_2 \left( G^{-1}(j_1(x)) \right) DG^{-1}(j_1(x)) Dj_1(x).$$

Der erste und der letzte Faktor sind einfach zu berechnen,

$$Dp_2(\cdot) = \begin{pmatrix} 0_{n,m} & E_n \end{pmatrix}, \quad Dj_1 = \begin{pmatrix} E_m \\ 0_{n,m} \end{pmatrix}.$$

Den mittleren Faktor erhalten wir aus dem Satz über die Umkehrfunktion (Satz ??).  $DG(x, y)$  haben wir bereits oben angegeben (wobei für  $(x_0, y_0)$  jetzt  $(x, y)$  zu setzen ist). Man sieht leicht, daß gilt

$$(DG(x, y))^{-1} = \begin{pmatrix} E_m & 0_{m,n} \\ -L_2^{-1}L_1 & L_2^{-1} \end{pmatrix} \quad \text{mit} \quad L_j := L_j(x, y),$$

also, mit der Formel für die Ableitung der Umkehrfunktion,

$$\begin{aligned} DG^{-1}(j_1(x)) &= DG^{-1}(x, 0) = \left( DG(G^{-1}(x, 0)) \right)^{-1} \\ &= \left( DG(x, f(x)) \right)^{-1} = \begin{pmatrix} E_m & 0_{m,n} \\ -M_2^{-1}M_1 & M_2^{-1} \end{pmatrix} \quad \text{mit} \quad M_j := L_j(x, f(x)), \end{aligned}$$

und damit erhalten wir

$$\begin{aligned} Df(x) &= \begin{pmatrix} 0_{n,m} & E_n \end{pmatrix} \begin{pmatrix} E_m & 0_{m,n} \\ -M_2^{-1}M_1 & M_2^{-1} \end{pmatrix} \begin{pmatrix} E_m \\ 0_{n,m} \end{pmatrix} \\ &= \begin{pmatrix} 0_{n,m} & E_n \end{pmatrix} \begin{pmatrix} E_m \\ -M_2^{-1}M_1 \end{pmatrix} = -M_2^{-1}M_1 \\ &= -L_2(x, f(x))^{-1}L_1(x, f(x)). \end{aligned}$$

Das ist die behauptete Formel. ■

Völlig analog zu Korollar ?? gilt:

**Korollar 7.8** Sei  $A \in \mathbb{R}^m$  offen,  $g : A \rightarrow \mathbb{R}$  stetig differenzierbar mit

$$\text{grad } g(x) \neq 0 \quad \text{für alle } x \in M := \{x \in A : g(x) = 0\}.$$

Dann ist  $M$  endliche Vereinigung von Graphen stetiger Funktionen  $f_j$  (jeweils eine Koordinate als Funktion der anderen  $m - 1$ ).

### 7.3 Bedingte Extrema

Häufig will man Extrema von Funktionen auf Flächen oder Kurven bestimmen, *bedingte Extrema*. Dies ist auch wichtig, wenn man Extrema auf Teilgebieten von  $\mathbb{R}^m$  einschließlich ihres Randes bestimmen will: Mit der früher beschriebenen Methode findet man lokale Extrema im Innern. Dies muß man dann noch vergleichen mit den lokalen Extrema auf dem Rand.

**Satz 7.9 (Methode der Lagrange–Multiplikatoren)** Sei  $A \subset \mathbb{R}^m$  offen,

$$g_j A \rightarrow \mathbb{R}^m \text{ stetig differenzierbar, } M_j := \{x \in A : g_j(x) = 0\} \text{ für } j = 1, \dots, k,$$

$$M := \bigcap_{j=1}^k M_j, \quad x_0 \in M \text{ und } \{\text{grad } g_j(x_0) : j = 1, \dots, k\} \text{ linear unabhängig.}$$

Ist  $f : A \rightarrow \mathbb{R}$  stetig differenzierbar und hat  $f$  in  $x_0$  ein bedingtes lokales Extremum auf  $M$ , so gilt

$$\text{grad } f(x_0) \in L\{\text{grad } g_j(x_0) : j = 1, \dots, k\},$$

d. h. es gibt reelle Zahlen  $\lambda_1, \dots, \lambda_k$  (Lagrange–Multiplikatoren) mit

$$\text{grad } f(x_0) = \sum_{j=1}^k \lambda_j \text{grad } g_j(x_0).$$

*Beweis.* Ist  $\gamma : [a, b] \rightarrow M$  eine (in  $x_0$ ) differenzierbare Kurve in  $M$  durch  $x_0$ ,  $\gamma(c) = x_0$  für ein  $c \in (a, b)$ , so hat die Funktion

$$h : [a, b] \rightarrow \mathbb{R}, \quad h(t) = f(\gamma(t))$$

in  $c$  ein lokales Extremum. Da  $h$  in  $c$  differenzierbar ist, gilt

$$0 = h'(c) = \langle \text{grad } f(\gamma(c)), \gamma'(c) \rangle,$$

d. h.  $\text{grad } f(x_0)$  steht in  $x_0$  senkrecht auf jeder Kurve in  $M$  durch  $x_0$ .

Da  $\gamma'(c)$  jede Richtung haben kann, die orthogonal zu  $\{\text{grad } g_j(x_0) : j = 1, \dots, k\}$  ist, bedeutet das, daß  $\text{grad } f(x_0)$  eine Linearkombination der  $\text{grad } g_j(x_0)$  sein muß. ■

**Bemerkung 7.10** Die Anwendung dieses Satzes geht so: Man löst das Gleichungssystem

$$g_j(x) = 0 \quad (j = 1, \dots, k)$$

$$\text{grad } f(x) - \sum_{j=1}^k \lambda_j \text{grad } g_j(x)$$

$k + m$  Gleichungen für die  $k + m$  Unbekannten  $\xi_1, \dots, \xi_m, \lambda_1, \dots, \lambda_k$  auf, wobei die  $\lambda_j$  eigentlich nicht interessieren. Dann überprüft man anhand der Funktionswerte in den gefundenen Punkten  $x$ , wo das Maximum bzw. Minimum liegt.

**Beispiel 7.11**  $A = \mathbb{R}^2$ ,  $g(x) = \xi_1^2 + \xi_2^2 - 1$ , also

$$M := \{x \in \mathbb{R}^2 : \xi_1^2 + \xi_2^2 = 1\} \quad \text{Einheitskreislinie.}$$

Es sollen die Extrema der Funktion

$$f(x) = \xi_1 + 2\xi_2$$

auf  $M$  bestimmt werden. Man hat also das Gleichungssystem

$$\xi_1^2 + \xi_2^2 - 1 = 0 \tag{1}$$

$$(1, 2) - \lambda(2\xi_1, 2\xi_2) = 0 (= \text{grad } f(x) + \lambda \text{grad } g(x)) \tag{2}$$

zu lösen. Aus (2) folgt  $\lambda \neq 0$  und

$$\xi_1 = \frac{1}{2\lambda}, \quad \xi_2 = \frac{1}{\lambda} \tag{3}$$

Einsetzen in (1) liefert

$$\frac{1}{4\lambda^2} + \frac{1}{\lambda^2} - 1 = 0, \quad \lambda^2 = \frac{5}{4}, \quad \lambda = \pm \frac{1}{2}\sqrt{5}.$$

Einsetzen in (3) liefert

$$\xi_1 = \pm \frac{1}{\sqrt{5}}, \quad \xi_2 = \pm \frac{2}{\sqrt{5}}.$$

Die Funktionswerte in diesen Punkten sind

$$f\left(\pm \frac{1}{\sqrt{5}}, \pm \frac{2}{\sqrt{5}}\right) = \pm\sqrt{5} \sim 2, 236 \dots$$

Also liegt in  $\left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$  das Maximum, in  $\left(-\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}}\right)$  das Minimum auf  $M$  vor.

Es gibt mindestens zwei weitere Möglichkeiten, dieses Problem zu lösen:

- Man stellt den Kreis in Parameterdarstellung dar und führt damit das Problem auf ein 1-dimensionales Problem zurück.
- Man stellt fest, daß  $f$  auf den Geraden  $\xi_2 = -\frac{1}{2}\xi_1 + c$  konstant ist (Wert =  $c$ ) und bestimmt die  $c$ , für die diese Geraden Tangenten sind.

□

## 8 Holomorphe Funktionen

Funktionen einer komplexen Variablen können natürlich als Funktionen von zwei reellen Variablen (Real- und Imaginärteil) aufgefaßt werden. Sie können dann partiell differenzierbar oder stetig partiell differenzierbar (und somit reell differenzierbar) sein. Dies ist jedoch nicht interessant, es bringt keine neue Struktur. Hier wird ein neuer Begriff wichtig:

Sei  $\Omega \subset \mathbb{C}$  offen,  $f : \Omega \rightarrow \mathbb{C}$  heißt *im Punkt*  $z_0 \in \Omega$  *komplex differenzierbar*, wenn der Grenzwert

$$f'(z_0) := \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

existiert.  $f'(z_0)$  heißt die (*komplexe*) *Ableitung* von  $f$  im Punkt  $z_0$ .  $f$  heißt *komplex differenzierbar*, wenn es in jedem Punkt von  $\Omega$  komplex differenzierbar ist. (Es wird sich zeigen, daß dann  $f$  bereits beliebig oft stetig komplex differenzierbar ist.)

Wenn  $f$  in  $z_0$  komplex differenzierbar ist, kann man natürlich auch die (reellen) partiellen Ableitungen

$$\begin{aligned} f_x(z_0) &:= \lim_{h \rightarrow 0} \frac{f(z_0 + h) - f(z_0)}{h} \\ f_y(z_0) &:= \lim_{h \rightarrow 0} \frac{f(z_0 + ih) - f(z_0)}{h} \end{aligned} \quad h \in \mathbb{R} \setminus \{0\}$$

betrachten. Es gilt der überraschende

**Satz 8.1 (Cauchy–Riemannsche Differentialgleichungen)** *Ist  $f$  in  $z_0$  komplex differenzierbar, so gilt*

$$f_x(z_0) = -if_y(z_0) = f'(z_0).$$

*Zerlegt man  $f$  in Real- und Imaginärteil,  $f = u + iv$  mit reellwertigen Funktionen  $u$  und  $v$  ( $u(z) := \operatorname{Re} f(z)$ ,  $v(z) := \operatorname{Im} f(z)$ ), so sind auch  $u$  und  $v$  in  $z_0$  partiell differenzierbar und es gelten die Cauchy–Riemanschen Differentialgleichungen*

$$u_x(z_0) = v_y(z_0), \quad v_x(z_0) = -u_y(z_0).$$

*Beweis.* Die Gleichung  $f_x(z_0) = f'(z_0)$  ist offensichtlich, bei  $f_x$  betrachtet man bei der Differentiation nur  $z$  mit  $\operatorname{Im} z = \operatorname{Im} z_0$ . Weiter gilt

$$f_y(z_0) = \lim_{h \rightarrow 0} \frac{f(z_0 + ih) - f(z_0)}{h} = i \lim_{h \rightarrow 0} \frac{f(z_0 + ih) - f(z_0)}{ih} = if'(z_0).$$

Damit folgt

$$\begin{aligned} u_x(z_0) + iv_x(z_0) &= (u + iv)_x(z_0) = f_x(z_0) = f'(z_0) \\ &= \frac{1}{i} f_y(z_0) = -i(u + iv)_y(z_0) = -iu_y(z_0) + v_y(z_0). \end{aligned}$$

Vergleich von Real- und Imaginärteil auf beiden Seiten liefert die Behauptung. ■

Für die Umkehrung des Satzes wird zusätzlich die stetig partielle Differenzierbarkeit benötigt:

**Satz 8.2** Sei  $\Omega \subset \mathbb{C}$  offen,  $f : \Omega \rightarrow \mathbb{C}$  stetig partiell differenzierbar (d.h. auch  $u$  und  $v$  sind stetig partiell differenzierbar) mit  $f_x(z) = -if_y(z)$  (bzw.  $u_x(z) = v_y(z)$  und  $v_x(z) = -u_y(z)$ ). Dann ist  $f$  komplex differenzierbar und  $f'$  stetig. (Tatsächlich genügt es, vgl. Beweis, in dem jeweils betrachteten Punkt die totale Differenzierbarkeit von  $f$  bzw.  $u$  und  $v$  vorauszusetzen.)

*Beweis.* Da  $u$  und  $v$  stetig partiell differenzierbar sind, sind sie total differenzierbar. Für  $z, z_0 \in \Omega$  ( $z = x + iy, z_0 = x_0 + iy_0$ ) gilt also

$$\begin{aligned} u(z) - u(z_0) &= u_x(z_0)(x - x_0) + u_y(z_0)(y - y_0) + \varphi_{u,z_0}(z) \\ v(z) - v(z_0) &= v_x(z_0)(x - x_0) + v_y(z_0)(y - y_0) + \varphi_{v,z_0}(z) \end{aligned}$$

mit

$$\varphi_{u,z_0}(z) \frac{1}{z - z_0} \rightarrow 0, \quad \varphi_{v,z_0}(z) \frac{1}{z - z_0} \rightarrow 0 \quad \text{für } z \rightarrow z_0.$$

Wegen  $u_x = v_y$  und  $v_x = -u_y$  folgt daraus

$$\begin{aligned} \frac{f(z) - f(z_0)}{z - z_0} &= \frac{1}{z - z_0} \left\{ u(z) - u(z_0) + i[v(z) - v(z_0)] \right\} \\ &= \frac{1}{z - z_0} \left\{ u_x(z_0)(x - x_0) + u_y(z_0)(y - y_0) + \varphi_{u,z_0}(z) \right. \\ &\quad \left. + i \left[ v_x(z_0)(x - x_0) + v_y(z_0)(y - y_0) + \varphi_{v,z_0}(z) \right] \right\} \\ &= \frac{1}{z - z_0} \left\{ u_x(z_0)(x - x_0) - v_x(z_0)(y - y_0) + \varphi_{u,z_0}(z) \right. \\ &\quad \left. + i \left[ v_x(z_0)(x - x_0) + u_x(z_0)(x - x_0) + \varphi_{v,z_0}(z) \right] \right\} \\ &= \frac{1}{z - z_0} \left\{ u_x(z_0) \left[ (x - x_0) + i(y - y_0) \right] + i v_x(z_0) \left[ (x - x_0) + i(y - y_0) \right] \right\} \\ &\quad + \varphi_{u,z_0}(z) \frac{1}{z - z_0} + \varphi_{v,z_0}(z) \frac{1}{z - z_0} \\ &\rightarrow u_x(z_0) + i v_x(z_0) = f_x(z_0) \quad \text{für } z \rightarrow z_0. \end{aligned}$$

Also ist  $f$  komplex differenzierbar. Die Stetigkeit von  $f'$  folgt aus der von  $f_x$ . ■

Sei  $\Omega \subset \mathbb{C}$  offen. Eine Funktion  $f : \Omega \rightarrow \mathbb{C}$  heißt *holomorph* (auch *analytisch*, *regulär*), wenn es in jedem Punkt komplex differenzierbar ist.

$f$  heißt *holomorph bei*  $z_0$ , wenn sie in einer Umgebung von  $z_0$  holomorph ist.

$f$  heißt *bei*  $\infty$  *holomorph*, wenn  $g(z) := f(1/z)$  mit einem geeigneten Wert für  $g(0)z : f(\infty)$  bei  $0$  holomorph ist. Z. B. ist die Funktion  $f(z) = 1/z$  bei  $\infty$  holomorph mit  $f(\infty) = 0$ .

Wie im reellen gilt:

**Satz 8.3**  $\Omega \in \mathbb{R}^m$  offen,  $z_0 \in \Omega$ .  $f$  ist bei  $z_0$  genau dann komplex differenzierbar, wenn es bei  $z_0$  (im komplexen Sinn) linear approximierbar ist.

*Beweis.*  $\Rightarrow$ : Ist  $f$  bei  $z_0$  komplex differenzierbar, so gilt

$$\varrho(z) := \frac{f(z) - f(z_0)}{z - z_0} - f'(z_0) \rightarrow 0 \quad \text{für } z \rightarrow z_0,$$

also

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + \varphi_{z_0}(z)$$

mit

$$\frac{\varphi_{z_0}(z)}{z - z_0} = \varrho(z) \rightarrow 0 \quad \text{für } z \rightarrow z_0.$$

$\Leftarrow$ : Ist  $f$  bei  $z_0$  linear approximierbar, so gilt

$$f(z) = f(z_0) + m(z - z_0) + \varphi_{z_0}(z) \quad \text{mit } \frac{\varphi_{z_0}(z)}{z - z_0} \rightarrow 0 \quad \text{für } z \rightarrow z_0,$$

also

$$m = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0},$$

d. h.  $f$  ist in  $z_0$  komplex differenzierbar mit  $f'(z_0) = m$ . ■

Für komplex differenzierbare Funktionen gelten alle aus dem Reellen bekannten **Differentiationsregeln** mit i. wes. unveränderten Beweisen: Linearität, Produktregel, Quotientenregel, Kettenregel und inverse Funktion.

Wir untersuchen zunächst eine wichtige (geometrische) Eigenschaft holomorpher Funktionen:

Sei  $f$  holomorph bei  $z_0$  mit  $f'(z_0) \neq \mathbf{0}$ ,  $(\Gamma, \gamma)$  eine reguläre Kurve durch  $z_0$ , d. h.  $\gamma : [a, b] \rightarrow \mathbb{C}$ ,  $\gamma(t_0) = z_0$  für ein  $t_0 \in (a, b)$  und  $\gamma'(t_0) \neq 0$ . Im Punkt  $z_0$  hat  $\Gamma$  die Richtung  $\gamma'(t_0)$  bzw.  $\gamma'(t_0)/|\gamma'(t_0)|$ . Bei Anwendung von  $f$  geht die Kurve  $(\Gamma, \gamma)$  über in eine Kurve  $(\tilde{\Gamma}, \tilde{\gamma})$  mit

$$\tilde{\gamma} : [a, b] \rightarrow \mathbb{C}, \quad \tilde{\gamma}(t) := f(\gamma(t)).$$

Diese Kurve hat in  $\tilde{\gamma}(t_0)$  die Richtung

$$\tilde{\gamma}'(t) = f'(\gamma(t_0))\gamma'(t_0) = f'(z_0)\gamma'(t_0).$$

Die Richtung von  $\tilde{\Gamma}$  in  $f(z_0)$  ist also gegenüber der Richtung von  $\Gamma$  in  $z_0$  um den Winkel  $\arg f'(z_0)$  verdreht.

Betrachtet man zwei Kurven, die sich in  $z_0$  schneiden, so werden diese auf zwei Kurven abgebildet, die sich in  $f(z_0)$  im gleichen Winkel schneiden.

**Satz 8.4** *Eine holomorphe Funktion  $f$  mit  $f'(z) \neq 0$  ist winkeltreu (dies gilt auch punktweise in jedem Punkt  $z_0$  mit  $f'(z_0) \neq 0$ ). (Solche Funktionen heißen auch konforme Abbildungen.)*

**Bemerkung 8.5** *Eine total differenzierbare Abbildung  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit  $\det Df(x) \neq 0$  ist i. allg. nicht winkeltreu, z. B.  $f(\xi_1, \xi_2) = \left(\frac{1}{2}\xi_1, \xi_2\right)$ .*

Wichtige elementare **Beispiele holomorpher Funktionen** sind:

1. *Monome*  $f_n(z) := z^n$  ( $n \in \mathbb{N}_0$ ) sind holomorph in ganz  $\mathbb{Z}$  mit  $f'_n(z) = n z^{n-1}$ . Entsprechendes gilt für beliebige komplexe Polynome.

2. *Rationale Funktionen*,  $p, q$  Polynome

$$f : \Omega = \{z \in \mathbb{C} : q(z) \neq 0\} \rightarrow \mathbb{C}, \quad f(z) = \frac{p(z)}{q(z)}$$

sind holomorph in  $\Omega$ . Differentiation mittels Quotientenregel.

Ist  $m := \text{grad } p \leq \text{grad } q =: n$ , so ist  $f = p/q$  holomorph bei  $\infty$  mit

$$f(\infty) = \begin{cases} \frac{a_m}{b_m} & \text{für } m = n, \\ 0 & \text{für } m < n, \end{cases}$$

denn, mit  $p(z) = \sum_{k=0}^m a_k z^k, q(z) = \sum_{k=0}^n b_k z^k$  ist

$$g(z) = f\left(\frac{1}{z}\right) = z^{n-m} \frac{a_m + a_{m-1}z + \dots + a_0 z^m}{b_n + \dots + b_0 z^n}$$

holomorph bei 0.

3. *Exponentialfunktion*  $\exp : \mathbb{C} \rightarrow \mathbb{C}$  ist holomorph in ganz  $\mathbb{C}$  mit  $\exp' = \exp$ :

$$\frac{e^w - 1}{w} = \frac{w + \frac{w^2}{2} + \dots}{w} = 1 + \frac{w}{2} + \frac{w^2}{3!} + \dots \rightarrow 1 \quad \text{für } w \rightarrow \infty,$$

$$\lim_{z \rightarrow z_0} \frac{e^z - e^{z_0}}{z - z_0} = \lim_{z \rightarrow z_0} \frac{e^{z-z_0} - 1}{z - z_0} e^{z_0} = 1 \cdot e^{z_0} = e^{z_0}.$$

4. *Sinus* und *Cosinus* sind holomorph in  $\mathbb{C}$ ,

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots = \frac{1}{2i}(e^{iz} - e^{-iz}), \quad \sin' = \cos,$$

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots = \frac{1}{2}(e^{iz} + e^{-iz}), \quad \cos' = -\sin.$$

Es gibt aber auch viele einfache Funktionen, die *nicht holomorph* sind:

1.  $\Omega \subset \mathbb{C}$  offen,  $f : \Omega \rightarrow \mathbb{R}$  (also  $f$  *reellwertig*) :  $f$  ist entweder

- holomorph: Dann ist  $f'(z) = 0$ , d. h.  $f$  ist konstant, oder
- nicht holomorph. wenn  $f$  nicht konstant ist, ist  $f$  nicht holomorph.

*Beweis.*  $f = u + iv \Rightarrow v \equiv 0 \Rightarrow v_x = v_y = 0$  (also auch  $u_y = u_x = 0$ )  $\Rightarrow f' = 0$ . ■

**17: 22. Juni 2006** Insbesondere sind  $z \mapsto \text{Re } z, z \mapsto \text{Im } z, z \mapsto |z|$  nicht holomorph.

2.  $f(z) = \bar{z} = \text{Re } z - i \text{Im } z = z - 2i \text{Im } z$ .

Da  $z \mapsto z$  holomorph ist und  $z \mapsto \text{Im } z$  nicht holomorph  $\Rightarrow z \mapsto \bar{z}$  ist nicht holomorph.



Ist  $f$  holomorph und 2 mal stetig differenzierbar, so gilt

$$\Delta u = u_{xx} + u_{yy} = v_{yx} - v_{xy} = 0,$$

$$\Delta v = v_{xx} + v_{yy} = -u_{yx} + u_{xy} = 0.$$

Dies ist (bei einer zusätzlichen Voraussetzung) umkehrbar:

**Satz 8.6**  $\Omega \subset \mathbb{C}$  offen und einfach zusammenhängend. Für eine Funktion  $u : \Omega \rightarrow \mathbb{R}$  sind die folgenden Aussagen äquivalent:

- (i)  $u$  ist zwei mal stetig differenzierbar mit  $\Delta u = 0$ .
- (ii) Es gibt eine in  $\Omega$  holomorphe Funktion  $f : \Omega \rightarrow \mathbb{C}$  mit  $u = \operatorname{Re} f$ .

Die gleiche Aussage gilt für den Imaginärteil.

*Beweis.* (ii)  $\Rightarrow$  (i): Hier benutzen wir (was erst später gezeigt wird), daß jede holomorphe Funktion zwei mal (tatsächlich sogar beliebig oft) stetig differenzierbar ist. Wie oben gezeigt wurde gilt dann  $\Delta u = 0$ .

(i)  $\Rightarrow$  (ii): Es ist ein  $v : \Omega \rightarrow \mathbb{R}$  zu finden so, daß  $f := u + iv$  holomorph ist, d. h. daß gilt

$$\operatorname{grad} v(x, y) = (v_x(x, y), v_y(x, y)) = (-u_y(x, y), u_x(x, y)).$$

Das Vektorfeld

$$k(x, y) := (-u_y(x, y), u_x(x, y))$$

ist rotationsfrei,

$$\operatorname{rot} k(x, y) = \frac{\partial}{\partial x} k_2(x, y) - \frac{\partial}{\partial y} k_1(x, y) = u_{xx}(x, y) + u_{yy}(x, y) = 0,$$

erfüllt also die Integrabilitätsbedingung. Also ist  $k(\cdot, \cdot)$  ein Gradientenfeld, d. h. es gibt ein Potential  $v(x, y)$  mit

$$(v_x(x, y), v_y(x, y)) = \operatorname{grad} v(x, y) = (u_y(x, y), u_x(x, y)).$$

Also ist  $f = u + iv$  holomorph. ■

Man kann dieses  $v(x, y)$  und damit  $f(z)$  *explizit berechnen* (natürlich nur bis auf eine additive Konstante): Wähle  $z_0 = x_0 + iy_0$  beliebig aus  $\Delta$ , und definieren

$$v(x, y) = \int_{(x_0, y_0)}^{(x, y)} (-u_y(\xi, \eta), u_x(\xi, \eta)) \, d(\xi, \eta),$$

wobei das Integral über eine beliebige Kurve in  $\Omega$  von  $(x_0, y_0)$  nach  $(x, y)$  erstreckt wird, z. B. – wenn es  $\Omega$  erlaubt – von  $(x, y)$  geradlinig nach  $(x_0, y_0)$  und von dort geradlinig nach  $(x, y)$ .

## 8.1 Möbiusfunktionen

Möbiusfunktionen (oder *-abbildungen*) haben die Form

$$f(z) = \frac{az + b}{cz + d} \quad \text{mit } \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \neq 0.$$

(Wäre  $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = 0$ , so wären die Zeilen linear abhängig und somit  $f$  konstant.) Um die Abbildungseigenschaften dieser Funktionen zu untersuchen sind i. wes. zwei Typen von Abbildungen zu untersuchen:

1. *Affine Funktionen*  $g(z) = az + b$  (entspricht der Möbiusfunktion mit  $c = 0$  und  $d = 1$ ) mit  $a \neq 0$ , denn sonst wäre  $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = 0$ . Diese besteht aus
  - (i) Drehung aus dem Winkel  $\varphi = \arg(a)$ ,
  - (ii) Streckung mit Faktor  $|a|$ ,
  - (iii) Translation um  $b$ .

Wegen  $a \neq 0$  ist  $g$  bijektiv,  $g^{-1}(w) = \frac{1}{a}(w - b) = \frac{1}{a}w - \frac{b}{a}$ , also wieder eine affine Funktion.

**Satz 8.7** *Jede affine Funktion  $g(z) = az + b$  mit  $a \neq 0$  bildet Geraden auf Geraden und Kreise auf Kreise ab. (Genauerer im Beweis!)*

*Beweis.* Die Gerade mit Richtung  $w$  durch  $z_0$  geht über in die Gerade mit Richtung  $aw$  durch  $az_0 + b$ . Der Kreis um  $z_0 = x_0 + iy_0$  mit Radius  $r$  hat die Form

$$K(z_0, r) = \{z \in \mathbb{C} : |z - z_0| = r\}.$$

Dann ist

$$\begin{aligned} g(K(z_0, r)) &= \left\{ w = az + b : z \in \mathbb{C}, |z - z_0| = r \right\} \\ &= \left\{ w \in \mathbb{C} : \left| \frac{1}{a}w - \frac{b}{a} - z_0 \right| = r \right\} \\ &= \left\{ w \in \mathbb{C} : |w - (b + az_0)| = r|a| \right\} = K(b + az_0, r|a|). \end{aligned}$$

■

2. Die Funktion  $h : \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C} \setminus \{0\}$ ,  $h(z) = \frac{1}{z}$ . Sie ist offenbar bijektiv (aber nicht auf  $\mathbb{C}$ , sondern auf  $\mathbb{C} \setminus \{0\}$ ). Man erweitert deshalb  $\mathbb{C}$  um einen Punkt „ $\infty$ “ = *Unendlich* und definiert

$$\tilde{h}(0) := \infty, \quad \tilde{h}(\infty) := 0.$$

$\tilde{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$  heißt die *erweiterte Zahlenebene*.  $\tilde{h} : \tilde{\mathbb{C}} \rightarrow \tilde{\mathbb{C}}$  ist bijektiv.

Zur weiteren Untersuchung ist es nützlich, die *stereographische Projektion* zu betrachten. Sie bildet  $\tilde{\mathbb{C}}$  bijektiv auf die *Riemannsche Zahlenkugel* ab,

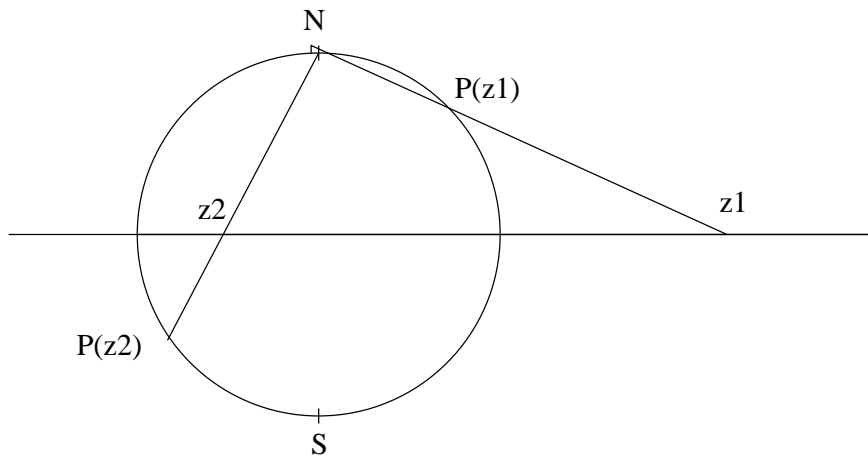


Abbildung 4: Stereographische Projektion

Dabei ist  $P(\infty) = N$  (Nordpol). In Formeln gilt

$$P(z) = \begin{cases} \frac{1}{|z|^2 + 1} (2 \operatorname{Re} z, 2 \operatorname{Im} z, |z|^2 - 1) & \text{für } z \in \mathbb{C}, \\ N = (0, 0, 1) & \text{für } z = \infty, \end{cases}$$

$$P^{-1}(x) = P^{-1}(\xi_1, \xi_2, \xi_3) = \begin{cases} \frac{1}{1 - \xi_3} (\xi_1 + i\xi_2) & \text{für } x \neq N, \\ \infty & \text{für } x = N. \end{cases}$$

Offensichtlich sind  $P$  und  $P^{-1}$  stetig, wobei dies bei  $N$  bzw.  $\infty$  so zu verstehen ist:

$$\text{aus } z \rightarrow \infty \text{ folgt } P(z) \rightarrow N, \quad \text{aus } x \rightarrow N \text{ folgt } P^{-1}(x) \rightarrow \infty.$$

(Mit Hilfe von Umgebungen heißt dies: Jede Menge in  $\tilde{\mathbb{C}}$ , die den gesamten Außenbereich eines Kreises um 0 enthält, ist eine Umgebung von  $\infty$ .) In diesem Sinn gilt: Stetige Funktionen auf  $\tilde{\mathbb{C}}$  entsprechen genau den stetigen Funktionen auf der Zahlenkugel.

**Satz 8.8** *Kreise auf der Zahlenkugel entsprechen unter stereographischer Projektion den Kreisen und Geraden in  $\tilde{\mathbb{C}}$ , diese werden als verallgemeinerte Kreise, kurz „Kreise“ bezeichnet. (Kreise durch  $N$  entsprechen offenbar Geraden [speziell Großkreise durch  $N$  entsprechen Geraden durch 0], alle anderen Kreise auf der Kugel gehen in Kreise über.)*

Der *Beweis* ist etwas mühsam (vgl. z.B. Skript zur Analysis III).

**Satz 8.9** *Die Abbildung  $\tilde{h} : \tilde{\mathbb{C}} \rightarrow \tilde{\mathbb{C}}$ ,  $\tilde{h}(z) = \frac{1}{z}$  ( $\tilde{h}(\infty) = 0, \tilde{h}(0) = \infty$ ) bildet „Kreise“ auf „Kreise“ ab. Genauer:*

- Geraden durch 0 auf Geraden durch 0,
- Geraden, die nicht durch 0 gehen auf Kreise durch 0,
- Kreise durch 0 auf Geraden,

– Kreise die nicht durch 0 gehen auf Kreise, die nicht durch 0 gehen.

Auch für diesen Beweis sei auf das Skript zur Analysis III verwiesen.

Damit sind wir in der Lage, die allgemeine Möbiustransformation zu studieren:

$c = 0$ : Wegen  $\det \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \neq 0$  folgt  $a \neq 0, d \neq 0$ , d. h.  $f(z) = \frac{a}{d}z + \frac{b}{d}$  ist eine affine Abbildung.

$c \neq 0$ : Für  $z \neq -\frac{d}{c}$  und  $z \neq \infty$  gilt

$$\begin{aligned} f(z) &= \frac{\frac{a}{c}(cz+d) + b - \frac{a}{c}d}{cz+d} = \frac{a}{c} + \frac{1}{c} \frac{bc-ad}{cz+d} \\ &= f_3 \circ f_2 \circ f_1 \end{aligned}$$

mit

$$\begin{aligned} f_1(z) &= cz + d \quad \text{affin,} \\ f_2(z) &= \frac{1}{z} \\ f_3(z) &= \frac{bc-ad}{c}z + \frac{a}{c} \quad \text{affin.} \end{aligned}$$

Durch Grenzübergang folgt: Für  $z = -\frac{d}{c}$  ist  $f\left(-\frac{d}{c}\right) = \infty$ , für  $z = \infty$  ist  $f(\infty) = \frac{a}{c}$ ; diese Werte erhält man auch mit  $f = f_3 \circ f_2 \circ f_1$ .

**Satz 8.10** Jede Möbiustransformation  $\left(A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)$

$$f_A(z) = \frac{az+b}{cz+d} \quad \text{für } z \in \mathbb{C} \setminus \left\{-\frac{d}{c}\right\}, \quad f_A\left(-\frac{d}{c}\right) = \infty, \quad f_A(\infty) = \frac{a}{c}$$

ist eine bijektive Abbildung von  $\tilde{\mathbb{C}}$  auf  $\tilde{\mathbb{C}}$ , die „Kreise“ auf „Kreise“ abbildet.

18: 27. Juni 2006

**Satz 8.11** Für  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$  mit  $\det A \neq 0, \det B \neq 0$  gilt

$$f_{AB} = f_A \circ f_B, \quad \text{speziell } f_E = \text{id}, \quad f_{A^{-1}} = f_A^{-1}.$$

*Beweis.* Nachrechnen

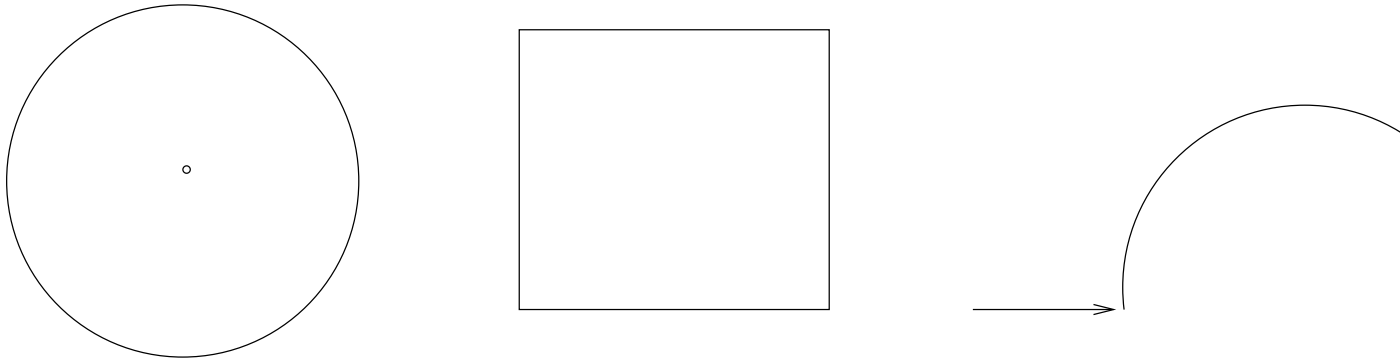
$$f_{AB}(z) = f_A\left(\frac{\alpha z + \beta}{\gamma z + \delta}\right) = \frac{a \frac{\alpha z + \beta}{\gamma z + \delta} + b}{c \frac{\alpha z + \beta}{\gamma z + \delta} + d} = \dots = \frac{(a\alpha + b\gamma)z + a\beta + b\delta}{(c\alpha + d\gamma)z + c\beta + d\delta}$$

und Ausnahmepunkte extra überprüfen. ■

**Beispiel 8.12** Die Abbildung  $f : z \mapsto \frac{1}{z}$  ( $\infty \mapsto 0, 0 \mapsto \infty$ ) bildet die Gerade  $x = 1$  ( $y \in \mathbb{R}$ ) auf den Kreis mit Radius  $\frac{1}{2}$  um  $\left(\frac{1}{2}, 0\right)$  ab. Die Halbebene rechts von  $x = 1$  geht in das Innere des Kreises, die Halbebene links von  $x = 1$  in das Äußere des Kreises über. □

## 9 Kurvenintegrale und Cauchyscher Integralsatz

Die im folgenden betrachteten *Kurven* oder Wege  $\Gamma$  in  $\mathbb{C}$  seien mit einer Parameterdarstellung  $\gamma : [a, b] \rightarrow \mathbb{C}$  gegeben, wobei  $\gamma$  zumindest stückweise stetig differenzierbar ist. (In der Regel wird  $\gamma$  auf einem beschränkten abgeschlossenen Intervall definiert sein; gelegentlich werden wir auch ein unbeschränktes Parameterintervall zulassen.) Typische Kurven sind etwa



$\Omega \in \mathbb{C}$  offen,  $f : \Omega \rightarrow \mathbb{C}$  stetig und  $\Gamma$  eine Kurve in  $\Omega$  mit Parameterdarstellung  $\gamma : [a, b] \rightarrow \mathbb{C}$ . Wie beim Integral über ein reelles Intervall betrachten wir Summen ( $a = t_0 < t_1 < \dots < t_k = b$ )

$$\sum_{n=1}^k f(\gamma(t_n))(\gamma(t_n) - \gamma(t_{n-1})) \quad \text{„} = \sum \text{ Funktionswert} \times \text{Wegstück“}.$$

Ist  $\gamma$  stetig differenzierbar, so ist die etwa gleich (und zwar umso besser, je feiner die Zerlegung  $a = t_0 < t_1 < \dots < t_k = b$  gewählt ist)

$$\dots \sim \sum_{n=1}^k f(\gamma(t_n))\gamma'(t_n)(t_n - t_{n-1}).$$

Das ist eine Riemannsumme für das Integral

$$\int_a^b f(\gamma(t))\gamma'(t) dt =: \int_{\Gamma} f(z) dz.$$

Letzteres ist dadurch definiert und heißt das *Kurvenintegral von  $f$  über die Kurve  $\Gamma$* .

- *Linearität:*  $\int_{\Gamma} (af + bg) dz = a \int_{\Gamma} f dz + b \int_{\Gamma} g dz.$
- *Abschätzung:*  $|\int f(z) dz| \leq |G| \max\{|f(z)| : z \in \Gamma\}.$
- Sind  $\Gamma_1, \Gamma_2$  zwei Wege mit Endpunkt von  $\Gamma_1 =$  Anfangspunkt von  $\Gamma_2$ , so gilt für  $\Gamma := \Gamma_1 + \Gamma_2$

$$\int_{\Gamma} = \int_{\Gamma_1} + \int_{\Gamma_2}.$$

– Ist  $\Gamma_-$  der inverse Weg zu  $\Gamma$ , d. h.  $\Gamma$  wird in umgekehrter Richtung durchlaufen,

$$\Gamma_- : [a, b] \rightarrow \mathbb{C}, \quad \gamma_-(t) = \gamma(b + a - t) \quad \text{und} \quad (\gamma_-(a) = \gamma(b), \quad \gamma_-(b) = \gamma(a)),$$

so gilt ( $\gamma'_-(t) = -\gamma'(b + a - t)$ )

$$\begin{aligned} \int_{\Gamma_-} f(z) dz &= \int_a^b f(\gamma_-(t)) \gamma'_-(t) dt = - \int_a^b f(\gamma(a + b - t)) \gamma'(b + a - t) dt \\ &= \int_b^a f(\gamma(s)) \gamma'(s) ds = - \int_a^b f(\gamma(s)) \gamma'(s) ds = - \int_{\Gamma} f(z) dz. \end{aligned}$$

**Beispiel 9.1** Wir integrieren  $f(z) = z^n$  ( $n \in \mathbb{N}_0$ ) über verschiedene Wege in  $\mathbb{C}$  von  $-1$  nach  $1$ : a) *Geradlinige Verbindung:*  $\Gamma_1, \gamma_1 : [-1, 1] \rightarrow \mathbb{C}, \gamma_1 = \text{id}$ . In diesem Fall ist das Integral

$$\int_{[-1,1]} z^n dz = \int_{-1}^1 t^n \cdot 1 dt = \frac{1}{n+1} t^{n+1} \Big|_{-1}^1 = \frac{1}{n+1} \{1 - (-1)^{n+1}\},$$

also das ganz normale „reelle“ Integral über  $[-1, 1]$ . Oder, mit der Parameterdarstellung  $\varphi_1 : [0, 1] \rightarrow \mathbb{C}, \varphi_1(t) = 2t - 1$

$$\int_{[-1,1]} z^n dz = \int_0^1 (2t - 1)^n \cdot 2 dt = \frac{1}{n+1} (2t - 1)^{n+1} \Big|_0^1 = \text{vgl. oben.}$$

b) Kreisbogen um  $0$  in der unteren Halbebene,

$$\begin{aligned} \Gamma_2, \gamma_2 : [0, 1] \rightarrow \mathbb{C}, \quad \gamma_2(t) &= e^{i\pi(t-1)} \\ \int_{\Gamma_2} z^n dz &= \int_0^1 \gamma_2(t)^n \gamma_2'(t) dt = \int_0^1 e^{in\pi(t-1)} i\pi e^{i\pi(t-1)} dt \\ &= i\pi \int_0^1 e^{i(n+1)\pi(t-1)} dt = \frac{1}{n+1} e^{i(n+1)\pi(t-1)} \Big|_0^1 \\ &= \frac{1}{n+1} \{1 - (-1)^{n+1}\} \stackrel{!}{=} \int_{\Gamma_1} z^n dz. \end{aligned}$$

Diese letzte Rechnung gilt übrigens (bitte nachprüfen!) auch für  $n \in \mathbb{Z}$  mit  $z < -1$ . c) Kreisbogen um  $0$  in der oberen Halbebene liefert (ebenfalls für alle  $z \in \mathbb{Z} \setminus \{-1\}$ ) das gleiche Resultat (Übung!).  
□

Wie in  $\mathbb{R}$  definieren wir: Sei  $\Omega \subset \mathbb{C}$  offen,  $f : \Omega \rightarrow \mathbb{C}$ . Dann heißt  $F : \Omega \rightarrow \mathbb{C}$  (eine) *Stammfunktion* von  $f$ , wenn  $F$  in  $\Omega$  komplex differenzierbar (also holomorph) ist mit  $F'(x) = f(x)$  für alle  $x \in \Omega$ .

Mit  $F$  ist immer auch  $F + c$  eine Stammfunktion. Zwei Stammfunktionen unterscheiden sich nur um eine Konstante (die allerdings auf verschiedenen Zusammenhangskomponenten verschieden sein kann).

Wann besitzt  $f$  eine Stammfunktion? Oder anders formuliert, wann ist  $f$  *integrierbar* (später mehr zu diesem Begriff).

Beispiele integrierbarer Funktionen sind jedenfalls

- $f(z) = z^n$  für  $n \in \mathbb{N}$  hat auf  $\mathbb{C}$  eine Stammfunktion  $F(x) = \frac{1}{n+1}z^{n+1}$ . Für  $n \in \mathbb{Z}$ ,  $n < -1$  ist  $f(z) = z^n$  auf  $\Omega \setminus \{0\}$  definiert und hat dort ebenfalls die Stammfunktion  $\frac{1}{n+1}z^{n+1}$ .
- $f(z) = \exp(z)$  hat auf  $\mathbb{C}$  die Stammfunktion  $\exp(z)$ ,
- $\sin z$  hat die Stammfunktion  $-\cos z$ .
- $\cos z$  hat die Stammfunktion  $\sin z$ .

Allgemein gilt der folgende Satz:

**Satz 9.2**  $\Omega \subset \mathbb{C}$  offen,  $f : \Omega \rightarrow \mathbb{C}$  stetig. Dann gilt:  $f$  besitzt genau dann eine Stammfunktion  $F$ , wenn das Kurvenintegral über  $f$  (in  $\Omega$ ) wegunabhängig ist. Es gilt dann

$$\int_{\Gamma} f(z) dz = F(\text{End-Pkt}) - F(\text{Anf-Pkt}),$$

insbesondere ist das Integral über jede geschlossene Kurve Null.

*Beweis.*  $\Rightarrow$ : Sei  $F$  eine Stammfunktion von  $f$ . Dann gilt

$$\begin{aligned} \int_{\Gamma} f(z) dz &= \int_a^b F(\gamma(t))\gamma'(t) dt = \int_a^b \frac{d}{dt} F(\gamma(t)) dt \\ &= F(\gamma(b)) - F(\gamma(a)) = F(\text{End-Pkt}) - F(\text{Anf-Pkt}). \end{aligned}$$

$\Leftarrow$ : Sei jetzt das Kurvenintegral wegunabhängig. Dann können wir offenbar definieren

$$F(z) = \int_{z_0}^z f(\xi) \quad \text{für alle } z \in \Omega,$$

wobei  $z_0 \in \Omega$  fest gewählt wird und das Integral über einen beliebigen Weg von  $z_0$  nach  $z$  erstreckt wird (wenn  $\Omega$  zusammenhängend ist, ist dies für alle  $z \in \Omega$  möglich; andernfalls muß man in jeder Zusammenhangskomponente entsprechend verfahren).

Wir wollen nun  $F(\cdot)$  in einem Punkt  $z$  differenzieren. Für jedes  $z' \in \Omega$ , daß in  $\Omega$  geradlinig mit  $z$  verbunden werden kann (da  $\Omega$  offen ist, sind das alle Punkte aus einem kleinen Kreis um  $z$ )

$$\begin{aligned} \frac{F(z') - F(z)}{z' - z} &= \frac{1}{z' - z} \int_{[z, z']} f(\xi) d\xi = f(z) + \frac{1}{z' - z} \int_{[z, z']} (f(\xi) - f(z)) d\xi \\ &\rightarrow f(z) \quad \text{für } z' \rightarrow z, \end{aligned}$$

da der Integrand des Integrals gegen 0 konvergiert. Also ist  $F$  eine Stammfunktion von  $f$ . ■

Wie schon in der reellen Analysis definieren wir: Ein Gebiet  $G \subset \mathbb{C}$  heißt *einfach zusammenhängend*, wenn jede geschlossene Kurve in  $G$  nur Punkte aus  $G$  umschließt oder, wenn jede geschlossene Kurve stetig auf einen Punkt zusammengezogen werden kann, ohne daß man dabei  $G$  verläßt. Das wichtigste *nicht einfach zusammenhängende Gebiet* ist  $\mathbb{C} \setminus \{0\}$  oder  $\mathbb{C} \setminus \{0\}$  mit einem beliebigen  $z_0 \in \mathbb{C}$ .

**Satz 9.3 (Cauchyscher Integralsatz für einfach zusammenhängende Gebiete)** *Ist  $G$  ein einfach zusammenhängendes Gebiet,  $f : G \rightarrow \mathbb{C}$  holomorph,  $\Gamma$  eine geschlossene Kurve in  $G$ . Dann gilt*

$$\int_{\Gamma} f(z) dz = 0, \quad \text{d. h. } f \text{ ist integrierbar.}$$

*Beweis 1.* Der Beweis kann ganz kurz mit uns bereits bekannten Mitteln der reellen Analysis geführt werden, wenn wir annehmen, daß  $f$  nicht nur differenzierbar, sondern sogar stetig differenzierbar ist. Zwar wird sich später zeigen, daß jede komplex differenzierbare Funktion sogar beliebig oft differenzierbar ist. Aber dazu brauchen wir gerade den hier zu beweisenden Satz.

Sei also  $f$  stetig komplex differenzierbar,  $f = u + iv$ . Die beiden Vektorfelder in  $G$

$$k = (u, -v) \quad \text{und} \quad \tilde{k} = (v, u)$$

sind Gradientenfelder, denn sie sind stetig differenzierbar und es gilt (Integrabilitätsbedingung)

$$\begin{aligned} \frac{\partial}{\partial x_1} k_2 - \frac{\partial}{\partial x_2} k_1 &\cong -v_x - u_y = -(v_x + u_y) = 0, \\ \frac{\partial}{\partial x_1} \tilde{k}_2 - \frac{\partial}{\partial x_2} \tilde{k}_1 &\cong u_x - v_y = 0. \end{aligned}$$

Damit folgt (mit der Parameterdarstellung  $\gamma(t) = x(t) + iy(t)$ ):

$$\begin{aligned} \int_{\Gamma} f(z) dz &= \int_a^b f(\gamma(t)) \gamma'(t) dt \\ &= \int_a^b \left( u(\gamma(t)) + iv(\gamma(t)) \right) (\dot{x}(t) + i\dot{y}(t)) dt \\ &= \int_a^b \left\{ u(x(t), y(t)) \dot{x}(t) - v(x(t), y(t)) \dot{y}(t) \right. \\ &\quad \left. + i \left[ u(x(t), y(t)) \dot{y}(t) + v(x(t), y(t)) \dot{x}(t) \right] \right\} dt \\ &= \int_{\Gamma} (u(x, y), -v(x, y)) d(x, y) + i \int_{\Gamma} (v(x, y), u(x, y)) d(x, y) \\ &= 0 \end{aligned}$$

■



**19: 29. Juni 2006** Für einen vollständigen Beweis werden einige Vorbereitungen benötigt. Entscheidend ist der

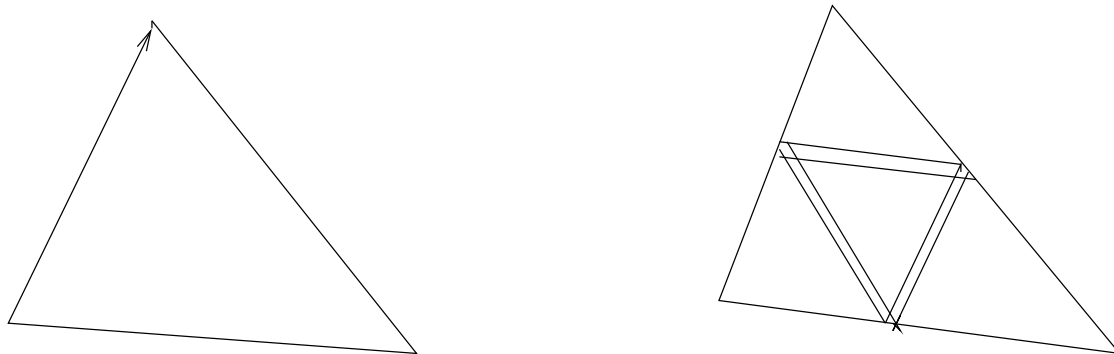
**Hilfssatz 9.4** Sei  $G$  ein einfach zusammenhängendes Gebiet,  $f : G \rightarrow \mathbb{C}$  holomorph. Dann gilt für jeden Dreiecksweg  $\Delta$  in  $G$ :

$$\int_{\Delta} f(z) dz = 0.$$

*Beweis.* Man beachte, daß wir nur die Differenzierbarkeit in jedem Punkt von  $G$  benutzen, aber *nicht* die Stetigkeit der Ableitung.

Sei  $\Delta_0$  ein Dreiecksweg in  $G$  (insbesondere liegt, da  $G$  einfach zusammenhängend ist, auch das Innere von  $\Delta$  in  $G$ ). Wir nehmen an, daß  $C := \left| \int_{\Delta_0} f(z) dz \right| > 0$  gilt, und führen dies zu einem Widerspruch zur Differenzierbarkeit.

Dazu zerlegen wir  $\Delta_0$  in 4 Dreieckswege, indem jede Seite halbiert wird; jede Strecke im Inneren wird 2 mal – in verschiedener Richtung – durchlaufen:



Dann gilt

$$C = \left| \int_{\Delta_0} f(z) dz \right| \leq \sum_{j=1}^4 \left| \int_{\Delta_j} f(z) dz \right|.$$

Also gibt es ein  $j_1 \in \{1, 2, 3, 4\}$  mit

$$\left| \int_{\Delta_1^{j_1}} f(z) dz \right| \geq \frac{C}{4}.$$

Sei nun  $\Delta_1 := \Delta_1^{j_1}$ . Mit diesem verfährt man entsprechend wie mit  $\Delta_0$  und findet ein  $\Delta_2 := \Delta_2^{j_2}$  mit

$$\left| \int_{\Delta_2} f(z) dz \right| \geq \frac{C}{4^2}$$

usw. für alle  $n \in \mathbb{N}$  ein  $\Delta_n$  mit

$$\left| \int_{\Delta_n} f(z) dz \right| \geq \frac{C}{4^n} \quad (*)$$

Ist  $D_n$  die größte Kantenlänge von  $\Delta_n$ ,  $L_n$  die Länge des gesamten Weges  $\Delta_n$ , so gilt

$$D_n = \frac{D_0}{2^n}, \quad L_n = \frac{L_0}{2^n}.$$

Da die Dreiecke ineinander enthalten sind ( $\Delta_{n+1} \subset \Delta_n$ ), gibt es genau einen Punkt  $z_0 \in \mathbb{C}$ , der in allen Dreiecken  $\Delta_n$  enthalten ist, und auf den sich die  $\Delta_n$  zusammenziehen. Da  $f$  in  $z_0$  differenzierbar ist <sup>7</sup>, gilt

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + \varphi_{z_0}(z) \quad \text{mit} \quad \frac{\varphi_{z_0}(z)}{z - z_0} \rightarrow 0 \quad \text{für} \quad z \rightarrow z_0,$$

also

$$\begin{aligned} \left| \int_{\Delta_n} f(z) \, dz \right| &= \left| \int_{\Delta_n} \left\{ f(z_0) + f'(z_0)(z - z_0) \right\} dz + \int_{\Delta_n} \left\{ (z - z_0) \frac{\varphi_{z_0}(z)}{z - z_0} dz \right\} \right| \\ &\quad (\text{der 1. Integrand hat die Stf. } f(z_0)z + \frac{1}{2}f'(z_0)(z - z_0)^2) \\ &\leq |\Delta_n| \max \left\{ |z - z_0| \frac{|\varphi_{z_0}(z)|}{|z - z_0|} : z \in \Delta_n \right\} \\ &\leq L_n D_n \max \left\{ \frac{|\varphi_{z_0}(z)|}{|z - z_0|} : z \in \Delta_n \right\} \\ &= \frac{1}{4^n} L_0 D_0 \max \left\{ \frac{|\varphi_{z_0}(z)|}{|z - z_0|} : z \in \Delta_n \right\}. \end{aligned}$$

Wegen  $\max\{\dots\} \rightarrow 0$  für  $n \rightarrow \infty$  ist dies ein Widerspruch zu (\*). ■

*Beweis 2 von Satz ??.* Aus dem Hilfssatz folgt für jeden geschlossenen Polygonzug  $\Gamma$ , indem man ihn in Dreieckswege zerlegt,  $\int_{\Gamma} f(z) \, dz = 0$ . Dann ist

$$F(z) := \int_{z_0}^z f(z) \, dz$$

wohldefiniert, wenn man als Integrationsweg beliebige Polygonzüge von  $z_0$  nach  $z$  zuläßt. Wie im Beweis des vorhergehenden Satzes folgt, daß  $F$  eine Stammfunktion von  $f$  ist. ■

Im Folgenden betrachten wir beliebige – also nicht notwendig einfach zusammenhängende Gebiete  $G$  in  $\mathbb{C}$ .

Einen geschlossenen Weg  $\Gamma$  in  $G$  nennen wir *nullhomotop* (bezüglich  $G$ ), wenn er sich stetig auf einen Punkt zusammenziehen läßt, ohne daß man dabei  $G$  verläßt. Im allgemeinen ist das sicher schwer zu entscheiden, ob ein Weg nullhomotop ist. In allen konkreten Fällen ist es leicht. (Man kann das auch etwas genauer so fassen:  $\Gamma$  ist nullhomotop (bezüglich  $G$ ), wenn es sich schreiben läßt in der Form  $\Gamma = \Gamma_1 + \dots + \Gamma_n$  mit geschlossenen Wegen  $\Gamma_j$ , die in einer einfach zusammenhängenden

<sup>7</sup>Wir benutzen die Differenzierbarkeit zwar nur in diesen Punkt, da man aber nicht weiß, welches der Punkt  $z_0$  sein wird, muß die Differenzierbarkeit in allen Punkten vorausgesetzt werden.

Teilmenge  $G_j$  von  $G$  liegen [bei der Zerlegung werden natürlich in der Regel gewisse Wegstücke hinzugenommen, die in beiden Richtungen durchlaufen werden].)

Am schwierigsten ist es natürlich, von einem Weg zu zeigen, daß er nicht nullhomotop ist; es könnte ja sein, daß man zu ungeschickt ist, es zu zeigen.

**Beispiel 9.5** Ein Kreisweg  $\Gamma$  um  $0$  (und entsprechend um jeden anderen Punkt  $z_0$ ) ist nicht nullhomotop bezüglich  $\mathbb{C} \setminus \{0\}$  (bzw.  $\mathbb{C} \setminus \{z_0\}$ ). Da  $f(z) = \frac{1}{z}$  in  $\mathbb{C} \setminus \{0\}$  holomorph ist, müßte sonst nach dem folgenden Satz  $\int_{\Gamma} \frac{1}{z} dz = 0$  gelten. Tatsächlich ist aber

$$\int_{\Gamma} \frac{1}{z} dz = \int_0^{2\pi} \frac{1}{e^{it}} i e^{it} dt = \int_0^{2\pi} i dt = 2\pi i \neq 0.$$

□

**Satz 9.6 (Cauchyscher Integralsatz, allgemeine Form)** Sei  $G \subset \mathbb{C}$  offen,  $f : G \rightarrow \mathbb{C}$  holomorph. Dann gilt:

- a) Für jeden geschlossenen bezüglich  $G$  nullhomotopen Weg  $\Gamma$  in  $G$  ist  $\int_{\Gamma} f(z) dz = 0$ .
- b) Sind  $\Gamma_1, \Gamma_2$  zwei Wege von  $z_0$  nach  $z_1$  so, daß  $\Gamma_1 + (-\Gamma_2)$  bezüglich  $G$  nullhomotop ist, so gilt

$$\int_{\Gamma_1} f(z) dz = \int_{\Gamma_2} f(z) dz.$$

*Beweis.* a) Sei  $\Gamma = \Gamma_1 + \dots + \Gamma_n$ ,  $\Gamma_j$  geschlossen mit  $\Gamma_j \subset G_j \subset G$  und  $G_j$  einfach zusammenhängend. Dann gilt nach dem vorhergehenden Satz

$$\int_{\Gamma} f(z) dz = \sum_{j=1}^n \int_{\Gamma_j} f(z) dz = 0.$$

b) Folgt aus Teil a. ■

**Satz 9.7** Sei  $G \subset \mathbb{C}$  offen und einfach zusammenhängend,  $A \subset G$ ,  $f : G \setminus A \rightarrow \mathbb{C}$  holomorph. Dann existiert ein  $C \in \mathbb{C}$  so, daß für jeden geschlossenen Weg  $\Gamma$  in  $G \setminus A$ , der  $A$  genau  $N$  mal im positiven Sinn umläuft (Umläufe im negativen Sinn zählen negativ) gilt

$$\int_{\Gamma} f(z) dz = NC.$$

Ist  $A = \{z_0\}$  ein einzelner Punkt, so nennt man  $z_0$  eine isolierte Singularität; die Zahl  $C$  heißt dann das Residuum von  $f$  bei  $z_0$ ,  $\text{Res}(f, z_0)$ .

*Beweis.* Sei  $\Gamma_0$  ein beliebiger (fest gewählter) geschlossener Weg um  $A$  in  $G \setminus A$ , der  $A$  genau 1 mal umläuft und

$$C := \int_{\Gamma_0} f(z) dz.$$

Durch Einfügen geeigneter doppelt durchlaufener Wege erkennt man, daß  $\Gamma - \Gamma_0 - \dots - \Gamma_0 = \Gamma - N\Gamma_0$  bezüglich  $G \setminus A$  nullhomotop ist, also

$$\int_{\Gamma} f(z) dz = N \int_{\Gamma_0} f(z) dz = NC.$$

■

**Beispiel 9.8**  $G = \mathbb{C} \setminus \{z_0\}$ ,  $f(z) = (z - z_0)^n$ ,  $n \in \mathbb{Z}$ ,  $\Gamma_r$  der Kreis mit Radius  $r$  um  $z_0$  in positiver Richtung,

$$\begin{aligned} \gamma_r : [0, 1] &\rightarrow \mathbb{C}, \quad \gamma_r(t) = z_0 + re^{2\pi it} \\ \int_{\gamma_r} (z - z_0)^n dz &= \int_0^1 r^n e^{2\pi int} 2\pi i r e^{2\pi it} dt = 2\pi i r^{n+1} \int_0^1 e^{2\pi i(n+1)t} dt. \end{aligned}$$

$n \geq 0$ :  $\dots = 0$ , folgt durch Rechnung oder aus dem Cauchyschen Integralsatz.

$n = -1$ :  $\dots = 2\pi i \int_0^1 1 dt = 2\pi i$ .

$n < -1$ : Mit der gleichen Rechnung, wie sie für  $n \geq 0$  möglich gewesen wäre folgt

$$\dots = \frac{2\pi i r^{n+1}}{2\pi i(n+1)} e^{2\pi i(n+1)t} \Big|_{t=0}^{t=1} = \frac{r^{n+1}}{n+1} (1 - 1) = 0.$$

Für alle  $n < 0$  ist  $z_0$  eine isolierte Singularität, aber nur für  $n = -1$  ist das Residuum  $\neq 0$ . □

**20: 4. Juli 2006**

## 10 Cauchysche Integralformel

Die Werte einer holomorphen Funktion im Inneren einer geschlossenen Kurve sind durch die Werte auf der Kurve festgelegt:

**Satz 10.1 (Cauchysche Integralformel)** Sei  $\Omega \subset \mathbb{C}$  offen,  $f : \Omega \rightarrow \mathbb{C}$  holomorph,  $z \in \Omega$ . Dann gilt für jede geschlossene Kurve  $\Gamma$  in  $\Omega$ , deren Inneres ganz in  $\Omega$  liegt und die  $z$  genau ein mal in mathematisch positivem Sinn umläuft, die Cauchysche Integralformel

$$f(z) = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(\xi)}{\xi - z} d\xi.$$

*Beweis.* Da  $\Omega$  offen ist, gibt es ein  $\delta > 0$  mit  $\overline{K}(z, \delta) \subset \Omega$ . Also liegen alle Kreiswege  $\Gamma_r$  mit  $0 < r \leq \delta$

$$\gamma_r : [0, 2\pi] \rightarrow \mathbb{C}, \quad \gamma_r(t) = z + re^{it}$$

und deren Inneres ganz in  $\Omega$ . Da  $\Gamma$  den Punkt  $z$  genau ein mal umläuft, ist  $\Gamma$  bezüglich  $\Omega$  homotop zu  $\Gamma_r$  (d. h.  $\Gamma - \Gamma_r$  ist nullhomotop bezüglich  $\Omega$ ), und die  $\Gamma_r$  sind alle zueinander homotop bezüglich  $\Omega$ . Also gilt für alle  $r \in (0, \delta]$

$$\begin{aligned} \int_{\Gamma} \frac{f(\xi)}{\xi - z} d\xi &= \int_{\Gamma_r} \frac{f(\xi)}{\xi - z} d\xi = \int_{\Gamma_r} \frac{f(\xi) - f(z)}{\xi - z} d\xi + f(z) \int_{\Gamma_r} \frac{1}{\xi - z} d\xi \\ &= \int_{\Gamma_r} \frac{f(\xi) - f(z)}{\xi - z} d\xi + 2\pi i f(z). \end{aligned}$$

Wegen

$$\left| \int_{\Gamma_r} \frac{f(\xi) - f(z)}{\xi - z} d\xi \right| \leq 2\pi r \max \left\{ |f(\xi) - f(z)| : |\xi - z| = r \right\} \rightarrow 0 \quad \text{für } r \rightarrow 0$$

folgt daraus

$$\int_{\Gamma} \frac{f(\xi)}{\xi - z} d\xi = 2\pi i f(z);$$

das ist die Behauptung. ■

**Satz 10.2 (Mittelwerteigenschaft)** Sei  $\Omega \subset \mathbb{C}$  offen,  $f : \Omega \rightarrow \mathbb{C}$  holomorph,  $z \in \Omega$ . Dann gilt für jedes  $r > 0$  mit  $\overline{K}(z, r) \subset \Omega$

$$\begin{aligned} f(z) &= \frac{1}{2\pi} \int_0^{2\pi} f(z + re^{it}) dt = \int_0^1 f(z + re^{2\pi it}) dt \\ &= \frac{1}{\pi r^2} \int_{|\xi - z| \leq r} f(x + iy) d(x, y) \quad (\xi = x + iy). \end{aligned}$$

Die erste Zeile ist das Mittel von  $f$  über die Kreislinie mit Radius  $r$  um  $z$ , die zweite das Mittel von  $f$  über die Kreisscheibe mit Radius  $r$  und Mittelpunkt  $z$ .

*Beweis.* Nach der Cauchyschen Integralformel gilt

$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \int_{\Gamma_r} \frac{f(\xi)}{\xi - z} d\xi = \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(z + re^{it})}{re^{it}} ire^{it} dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(z + re^{it}) dt = \int_0^1 f(z + re^{2\pi is}) ds. \end{aligned}$$

Da diese Formel mit  $\varrho$  statt  $r$  für alle  $\varrho \leq r$  gilt, folgt

$$\begin{aligned} f(z) &= f(z) \frac{2}{r^2} \int_0^r \varrho d\varrho = \frac{2}{r^2} \int_0^r \varrho f(z) d\varrho \\ &= \frac{2}{2\pi r^2} \int_0^r \varrho \int_0^{2\pi} f(z + \varrho e^{it}) dt d\varrho = \frac{1}{\pi r^2} \int_{|\xi - z| \leq r} f(x + iy) d(x, y). \end{aligned}$$

Die letzte Gleichheit setzt etwas Kenntnis der mehrdimensionalen Integration voraus, die in diesem Vorlesungszyklus noch nicht behandelt werden. (Man kann dies auch als zusätzliche Mittelung über die verschiedenen Kreise betrachten.) ■

**Korollar 10.3** Sei  $\Omega \subset \mathbb{R}^2$  offen. Jede harmonische Funktion  $u : \Omega \rightarrow \mathbb{R}$  erfüllt die obigen Mittelwertesigenschaften. (Dies gilt auch für harmonische Funktionen in  $\mathbb{R}^m$ , was aber ganz anders bewiesen wird.)

*Beweis.* Jede harmonische Funktion ist (z. B.) Realteil einer holomorphen Funktion. Da die obigen Eigenschaften offenbar auch für Real- und Imaginärteile von  $f$  gelten, folgt damit die Behauptung. ■

**Satz 10.4 (Maximumsprinzip)** Sei  $\Omega \subset \mathbb{C}$  offen,  $f : \Omega \rightarrow \mathbb{C}$  holomorph.

- Ist  $\overline{K}(z, r) \subset \Omega$  und  $|f(z)| = \max\{|f(\xi)| : |\xi - z| \leq r\}$ , so ist  $f$  konstant auf  $\overline{K}(z, r)$ .
- Ist  $\Omega$  zusammenhängend und  $f$  nicht konstant, so besitzt  $|f|$  in  $\Omega$  kein globales Maximum.

*Beweis.* a) Nehmen wir an, daß  $f$  auf  $K(z, r)$  nicht konstant ist. Dann ist auch  $|f|$  nicht konstant. Wäre nämlich  $|f|$  konstant, so lägen alle Werte auf einer Kreislinie um 0, könnten also mit einer Möbiustransformation  $\varphi$  auf die reelle Gerade abgebildet werden, d. h.  $\varphi \circ f$  wäre reellwertig und somit konstant.

Aus der Voraussetzung folgt

$$|f(\xi)| \leq |f(z)| \quad \text{für alle } \xi \in K(z, r),$$

aber *nicht*  $|f(\xi)| = |f(z)|$  für alle  $\xi \in K(z, r)$ , da sonst  $f$  konstant wäre. Damit ergibt sich aus der Mittelwerteigenschaft

$$\begin{aligned} |f(z)| &= \frac{1}{\pi r^2} \left| \int_{|\xi-z|\leq r} f(x+iy) d(x,y) \right| \leq \frac{1}{\pi r^2} \int_{|\xi-z|\leq r} |f(x+iy)| d(x,y) \\ &< \max\{|f(\xi)| : |\xi-z| \leq r\} = |f(z)|, \end{aligned}$$

ein Widerspruch.

b) Nehmen wir an, daß ein  $z_0 \in \Omega$  existiert mit

$$|f(z_0)| \geq |f(z)| \quad \text{für alle } z \in \Omega.$$

Da  $|f|$  nicht konstant ist, existiert ein  $z_1$  mit  $|f(z_1)| < |f(z_0)|$ . Sei nun  $\Gamma$  eine Kurve in  $\Omega$  mit

$$\gamma[0, 1] \rightarrow \Omega, \quad \gamma(0) = z_0, \quad \gamma(1) = z_1.$$

Dann existiert ein  $t_0 \in [0, 1)$  mit

$$|f(\gamma(t_0))| = |f(z_0)| \quad \text{und} \quad |f(\gamma(t))| < |f(z_0)| \quad \text{für } t \in (t_0, 1].$$

Sei nun  $r > 0$  so, daß  $K(\gamma(t_0), r) \subset \Omega$  gilt. Dann nimmt  $|f|$  in  $K(\gamma(t_0), r)$  sein Maximum im Zentrum an, obwohl  $f$  nicht konstant ist; Widerspruch zu Teil a. ■

**Korollar 10.5** Sei  $\Omega$  offen und zusammenhängend,  $f : \Omega \rightarrow \mathbb{C}$  holomorph und nicht konstant,  $f(z) \neq 0$  für alle  $z \in \Omega$ . Dann nimmt  $|f|$  in  $\Omega$  kein Minimum an.

*Beweis.* Maximumprinzip für  $\frac{1}{f}$ . ■

**Korollar 10.6** Sei  $\Omega$  offen,  $f : \Omega \rightarrow \mathbb{C}$  holomorph und  $\overline{K}(z_0, r) \subset \Omega$ . Dann nimmt  $|f|$  auf  $\overline{K}(z_0, r)$  sein Maximum nur auf dem Rand an. Ist  $f(z) \neq 0$  in  $K(z_0, r)$ , so nimmt  $|f|$  auch sein Minimum nur auf dem Rand an.

Es gibt auch eine der Cauchyschen Integralformel entsprechende Formel für die Ableitungen. Daraus ergibt sich dann *endlich*, daß jede holomorphe Funktion in Wirklichkeit beliebig oft stetig differenzierbar ist. Zunächst eine (leider etwas technische) Vorbereitung:

**Satz 10.7** Sei  $\Gamma$  eine rektifizierbare Kurve in  $\mathbb{C}$  (also insbesondere mit endlicher Länge),  $g : \Gamma \rightarrow \mathbb{C}$  stetig. Dann ist für jedes  $n \in \mathbb{N}_0$  die Funktion

$$f_n : \mathbb{C} \setminus \Gamma \rightarrow \mathbb{C}, \quad f_n(z) := \frac{n!}{2\pi i} \int_{\Gamma} \frac{g(\zeta)}{(\zeta - z)^{n+1}} d\zeta$$

holomorph, und es gilt

$$f'_n(z) = f_{n+1}(z) \quad \text{für alle } z \in \mathbb{C} \setminus \Gamma \quad \text{und } n \in \mathbb{N}_0.$$

Inbesondere sind alle  $f_n$  beliebig oft (stetig) differenzierbar.

*Beweis.* Für  $n \in \mathbb{N}_0$  und  $z \in \mathbb{C} \setminus \Gamma$  ist zu zeigen

$$\lim_{w \rightarrow z} \frac{f_n(w) - f_n(z)}{w - z} = f_{n+1}(z).$$

Zunächst gilt

$$\begin{aligned} \frac{1}{(\zeta - w)^{n+1}} - \frac{1}{(\zeta - z)^{n+1}} &= \frac{(\zeta - z)^{n+1} - (\zeta - w)^{n+1}}{(\zeta - w)^{n+1}(\zeta - z)^{n+1}} \\ &\quad \left( \text{mit } a^{n+1} - b^{n+1} = (a - b)(a^n + a^{n-1}b + \dots + b^n) \right) \\ &= \left\{ (\zeta - z) - (\zeta - w) \right\} \frac{(\zeta - z)^n + (\zeta - z)^{n-1}(\zeta - w) + \dots + (\zeta - w)^n}{(\zeta - w)^{n+1}(\zeta - z)^{n+1}} \\ &= (w - z) \frac{(\zeta - z)^n + (\zeta - z)^{n-1}(\zeta - w) + \dots + (\zeta - w)^n}{(\zeta - w)^{n+1}(\zeta - z)^{n+1}}, \end{aligned}$$

also

$$\frac{f_n(w) - f_n(z)}{w - z} = \frac{n!}{2\pi i} \int_{\Gamma} g(\zeta) \frac{(\zeta - z)^n + (\zeta - z)^{n-1}(\zeta - w) + \dots + (\zeta - w)^n}{(\zeta - w)^{n+1}(\zeta - z)^{n+1}} d\zeta.$$

Der Bruch unter dem Integral strebt für  $z \in \mathbb{C} \setminus \Gamma$  und  $w \rightarrow z$  gleichmäßig auf  $\Gamma$  gegen  $\frac{n+1}{(\zeta - z)^{n+1}}$ . Also strebt der Integrand (da  $g$  auf  $\Gamma$  beschränkt ist) gleichmäßig gegen  $(n+1)g(\zeta)/(\zeta - z)^{n+1}$ , d. h. es gilt

$$\frac{f_n(w) - f_n(z)}{w - z} \rightarrow \frac{(n+1)!}{2\pi i} \int_{\Gamma} \frac{g(\zeta)}{(\zeta - z)^{n+2}} d\zeta = f_{n+1}(z).$$

8 ■

## 21:6. Juli 2006

8

$$\begin{aligned} &\left| \frac{(\zeta - z)^n + (\zeta - z)^{n-1}(\zeta - w) + \dots + (\zeta - w)^n}{(\zeta - w)^{n+1}(\zeta - z)^{n+1}} - \frac{n+1}{(\zeta - z)^{n+2}} \right| \\ &\leq \sum_{j=0}^n \left| \frac{(\zeta - z)^{n-j}(\zeta - w)^j}{(\zeta - w)^{n+1}(\zeta - z)^{n+1}} - \frac{1}{(\zeta - z)^{n+2}} \right| \\ &= \sum_{j=0}^n \left| \frac{1}{(\zeta - w)^{n+1-j}(\zeta - z)^{j+1}} - \frac{1}{(\zeta - z)^{n+2}} \right| \\ &= \sum_{j=0}^n \left| \frac{(\zeta - z)^{n+1-j} - (\zeta - w)^{n+1-j}}{(\zeta - w)^{n+1-j}(\zeta - z)^{n+2}} \right| \\ &\quad \left( \text{mit } a^{k+1} - b^{k+1} = (a - b)(a^k + a^{k-1}b + \dots + b^k) \right) \\ &\leq |z - w| \sum_{j=0}^n \frac{|\zeta - z|^{n-j} + |\zeta - z|^{n-j-1}|\zeta - w| + \dots + |\zeta - w|^{n-j}}{|\zeta - w|^{n+1-j}|\zeta - z|^{n+2}}. \end{aligned}$$

Für (festes)  $z \in \mathbb{C} \setminus \Gamma$ ,  $w \in K(z, \varrho)$  mit  $K(z, 2\varrho) \cap \Gamma = \emptyset$  ist die Summe gleichmäßig beschränkt auf  $\gamma$ , also konvergiert der gesamte Ausdruck für  $z \rightarrow w$  gleichmäßig gegen 0.



**Satz 10.8 (Cauchysche Integralformel für Ableitungen)** Sei  $\Omega \subset \mathbb{C}$  offen,  $f : \Omega \rightarrow \mathbb{C}$  holomorph. Dann sind auch die Ableitungen  $f^{(1)} = f'$ ,  $f^{(2)} = f''$ ,  $f^{(3)}, \dots$  holomorph und es gilt

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_{\Gamma} \frac{f(\zeta)}{(\zeta - z)^{n+1}} d\zeta \quad (n \in \mathbb{N}_0)$$

für jede zu einem hinreichend kleinen Kreis  $\Gamma_r$  um  $z$  bezüglich  $\Omega \setminus \{z\}$  homotope Kurve  $\Gamma$ .

Der Beweis ergibt sich mit Hilfe der Cauchyschen Integralformel sofort aus dem vorhergehenden Satz.

**Satz 10.9 (Satz von Morera)** Ist  $\Omega \subset \mathbb{C}$  offen um  $f : \Omega \rightarrow \mathbb{C}$  stetig, so sind die folgenden Aussagen äquivalent:

- (i)  $f$  ist holomorph in  $\Omega$ ,
- (ii) das Kurvenintegral über jede bezüglich  $\Omega$  reelle nullhomotope Kurve verschwindet.

*Beweis.* Nur (ii)  $\Rightarrow$  (i) ist zu beweisen: Zu jedem  $z_0 \in \Omega$  gibt es ein  $\varrho > 0$  so, daß  $K(z_0, \varrho) \subset \Omega$  ist. Wir zeigen, daß  $f$  in  $K(z_0, \varrho)$  holomorph ist. Dazu definieren wir für  $z \in K(z_0, \varrho)$  mit einem beliebigen Weg von  $z_0$  nach  $z$

$$F(z) := \int_{z_0}^z f(\xi) d\xi.$$

$F$  ist wohldefiniert, und wie schon zwei mal gezeigt, ist  $F$  in  $K(z_0, \varrho)$  komplex differenzierbar. Also ist  $F'$ , und damit auch  $f = F'$  holomorph in  $K(z_0, \varrho)$ . ■

Sei  $\Omega \subset \mathbb{C}$  offen,  $f_n (n \in \mathbb{N})$  und  $f$  komplexe Funktionen auf  $\Omega$ . Man sagt  $(f_n)$  konvergiert lokal gleichmäßig gegen  $f$ , wenn  $(f_n)$  auf jeder kompakten Teilmenge von  $\Omega$  gleichmäßig gegen  $f$  konvergiert. Man spricht deshalb auch von kompakter Konvergenz.

**Satz 10.10**  $\Omega \subset \mathbb{C}$  offen,  $(f_n)$  eine Folge holomorpher Funktionen auf  $\Omega$ , die lokal gleichmäßig gegen  $f$  konvergiert. Dann ist auch  $f$  holomorph und es gilt  $f_n^{(k)} \rightarrow f^{(k)}$  ebenfalls lokal gleichmäßig.

*Beweis.* Für jede nullhomotope Kurve  $\Gamma$  in  $\Omega$  gilt

$$\int_{\Gamma} f(z) dz = \lim_{n \rightarrow \infty} \int_{\Gamma} f_n(z) dz = 0.$$

Nach dem Satz von Morera ist also  $f$  holomorph.

Bleibt zu zeigen, daß  $(f^{(k)})$  für jedes  $k$  in jeder kompakten Teilmenge  $K$  gleichmäßig gegen  $f^{(k)}$  konvergiert. Dazu genügt es zu zeigen, daß zu jedem  $x_0 \in \Omega$  ein  $r > 0$  existiert so, daß  $(f^{(k)})$  in

$K(z_0, r)$  gleichmäßig gegen  $f^{(k)}$  konvergiert. Wir wählen  $r > 0$  so, daß  $K(z_0, 2r)$  in  $\Omega$  liegt. Da  $f_n$  und  $f$  holomorph sind, gilt für  $z \in K(z_0, 2r)$  und den Kreisweg  $\Gamma_{z_0, 2r}$  um  $z_0$  mit Radius  $2r$

$$f_n^{(k)}(z) - f^{(k)}(z) = \frac{k!}{2\pi i} \int_{\Gamma_{z_0, 2r}} \frac{f_n(\zeta) - f(\zeta)}{(\zeta - z)^{k+1}} d\zeta.$$

Da  $(f_n)$  auf  $\Gamma_{z_0, 2r}$  gleichmäßig gegen  $f$  konvergiert, folgt  $f_n^{(k)} \rightarrow f^{(k)}$  gleichmäßig in  $K(z_0, r)$  für jedes  $k$ . ■

## 11 Taylorentwicklung holomorpher Funktionen

Wir erinnern an *Potenzreihen*: Eine Potenzreihen-Entwicklung um den *Entwicklungspunkt*  $z_0$  hat die Form

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n.$$

Sie konvergiert jedenfalls im Innern des *Konvergenzkreises* um  $z_0$  mit Radius

$$\varrho = \frac{1}{\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}} = \liminf_{n \rightarrow \infty} \frac{1}{\sqrt[n]{|a_n|}}.$$

Für  $z$  mit  $|z - z_0| = \varrho$  kann i. allg. keine Aussage gemacht werden (außer, wenn  $\varrho = 0$  ist: für  $z = z_0$  ist die Reihe trivialerweise konvergent).

Ist  $\tilde{\varrho} < \varrho$ , so ist die Reihe *gleichmäßig konvergent* in  $\overline{K}(z_0, \tilde{\varrho}) = \{z \in \mathbb{C} : |z - z_0| \leq \tilde{\varrho}\}$ .

**Satz 11.1 (Taylorentwicklung)** *Ist  $f$  holomorph in*

$$\{z \in \mathbb{C} : |z - z_0| < r\}$$

*so gibt es eine eindeutig bestimmte Potenzreihe  $\sum_{n=0}^{\infty} a_n (z - z_0)^n$  mit Konvergenzradius  $\varrho \geq r$  und*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \quad \text{für } |z - z_0| < r.$$

*Dabei ist*

$$a_n = \frac{1}{n!} f^{(n)}(z_0)$$

*und es gilt die Cauchysche Abschätzung für die Koeffizienten:*

$$|a_n| \leq \frac{1}{\eta^n} \max \left\{ |f(z)| : z \in \mathbb{C}, |z - z_0| = \eta \right\} \quad \text{für alle } \eta \in (0, r)$$

*(man beachte, daß dies eine ganze Schar von Abschätzungen ist, aus der man sich ggf. die beste auswählen kann).*

*Beweis.* Für  $\eta < r$  sei  $\Gamma_\eta$  die Kreislinie mit Radius  $\eta$  um  $z_0$ . Für  $\zeta \in \Gamma_\eta$  und  $z \in K(z_0, \eta)$  gilt

$$\left| \frac{z - z_0}{\zeta - z_0} \right| = \left| \frac{z - z_0}{\eta} \right| =: q_z < 1,$$

also (geometrische Reihe)

$$\frac{1}{\zeta - z} = \frac{1}{\zeta - z_0} \frac{1}{1 - \frac{z - z_0}{\zeta - z_0}} = \frac{1}{\zeta - z_0} \sum_{n=0}^{\infty} \left( \frac{z - z_0}{\zeta - z_0} \right)^n.$$

Bei festem  $z \in K(z_0, \eta)$  hat diese Reihe für alle  $\zeta \in \Gamma_\eta$  die Majorante  $\sum q_z^n$ , ist also gleichmäßig konvergent für  $\zeta \in \Gamma_\eta$ . Deshalb sind folgende Umformungen möglich:

$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \int_{\Gamma_\eta} \frac{f(\zeta)}{\zeta - z} d\zeta = \frac{1}{2\pi i} \int_{\Gamma_\eta} f(\zeta) \frac{1}{\zeta - z_0} \sum_{n=0}^{\infty} \left( \frac{z - z_0}{\zeta - z_0} \right)^n d\zeta \\ &= \frac{1}{2\pi i} \int_{\Gamma_\eta} f(\zeta) \frac{1}{\zeta - z_0} \lim_{m \rightarrow \infty} \sum_{n=0}^m \left( \frac{z - z_0}{\zeta - z_0} \right)^n d\zeta \\ &\quad \text{Integral und Limes vertauschbar wegen gleichmäßiger Konvergenz} \\ &= \frac{1}{2\pi i} \lim_{m \rightarrow \infty} \sum_{n=0}^m \int_{\Gamma_\eta} f(\zeta) \frac{1}{\zeta - z_0} \left( \frac{z - z_0}{\zeta - z_0} \right)^n d\zeta \\ &= \frac{1}{2\pi i} \sum_{n=0}^{\infty} (z - z_0)^n \int_{\Gamma_\eta} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} d\zeta =: \sum_{n=0}^{\infty} a_n (z - z_0)^n \end{aligned}$$

mit

$$a_n = \frac{1}{2\pi i} \int_{\Gamma_\eta} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} d\zeta = \frac{f^{(n)}(z_0)}{n!}.$$

Die Cauchysche Abschätzung folgt aus

$$a_n = \frac{1}{2\pi i} \int_{\Gamma_\eta} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} d\zeta,$$

die Eindeutigkeit aus

$$f^{(k)}(z_0) = \sum_{n=k}^{\infty} a_n n(n-1) \dots (n-k+1) (z - z_0)^{n-k} = k! a_k.$$

■

**Folgerung 11.2** Ist  $r$  der Konvergenzradius der Potenzreihe  $\sum_{n=0}^{\infty} a_n (x - x_0)^n$  und  $\sigma > r$ , so gibt es keine in  $K(z_0, \sigma)$  holomorphe Funktion, die in  $K(z_0, \varrho)$  mit dieser Reihe übereinstimmt. Mit anderen Worten: Es gibt mindestens einen Punkt auf dem Rand des Konvergenzkreises, über den hinaus die Funktion nicht holomorph fortgesetzt werden kann.

**Satz 11.3 (Identitätssatz)** Sei  $\Omega \subset \mathbb{C}$  offen und zusammenhängend,  $f : \Omega \rightarrow \mathbb{C}$  holomorph,  $(z_j)$  eine Folge aus  $\mathbb{C}$  mit Häufungspunkt  $z_0$  in  $\Omega$  (O. E.  $z_j \rightarrow z_0$ ) und  $f(z_j) = 0$  für alle  $j$ . Dann ist  $f(z) = 0$  für alle  $z \in \Omega$ .

*Beweis.* Ist  $r > 0$  so, daß  $K(z_0, r) \subset \Omega$  gilt, so läßt sich  $f$  in  $K(z_0, r)$  schreiben in der Form

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n.$$

Wir zeigen  $a_n = 0$  für alle  $n \in \mathbb{N}_0$  (also  $f(z) \equiv 0$  in  $K(z_0, r)$ ). Beweis durch Induktion:

$$\mathbf{n = 0:} \quad a_0 = f(z_0) = \lim_{j \rightarrow \infty} f(z_j) = 0.$$

*Induktionsannahme:*  $a_m = 0$  für  $m = 0, \dots, n-1$ , also

$$f(z) = \sum_{m=n}^{\infty} a_m (z - z_0)^m,$$

$$f_n(z) := \sum_{m=n}^{\infty} a_m (z - z_0)^{m-n} = \frac{f(z)}{(z - z_0)^n}.$$

Wegen  $f(z_j) = 0$  folgt daraus  $f_n(z_j) = 0$  (das gilt offensichtlich auch für  $z_j = z_0$ ). Damit folgt

$$a_n = f_n(z_0) = \lim_{i \rightarrow \infty} f_n(z_j) = 0.$$

Also ist  $f(z) \equiv 0$  für  $|z - z_0| < r$ .

Sei nun  $z \in \Omega$  beliebig,  $\Gamma$  eine Kurve von  $z_0$  nach  $z$ . Da es ein  $\sigma > 0$  gibt mit  $K(\zeta, \sigma) \subset \Omega$  für jedes  $\zeta \in \Gamma$ , läßt sich  $\Gamma$  durch endlich viele Kreise in  $\Omega$  um Punkte auf  $\Gamma$  überdecken, wobei der erste das Zentrum  $z_0$  hat und der  $(n+1)$ -te sein Zentrum im  $n$ -ten Kreis hat. Die obige Überlegung zeigt sukzessive, daß  $f$  in allen diesen Kreisen verschwindet, insbesondere ist  $f(z) = 0$ . ■

## 22: 11. Juli 2006

**Beispiel 11.4**  $f(z) = \frac{1}{1-z}$  ist im Kreis  $K(0, 1)$  holomorph,

$$f'(z) = \frac{1}{(1-z)^2}, \dots, f^n(z) = \frac{n!}{(1-z)^{n+1}}, \quad \text{also } \frac{f^{(w)}(0)}{w!} = 1.$$

Damit folgt

$$f(z) = \sum_{n=0}^{\infty} z^n \quad \text{für } z \in K(0, 1).$$

Der Konvergenzradius kann aber auch nicht größer sein, da  $f$  bei 1 nicht fortsetzbar ist. □

**Beispiel 11.5** Der Logarithmus  $\ln$  als Umkehrfunktion von

$$\exp : \{z \in \mathbb{C} : -\pi < \operatorname{Im} z < \pi\} \rightarrow \mathbb{C},$$

also

$$\ln z = \ln |z| + i \arg(z)$$

ist holomorph in  $K(1, 1)$ . Die Ableitungen im Punkt  $z = 1$  sind (natürlich wie im Reellen)

$$\ln(1) = 0, \quad \ln'(1) = 1, \quad \ln''(1) = -1, \dots, \quad \ln^{(n)}(1) = (-1)^{n+1} (n-1)!$$

Damit folgt

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

Schon aus dem Reellen ist klar, daß der Konvergenzradius 1 ist. □

Eine auf ganz  $\mathbb{C}$  holomorphe Funktion heißt auch *ganze Funktion* (oder *ganze analytische Funktion*). In diesem Fall ist also die Taylorreihe um jedes  $z_0 \in \mathbb{C}$  eine auf ganz  $\mathbb{C}$  konvergente Reihe.

**Satz 11.6** a) Sei  $f : \mathbb{C} \rightarrow \mathbb{C}$  eine ganze Funktion. Mit einem  $m \in \mathbb{N}_0$ , einem  $C \geq 0$  und einem  $r_0 > 0$  gelte

$$|f(z)| \leq C|z|^m \quad \text{für } |z| > r_0.$$

Dann ist  $f$  ein Polynom vom Grad  $\leq m$ .

b) Erster Liouvillescher Satz: Eine beschränkte ganze Funktion ist konstant.

*Beweis.* a) Für  $n > m$  liefert die Cauchysche Abschätzung der Koeffizienten der Taylorreihe um den Nullpunkt

$$|a_n| \leq \eta^{-n}(C\eta^m) \rightarrow 0 \quad \text{für } \eta \rightarrow \infty,$$

also  $a_n = 0$  für  $n > m$ .

b) Dies folgt aus Teil a mit  $m = 0$ . ■

Ist  $f$  bei  $z_0$  holomorph, so sagt man „ $f$  hat bei  $z_0$  eine  $m$ -fache Nullstelle“, wenn gilt  $f(z) = (z - z_0)^m g(z)$ , wobei  $g$  eine bei  $z_0$  holomorphe Funktion ist mit  $g(z_0) \neq 0$ . Die kann auch anders charakterisiert werden:

**Satz 11.7** Sei  $f$  bei  $z_0$  holomorph. Dann ist  $z_0$  genau dann eine  $m$ -fache Nullstelle von  $f$ , wenn  $f^{(n)}(z_0) = 0$  gilt für  $n = 0, \dots, m - 1$  und  $f^{(m)}(z_0) \neq 0$ .

*Beweis.* Dies kann unmittelbar der Taylorentwicklung um  $z_0$  entnommen werden. ■

**Satz 11.8 (Fundamentalsatz der Algebra)** Jedes Polynom  $p$  vom Grad  $m \geq 1$  hat (mit Vielfachheit gezählt) genau  $m$  Nullstellen.

*Beweis.* Hätte  $p$  keine Nullstelle, so wäre  $\frac{1}{p}$  eine beschränkte (Beweis!) ganze Funktion und somit konstant (also ein Polynom vom Grad 0). Also gibt es ein  $z_0$  mit  $p(z_0) = 0$ ,  $p(z) = (z - z_0)p_1(z)$ . Ist  $m \geq 2$ , so hat  $p_1$  wieder eine Nullstelle usw. ■

## 12 Laurententwicklung und Residuensatz

Eine *Laurent-Reihe* ist eine „Potenz“-Reihe der Form

$$\sum_{n=-\infty}^{\infty} a_n (z - z_0)^n.$$

Sie ist *konvergent* (oder: man bezeichnet sie als konvergent), wenn

$$\sum_{n=-\infty}^{-1} a_n (z - z_0)^n = \sum_{m=1}^{\infty} a_{-m} (z - z_0)^{-m}$$

und

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n$$

konvergieren. Die zweite Reihe konvergiert für  $z$  mit

$$|z - z_0| < \varrho_2 \quad \text{mit} \quad \varrho_2 = \frac{1}{\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}};$$

sie divergiert für  $|z - z_0| > \varrho_2$ . Die erste konvergiert für  $z$  mit

$$\frac{1}{|z - z_0|} < \tilde{\varrho}_1 = \frac{1}{\limsup_{n \rightarrow \infty} \sqrt[n]{|a_{-n}|}}$$

bzw.

$$|z - z_0| > \varrho_1 := \frac{1}{\tilde{\varrho}_1} = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_{-n}|};$$

sie divergiert für  $|z - z_0| < \varrho_1$ . Laurentreihen konvergieren also in einem Ring

$$\varrho_1 < |z - z_0| < \varrho_2 \quad (\text{falls } \varrho_1 < \varrho_2 \text{ gilt})$$

und divergieren außerhalb von  $\varrho_1 \leq |z - z_0| \leq \varrho_2$ . Sie konvergieren gleichmäßig in jedem kleineren Ring

$$\sigma_1 \leq |z - z_0| \leq \sigma_2 \quad \text{mit} \quad \varrho_1 < \sigma_1 \leq \sigma_2 < \varrho_2.$$

**Satz 12.1 (Laurent-Entwicklung)** *Sei  $f$  holomorph in  $\{z \in \mathbb{C} : r_1 < |z - z_0| < r_2\}$  (besonders wichtig ist der Fall  $r_1 = 0$ ). Dann gilt*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n \quad \text{für} \quad r_1 < |z - z_0| < r_2$$

mit

$$a_n = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(\zeta)}{(\zeta - z)^{n+1}} \quad (n \in \mathbb{Z}),$$

wobei  $\Gamma$  eine geschlossene Kurve zwischen den Kreisen mit Radien  $r_1$  und  $r_2$  ist, die  $z_0$  genau 1 mal in positiver Richtung umläuft, also z. B.

$$\Gamma = \Gamma_r := \{z = z_0 + re^{it} : 0 \leq t \leq 1\}, \quad r_1 < r < r_2.$$

Die Reihe konvergiert in jedem kleineren Kreisring  $\{z \in \mathbb{C} : \varrho_1 \leq |z - z_0| \leq \varrho_2\}$  mit  $r_1 < \varrho_1 \leq \varrho_2 < r_2$  gleichmäßig.

*Beweis.* Für  $z \in \mathbb{C}$  mit  $r_1 < |z - z_0| < r_2$  gibt es  $\varrho_1$  und  $\varrho_2$  mit

$$r_1 < \varrho_1 < |z - z_0| < \varrho_2 < r_2.$$

Und für hinreichend kleines  $\varrho > 0$  gilt mit  $\tilde{\Gamma} := \{\zeta = z + \varrho e^{it} : 0 \leq t \leq 2\pi\}$ ,  $\tilde{\Gamma} = \Gamma_{\varrho_2} - \Gamma_{\varrho_1}$  bezüglich der Menge  $\{z \in \mathbb{C} : r_1 < |z - z_0| < r_2\}$ , also

$$f(z) = \frac{1}{2\pi i} \int_{\tilde{\Gamma}} \frac{f(\zeta - z)}{\zeta - z} d\zeta = \frac{1}{2\pi i} \left\{ \int_{\Gamma_{\varrho_2}} - \int_{\Gamma_{\varrho_1}} \frac{f(\zeta)}{\zeta - z} d\zeta \right\}.$$

Für  $\zeta \in \Gamma_{\varrho_2}$  gilt

$$\left| \frac{z - z_0}{\zeta - z_0} \right| = \frac{|z - z_0|}{\varrho_2} = q_{z,2} < 1,$$

also (wie im Beweis des Taylorschen Satzes, Satz ??)

$$\frac{1}{\zeta - z} = \frac{1}{\zeta - z_0} \frac{1}{1 - \frac{z - z_0}{\zeta - z_0}} = \frac{1}{\zeta - z_0} \sum_{n=0}^{\infty} \left( \frac{z - z_0}{\zeta - z_0} \right)^n = \sum_{n=0}^{\infty} \frac{(z - z_0)^n}{(\zeta - z_0)^{n+1}},$$

wobei diese Reihe bei festem  $z$  gleichmäßig für  $\zeta \in \Gamma_2$  konvergiert. Damit folgt

$$\frac{1}{2\pi i} \int_{\Gamma_{\varrho_2}} \frac{f(\zeta)}{\zeta - z} d\zeta = \sum_{n=0}^{\infty} a_n (z - z_0)^n \quad \text{mit obigen } a_n \text{ für } n \geq 0.$$

Für  $\zeta \in \Gamma_{\varrho_1}$  gilt

$$\left| \frac{z - z_0}{\zeta - z_0} \right| = \frac{|z - z_0|}{\varrho_1} = \frac{1}{q_{z,1}} > 1,$$

also

$$\frac{1}{\zeta - z} = -\frac{1}{z - z_0} \frac{1}{1 - \frac{\zeta - z_0}{z - z_0}} = -\frac{1}{z - z_0} \sum_{n=0}^{\infty} \left( \frac{\zeta - z_0}{z - z_0} \right)^n = -\sum_{n=0}^{\infty} \frac{(\zeta - z_0)^n}{(z - z_0)^{n+1}},$$

wobei diese Reihe bei festem  $z$  gleichmäßig bezüglich  $\zeta \in \Gamma_1$  konvergiert. Damit folgt

$$\begin{aligned} \frac{1}{2\pi i} \int_{\Gamma_{\varrho_1}} \frac{f(\zeta)}{\zeta - z} d\zeta &= -\sum_{n=0}^{\infty} (z - z_0)^{-n-1} \frac{1}{2\pi i} \int_{\Gamma_{\varrho_1}} f(\zeta) (\zeta - z_0)^n d\zeta \\ &= -\sum_{m=-\infty}^{-1} a_m (z - z_0)^m \quad \text{mit obigen } a_m \text{ für } m < 0. \end{aligned}$$

Insgesamt folgt damit die Behauptung. ■

Sei  $\Omega = \{z \in \mathbb{C} : 0 < |z - z_0| < r\}$  und  $f : \Omega \rightarrow \mathbb{C}$  holomorph (i. allg. nicht definiert für  $z = z_0$ ). Dann heißt  $z_0$  eine *isolierte Singularität* von  $f$ . Entsprechend der (in diesem Fall existierenden) Laurentreihe unterscheidet man:

- $f$  hat bei  $z_0$  eine *hebbare Singularität*, wenn in der Laurentreihe  $a_n = 0$  gilt für  $n < 0$ . Die Laurentreihe ist also eine Taylorreihe;  $f$  ist in den Punkt  $x_0$  holomorph fortsetzbar (z. B.  $f(z) = \frac{z^2 + 1}{z - i} = \frac{(z + i)(z - i)}{z - i} = z + i$  für  $z \neq 0$ ).



- $f$  hat in  $z_0$  einen *Pol der Ordnung  $m$* , wenn  $a_{-m} \neq 0$  und  $a_n = 0$  für  $n < -m$  gilt (z. B.  $f(z) = (z - z_0)^{-m}$ ).
- $f$  hat in  $z_0$  eine *wesentliche Singularität*, wenn  $a_n \neq 0$  für unendlich viele negative  $n$  gilt (z. B.  $f(z) = \exp(1/z)$ ).

### 23: 13. Juli 2006

Der Koeffizient  $a_{-1}$ , für den gilt

$$a_{-1} = \frac{1}{2\pi i} \int_{\Gamma} f(\zeta) d\zeta$$

(wobei  $\Gamma$  eine geschlossene Kurve ist, die  $z_0$  genau 1 mal im positiven Sinn umläuft und in deren Inneren nur Punkte von  $\Omega$  liegen), heißt das *Residuum* von  $f$  bei  $z_0$ :  $\text{Res}(f, z_0)$ . Dieses wird für die explizite Berechnung von Integralen wichtig sein.

Wir untersuchen nun, wie das Verhalten der Funktion in der Umgebung einer isolierten Singularität mit den verschiedenen Singularitäten korrespondiert. Wir werden die drei Typen von Singularitäten durch das Verhalten der Funktion in ihrer Umgebung charakterisieren.

**Satz 12.2** *Hat  $f$  bei  $z_0$  einen Pol (der Ordnung  $m > 0$ ), so gilt  $f(z) \rightarrow \infty$  für  $z \rightarrow z_0$  (man beachte, daß  $f(z) \rightarrow \infty$  im Sinne der erweiterten komplexen Ebene bzw. der Riemannschen Zahlenkugel zu verstehen ist, d. h.  $|f(z)| \rightarrow \infty$  bzw.  $Pf(z) \rightarrow N$ ).*

*Beweis.* Wegen  $a_{-m} \neq 0$  gilt

$$\begin{aligned} f(z) &= \sum_{n=-m}^{\infty} a_n (z - z_0)^n = \frac{1}{(z - z_0)^m} \left\{ a_{-m} + \sum_{n=1-m}^{\infty} a_n (z - z_0)^{n+m} \right\} \\ &\rightarrow \infty \text{ für } z \rightarrow z_0, \end{aligned}$$

da die Summe  $\sum_{n=1-m}^{\infty} \dots$  für  $z \rightarrow z_0$  gegen 0 strebt, also der Ausdruck  $\{\dots\}$  gegen  $a_{-m} \neq 0$ . ■

**Satz 12.3** *Ist  $f$  holomorph bei  $z_0$ ,  $z_0$   $m$ -fache Nullstelle (d. h.  $f(z_0) = \dots = f^{(m-1)}(z_0) = 0$ ,  $f^{(m)}(z_0) \neq 0$ ), so existiert ein  $r > 0$  mit  $f(z) \neq 0$  für  $0 < |z - z_0| < r$  und  $g := 1/f$  holomorph in  $\{z \in \mathbb{C} : 0 < |z - z_0| < r\}$  mit einem Pol der Ordnung  $m$  bei  $z_0$ . Entsprechend gilt: Hat  $f$  bei  $z_0$  einen Pol der Ordnung  $m$ , so ist  $g := 1/f$  holomorph in einer Umgebung von  $z_0$  (ohne  $z_0$ ) und holomorph nach  $z_0$  fortsetzbar mit einer  $m$ -fachen Nullstelle  $z_0$ .*

*Beweis.* Aus der Voraussetzung folgt

$$\begin{aligned} f(z) &= \sum_{n=m}^{\infty} a_n (z - z_0)^n = (z - z_0)^m \sum_{k=0}^{\infty} a_{k+m} (z - z_0)^k \quad (a_m \neq 0) \\ &= (z - z_0)^m h(z) \end{aligned}$$

mit einer bei  $z_0$  holomorphen Funktion  $h$ ,  $h(z_0) \neq 0$ . Also existiert ein  $r > 0$  mit

$$h(z) \neq 0 \text{ für } |z - z_0| < r.$$

Daraus folgt, daß  $1/h$  bei  $z_0$  holomorph ist

$$\frac{1}{h(z)} = \sum_{h=0}^{\infty} b_n (z - z_0)^n \text{ für } |z - z_0| < r \text{ mit } b_0 \neq 0.$$

Also ist für  $z \neq z_0$

$$g(z) := \frac{1}{f(z)} = \frac{1}{(z - z_0)^m} \sum_{n=0}^{\infty} b_n (z - z_0)^n = \sum_{k=-m}^{\infty} b_{k+m} (z - z_0)^k,$$

d. h.  $g$  hat einen Pol der Ordnung  $m$  bei  $z_0$ .

Die Umkehrung kann dem Leser überlassen werden. ■

Das Residuum an einer Polstelle läßt sich leicht bestimmen:

**Satz 12.4** Sei  $f$  in  $K(z_0)$ ,  $r \setminus \{z_0\}$  holomorph mit einem Pol der Ordnung  $m$  bei  $z_0$ . Dann gilt

$$\text{Res}(f, z_0) = a_{-1} = \lim_{z \rightarrow z_0} \frac{1}{(m-1)!} \left\{ \frac{d^{m-1}}{dz^{m-1}} [f(z)(z - z_0)^m] \right\}.$$

Speziell für  $m = 1$  gilt

$$\text{Res}(f, z_0) = \lim_{z \rightarrow z_0} (z - z_0) f(z).$$

*Beweis.*  $m = 1$  (folgt auch als Spezialfall aus dem allgemeinen Fall): Wegen

$$f(z) = \sum_{n=-1}^{\infty} a_n (z - z_0)^n \quad (\text{mit } a_{-1} \neq 0)$$

folgt

$$(z - z_0) f(z) = \sum_{n=0}^{\infty} a_{n-1} (z - z_0)^n \rightarrow a_{-1} \text{ für } z \rightarrow z_0.$$

$m \geq 1$ : Jetzt gilt

$$\begin{aligned} f(z) &= \sum_{n=-m}^{\infty} a_n (z - z_0)^n, \\ (z - z_0)^m f(z) &= \sum_{n=0}^{\infty} a_{n-m} (z - z_0)^n, \\ \frac{d^{m-1}}{dz^{m-1}} [f(z)(z - z_0)^m] &= \sum_{n=m-1}^{\infty} a_{n-m} n(n-1) \dots (n-m+2) (z - z_0)^{n-m+1} \\ &\rightarrow a_{m-1-m} (m-1)! = a_{-1} (m-1)! \text{ für } z \rightarrow z_0. \end{aligned}$$

Das ist die Behauptung. ■

**Beispiel 12.5** Die Funktion  $\cot = \frac{1}{\tan} = \frac{\cos}{\sin}$  hat isolierte Singularitäten bei  $z_n = n\pi$  ( $n \in \mathbb{Z}$ ).  $\sin$  hat bei 0 eine einfache Nullstelle

$$\sin z = z - \frac{z^3}{3!} + \dots$$

Wegen der  $2\pi$ -Periodizität gilt dies auch bei  $z_n = n\pi$  ( $n \in \mathbb{Z}$ ). Da  $\cos$  dort holomorph ist mit  $\cos z_n \neq 0$ , hat  $\tan = \frac{\sin}{\cos}$  dort einfache Nullstellen,  $\cot$  einfache Pole. Es gilt

$$\operatorname{Res}(\cot, 0) = \lim_{z \rightarrow 0} (z - 0) \frac{\cos z}{\sin z} = \lim_{z \rightarrow 0} \frac{\cos z}{\frac{1}{z} \sin z} = \frac{\cos 0}{1} = 1.$$

Das Gleiche ergibt sich für die anderen Pole von  $\cot$ . □

**Satz 12.6 (Riemannscher Hebbarkeitssatz)** Sei  $f$  holomorph in  $K(z_0, r) \setminus \{z_0\}$  und beschränkt bei  $z_0$ .<sup>9</sup> Dann ist  $f$  holomorph nach  $z_0$  fortsetzbar, d. h. es gibt eine bei  $z_0$  holomorphe Funktion  $\tilde{f}$  mit  $f(z) = \tilde{f}(z)$  für  $z \in K(z_0, r) \setminus \{z_0\}$ ; die Laurentreihe von  $f$  um  $z_0$  ist also gleich der Taylorreihe von  $\tilde{f}$  um  $z_0$ . ( $z_0$  ist also eine hebbare Singularität; es ist eigentlich keine Singularität, sondern war nur eine „Definitionslücke“.)

*Beweis.*  $f$  besitzt eine Laurent-Entwicklung um  $z_0$

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} b_n (z - z_0)^{-n}.$$

Die erste Summe stellt eine bei  $z_0$  holomorphe Funktion dar. Die zweite Summe stellt eine in  $\mathbb{C} \setminus \{z_0\}$  holomorphe Funktion dar.

Wäre die Singularität nicht hebbar, so gäbe es (mindestens) ein  $n \in \mathbb{N}$  mit  $b_n \neq 0$ , und

$$g(w) := \sum_{n=1}^{\infty} b_n w^n$$

wäre eine ganze, nicht konstante, Funktion. also wäre  $g$  unbeschränkt, d. h. es gäbe eine Folge

$$(w_m) \text{ aus } \mathbb{C} \text{ mit } w_m \rightarrow \infty \text{ und } g(w_m) \rightarrow \infty.$$

Mit  $z_m = z_0 + 1/w_m$  würde also gelten

$$\begin{aligned} f(z_m) &= \sum_{n=0}^{\infty} a_n (z_m - z_0)^n + \sum_{n=1}^{\infty} b_n (z_m - z_0)^{-n} \\ &= \sum_{n=0}^{\infty} a_n (z_m - z_0)^n + g(w_m) \rightarrow \infty \text{ für } n \rightarrow 0, \end{aligned}$$

da die Summe  $\sum_{n=0}^{\infty} a_n (z - z_0)^n$  bei  $z_0$  holomorph (und somit beschränkt) ist. Dies ist aber ein Widerspruch zur Beschränktheit von  $f$  bei  $z_0$ . ■

<sup>9</sup>Das bedeutet, es gibt ein  $\varepsilon > 0$  und ein  $C \geq 0$  mit  $|f(z)| \leq C$  für alle  $z$  mit  $0 < |z - z_0| < \varepsilon$ .

**Satz 12.7 (Casorati–Weierstraß)** Sei  $f$  holomorph in  $K(r, z_0) \setminus \{z_0\}$ . Dann ist  $z_0$  genau dann eine wesentliche Singularität, wenn für jedes  $c \in \mathbb{C}$  eine Folge  $(z_n)$  mit  $z_n \rightarrow z_0$  existiert, für die gilt  $f(z_n) \rightarrow c$  (das heißt, die Bildmenge jedes – beliebig kleinen – Kreises um  $z_0$  ist dicht in  $\mathbb{C}$ ).

**Bemerkung 12.8** Ein Satz von PICARD besagt wesentlich mehr: Ist  $z_0$  eine wesentliche Singularität, so gilt: Für jedes  $\varepsilon > 0$  ist  $f(K(z_0, \varepsilon) \setminus \{z_0\})$  entweder gleich  $\mathbb{C}$  oder gleich  $\mathbb{C} \setminus \{w_0\}$  mit einem geeigneten  $w_0 \in \mathbb{C}$ . Beide Fälle treten tatsächlich auf:

- $f(z) = \exp(1/z)$  nimmt den Wert 0 nicht an,
- $f(z) = \cosh(1/z)$  nimmt in jeder Umgebung von 0 jeden Wert an. Dazu löst man die Gleichung

$$\frac{1}{2}(e^{1/z} + e^{-1/z}) = c \quad \text{bzw.} \quad \exp(2/z) - 2c \exp(1/z) + 1 = 0 \quad \text{nach}$$

$\exp(1/z)$  auf,

$$\exp(1/z) = c \pm \sqrt{c^2 - 1}, \quad z = \left\{ \ln(c \pm \sqrt{c^2 - 1}) + 2\pi i n \right\}^{-1} \quad (n \in \mathbb{Z}).$$

Es gibt also beliebig kleine  $z$  mit  $\cosh z = c$ .

*Beweis von Satz ??.*  $\Leftarrow$ : Ist diese Eigenschaft erfüllt, so ist die Singularität nicht hebbar und auch kein Pol, also wesentlich.

$\Rightarrow$ : Sei also  $z_0$  eine wesentliche Singularität. Wir nehmen an, daß ein  $c \in \mathbb{C}$  existiert, für das keine Folge  $(z_n)$  mit  $z_n \rightarrow z_0$  und  $f(z_n) \rightarrow c$  existiert. Dann existieren ein  $r > 0$  und ein  $\varepsilon > 0$  mit

$$|f(z) - c| > \varepsilon \quad \text{für } z \in K(z_0, r).$$

Also ist

$$g(z) := \frac{1}{f(z) - c} \quad \text{holomorph in } K(z_0, r) \setminus \{z_0\}$$

und beschränkt

$$|g(z)| \leq \frac{1}{\varepsilon}.$$

Also ist  $g$  holomorph nach  $z_0$  fortsetzbar mit  $|g(z_0)| \leq \frac{1}{\varepsilon}$ . Da  $f$  bei  $z_0$  unbeschränkt ist, muß  $g$  die Nullstelle  $z_0$  haben. Wegen  $g(z) \not\equiv 0$  ist  $z_0$  Nullstelle endlicher Vielfachheit, d. h.  $z_0$  ist Pol von  $\frac{1}{g} = f(z) - c$ , also auch von  $f = \frac{1}{g} + c$ , ein Widerspruch zur Voraussetzung. ■

Damit kann jetzt die Charakterisierung der Pole vervollständigt werden:

**Korollar 12.9**  $z_0$  ist genau dann ein Pol von  $f$  (mit Ordnung  $m \geq 1$ ), wenn gilt  $f(x) \rightarrow \infty$  für  $x \rightarrow x_0$ .

*Beweis.*  $\Rightarrow$ : Ist bereits bekannt.

$\Leftarrow$ : Die Singularität ist auf Grund der obigen Sätze nicht hebbar und nicht wesentlich, also ein Pol. ■

Eine ganze Funktion mit unendlicher Potenzreihe (d. h. unendlich viele  $a_n \neq 0$ ) heißt *ganz transzendent*.

**Folgerung 12.10** *Ist  $f$  ganz transzendent und  $c \in \mathbb{C}$  beliebig, so gibt es eine Folge  $(z_n)$  aus  $\mathbb{C}$  mit  $z_n \rightarrow \infty$  und  $f(z_n) \rightarrow c$ . (Man sagt auch:  $f$  hat bei  $\infty$  eine wesentliche Singularität.)*

*Beweis.*  $f\left(\frac{1}{z}\right)$  hat bei 0 eine wesentliche Singularität. ■

**Beispiel 12.11**  $f(z) = z + e^z$  hat offenbar bei  $\infty$  eine wesentliche Singularität. Es gilt aber für jede Zahl  $w = a + ib \in \mathbb{C} \setminus \{0\}$

$$f(rw) \rightarrow \infty \quad \text{für } r \rightarrow \infty \quad (r \in \mathbb{R}).$$

Aus  $|f(rw)| = |ra + rib + e^{ra} e^{rib}|$  folgt:

- Ist  $a > 0$ , so ist  $|e^{ra} e^{rib}| = e^{ra}$  für große  $r$  wesentlich größer als  $ra + rib$ , also  $f(rw) \rightarrow \infty$  für  $r \rightarrow \infty$ .
- Ist  $a \leq 0$ , so gilt  $|ra + rib| = r\sqrt{a^2 + b^2} \rightarrow \infty$ ,  $|e^{ra} e^{rib}| \rightarrow 0$  für  $r \rightarrow \infty$ .

□

**Satz 12.12 (Residuensatz)**  $\Omega \subset \mathbb{C}$  offen,  $f$  in  $\Omega$  holomorph bis auf isolierte Singularitäten  $\{z_j : j \in I\}$  (die sich natürlich nicht in  $\Omega$  häufen können),  $\Gamma$  eine positiv orientierte geschlossene Kurve in  $\Omega \setminus \{z_j : j \in I\}$ , deren Inneres  $O$  ganz in  $\Omega$  liegt. Dann gilt

$$\int_{\Gamma} f(z) dz = 2\pi i \sum_{z_j \in O} \text{Res}(f, z_j).$$

*Beweis.*  $\Gamma$  ist bezüglich  $\Omega \setminus \{z_j : j \in I\}$  homotop zur Summe (hinreichend) kleiner positiv orientierter Kreise um die  $z_j \in O$ . Damit folgt die Behauptung aus dem Satz über die Laurent-Entwicklung. ■

Eine typische Anwendung des Residuensatzes sieht wie folgt aus:

**Beispiel 12.13** Es soll das uneigentliche Integral

$$\int_{-\infty}^{\infty} \frac{x-1}{(x^2+1)(x+1-i)} dx$$

berechnet werden. Das Integral existiert jedenfalls, da sich der Integrand für große  $|x|$  wie  $|x|^{-2}$  verhält.

Natürlich könnte man auch eine Partialbruchzerlegung versuchen, damit eine Stammfunktion finden, und so das Integral auswerten. Mit Hilfe des Residuensatzes geht das einfacher.

Die Nullstellen des Nenners sind alle einfach:  $i$ ,  $-i$  und  $i-1$ . Keine dieser Nullstellen ist Nullstelle des Zählers, d. h. der Integrand  $f$  hat bei  $i$ ,  $-i$  und  $1-i$  jeweils einen Pol der Ordnung 1. Ist  $r$  so groß, daß  $K(0, r)$  alle Pole enthält, und  $\Gamma_r$  der Halbkreis in der oberen Halbebene von  $+r$  nach  $-r$  plus die Strecke von  $-r$  nach  $r$ , so gilt

$$\int_{\Gamma_r} \frac{z-1}{(z+i)(z-i)(z+1-i)} dz = 2\pi i \left\{ \operatorname{Res}(f, i) + \operatorname{Res}(f, i-1) \right\}.$$

Die Residuen sind

$$\begin{aligned} \operatorname{Res}(f, i) &= \lim_{z \rightarrow i} \frac{z-1}{(z+i)(z+1-i)} = \frac{i-1}{2i \cdot 1} = \frac{1}{2}(1+i), \\ \operatorname{Res}(f, i-1) &= \lim_{z \rightarrow i-1} \frac{z-1}{(z+i)(z-i)} = \frac{i-2}{(2i-1)(-1)} = \frac{i-2}{1-2i} \\ &= \frac{(i-2)(1+2i)}{5} = \frac{-4-3i}{5}. \end{aligned}$$

Da das Integral über den Kreisbogen für  $r \rightarrow \infty$  gegen 0 strebt (Länge des Weges  $\pi r$ , Integrand  $\sim r^{-2}$ ) folgt

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{x-1}{(x^2+1)(x+1-i)} dx &= \lim_{r \rightarrow \infty} \int_{\Gamma_r} \frac{z-1}{(z^2+1)(z+1-i)} dz \\ &= 2\pi i \left\{ \frac{1+i}{2} + \frac{-4-3i}{5} \right\} = 2\pi i \frac{5+5i-8-6i}{10} \\ &= \pi \frac{1-3i}{5}. \end{aligned}$$

Natürlich hätte man das Ganze bequemer mit dem Halbkreisbogen in der unteren Halbebene machen können, da dort nur ein Residuum liegt.

$$\begin{aligned} \operatorname{Res}(f, -i) &= \lim_{z \rightarrow -i} \frac{z-1}{(z-i)(z+1-i)} = \frac{-i-1}{-2i(-2i+1)} = \frac{1+i}{2i(1-2i)} \\ &= \frac{(1+i)(1+2i)}{2i \cdot 5} = \frac{-1+3i}{10i} = \frac{3+i}{10}, \end{aligned}$$

$$\int_{-\infty}^{\infty} \frac{x-1}{(x^2+1)(x+1-i)} dx = -2\pi i \operatorname{Res}(f, -i) = \pi \frac{1-3i}{5} \quad \square$$

**Beispiel 12.14** Das uneigentliche Integral  $\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx$  ist leicht mit Hilfe reeller Analysis auszurechnen:

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = \lim_{A \rightarrow \infty} \int_{-A}^A \frac{1}{1+x^2} dx = \lim_{A \rightarrow \infty} \arctan x \Big|_{-A}^A = \pi.$$

Aber auch dieses Integral ist leicht mit dem Residuensatz zu bestimmen. Das kann dem Leser als Übung überlassen werden.  $\square$

<sup>10</sup>woher kommt das Minuszeichen nach dem ersten Gleichheitszeichen in dieser Zeile?

**Beispiel 12.15** Man bestimme das uneigentliche Integral

$$\int_{-\infty}^{\infty} \frac{e^{ix}}{1+x^2} dx.$$

Es existiert offenbar (warum?). Für  $\text{Im } z > 0$  gilt

$$\left| \frac{e^{iz}}{1+z^2} \right| \leq \frac{1}{|1+z^2|} \sim \frac{1}{|z|^2},$$

d. h. das Integral über den Halbkreisbogen in der oberen Halbebene von  $R$  nach  $-R$  geht für  $R \rightarrow \infty$  gegen 0. Also gilt, da  $e^{iz}/(1+z^2)$  in der oberen Halbebene nur die Singularität  $i$  hat,

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{e^{ix}}{1+x^2} dx &= 2\pi i \text{Res}\left(\frac{e^{iz}}{1+z^2}, i\right) = 2\pi i \lim_{z \rightarrow i} \frac{e^{iz}}{z+i} \\ &= 2\pi i \frac{e^{-1}}{2i} = \frac{\pi}{e}. \end{aligned}$$

Den Bogen in der unteren Halbebene kann man nicht benutzen, da dort der Integrand nicht gegen 0 geht.  $\square$

Zum Abschluß ein interessantes Resultat zur Abzählung von Nullstellen und Polstellen:

Sei  $\Omega \subset \mathbb{C}$  offen.  $f$  heißt *meromorph* in  $\Omega$ , wenn es bis auf Pole, die sich in  $\Omega$  nicht häufen, holomorph ist.

Eine in  $\Omega$  meromorphe Funktion hat also in jeder kompakten Teilmenge  $K$  von  $\Omega$  nur endlich viele Pole  $z_1, \dots, z_n$  mit Ordnungen  $p_1, \dots, p_n$  ( $n = n(K)$ ). Also ist

$$g(z) := f(z) \prod_{j=1}^n (z - z_j)^{p_j}$$

holomorph auf  $K$  fortsetzbar; d. h.  $f$  ist in  $K$  Quotient einer holomorphen Funktion  $g$  und eines Polynoms  $p(z) = \prod_{j=1}^n (z - z_j)^{p_j}$ . (Mit Hilfe des Weierstraßschen Produktsatzes kann man sogar zeigen: eine Funktion ist genau dann meromorph, wenn sie Quotient zweier holomorpher Funktionen ist.)

**Satz 12.16 (Null- und Polstellen zählendes Integral)** Sei  $\Omega \subset \mathbb{C}$  offen,  $f$  meromorph in  $\Omega$ ,  $\Gamma$  eine doppelpunktfreie positiv orientierte geschlossene Kurve in  $\Omega$ , die keine Nullstelle und keinen Pol von  $f$  trifft und deren Inneres ganz in  $\Omega$  liegt. Dann gilt

$$\frac{1}{2\pi i} \int_{\Gamma} \frac{f'(z)}{f(z)} dz = N - P,$$

mit

$N =$  Zahl der mit Vielfachheit gezählten Nullstellen von  $f$  im Inneren von  $\Gamma$ ,

$P =$  Zahl der mit Ordnung gezählten Postellen von  $f$  im Inneren von  $\Gamma$ .

*Beweis.*  $\Gamma$  ist homotop zur Summe kleiner Kreise um die Nullstellen und Polstellen im Innern von  $\Gamma$ . Es genügt also, die Formel für einen solchen Kreis zu beweisen:

(i)  $z_0$  Nullstelle mit Vielfachheit  $n$ , also

$$f(z) = (z - z_0)^n f_0(z) \quad \text{mit } f_0(z_0) \neq 0,$$

$$f'(z) = n(z - z_0)^{n-1} f_0(z) + (z - z_0)^n f_0'(z),$$

$$\begin{aligned} \frac{1}{2\pi i} \int_{\Gamma} \frac{f'(z)}{f(z)} dz &= \frac{1}{2\pi i} \int_{\Gamma} \left\{ \frac{n}{z - z_0} + \frac{f_0'(z)}{f_0(z)} \right\} dz \\ &\quad \text{(zweiter Summand holomorph nahe } z_0) \\ &= \frac{1}{2\pi i} \int_{\Gamma} \frac{n}{z - z_0} dz = \frac{1}{2\pi i} n 2\pi i = n. \end{aligned}$$

(ii)  $z_0$  Polstelle der Ordnung  $p$ , also

$$f(z) = (z - z_0)^{-p} f_0(z) \quad \text{mit } f_0(z_0) \neq 0.$$

Die gleiche Rechnung wie oben liefert

$$\frac{1}{2\pi i} \int_{\Gamma} \frac{f'(z)}{f(z)} dz = -p.$$

■

**Satz 12.17 (Rouché)** Sei  $\Omega \subset \mathbb{C}$  offen,  $f : \Omega \rightarrow \mathbb{C}$  holomorph,  $\Gamma$  eine doppelunktfreie geschlossene Kurve, deren Inneres ganz in  $\Omega$  liegt,  $|g(z)| < |f(z)|$  auf  $\Gamma$ . Dann haben  $f$  und  $f + g$  gleich viele Nullstellen (mit Vielfachheit gezählt) im Inneren von  $\Gamma$ . (Die Voraussetzung  $|g(z)| < |f(z)|$  auf  $\Gamma$  kann ersetzt werden durch  $f(z) + \lambda g(z) \neq 0$  auf  $\Gamma$  für alle  $\lambda \in [0, 1]$ .)

*Beweis.* Für  $\lambda \in [0, 1]$  sei  $N_\lambda$  die Zahl der Nullstellen (mit Vielfachheit) von  $f + \lambda g$  im Inneren von  $\Gamma$ . Dann gilt nach obigem Satz (da  $P_\lambda = 0$  ist)

$$N_\lambda = \frac{1}{2\pi i} \int_{\Gamma} \frac{f'(z) + \lambda g'(z)}{f(z) + \lambda g(z)} dz.$$

Also ist  $\lambda \mapsto N_\lambda$  stetig auf  $[0, 1]$ . Da es nur ganzzahlige Werte annimmt, ist  $N_\lambda$  konstant, also  $N_1 = N_0$ . ■

Damit kann man einen sehr einfachen *Beweis des Fundamentalsatzes der Algebra* angeben: Sei

$$p(z) = z^n + \sum_{j=0}^{n-1} a_j z^j \quad (\text{o. E. } a_n = 1).$$



Die Funktion  $f(z) = z^n$  hat die  $n$ -fache Nullstelle  $0$  (und sonst keine Nullstelle). Mit

$$g(z) := \sum_{j=0}^{n-1} a_j z^j$$

gilt für hinreichend große  $z$

$$|f(z)| = |z|^n > \left| \sum_{j=0}^{n-1} a_j z^j \right| = |g(z)|.$$

Also haben  $f$  und  $p = f + g$  für große  $r$  gleich viele Nullstellen (mit Vielfachheit gezählt) in  $K(0, r)$ . Da das für alle großen  $r$  gilt, hat  $f$  insgesamt genau  $n$  Nullstellen in  $\mathbb{C}$ .  $\square$

## Inhaltsverzeichnis

<b>1</b>	<b>Lineare Gleichungssysteme und Determinanten</b>	<b>2</b>
<b>2</b>	<b>Eigenwerte und Eigenvektoren</b>	<b>15</b>
<b>3</b>	<b>Stetige Funktionen mehrerer Variablen</b>	<b>26</b>
<b>4</b>	<b>Differenzierbare Funktionen mehrerer Variablen</b>	<b>32</b>
<b>5</b>	<b>Höhere Ableitungen, der Satz von Taylor, Extrema</b>	<b>43</b>
<b>6</b>	<b>Kurven, Kurvenlänge und Kurvenintegrale</b>	<b>50</b>
6.1	Parameterdarstellung . . . . .	50
6.2	Rektifizierbare Kurven, Kurvenlänge . . . . .	52
6.3	Kurvenintegrale . . . . .	54
<b>7</b>	<b>Umkehrfunktion und implizit definierte Funktion</b>	<b>58</b>
7.1	Umkehrfunktion . . . . .	58
7.2	Implizit definierte Funktionen . . . . .	62
7.3	Bedingte Extrema . . . . .	67
<b>8</b>	<b>Holomorphe Funktionen</b>	<b>69</b>
8.1	Möbiusfunktionen . . . . .	74
<b>9</b>	<b>Kurvenintegrale und Cauchyscher Integralsatz</b>	<b>77</b>
<b>10</b>	<b>Cauchysche Integralformel</b>	<b>85</b>
<b>11</b>	<b>Taylorentwicklung holomorpher Funktionen</b>	<b>91</b>
<b>12</b>	<b>Laurententwicklung und Residuensatz</b>	<b>95</b>