

**A. INTEGRATION**

**A1. Einleitung**

In diesem Semester fangen wir mit Integration an. Es gibt viele Möglichkeiten, das Integral einer Funktion genau zu definieren; wir benutzen das sogenannte Regelintegral.

Danach definieren wir metrische Räume. Wichtige Beispiele davon sind die komplexe Zahlen  $\mathbb{C}$  und der euklidische Raum  $\mathbb{R}^n$ . Wir übertragen die topologische Begriffe (Offenheit, Stetigkeit, Kompaktheit, Konvergenz, usw.) von  $\mathbb{R}$  auf beliebige metrische Räume.

Mit diesen Grundlagen können wir Analysis auch mehrdimensional machen. Insbesondere werden wir Ableitungen von Funktionen  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  untersuchen und die wichtigsten Sätze dazu (Umkehrsatz, Rangatz) beweisen.

**A2. Treppenfunktionen und Regelfunktionen**

Wir betrachten ein kompaktes Intervall  $[a, b] \subset \mathbb{R}$  und Funktionen  $f: [a, b] \rightarrow \mathbb{R}$ .

Wir wollen den Flächeninhalt zwischen dem Graphen der Funktion  $f$  und der Achse messen; wir nennen diesen das *Integral*

$$\int_a^b f = \int_a^b f(x) dx.$$

(Wenn  $f$  negative Werte annimmt wird hierbei vereinbart, dass Flächen unterhalb der Achse negativ berechnet werden.) Offensichtlich gilt im Falle einer konstanten Funktion

$$\int_a^b c = c(b - a).$$

Es muss nicht sein, dass  $f$  stetig ist. Hingegen sind es die Treppenfunktionen, die man am einfachsten integrieren kann.

**Definition A2.1.** Eine Zerlegung  $Z$  von  $[a, b]$  in  $k$  Teile ( $k \in \mathbb{N}^+$ ) ist durch

$$a = z_0 < z_1 < \dots < z_{k-1} < z_k = b$$

gegeben. Für  $1 \leq j \leq k$  schreiben wir  $\Delta z_j = z_j - z_{j-1}$ ; wir nennen  $\Delta(Z) := \max_j \Delta z_j$  die *Maschenweite* von  $Z$ . Eine *markierte Zerlegung*  $\tilde{Z}$  ist eine Zerlegung  $Z$  wo man zusätzlich für jedes  $1 \leq j \leq k$  einen Punkt  $\xi_j \in (z_{j-1}, z_j)$  ausgewählt hat.

**Definition A2.2.** Eine Funktion  $f: [a, b] \rightarrow \mathbb{R}$  heißt dann *stückweise konstant* bzw. *stückweise stetig* bzw. *stückweise monoton*, falls es eine Zerlegung  $Z$  von  $[a, b]$  und konstante bzw. stetige bzw. monotone Funktionen  $f_j: [z_{j-1}, z_j] \rightarrow \mathbb{R}$  so gibt, dass  $f_j(x) = f(x)$  für alle  $x \in (z_{j-1}, z_j)$ .

Eine stückweise konstante Funktion  $f$  heißt auch *Treppenfunktion*; eine Zerlegung, die zeigt, dass  $f$  eine Treppenfunktion ist, heißt eine *Treppenzerlegung* für  $f$ .

*Bemerkung A2.3.* Es folgt, dass  $f$  selbst auf jedem offenen Intervall  $(z_{j-1}, z_j)$  konstant bzw. stetig bzw. monoton ist. Die Werte  $f(z_j)$  hingegen dürfen beliebig sein.

*Bemerkung A2.4.* Eine Funktion  $f$  ist genau dann stückweise stetig, wenn es eine Zerlegung so gibt, dass für jedes  $j$  die Einschränkung  $f|_{(z_{j-1}, z_j)}$  eine stetige Erweiterung auf  $[z_{j-1}, z_j]$  hat.

*Beispiel A2.5.* Seien  $f, g: [-1, 1] \rightarrow \mathbb{R}$  wie folgt definiert:  $f(x) := \sin 1/x$  und  $g(x) := 1/\sin x$  für  $x \neq 0$  und  $f(0) := g(0) := 0$ . Weder  $f$  noch  $g$  ist stückweise stetig. Betrachten wir die Zerlegung

$$-1 = z_0 < 0 = z_1 < 1 = z_2,$$

sind die Einschränkungen  $f|_{(z_{j-1}, z_j)}$  und  $g|_{(z_{j-1}, z_j)}$  jeweils stetig, erlauben aber keine stetige Erweiterungen in 0.

**Lemma A2.6.** Eine Funktion  $f$ , die stückweise konstant oder stetig oder monoton ist, ist beschränkt.

*Beweis.* Jede Funktion  $f_j$  (in der Definition) ist beschränkt, weil sie konstant bzw. stetig bzw. monoton auf dem kompakten Intervall  $[z_{j-1}, z_j]$  ist. Schreiben wir die Schranke als  $|f_j| \leq C_j$ , dann gilt

$$|f(x)| \leq \max\left\{ \max_j C_j, \max_j |f(z_j)| \right\}. \quad \square$$

*Bemerkung A2.7.* Eine stückweise stetige Funktion  $f$  hat in jedem Punkt  $p \in [a, b]$  einseitige Grenzwerte

$$f(p^\pm) := \lim_{x \rightarrow p^\pm} f(x).$$

In den Punkten  $z_j$  dürfen die Werte  $f(z_j^-)$ ,  $f(z_j)$  und  $f(z_j^+)$  alle unterschiedlich sein.

**Definition A2.8.** Eine *Regelfunktion*  $f: [a, b] \rightarrow \mathbb{R}$  ist eine Funktion, von der alle möglichen einseitigen Grenzwerte existieren. D.h., für jedes  $p \in [a, b]$  existiert  $f(p^+)$  und für jedes  $p \in (a, b]$  existiert  $f(p^-)$ .

*Beispiel A2.9.* Stückweise stetige Funktionen (insbesondere Treppenfunktionen) sind Regelfunktionen.

**Lemma A2.10.** Eine Regelfunktion ist beschränkt.

*Beweis.* (Aufgabe.) □

*Beispiel A2.11.* Sei  $W: [0, 1] \rightarrow \mathbb{R}$  die Funktion

$$W(x) = \begin{cases} 1/q, & x = p/q \in \mathbb{Q} \text{ mit } p, q \text{ teilerfremd,} \\ 0, & x \notin \mathbb{Q}, \end{cases}$$

die wir schon im letzten Semester gesehen haben. Dann ist  $W$  eine Regelfunktion, weil in jedem Punkt die einseitigen Grenzwerte 0 sind. (Hingegen ist  $\chi_{\mathbb{Q}}: x \rightarrow \text{sgn } W(x)$  keine Regelfunktion.)

**Definition A2.12.** Wir schreiben  $\mathcal{R}_{[a,b]}$  für die Menge aller Regelfunktionen auf  $[a, b]$  und  $\mathcal{T}_{[a,b]}$  für die Teilmenge aller Treppenfunktionen.

**Lemma A2.13.** Seien  $f, g: [a, b] \rightarrow \mathbb{R}$  Treppen- bzw. Regelfunktionen und sei  $c \in \mathbb{R}$ . Dann ist auch  $f + cg$  eine Treppen- bzw. Regelfunktion. D.h.,  $\mathcal{R}_{[a,b]}$  und  $\mathcal{T}_{[a,b]}$  sind Vektorräume.

*Beweis.* (Aufgabe.) □

**A3. Konvergenz von Funktionen**

Wir betrachten nach wie vor Funktionen auf einem kompakten Intervall  $[a, b] \subset \mathbb{R}$ .

**Definition A3.1.** Für jedes  $k \in \mathbb{N}$  sei  $f_k: [a, b] \rightarrow \mathbb{R}$  eine Funktion. Falls für jedes  $x \in [a, b]$  die Folge  $(f_k(x))$  einen Grenzwert in  $\mathbb{R}$  hat, definieren wir  $f: [a, b] \rightarrow \mathbb{R}$  durch

$$f(x) := \lim_{k \rightarrow \infty} f_k(x).$$

Wir sagen, die Folge  $(f_k)$  von Funktionen *konvergiert punktweise* gegen den Limes  $f$ .

*Beispiel A3.2.* Für  $k \in \mathbb{N}$  sei  $f_k$  die Potenzfunktion  $f_k: x \mapsto x^k$  auf  $[0, 1]$ . Dann konvergiert  $(f_k)$  punktweise gegen die Funktion  $f = \chi_{\{1\}}$  mit  $f(1) = 1$  und  $f(x) = 0$  für  $x < 1$ .

**Definition A3.3.** Seien  $f$  und  $f_k$  ( $k \in \mathbb{N}$ ) Funktionen auf  $[a, b]$ . Für jedes  $k$  setzen wir

$$\delta_k := \sup_{x \in [a, b]} |f(x) - f_k(x)| \in [0, +\infty).$$

Falls  $\delta_k \rightarrow 0$ , sagen wir,  $(f_k)$  *konvergiert gleichmäßig* gegen  $f$ .

*Bemerkung A3.4.* Eine Folge, die gleichmäßig konvergiert, konvergiert auch punktweise aber nicht umgekehrt. (Im letzten Beispiel gilt  $\delta_k = 1$  für alle  $k$ , d.h.  $(f_k)$  konvergiert nicht gleichmäßig.)

---

Ende der Vorlesung 2008 Oktober 13

---

**Satz A3.5.** Eine Funktion  $f: [a, b] \rightarrow \mathbb{R}$  ist genau dann eine Regelfunktion, wenn es eine Folge  $(\varphi_k)$  von Treppenfunktionen gibt, die gegen  $f$  gleichmäßig konvergiert.

*Bemerkung A3.6.* Falls  $\varphi_k$  eine Treppenfunktion ist und  $\sup |f - \varphi_k| < \varepsilon$ , sagen wir,  $\varphi_k$  ist eine  $\varepsilon$ -Treppenapproximation für  $f$ . Der Satz sagt, die Regelfunktionen sind genau die, die für jedes  $\varepsilon > 0$  eine  $\varepsilon$ -Treppenapproximation haben.

Für den Beweis brauchen wir zwei Lemmata.

**Lemma A3.7.** Sei  $p \in [a, b]$  und sei  $f: [a, b] \rightarrow \mathbb{R}$  eine Funktion, deren einseitige Grenzwerte in  $p$  existieren. Zu jedem  $\varepsilon > 0$  gibt es dann  $\delta > 0$ , so dass  $f|_{I_\delta}$  eine  $\varepsilon$ -Treppenapproximation hat, wo  $I_\delta$  das Intervall  $I_\delta := [a, b] \cap [p - \delta, p + \delta]$  um  $p$  bedeutet.

*Beweis.* Wir betrachten den Fall  $p \in (a, b)$ . (Falls  $p$  ein Endpunkt ist, wird alles nur einfacher, weil wir nur eine Seite betrachten müssen.) Seien  $f(p^-)$  und  $f(p^+)$  die beiden Grenzwerte und sei  $\varphi$  die Treppenfunktion

$$\varphi(x) = \begin{cases} f(p^-), & x < p, \\ f(p), & x = p, \\ f(p^+), & x > p. \end{cases}$$

Existenz der einseitigen Grenzwerte heißt, dass zu jedem  $\varepsilon > 0$  es  $\delta^\pm > 0$  so gibt, dass auf  $[p - \delta^-, p]$  bzw. auf  $[p, p + \delta^+]$  gilt  $|f(x) - f(p^\pm)| < \varepsilon$ . Mit  $\delta = \min(\delta^-, \delta^+)$  folgt das Lemma.  $\square$

**Lemma A3.8.** Sei  $f$  eine Funktion auf  $[a, b]$ . Sie hat in  $b$  genau dann einen linksseitigen Grenzwert  $f(b^-)$ , wenn zu jedem  $\varepsilon > 0$  es ein  $\delta > 0$  so gibt, dass  $|f(x) - f(y)| < \varepsilon$  für alle  $x, y \in (b - \delta, b)$ . (Ähnliches gilt für rechtsseitige Grenzwerte.)

*Beweis.* Eine Richtung ist einfach: falls der Grenzwert  $f(b^-)$  existiert, gibt es  $\delta > 0$ , so dass  $|f(x) - f(b^-)| < \varepsilon/2$  für alle  $x \in (b - \delta, b)$ . Dann folgt  $|f(x) - f(y)| < \varepsilon$  aus der Dreiecksungleichung.

Umgekehrt habe wir – zu jedem  $\varepsilon > 0$  – ein  $\delta > 0$ , so dass  $|f(x) - f(y)| < \varepsilon$  für alle  $x, y \in (b - \delta, b)$ . Aus dem letzten Semester (Lemma I.D3.8, leicht an einseitige Grenzwerte angepasst) wissen wir, dass  $f(b^-)$  genau dann existiert, wenn für jede Folge  $x_n \rightarrow b$  mit  $x_n < b$  die Folge  $f(x_n)$  auch konvergiert. Weil  $x_n \rightarrow b$  gibt es  $n_0$ , so dass  $x_n \in (b - \delta, b)$  für alle  $n \geq n_0$ . Dann gilt  $|f(x_n) - f(x_m)| < \varepsilon$  für alle  $n, m \geq n_0$ . Weil  $\varepsilon > 0$  beliebig war, ist  $f(x_n)$  eine Cauchyfolge und hat deshalb einen Grenzwert.  $\square$

*Beweis des Satzes.* Sei  $(\varphi_k)$  eine Folge von Treppenfunktionen, die gegen  $f$  gleichmäßig konvergiert. Sei  $p \in (a, b)$ . Wir wollen zeigen, dass der Grenzwert  $f(p^-)$  existiert. Dazu sei  $\varepsilon > 0$  gegeben. Wir wählen  $k \in \mathbb{N}$ , so dass

$$\sup_{[a, b]} |f(x) - \varphi_k(x)| < \varepsilon/2.$$

Weil  $\varphi_k$  eine Treppenfunktion ist, gibt es  $\delta > 0$ , so dass  $\varphi_k$  auf  $(p - \delta, p)$  konstant ist. Dann gilt

$$\begin{aligned} |f(x) - f(y)| & \leq |f(x) - \varphi_k(x)| + |\varphi_k(x) - \varphi_k(y)| + |\varphi_k(y) - f(y)| \\ & < \varepsilon/2 + 0 + \varepsilon/2 = \varepsilon \end{aligned}$$

für alle  $x, y \in (p - \delta, p)$ . Dann existiert  $f(p^-)$  aus Lemma A3.8.

Umgekehrt, sei  $\varepsilon > 0$  und sei  $f$  eine Funktion, die keine  $\varepsilon$ -Treppenapproximation auf  $[a, b] =: [a_0, b_0]$  hat. Wir halbieren das Intervall:

$$[a_0, b_0] = [a_0, m_0] \cup [m_0, b_0], \quad m_0 = \frac{a_0 + b_0}{2}.$$

Auf mindestens einer Hälfte (die wir  $[a_1, b_1]$  nennen) hat  $f$  wieder keine  $\varepsilon$ -Treppenapproximation, sonst könnten wir die beiden Approximationen zusammenkleben. Dieses Halbieren wiederholen wir per Induktion; wir bekommen ineinander geschachtelte Intervalle  $[a_n, b_n]$  von Länge  $(b - a)/2^n$ . Die Cauchyfolgen  $(a_n)$  und  $(b_n)$  haben einen gemeinsamen Grenzwert  $p \in [a, b]$ . Aus Lemma A3.7 folgt nun, dass  $f$  keine Regelfunktion ist: hätte  $f$  einseitige Grenzwerte in  $p$ , gäbe es für  $f$  eine  $\varepsilon$ -Treppenapproximation auf einer Umgebung  $I_\delta$  von  $p$ ; diese Umgebung enthielte aber für großes  $n$  das Intervall  $[a_n, b_n]$ , wo es keine Approximation gibt.  $\square$

**A4. Integrale**

a. Treppenintegral

**Definition A4.1.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  eine Funktion und  $Z$  eine markierte Zerlegung von  $[a, b]$ . Die Riemann'sche Summe von

$f$  bezüglich  $\dot{Z}$  ist

$$S(f, \dot{Z}) := \sum_{j=1}^k f(\xi_j) \Delta z_j = \sum_{j=1}^k f(\xi_j) (z_j - z_{j-1}).$$

**Lemma A4.2.** Sei  $\varphi: [a, b] \rightarrow \mathbb{R}$  eine Treppenfunktion und seien  $\dot{Z}, \dot{Z}'$  zwei markierte Treppenzersetzungen für  $\varphi$ . Dann gilt  $S(\varphi, \dot{Z}) = S(\varphi, \dot{Z}')$ .

*Beweisskizze.* Weil jedes  $\varphi_j$  konstant ist, hängt  $S(\varphi, \dot{Z})$  nicht von den ausgewählten  $\xi_j \in [z_{j-1}, z_j]$  ab, d.h., höchstens von der Zerlegung  $Z$ . Die zwei Zerlegungen  $Z$  und  $Z'$  haben eine gemeinsame Verfeinerung  $Z \cup Z'$ . Man muss nur noch zeigen, dass die Riemann'sche Summe unverändert bleibt, wenn man einer Treppenzersetzung einen neuen Punkt hinzufügt.  $\square$

**Definition A4.3.** Sei  $\varphi: [a, b] \rightarrow \mathbb{R}$  eine Treppenfunktion. Das *Integral* von  $\varphi$  definieren wir wie folgt:

$$\int_a^b \varphi := S(\varphi, \dot{Z}),$$

wo  $\dot{Z}$  eine beliebige Treppenzersetzung für  $\varphi$  ist.

**Lemma A4.4.** Das Integral  $\varphi \mapsto \int_a^b \varphi$  ist eine lineare Abbildung  $\mathcal{T}_{[a,b]} \rightarrow \mathbb{R}$ . Es ist monoton im Sinne, dass

$$\varphi \geq 0 \implies \int_a^b \varphi \geq 0.$$

*Beweis.* (Aufgabe.)  $\square$

Wie in jedem Fall, wo ein Integral linear und monoton ist, haben wir gleich ein Korollar:

**Korollar A4.5.** Seien  $f, g$  Treppenfunktionen auf  $[a, b]$ . Falls  $f \leq g$ , gilt  $\int_a^b f \leq \int_a^b g$ . Es folgt, dass

$$\left| \int_a^b f \right| \leq \int_a^b |f|.$$

*Beweis.* Die erste Ungleichung folgt direkt aus Linearität und Monotonie des Integrals. (Aufgabe.) Die Zweite folgt dann aus  $-|f| \leq f \leq |f|$ .  $\square$

*Bemerkung A4.6.* Die Einschränkung einer Treppenfunktion  $\varphi: [a, b] \rightarrow \mathbb{R}$  auf ein beliebiges kompaktes Teilintervall ist auch eine Treppenfunktion.

**Lemma A4.7.** Sei  $\varphi$  eine Treppenfunktion auf  $[a, b]$  und sei  $p \in (a, b)$ . Dann gilt Intervalladditivität:

$$\int_a^b \varphi = \int_a^p \varphi + \int_p^b \varphi.$$

*Beweis.* (Aufgabe.)  $\square$

b. *Regelintegral*

Natürlich wollen wir nicht nur Treppenfunktionen integrieren können. Es gibt viele Möglichkeiten, das Integral zu erweitern (z.B. Riemann-Integral, Lebesgue-Integral, Gauge-Integral). Die unterscheiden sich dadurch, auf welchem Vektorraum  $V \supset \mathcal{T}_{[a,b]}$  von Funktionen das Integral definiert ist. Das Integral ist aber immer eine monotone lineare Abbildung mit Intervalladditivität. Das heißt, auf unserem Wunschzettel für ein Integral steht:

$$\begin{aligned} \int_a^b (f + cg) &= \int_a^b f + c \int_a^b g, \\ f \geq 0 &\implies \int_a^b f \geq 0, \\ \int_a^b f &= \int_a^p f + \int_p^b f. \end{aligned}$$

Wie im Korollar A4.5 folgt, dass  $f \leq g \implies \int f \leq \int g$  und dass  $\left| \int f \right| \leq \int |f|$ .

**Lemma A4.8.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  eine Regelfunktion, und  $(\varphi_n)$  eine Folge von Treppenfunktionen, die gleichmäßig gegen  $f$  konvergiert. Dann konvergiert in  $\mathbb{R}$  die Folge  $(\int_a^b \varphi_n)$ . Der Grenzwert ist unabhängig von der gewählten Folge  $(\varphi_n)$ .

**Definition A4.9.** Den Grenzwert aus diesem Lemma nennen wir das (*Regel-*)Integral  $\int_a^b f$ .

---

Ende der Vorlesung 2008 Oktober 16

---

*Beweis.* Wir setzen  $\delta_n := \sup_x |f(x) - \varphi_n(x)|$ . Für jedes  $x$  gilt aus der Dreiecksungleichung

$$|\varphi_n(x) - \varphi_m(x)| \leq \delta_n + \delta_m.$$

Daher gilt

$$\left| \int_a^b \varphi_n - \int_a^b \varphi_m \right| \leq \int_a^b |\varphi_n - \varphi_m| \leq (b-a)(\delta_n + \delta_m).$$

Weil  $\delta_n \rightarrow 0$ , heißt das, dass  $(\int_a^b \varphi_n)$  eine Cauchyfolge ist. Deshalb konvergiert sie.

Konvergieren  $(\varphi_n)$  und  $(\psi_n)$  gleichmäßig gegen  $f$ , dann auch die Folge

$$(\varphi_0, \psi_0, \varphi_1, \psi_1, \dots).$$

Weil die Integrale zu dieser Folge konvergieren, haben  $(\int_a^b \varphi_n)$  und  $(\int_a^b \psi_n)$  den selben Grenzwert.  $\square$

**Lemma A4.10.** Das Regelintegral ist linear, monoton und intervalladditiv, erfüllt also unsere Wünsche.

*Beweis.* (Aufgabe.)  $\square$

**Definition A4.11.** Der *Mittelwert* von  $f$  auf  $[a, b]$  ist

$$\int_a^b f := \frac{\int_a^b f}{\int_a^b 1} = \frac{1}{b-a} \int_a^b f.$$

**Satz A4.12 (Mittelwertsatz der Integralrechnung).** Sei  $f$  eine Regelfunktion auf  $[a, b]$ . Falls  $m \leq f(x) \leq M$  für alle  $x$ , dann gilt  $m \leq \int_a^b f \leq M$ . Falls  $f$  stetig ist, wird der Mittelwert angenommen, d.h., es gibt ein  $p \in [a, b]$  mit  $f(p) = \int_a^b f$ .

*Beweis.* Betrachten wir  $m$  und  $M$  als konstante Funktionen, folgt  $m \leq \int_a^b f \leq M$  direkt aus dem letzten Korollar. Für stetige  $f$  nehmen wir  $m := \inf f$ ,  $M := \sup f$ . Nach dem Zwischenwertsatz (I.D6.1) nimmt  $f$  jeden Wert in  $[m, M]$  an.  $\square$

**Definition A4.13.** Sei  $f$  eine Funktion und  $\dot{Z}$  eine markierte Zerlegung. Wir definieren wie folgt eine Treppenfunktion  $\varphi(f, \dot{Z})$ , für die  $Z$  eine Treppenzzerlegung ist:

$$(\varphi(f, \dot{Z}))(x) := \begin{cases} f(\xi_j), & x \in (z_{j-1}, z_j), \\ f(z_j), & x = z_j. \end{cases}$$

Es gilt dann

$$\int_a^b \varphi(f, \dot{Z}) = S(f, \dot{Z}).$$

**Satz A4.14.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  eine Regelfunktion. Zu jedem  $\varepsilon > 0$  gibt es dann ein  $\delta > 0$ , so dass für eine beliebige markierte Zerlegung  $\dot{Z}$  mit Maschenweite  $\Delta(Z) < \delta$  gilt

$$\left| S(f, \dot{Z}) - \int_a^b f \right| < \varepsilon.$$

Kurz gefasst:

$$\int_a^b f = \lim_{\Delta(Z) \rightarrow 0} S(f, \dot{Z}).$$

*Bemerkung A4.15.* Für stetige  $f$  und  $\Delta(Z) \rightarrow 0$  konvergieren die Treppenfunktionen  $\varphi(f, \dot{Z})$  gleichmäßig gegen  $f$ . Das gilt nicht mehr, wenn  $f$  einen Sprung hat. Im Satz behaupten wir aber nur, dass die Integrale dieser Treppenfunktionen konvergieren.

**Definition A4.16.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  eine (beschränkte) Funktion. Falls der Grenzwert  $\lim_{\Delta(Z) \rightarrow 0} S(f, \dot{Z})$  existiert, nennen wir ihn das *Riemann-Integral* von  $f$  und sagen,  $f$  ist *Riemann-integrierbar*. (Der Satz sagt, jede Regelfunktion ist Riemann-integrierbar und das Riemann-Integral stimmt mit dem Regelintegral überein.)

**Lemma A4.17.** Der Satz gilt im Falle, dass  $f = \psi$  eine Treppenfunktion ist.

*Beweis.* Sei  $Z'$  eine Treppenzzerlegung für  $\psi$  und (für jedes  $1 \leq j \leq k'$ ) sei  $c_j$  der konstante Wert von  $\psi|_{(c'_{j-1}, c'_j)}$ . In  $z'_j$  hat  $\psi$  einen Sprung; in einer kleinen Umgebung nimmt  $\psi$  die drei

Werte  $c_j, \psi(z'_j), c_{j+1}$  an. Der maximale Unterschied dazwischen nennen wir

$$s_j := \max\{|c_j - c_{j+1}|, |c_j - \psi(z'_j)|, |c_{j+1} - \psi(z'_j)|\}.$$

Wir vereinbaren zunächst, dass  $\delta \leq \min_j \Delta z'_j$ . Sei nun  $\dot{Z}$  eine Zerlegung mit Maschenweite  $\Delta(\dot{Z}) < \delta$  und sei  $\varphi := \varphi(\psi, \dot{Z})$ . Jedes Teilintervall in  $Z$  ist kleiner als jedes in  $Z'$ , d.h., jedes  $(z_{i-1}, z_i)$  beinhaltet entweder kein  $z'_j$  oder genau eins.

Im ersten Fall gibt es  $j$ , so dass  $(z_{i-1}, z_i) \subset (z'_{j-1}, z'_j)$ . Dann gilt  $\psi \equiv \varphi \equiv c_j$  auf  $(z_{i-1}, z_i)$  und deshalb

$$\int_{z_{i-1}}^{z_i} |\psi - \varphi| = 0.$$

Im zweiten Fall nimmt  $\psi$  auf  $(z_{i-1}, z_i)$  die drei Werte  $c_j, \psi(z'_j), c_{j+1}$  an; hingegen ist  $\varphi$  konstant mit einer dieser drei Werte. D.h.,  $|\psi - \varphi| \leq s_j$  und

$$\int_{z_{i-1}}^{z_i} |\psi - \varphi| \leq s_j \Delta z_i < s_j \delta.$$

Jedes  $j$  taucht höchstens einmal auf, also

$$\int_a^b |\psi - \varphi| < \delta \sum_{j=1}^{k'-1} s_j.$$

Jetzt nehmen wir  $\delta < \varepsilon / \sum s_j$  (und immer noch  $\delta \leq \Delta z'_j$ ). Dann gilt

$$\left| S(\psi, \dot{Z}) - \int_a^b \psi \right| \leq \left| \int_a^b \varphi - \int_a^b \psi \right| \leq \int_a^b |\varphi - \psi| < \varepsilon. \quad \square$$

Ende der Vorlesung 2008 Oktober 20

*Beweis des Satzes.* Sei  $\varepsilon > 0$  gegeben. Wir setzen  $\varepsilon' := \frac{\varepsilon}{3(b-a)}$  und wählen eine  $\varepsilon'$ -Treppennäherung  $\psi$  für  $f$ .

Sei nun eine markierte Zerlegung  $\dot{Z}$  mit Maschenweite  $\Delta(Z)$  gegeben. Wir setzen  $\varphi := \varphi(f, \dot{Z})$  und  $\varphi' := \varphi(\psi, \dot{Z})$ . Aus  $\sup |f - \psi| \leq \varepsilon'$  folgt  $\sup |\varphi - \varphi'| \leq \varepsilon'$ , weil  $\varphi$  bzw.  $\varphi'$  die Werte  $f(\xi_j)$  bzw.  $\psi(\xi_j)$  annehmen. Deshalb gilt

$$\begin{aligned} \left| \int_a^b f - S(f, \dot{Z}) \right| &= \left| \int_a^b f - \int_a^b \varphi \right| \\ &\leq \int_a^b |f - \psi| + \left| \int_a^b (\psi - \varphi') \right| + \int_a^b |\varphi' - \varphi| \\ &\leq \frac{\varepsilon}{3} + \left| \int_a^b (\psi - \varphi') \right| + \frac{\varepsilon}{3}. \end{aligned}$$

Aus dem Lemma können wir  $\delta$  so wählen, dass für  $\Delta(Z) < \delta$  gilt  $\left| \int_a^b (\psi - \varphi') \right| \leq \varepsilon/3$ . Dann gilt auch  $\left| \int_a^b f - S(f, \dot{Z}) \right| < \varepsilon. \quad \square$

**A5. Unbestimmte Integrale**

**Definition A5.1.** Bis jetzt hat  $\int_a^b f$  nur dann eine Bedeutung, wenn  $a < b$ . Wir ergänzen jetzt diese Notation, in dem wir  $\int_a^a f := 0$  und  $\int_b^a f := -\int_a^b f$  setzen.

**Lemma A5.2.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  eine Regelfunktion und seien  $p, q, r \in [a, b]$  drei Punkte in beliebiger Reihenfolge. Dann gilt Intervalladditivität

$$\int_p^q f + \int_q^r f = \int_p^r f.$$

Beweis. (Aufgabe.) □

**Definition A5.3.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  eine Regelfunktion und sei  $c \in [a, b]$ . Die Funktion  $F: [a, b] \rightarrow \mathbb{R}$ , definiert durch

$$F(p) := \int_c^p f(x) dx,$$

heißt ein *unbestimmtes Integral* von  $f$ .

Einseitige Ableitungen wurden im letzten Semester kurz eingeführt (siehe Beispiel I.E1.6).

**Satz A5.4.** Sei  $f$  eine Regelfunktion auf  $[a, b]$ . Ein unbestimmtes Integral  $F$  von  $f$  ist stetig und hat in jedem  $p \in [a, b]$  einseitige Ableitungen  $F'_\pm(p) = f(p^\pm)$ . Falls  $f$  in  $p$  stetig ist, folgt, dass  $F$  in  $p$  differenzierbar ist mit  $F'(p) = f(p)$ .

Beweis. Wir betrachten den Fall der rechtsseitige Ableitung. Deren Definition (mit der von  $F$ ) gibt:

$$\begin{aligned} F'_+(p) &:= \lim_{h \rightarrow 0^+} \frac{F(p+h) - F(p)}{h} = \lim_{h \rightarrow 0^+} \frac{\int_c^{p+h} f - \int_c^p f}{h} \\ &= \lim_{h \rightarrow 0^+} \int_p^{p+h} f. \end{aligned}$$

Weil  $f$  eine Regelfunktion ist, existiert  $f(p^+)$ . D.h., zu jedem  $\varepsilon > 0$  gibt es  $\delta > 0$  so, dass  $|f(x) - f(p^+)| < \varepsilon$  für  $x \in (p, p + \delta)$ . Für  $h < \delta$  gilt dann aus dem Mittelwertsatz A4.12,  $|\int_p^{p+h} f - f(p^+)h| < \varepsilon$ . Es folgt,

$$F'_+(p) = \lim_{h \rightarrow 0^+} \int_p^{p+h} f = f(p^+).$$

Ähnliches gilt für die linksseitige Ableitung. (Mit der ergänzten Notation sieht der Beweis sogar fast gleich aus, z.B. auch im Falle  $h < 0$  können wir  $\int_p^{p+h} f$  schreiben.)

Schon die Existenz der einseitigen Ableitungen von  $F$  in  $p$  genügt, um zu zeigen,  $F$  ist in  $p$  stetig. Hier heißt das,  $F$  ist auf  $[a, b]$  stetig. Falls  $f$  in  $p$  stetig ist, gilt

$$F'_-(p) = f(p^-) = f(p) = f(p^+) = F'_+(p),$$

d.h.,  $F$  ist in  $p$  differenzierbar. □

Wir werden den Hauptsatz der Differential- und Integralrechnung nur für stetige Funktionen betrachten. Man kann diesen auch auf beliebigen Regelfunktionen erweitern, in dem man vorsichtig mit den abzählbaren Mengen der Unstetigkeit umgeht. (Siehe z.B. Königsberger, *Analysis I*, §11.4.)

**Definition A5.5.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  stetig. Eine stetig differenzierbare Funktion  $\Phi: [a, b] \rightarrow \mathbb{R}$  mit  $\Phi' = f$  heißt eine *Stammfunktion* für  $f$ .

**Bemerkung A5.6.** Seien  $\Phi$  und  $\Psi$  zwei Stammfunktionen für  $f$ . Korollar I.E2.8 sagt dann,  $\Phi - \Psi \equiv C \in \mathbb{R}$ .

**Satz A5.7 (Hauptsatz).** Sei  $f: [a, b] \rightarrow \mathbb{R}$  stetig. Für jedes  $c \in [a, b]$  ist das unbestimmte Integral  $F(x) := \int_c^x f$  eine Stammfunktion für  $f$ . Sei nun  $\Phi$  eine beliebige Stammfunktion für  $f$ . Für  $p, q \in [a, b]$  gilt

$$\int_p^q f = \Phi(q) - \Phi(p) =: \Phi \Big|_p^q.$$

Beweis. Weil  $f$  stetig ist, sagt Satz A5.4, dass  $F' = f$  in jedem Punkt. Deshalb ist  $F$  eine Stammfunktion. Jede Stammfunktion ist  $\Phi = F + C$  für  $C \in \mathbb{R}$ . Aus Intervalladditivität folgt, dass

$$\Phi \Big|_p^q = F \Big|_p^q = \int_p^q f. \quad \square$$

Der Satz sagt, unbestimmte Integrale sind Stammfunktionen. Wenn  $\Phi$  eine Stammfunktion für  $f$  ist, schreiben wir deshalb  $\int f = \Phi$ .

Jede Ableitung, die wir kennen, gibt uns jetzt eine Integrationsformel. Es gelten z.B.

$$\begin{aligned} \int \exp &= \exp, & \int \frac{dx}{1+x^2} &= \arctan x, \\ \int \sin &= -\cos, & \int \cos &= \sin. \end{aligned}$$

Für  $-1 \neq \alpha \in \mathbb{R}$  gilt

$$\int x^\alpha dx = \frac{1}{\alpha+1} x^{\alpha+1},$$

womit man leicht jedes Polynom integrieren kann. Für  $\alpha = -1$  und  $x > 0$  gilt  $\int 1/x dx = \log x$ .

Manche Autoren schreiben solche Integrationsformeln eher in der Form  $\int f = \Phi + C$ , um zu betonen, dass die Stammfunktion  $\Phi$  nicht eindeutig ist. Hier heißt  $C$  die *Integrationskonstante*, eine beliebige reelle Zahl. Wenn wir diese Konstante weglassen, müssen wir sorgfältig sein. Es gilt z.B. sowohl

$$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x$$

als auch

$$\int \frac{dx}{\sqrt{1-x^2}} = -\arccos x.$$

Man darf aber nicht schließen, dass  $\arcsin x = -\arccos x$ ; hingegen gilt  $\arcsin x = \pi/2 - \arccos x$ .

Eine Funktion, für die man eine Formel schreiben kann, in der nur die Funktionen  $\exp$ ,  $\log$ ,  $\sin$ ,  $\cos$ , die gewöhnlichen arithmetischen Verknüpfungen und Wurzeln auftauchen, heißt eine *elementare Funktion*. (Wir versuchen hier nicht, eine genaue Definition zu formulieren. Siehe Behrends, *Analysis 2*, §6.6.)

Aus den Rechenregeln für Ableitungen und der Kettenregel, ist es klar, wie man jede elementare Funktion ableiten kann; die Ableitung ist wieder elementar. Hingegen kann

es sehr schwierig sein, eine Stammfunktion zu finden; die Stammfunktion einer elementaren Funktion ist nicht immer elementar. Positiv ausgedrückt, liefert uns Integration viele neue Funktionen. Inzwischen haben hunderte solcher *speziellen Funktionen* ihren eigenen Namen. Formelsammlungen dazu findet man in dicken Büchern oder auch jetzt im Web (z.B. [functions.wolfram.com](http://functions.wolfram.com)). Als Beispiel nennen wir die (*Gauß'sche*) Fehlerfunktion

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt,$$

die sehr wichtig in der Statistik ist.

Ende der Vorlesung 2008 Oktober 23

### A6. Integrationstechniken

Zwei ganz wichtige Integrationstechniken sind einfache Umformulierungen der Kettenregel bzw. der Produktregel für Ableitungen.

**Satz A6.1 (Substitutionsregel).** Sei  $F$  eine Stammfunktion für eine stetige Funktion  $f: [a, b] \rightarrow \mathbb{R}$ . Weiter sei  $t: [c, d] \rightarrow [a, b]$  stetig differenzierbar und streng monoton. Dann ist  $F \circ t$  eine Stammfunktion für  $(f \circ t)t'$  und es gilt,

$$\int_a^b f(t(x))t'(x) dx = F(t(b)) - F(t(a)) = \int_{t(a)}^{t(b)} f(t) dt$$

*Beweis.* Die Kettenregel gibt  $(F \circ t)' = (f \circ t)t'$ ; die letzte Formel folgt dann aus dem Hauptsatz.  $\square$

*Bemerkung A6.2.* Hoffentlich ist es nicht verwirrend sondern sogar hilfreich, dass wir in der letzten Formel  $t$  nicht nur als Namen der Funktion benutzt haben, sondern auch als Integrationsvariable. Natürlich kann man auch eine beliebige Variable benutzen, etwa  $\int_{t(a)}^{t(b)} f(\zeta) d\zeta$ . Natürlich geht es auch ohne:  $\int_{t(a)}^{t(b)} f$  bedeutet dasselbe.

*Beispiel A6.3.* Wir haben schon angedeutet, das Integral  $\int e^{-x^2} dx$  lässt sich nicht als elementare Funktion ausdrücken, sondern liefert die neue Fehlerfunktion  $\operatorname{erf}(x)$ . Hingegen lässt sich das anscheinend kompliziertere Integral  $\int xe^{-x^2} dx$  mithilfe der Substitutionsregel leicht berechnen. Der Faktor  $e^{-x^2}$  ist eine Verkettung; wenn wir  $t(x) = -x^2$  und  $f = \exp$  (mit Stammfunktion  $F = \exp$ ) ausprobieren, finden wir

$$\int (-2x)e^{-x^2} dx = e^{-x^2}.$$

Damit ist unser Integral  $-e^{-x^2}/2$ .

**Satz A6.4 (Partielle Integration).** Seien  $u, v: [a, b] \rightarrow \mathbb{R}$  stetig differenzierbare Funktionen. Dann gilt

$$\int uv' = uv - \int u'v, \quad \int_a^b uv' = uv \Big|_a^b - \int_a^b u'v.$$

*Beweis.* Die Funktion  $uv' + u'v$  hat Stammfunktion  $uv$ .  $\square$

Die Schwierigkeit besteht darin, dass es nicht immer klar ist, wie man Funktionen  $f, t$  bzw.  $u, v$  wählen sollte, um einen gegebenen Integranden in die Form  $(f \circ t)t'$  bzw.  $uv'$  vernünftig umzuschreiben.

*Beispiel A6.5.* Wollen wir  $\int e^x x dx$  bestimmen und wählen wir  $u = e^x, v' = x$ , hilft das nicht: es taucht ein „komplizierteres“ Integral  $\int e^x x^2 dx$  auf. Wählen wir hingegen  $u = x, v' = e^x$ , finden wir,

$$\int xe^x dx = xe^x - \int 1e^x dx = (x - 1)e^x.$$

*Beispiel A6.6.* Für  $\int \sqrt{1 - x^2} dx$  hilft (vielleicht erstaunlicher-weise) die Wahl  $v' = 1$  (und noch ein bisschen Tricksen):

$$\begin{aligned} \int \sqrt{1 - x^2} dx &= x\sqrt{1 - x^2} - \int x \frac{-x}{\sqrt{1 - x^2}} dx \\ &= x\sqrt{1 - x^2} - \int \frac{dx}{\sqrt{1 - x^2}} - \int \frac{1 - x^2}{\sqrt{1 - x^2}} dx \\ &= x\sqrt{1 - x^2} - \arccos x - \int \sqrt{1 - x^2} dx. \end{aligned}$$

Damit gilt

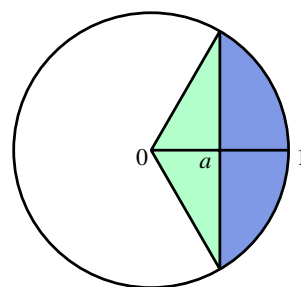
$$2 \int \sqrt{1 - x^2} dx = x\sqrt{1 - x^2} - \arccos x$$

zunächst auf jedem kompakten Teilintervall von  $(-1, 1)$ . D.h., es gilt auf  $(-1, 1)$  aber dann (weil beide Seiten stetig sind) auch in den Endpunkten, d.h. auf dem ganzen Intervall  $[-1, 1]$ .

*Beispiel A6.7.* Diese Formel hat eine schöne geometrische Interpretation als Flächeninhalt eines Kreissegments. Für das Integral auf  $[a, 1]$  haben wir

$$\begin{aligned} 2 \int_a^1 \sqrt{1 - x^2} dx &= x\sqrt{1 - x^2} \Big|_a^1 - \arccos x \Big|_a^1 \\ &= -a\sqrt{1 - a^2} + \arccos a. \end{aligned}$$

Der Flächeninhalt eines Kreissegments ist der Flächeninhalt eines Kreissektors ( $\arccos a$ ) minus der Flächeninhalt eines Dreiecks ( $a\sqrt{1 - a^2}$ ).



Für  $a = -1$  finden wir, dass der Flächeninhalt des Einheitskreises  $\pi$  ist. Das heißt, unsere Definition von  $\pi$  (im letzten Semester als das 2-fache einer Nullstelle von der Potenzreihe  $\cos$ ) stimmt mit der geometrischen Definition überein.

**Beispiel A6.8.** Dieses Integral kann man auch einfach mit der Substitutionsregel berechnen. Sieht man  $\sqrt{1-x^2}$ , darf man immer hoffen, dass  $\theta = \arcsin x$  hilfreich ist. Damit gilt  $x = \sin \theta$ ,  $\sqrt{1-x^2} = \cos \theta$  und  $dx = \cos \theta d\theta$ . Das ergibt

$$\int \sqrt{1-x^2} dx = \int \cos^2 \theta d\theta.$$

Wenn man sich daran erinnert, dass  $2 \cos^2 \theta = 1 + \cos 2\theta$ , hat man

$$2 \int \cos^2 \theta d\theta = \theta + \frac{1}{2} \sin 2\theta = \theta + \sin \theta \cos \theta.$$

**Satz A6.9.** Die Zahl  $\pi$  ist irrational.

Der erste Beweis dafür wurde schon im 18. Jahrhundert gefunden. Wir stellen einen Beweis dar, den Ivan Niven 1947 nur auf einer einzigen Seite veröffentlicht hat.

*Beweis.* Angenommen, es sei  $\pi = a/b$  mit  $a, b \in \mathbb{N}^+$ . Für ein  $n \in \mathbb{N}$ , das wir erst später wählen, definieren wir das Polynom

$$f(x) := \frac{x^n(a-bx)^n}{n!} = \frac{b^n}{n!} (x(\pi-x))^n,$$

in dem nur Monome von Grad zwischen  $n$  und  $2n$  auftauchen. Die Koeffizienten von  $n!f(x)$  sind ganze Zahlen. Deshalb gilt  $f^{(j)}(0) = 0$  für  $j < n$  bzw.  $j > 2n$ ; für  $n \leq j \leq 2n$  ist  $f^{(j)}$  eine ganze Zahl. Das heißt, in 0 ist jede Ableitung von  $f$  eine ganze Zahl. Dasselbe gilt in  $\pi$ , weil  $f(\pi-x) = f(x)$ .

Wir definieren nun ein zweites Polynom

$$F(x) := \sum_{j=0}^n (-1)^j f^{(2j)}(x) = f(x) - f''(x) + \dots \pm f^{(2n)}(x).$$

Dabei merken wir, dass  $F(0)$  und  $F(\pi)$  ganze Zahlen sind. Die Funktion  $G(x) := F'(x) \sin x - F(x) \cos x$  hat die Ableitung

$$G'(x) = (F''(x) + F(x)) \sin x = f(x) \sin x.$$

Aus dem Hauptsatz folgt dann, dass

$$\int_0^\pi f(x) \sin x dx = G(x) \Big|_0^\pi = F(\pi) + F(0).$$

D.h., dieses Integral ist eine ganze Zahl. Für  $0 < x < \pi$  gelten  $0 < \sin x \leq 1$  und  $0 < n!f(x) < a^n \pi^n$ . Deshalb gilt  $0 < f(x) \sin x < a^n \pi^n / n!$ . Jetzt wählen wir  $n$  groß genug, dass Letzteres kleiner als  $1/\pi$  ist. Damit gilt  $0 < \int_0^\pi f(x) \sin x dx < 1$ , ein Widerspruch.  $\square$

*Bemerkung A6.10.* Wie konnte Niven diesen Beweis finden? Will man das Integral  $\int f(x) \sin x dx$  berechnen, kann man partielle Integration (insgesamt  $(2n+2)$ -mal) anwenden und es kommt daraus die Formel für  $F$  bzw.  $G$ .

### A7. Uneigentliche Integrale

Per Definition ist eine Regelfunktion beschränkt und auf einem kompakten Intervall definiert. Oft will man auch unbeschränkte Funktionen bzw. Funktionen auf unendlichen Intervallen integrieren. Darum definiert man das „uneigentliche Integral“ als Grenzwert.

**Definition A7.1.** Sei  $[a, b) \subset \mathbb{R}$  ein halboffenes Intervall, wobei  $b = +\infty$  erlaubt ist. Weiter sei  $f: [a, b) \rightarrow \mathbb{R}$  eine Funktion mit folgender Eigenschaft: für jedes  $c \in [a, b)$  ist die Einschränkung  $f|_{[a,c]}$  eine Regelfunktion. Wir definieren das *uneigentliche Integral* als

$$\int_a^b f := \lim_{c \rightarrow b^-} \int_a^c f,$$

falls dieser Grenzwert existiert.

*Beispiel A7.2.*

$$\begin{aligned} \int_0^{+\infty} \frac{dx}{1+x^2} &= \lim_{c \rightarrow +\infty} \int_0^c \frac{dx}{1+x^2} \\ &= \lim_{c \rightarrow +\infty} \arctan x \Big|_0^c = \lim_{c \rightarrow +\infty} \arctan c = \frac{\pi}{2}. \end{aligned}$$

*Beispiel A7.3.* Für  $\alpha > 1$  gilt  $\int_1^\infty dx/x^\alpha = 1/(\alpha-1)$ . Für  $\alpha \leq 1$  existiert das Integral nicht (bzw. ist  $+\infty$ ).

*Bemerkung A7.4.* Ähnliches gilt für eine Funktion auf  $(a, b)$ : wir setzen  $\int_a^b f := \lim_{c \rightarrow a^+} \int_c^b f$ . Für eine Funktion auf  $(a, b)$  wählen wir ein beliebiges  $p \in (a, b)$  und setzen

$$\int_a^b f := \int_a^p f + \int_p^b f,$$

falls diese letzten beiden uneigentlichen Integrale existieren.

*Beispiel A7.5.* Für  $\alpha \in (0, 1)$  gilt  $\int_0^1 dx/x^\alpha = 1/(1-\alpha)$ . Das Integral  $\int_0^\infty dx/x^\alpha$  existiert (in  $\mathbb{R}$ ) für kein  $\alpha \in \mathbb{R}$ .

*Beispiel A7.6.* Man kann zeigen (am besten mit mehrdimensionalen Integralen), dass  $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$ .

*Beispiel A7.7.* Das Integral  $\int_{-\infty}^{+\infty} x dx$  existiert nicht, obwohl

$$\lim_{c \rightarrow +\infty} \int_{-c}^c x dx = 0.$$

### A8. Weitere Integrale

Wir haben das Integral zunächst für Regelfunktionen definiert und dann auf uneigentliche Integrale erweitert. Wir haben auch schon das Riemann-Integral erwähnt. Jetzt wollen wir diese und andere Integrale vergleichen. (Hier verzichten wir größtenteils auf Beweise.)

*Beispiel A8.1.* Sei  $(a_j)$  eine Folge in  $\mathbb{R}$ . Wir definieren wie folgt eine Funktion  $f^a: [0, 1] \rightarrow \mathbb{R}$ :

$$f^a(x) := \begin{cases} 0, & x = 0, \\ a_j, & x \in (2^{-j-1}, 2^{-j}], \quad j \in \mathbb{N}. \end{cases}$$

Die Einschränkung von  $f^a$  auf ein beliebiges Teilintervall  $[b, 1]$  (für  $b > 0$ ) ist eine Treppenfunktion. Die Funktion  $f^a$  selbst ist genau dann eine Treppenfunktion, wenn  $(a_j)$  „am Ende konstant“ ist, d.h., wenn es  $a \in \mathbb{R}$  so gibt, dass  $a_j = a$  für fast alle  $j$ . Sie ist genau dann eine Regelfunktion, wenn  $(a_j)$  einen Grenzwert (in  $\mathbb{R}$ ) hat. (Was sind dann die approximierende Treppenfunktionen?) Sie ist genau dann Riemann-integrierbar, wenn  $(a_j)$  beschränkt ist.

**Beispiel A8.2.** Für  $(a_j) = (0, 1, 0, 1, \dots)$  ist  $f^a$  Riemann-integrierbar obwohl keine Regelfunktion. Das Integral kann man berechnen:

$$\int_0^1 f^a = \sum_{k=1}^{\infty} 4^{-k} = \frac{1}{3}.$$

Dieses ist entweder als uneigentliches Regelintegral oder als (eigentliches) Riemann-Integral zu verstehen.

Für  $b_j = j$  ist  $f^b$  nicht mal Riemann-integrierbar (weil nicht beschränkt). Trotzdem existiert das uneigentliche (Regel- oder Riemann-) Integral:

$$\int_0^1 f^b = \sum_{j=0}^{\infty} j 2^{-j-1} = 1.$$

Für  $c_j = (-2)^j$  existiert das Integral  $\int_0^1 f^c$  in keinem vernünftigen Sinne. Es gilt

$$\int_{-2^{-j}}^1 f^c = \frac{1}{4}(1 + (-1)^j).$$

Das heißt, diese Integrale sind alternierend  $1/2$  und  $0$ .

**Definition A8.3.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  eine (beschränkte) Funktion. Die obere bzw. untere Summe von  $f$  bezüglich einer Zerlegung  $Z$  von  $[a, b]$  ist:

$$\begin{aligned} \overline{S}(f, Z) &:= \sum_j \Delta z_j \sup_{x \in [z_{j-1}, z_j]} f(x). \\ \underline{S}(f, Z) &:= \sum_j \Delta z_j \inf_{x \in [z_{j-1}, z_j]} f(x). \end{aligned}$$

Das obere bzw. untere Integral von  $f$  ist

$$\begin{aligned} \overline{\int_a^b f} &:= \inf_Z \overline{S}(f, Z) = \lim_{\Delta(Z) \rightarrow 0} \overline{S}(f, Z) \\ \underline{\int_a^b f} &:= \sup_Z \underline{S}(f, Z) = \lim_{\Delta(Z) \rightarrow 0} \underline{S}(f, Z). \end{aligned}$$

**Satz A8.4.** Eine Funktion  $f$  ist genau dann Riemann-integrierbar, wenn  $\overline{\int_a^b f} = \underline{\int_a^b f}$ . □

**Satz A8.5.** Eine Regelfunktion hat abzählbar viele Unstetigkeiten.

*Beweisskizze.* Sei  $f$  der gleichmäßige Grenzwert von Treppenfunktionen  $\varphi_n$ . Ist  $p$  ein Punkt, in dem jedes  $\varphi_n$  stetig ist, dann ist  $f$  in  $p$  stetig. (Warum?) Jedes  $\varphi_n$  hat aber nur endlich viele Unstetigkeiten. □

**Beispiel A8.6.** Die Regelfunktion  $W$  aus Beispiel A2.11 hat abzählbar unendlich viele Unstetigkeiten (in den rationalen Zahlen). Die Unstetigkeiten der Funktion  $f^a$  aus A8.2 sind abzählbar obwohl  $f^a$  keine Regelfunktion ist.

**Satz A8.7.** Eine beschränkte Funktion  $f: [a, b] \rightarrow \mathbb{R}$  ist genau dann Riemann-integrierbar, wenn die Unstetigkeiten von  $f$  eine Nullmenge (vgl. I.G2.2) bilden. □

**Beispiel A8.8.** Die Funktion  $\chi_{\mathbb{Q}}$  ist nicht Riemann-integrierbar, weil sie in jedem Punkt unstetig ist. Sie ist aber 0 außer der abzählbaren Nullmenge  $\mathbb{Q}$ .

Lebesgue hat eine allgemeinere Definition für das Integral gefunden, die keine Nullmengen berücksichtigt. Falls  $f$  Lebesgue-integrierbar ist und  $f = g$  fast überall (d.h., außer in einer Nullmenge), dann ist  $g$  auch integrierbar und es gilt  $\int f = \int g$ . Wir haben z.B.  $\int_0^1 \chi_{\mathbb{Q}} = \int_0^1 0 = 0$ . Das Lebesgue-Integral basiert auf Maßtheorie, welche wir hier nicht erklären können.

Noch allgemeiner ist das sogenannte Gauge-Integral (von Henstock, Kurzweil und anderen). Dieses kann man leicht mit Zerlegungen definieren.

**Definition A8.9.** Sei  $g: [a, b] \rightarrow (0, +\infty)$  eine beliebige positive Funktion, die wir ein Gauge nennen. Eine markierte Zerlegung  $Z$  von  $[a, b]$  heißt dann  $g$ -fein, wenn  $\Delta z_j < g(\zeta_j)$  für jedes  $j$ .

**Definition A8.10.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  eine Funktion und sei  $I \in \mathbb{R}$ . Wir sagen,  $I$  ist das Gauge-Integral von  $f$  auf  $[a, b]$ , falls zu jedem  $\varepsilon > 0$  es ein Gauge  $g$  auf  $[a, b]$  so gibt, dass für jede  $g$ -feine markierte Zerlegung  $Z$  gilt

$$|I - S(f, Z)| < \varepsilon.$$

*Bemerkung A8.11.* Der Fall, wo es (zu jedem  $\varepsilon$ ) ein konstantes Gauge  $g \equiv \delta > 0$  gibt, ist genau die Definition vom Riemann-Integral.

*Bemerkung A8.12.* Wenn man mit dem Gauge-Integral anfängt, muss man keine Erweiterung auf uneigentliche Integrale betrachten.

**Satz A8.13.** Jede Funktion, die ein Lebesgue-Integral oder ein (auch uneigentliches) Riemann-Integral hat, ist Gauge-integrierbar. Eine Funktion  $f$  ist genau dann Lebesgue-integrierbar, wenn  $f$  und  $|f|$  Gauge-integrierbar sind. □



**B. METRISCHE RÄUME**

**B1. Definition**

**Definition B1.1.** Sei  $X$  eine Menge. Eine Funktion oder Abbildung  $d: X \times X \rightarrow \mathbb{R}$  heißt dann eine *Metrik* auf  $X$ , falls für alle  $x, y, z \in X$  die folgenden (axiomatischen) Bedingungen erfüllt sind:

- Definitheit:  $d(x, y) = 0 \iff x = y$ ,
- Symmetrie:  $d(x, y) = d(y, x)$ ,
- Dreiecksungleichung:  $d(x, z) \leq d(x, y) + d(y, z)$ .

Wir sagen,  $X$  mit der Metrik  $d$  ist ein *metrischer Raum*  $(X, d)$ . (Wenn klar ist, welche Metrik gemeint ist, schreiben wir oft einfach  $X$  für den metrischen Raum  $(X, d)$ .)

**Lemma B1.2.** In einem metrischen Raum  $(X, d)$  gilt  $d(x, y) \geq 0$  für alle  $x, y \in X$ .

*Beweis.* Wir nehmen  $z = x$  in der Dreiecksungleichung. Mit Definitheit und Symmetrie gilt dann

$$0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y). \quad \square$$

**Beispiel B1.3.** Die *Standardmetrik* auf den reellen Zahlen  $\mathbb{R}$  ist  $d(x, y) := |x - y|$ .

**Definition B1.4.** Sei  $X$  eine beliebige Menge. Die *diskrete Metrik* auf  $X$  ist

$$d(x, y) := \begin{cases} 0, & x = y, \\ 1, & x \neq y. \end{cases}$$

(Offensichtlich gelten Definitheit und Symmetrie. Die Dreiecksungleichung muss man in Fällen betrachten. Es gilt z.B.  $1 \leq 1 + 1$  im Fall, wo  $x, y, z$  alle unterschiedlich sind.)

**Bemerkung B1.5.** Sei  $(X, d)$  ein metrischer Raum und sei  $Y \subset X$  eine beliebige Teilmenge. Dann ist  $(Y, d)$  auch ein metrischer Raum, ein *Unterraum* von  $X$ . (Streng gesehen müssten wir  $(Y, d|_{Y \times Y})$  schreiben.) Dies ist klar, weil die drei Bedingungen nicht nur für alle Punkte in  $Y$  sondern schon für alle Punkte in  $X$  erfüllt sind.

Die Standardmetrik auf  $\mathbb{R}$  passt mit der Körperstruktur zusammen im Sinne, dass  $d(x+a, y+a) = d(x, y)$  und  $d(\lambda x, \lambda y) = |\lambda|d(x, y)$ . Auf einem Vektorraum  $V$  sind die wichtigsten Metriken diejenigen, die ähnliche Eigenschaften haben, weil sie aus einer Norm auf  $V$  entstehen. Die Definition einer Norm ähnelt der Definition einer Metrik, wie wir gleich sehen.

**Definition B1.6.** Sei  $V$  ein (reeller) Vektorraum. Eine Funktion  $\|\cdot\|: V \rightarrow \mathbb{R}$  heißt eine *Norm* auf  $V$ , falls für alle  $v, w \in V$  und alle  $\lambda \in \mathbb{R}$  die folgenden Bedingungen erfüllt sind:

- Definitheit:  $\|v\| = 0 \iff v = 0$ ,
- Homogenität:  $\|\lambda v\| = |\lambda| \|v\|$ ,
- Dreiecksungleichung:  $\|v + w\| \leq \|v\| + \|w\|$ .

**Lemma B1.7.** Jeder normierte Vektorraum  $(V, \|\cdot\|)$  ist ein metrischer Raum mit der Metrik  $d(v, w) := \|v - w\|$ .

*Beweis.* (Aufgabe.) □

**Bemerkung B1.8.** Im Lemma braucht man Homogenität nur im Falle  $\lambda = -1$ .

**Definition B1.9.** Sei  $n \in \mathbb{N}$ . Auf dem Vektorraum  $\mathbb{R}^n$  definieren wir für jedes  $p \in [1, \infty)$  die  $\ell^p$ -Norm  $\|\cdot\|_p$  durch

$$\|(x_1, \dots, x_n)\|_p := \sqrt[p]{\sum_{i=1}^n |x_i|^p}.$$

Für  $p = \infty$  definieren wir die  $\ell^\infty$ -Norm durch

$$\|(x_1, \dots, x_n)\|_\infty := \max_i |x_i|.$$

**Bemerkung B1.10.** Für jedes  $x \in \mathbb{R}^n$  ist  $p \mapsto \|x\|_p$  eine stetige Abbildung  $(1, \infty) \rightarrow \mathbb{R}$  und es gilt  $\|x\|_p \rightarrow \|x\|_\infty$  für  $p \rightarrow \infty$ .

**Satz B1.11 (Hölder'sche Ungleichung).** Seien  $1 < p, q < \infty$  mit  $1/p + 1/q = 1$  und seien  $v, w \in \mathbb{R}^n$ . Es gilt

$$\sum_{i=1}^n |v_i w_i| \leq \|v\|_p \|w\|_q.$$

*Beweis.* (Aufgabe.) □

**Korollar B1.12.** Sei  $1 < p < \infty$  und seien  $a, b \in \mathbb{R}^n$ . Es gilt

$$\sum_{i=1}^n |a_i| |b_i|^{p-1} \leq \|a\|_p (\|b\|_p)^{p-1}.$$

*Beweis.* Wir setzen  $1/q := 1 - 1/p$  und definieren  $v, w \in \mathbb{R}^n$  durch  $v := a$  und  $w_i := |b_i|^{p-1}$ . Wegen  $p/q = p - 1$  ist die Hölder'sche Ungleichung in diesem Fall genau die Gewünschte. □

**Satz B1.13.** Die  $\ell^p$ -Norm auf  $\mathbb{R}^n$  ist für jedes  $p \in [1, \infty]$  eine Norm.

*Beweis.* Definitheit und Homogenität sind klar. Wir betrachten die Dreiecksungleichung im Falle  $p \in (1, \infty)$ , und lassen  $p = 1$  und  $p = \infty$  als Aufgaben. Seien  $v, w \in \mathbb{R}^n$ . Wir müssen zeigen,  $\|v + w\|_p \leq \|v\|_p + \|w\|_p$ . Mit  $b := v + w$  gilt

$$(\|b\|_p)^p = \sum |b_i|^p \leq \sum |v_i| |b_i|^{p-1} + \sum |w_i| |b_i|^{p-1}.$$

Jetzt wenden wir das Korollar zweimal an:

$$\dots \leq \|v\|_p (\|b\|_p)^{p-1} + \|w\|_p (\|b\|_p)^{p-1}.$$

Wenn wir durch  $(\|b\|_p)^{p-1}$  teilen, folgt die gewünschte Ungleichung  $\|b\|_p \leq \|v\|_p + \|w\|_p$ . □

**Bemerkung B1.14.** Man kann natürlich  $\|v\|_p$  mittels der selben Formel auch für  $0 < p < 1$  definieren; die Dreiecksungleichung gilt dann aber nicht.

**Bemerkung B1.15.** Im Falle  $n = 1$  gilt  $\|x\|_p = |x|$  für jedes  $0 < p \leq \infty$ .

**Definition B1.16.** Der  $n$ -dimensionale *euklidische Raum*  $\mathbb{R}^n$  ist  $\mathbb{R}^n$  mit der euklidischen Metrik  $d(x, y) = \|x - y\|_2$ .

**B2. Folgen in einem metrischen Raum**

Wir betrachten jetzt Folgen in einem metrischen Raum  $(X, d)$ . Im letzten Semester (I.B3, I.B6) haben wir Konvergenz, Grenzwerte, Cauchyfolgen usw. für Folgen in  $\mathbb{R}$  (oder in einem anderen angeordneten Körper) betrachtet. Hier müssen wir nur  $|x - y|$  in  $d(x, y)$  umschreiben, sonst bleibt alles wie damals.

**Definition B2.1.** [Vgl. I.B3.1.] Sei  $(a_n)$  eine Folge im metrischen Raum  $(X, d)$ . Die Folge *konvergiert* gegen  $x \in X$  (geschrieben  $a_n \rightarrow x$ ), falls für jedes  $\varepsilon \in \mathbb{R}^+$  gilt  $d(a_n, x) < \varepsilon$  für fast alle  $n$ . Das heißt,

$$a_n \rightarrow x \iff \forall \varepsilon > 0 \exists n_0 \forall n \geq n_0 \ d(a_n, x) < \varepsilon.$$

**Lemma B2.2.** [Vgl. I.B3.4.] Falls  $a_n \rightarrow x$  und  $a_n \rightarrow y$ , dann folgt  $x = y$ .

*Beweis.* Falls  $x \neq y$ , setzen wir  $\varepsilon := d(x, y)/2$ . Aus Definitheit folgt  $\varepsilon > 0$ . Deshalb können wir  $n$  groß genug wählen, dass  $d(a_n, x) < \varepsilon$  und  $d(a_n, y) < \varepsilon$ . Dann widerspricht  $d(a_n, x) + d(a_n, y) < 2\varepsilon = d(x, y)$  der Dreiecksungleichung.  $\square$

**Definition B2.3.** [Vgl. I.B3.5.] Falls  $a_n \rightarrow x$ , dann heißt  $x$  der *Limes* oder *Grenzwert* der Folge  $(a_n)$ . Wir schreiben

$$\lim_{n \rightarrow \infty} a_n = \lim a_n = x.$$

**Beispiel B2.4.** [Vgl. I.B3.6.1.] Sei  $(a_n)$  am Ende konstant, d.h., es gibt  $x \in X$  so, dass  $a_n = x$  für fast alle  $n$ . Dann konvergiert  $a_n \rightarrow x$ .

**Lemma B2.5.** Sei  $d$  die diskrete Metrik auf einer Menge  $X$  und sei  $(a_n)$  eine Folge in  $X$ . Die Folge  $(a_n)$  konvergiert genau dann, wenn sie am Ende konstant ist.

*Beweis.* Wir nehmen an,  $a_n \rightarrow x$ , und wählen  $\varepsilon \leq 1$ . Dann gilt  $d(x, a_n) < \varepsilon$  für fast alle  $n$ . Für die diskrete Metrik gilt  $d(x, y) < 1$  nur dann, wenn  $x = y$ . Das heißt, wir haben  $x = a_n$  für fast alle  $n$ .  $\square$

---

Ende der Vorlesung 2008 November 3

---

**Definition B2.6.** [Vgl. I.B6.1.] Eine Folge  $(a_n)$  in einem metrischen Raum  $(X, d)$  heißt dann *Cauchyfolge*, wenn

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n, m \geq n_0 \ d(a_m, a_n) < \varepsilon.$$

Die folgenden Aussagen haben genau dieselben Beweise wie im letzten Semester.

**Lemma B2.7.** [Vgl. I.B6.2.] Eine Folge  $(a_n)$  ist genau dann eine Cauchyfolge, wenn es zu jedem  $\varepsilon > 0$  ein  $x \in X$  so gibt, dass  $d(x, a_n) < \varepsilon$  für fast alle  $n$ .  $\square$

**Korollar B2.8.** [Vgl. I.B6.4.5.] Jede konvergente Folge ist eine Cauchyfolge. Hat eine Cauchyfolge eine konvergente Teilfolge, so ist sie selbst konvergent.  $\square$

**Definition B2.9.** [Vgl. I.B8.6.] Ein metrischer Raum  $(X, d)$  heißt dann *vollständig*, falls jede Cauchyfolge aus  $X$  gegen ein  $x \in X$  konvergiert.

**Beispiel B2.10.** Die euklidische Gerade  $\mathbb{B}^1$  (d.h., die reellen Zahlen mit der Standardmetrik) ist vollständig. Die Teilmengen  $\mathbb{Q}$  und  $(0, 1)$  sind nicht vollständig. Jeder diskrete metrische Raum ist vollständig.

**Bemerkung B2.11.** Es ist natürlich sehr wichtig, dass der Grenzwert im Raum  $X$  liegt. Wenn wir nur von einem Raum  $X$  reden, könnte man zurecht fragen „wo sonst?“. In vielen Fällen ist es aber verlockend, einen nicht vollständigen Raum  $X$  als Unterraum von einem größerem Raum zu sehen, wo ein Grenzwert doch existiert. Eigentlich geht das immer: man kann jeden metrischen Raum  $X$  *vervollständigen*, genau wie wir im letzten Semester  $\mathbb{R}$  aus  $\mathbb{Q}$  konstruierten.

**B3. Offene Mengen**

**Definition B3.1.** [Vgl. I.D1.1.] Sei  $(X, d)$  ein metrischer Raum. Für  $\varepsilon > 0$  und  $x \in X$  heißt

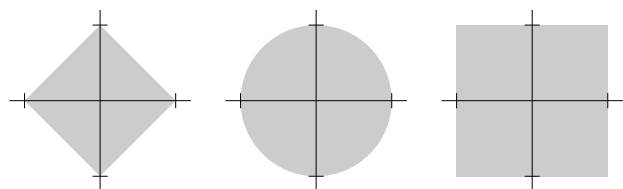
$$B_\varepsilon(x) := \{p \in X : d(x, p) < \varepsilon\}$$

der (offene)  $\varepsilon$ -Ball um  $x$  oder die  $\varepsilon$ -Umgebung von  $x$  in  $X$ .

**Definition B3.2.** [Vgl. I.D1.2.] Seien  $(X, d)$  ein metrischer Raum und  $U \subset X$  eine Teilmenge. Ein Punkt  $p \in U$  heißt ein *innerer Punkt* von  $U$ , falls es ein  $\varepsilon > 0$  gibt, so dass  $B_\varepsilon(p) \subset U$ . Die Teilmenge  $U$  heißt *offen* (in  $X$ ), falls jeder Punkt  $p \in U$  ein innerer Punkt ist. Eine Teilmenge  $A \subset X$  heißt *abgeschlossen*, falls das Komplement  $X \setminus A$  offen ist.

**Beispiel B3.3.** Sei  $d$  die diskrete Metrik auf einer Menge  $X$ . Für jedes  $x \in X$  gilt  $B_\varepsilon(x) = \{x\}$  für  $\varepsilon \leq 1$  und  $B_\varepsilon(x) = X$  für  $\varepsilon > 1$ . Jedes  $U \subset X$  ist offen und auch abgeschlossen.

**Beispiel B3.4.** Betrachten wir  $\mathbb{R}^2$  mit drei verschiedenen Metriken (aus den  $\ell^1$ -,  $\ell^2$ - und  $\ell^\infty$ -Normen), hat der Ball  $B_1(0)$  drei verschiedene Gestalten. Offene Mengen bezüglich der drei Metriken sind aber dieselben, wie wir später erfahren.



**Beispiel B3.5.** Sei  $X = [-1, 1)$  mit der Metrik induziert aus der Standardmetrik auf  $\mathbb{R}$ . Dann ist  $A = [-1, 0)$  offen in  $X$  (obwohl nicht in  $\mathbb{R}$ ). Das Komplement  $X \setminus A = [0, 1)$  ist dann abgeschlossen in  $X$  (obwohl nicht in  $\mathbb{R}$ ).

**Lemma B3.6.** [Vgl. I.D1.3.] Jeder Ball  $B_\varepsilon(x)$  ist offen.

*Beweis.* Zu  $p \in B_\varepsilon(x)$  setzen wir  $\delta := \varepsilon - d(p, x)$ . Dann gilt  $B_\delta(p) \subset B_\varepsilon(x)$ , weil für  $q \in B_\delta(p)$  aus der Dreiecksungleichung folgt

$$d(q, x) \leq d(q, p) + d(p, x) < \delta + d(p, x) = \varepsilon. \quad \square$$

**Lemma B3.7.** Sei  $(X, d)$  ein metrischer Raum. Die offenen Teilmengen von  $X$  haben folgende Eigenschaften:

- $X$  und  $\emptyset$  sind offen.
- Sind  $U$  und  $U'$  offen in  $X$ , so ist auch  $U \cap U'$  offen.
- Ist  $A$  eine Indexmenge und ist  $U_\alpha$  offen für jedes  $\alpha \in A$ , so ist auch  $\bigcup_\alpha U_\alpha$  offen.

*Beweis.* Die leere Menge ist offen, weil es kein  $x \in \emptyset$  nachzuprüfen gibt; der ganze Raum  $X$  ist offen, weil  $X$  jeden Ball enthält.

Für  $x \in U \cap U'$  haben wir  $B_\delta(x) \subset U$  und  $B_{\delta'}(x) \subset U'$ . Mit  $\varepsilon := \min(\delta, \delta')$  gilt dann  $B_\varepsilon(x) \subset U \cap U'$ .

Ein Punkt  $x \in \bigcup U_\alpha$  liegt in mindestens einem  $U_\beta$ . Dann gibt es einen Ball  $B_\varepsilon(x) \subset U_\beta \subset \bigcup U_\alpha$ .  $\square$

*Bemerkung B3.8.* Sei  $X$  eine Menge. Eine Topologie auf  $X$  ist eine Entscheidung, welche Teilmengen  $U \subset X$  für *offen* erklärt werden, wobei die drei Eigenschaften aus Lemma B3.7 gelten müssen. Das Lemma sagt, jede Metrik induziert eine Topologie, jeder metrische Raum ist auf natürliche Weise ein topologischer Raum. (Es gibt auch Topologien – z.B. diejenige, wo nur  $\emptyset$  und  $X$  offen sind –, die von keiner Metrik induziert sind.)

Die folgenden beiden Bemerkungen zeigen, dass die Begriffe „innerer Punkt“ und „Konvergenz“ *topologisch* sind, im Sinne, dass wir diese mittels offener Mengen hätten definieren können. (Für beide Begriffe hatten wir zunächst ähnliche Definitionen gegeben, die aber nur den Fall  $U = B_\varepsilon(x)$  betrachteten.)

*Bemerkung B3.9.* Sei  $(X, d)$  ein metrischer Raum und sei  $T \subset X$ . Ein Punkt  $p \in T$  ist genau dann ein innerer Punkt von  $T$ , wenn es ein offenes  $U$  in  $X$  gibt mit  $p \in U \subset T$ .

*Bemerkung B3.10.* Eine Folge  $(a_n)$  in einem metrischen Raum  $(X, d)$  konvergiert genau dann gegen  $x \in X$ , wenn für jedes offene  $U \subset X$  mit  $x \in U$  gilt  $a_n \in U$  für fast alle  $n$ .

**Definition B3.11.** [Vgl. I.D1.4.] Sei  $(X, d)$  ein metrischer Raum. Ein Punkt  $p \in X$  heißt *Randpunkt* der Teilmenge  $T \subset X$ , falls  $p$  weder ein innerer Punkt von  $T$  noch vom Komplement  $X \setminus T$  ist. Der *Rand von  $T$*  ist die Menge  $\partial T \subset X$  aller Randpunkte von  $T$ .

**Lemma B3.12.** [Vgl. I.D1.5.] Eine Teilmenge  $A \subset X$  ist genau dann abgeschlossen, wenn  $\partial A \subset A$ .  $\square$

---

Ende der Vorlesung 2008 November 6

---

**Lemma B3.13.** Sei  $(X, d)$  ein metrischer Raum. Für  $p \in X$  und  $\varepsilon > 0$  gilt

$$\partial B_\varepsilon(p) \subset \{x \in X : d(x, p) = \varepsilon\}.$$

*Beispiel B3.14.* Oft (z.B. in jeder aus einer Norm induzierten Metrik) gilt hier Gleichheit. Das ist aber bei weitem nicht immer der Fall: mit der diskreten Metrik z.B. ist  $\partial B_1(p)$  leer, obwohl  $d(x, p) = 1$  für alle  $x \neq p$ .

*Beweis.* Falls  $d(x, p) < \varepsilon$ , wissen wir schon,  $x$  ist innerer Punkt von  $B_\varepsilon(p)$ , d.h. kein Randpunkt. Falls  $d(x, p) > \varepsilon$ , setzen wir  $\delta = d(x, p) - \varepsilon$ ; aus der Dreiecksungleichung folgt,  $B_\delta(x) \cap B_\varepsilon(p) = \emptyset$ , d.h.,  $x$  ist innerer Punkt vom Komplement. Es bleiben nur die Punkte  $x$  mit  $d(x, p) = \varepsilon$  als mögliche Randpunkte.  $\square$

## B4. Äquivalente Metriken

**Definition B4.1.** Seien  $d$  und  $d'$  zwei Metriken auf einer Menge  $X$ . Wir sagen,  $d$  und  $d'$  sind *äquivalent*,  $d \sim d'$ , falls zu jedem  $x \in X$  und jedem  $\varepsilon > 0$  es positive Zahlen  $\delta, \delta' > 0$  so gibt, dass

$$B_{\delta'}(x) \subset B'_\varepsilon(x), \quad B'_\delta(x) \subset B_\varepsilon(x).$$

**Lemma B4.2.** Diese Relation  $\sim$  ist eine Äquivalenzrelation auf der Familie aller Metriken auf  $X$ .

*Beweisskizze.* Die Definition ist offensichtlich symmetrisch in  $d$  und  $d'$ . Reflexivität ( $d \sim d$ ) folgt mit  $\delta = \delta' = \varepsilon$ . Zur Transitivität nehmen wir an, dass  $d \sim d' \sim d''$ ; wir wollen zeigen,  $d \sim d''$ . Zu  $\varepsilon > 0$  wählen wir zunächst  $\delta > 0$  und zu  $\delta$  danach  $\delta' > 0$  so, dass

$$B''_\delta(x) \subset B'_\delta(x) \subset B_\varepsilon(x). \quad \square$$

**Satz B4.3.** Zwei Metriken  $d$  und  $d'$  auf  $X$  sind genau dann äquivalent, wenn sie dieselbe Topologie auf  $X$  induzieren, d.h., wenn für jede Teilmenge  $U \subset X$  gilt,

$$U \text{ offen in } (X, d) \iff U \text{ offen in } (X, d').$$

*Beweis.* Sei  $d \sim d'$  und sei  $U$  offen in  $(X, d)$ . Zu jedem  $x \in U$  gibt es  $\varepsilon > 0$  so, dass  $B_\varepsilon(x) \subset U$ . Zu diesem  $\varepsilon$  gibt es  $\delta > 0$  so, dass  $B'_\delta(x) \subset B_\varepsilon(x) \subset U$ . Weil  $x \in U$  beliebig war, heißt das,  $U$  ist offen in  $(X, d')$ .

Umgekehrt, seien  $d$  und  $d'$  zwei Metriken mit denselben offenen Mengen. Für jedes  $x \in X$  und jedes  $\varepsilon > 0$  ist der Ball  $U := B_\varepsilon(x)$  offen in  $(X, d)$  und deshalb auch in  $(X, d')$ . Insbesondere heißt das, es gibt  $\delta > 0$  so, dass  $B'_\delta(x) \subset U = B_\varepsilon(x)$ . Ähnlich gibt es auch  $\delta'$  so, dass  $B_\delta(x) \subset B'_\delta(x)$ . Das bedeutet,  $d \sim d'$ .  $\square$

*Bemerkung B4.4.* Dieses Lemma bedeutet, wenn wir eine Metrik  $d$  mit einer äquivalenten Metrik  $d'$  ersetzen, bleiben alle topologischen Begriffe (offen, abgeschlossen, kompakt, konvergent, stetig, usw.) unverändert.

**Lemma B4.5.** Alle Metriken auf einer endlichen Menge sind äquivalent.

*Beweis.* Sei  $X$  eine endliche Menge und sei  $d$  eine Metrik auf  $X$ . Wir zeigen,  $d$  ist zur diskreten Metrik  $d'$  äquivalent; aus Transitivität folgt das Lemma.

Sei nun  $x \in X$ . Wir setzen

$$\delta := \min\{d(y, x) : x \neq y \in X\} > 0.$$

(Hier benutzen wir, dass  $X$  endlich ist: nur deshalb können wir Minimum anstelle von Infimum schreiben; daraus folgt die Positivität.) Es folgt,  $B_\delta(x) = \{x\} = B'_1(x)$ . Deshalb gilt

$$B_\delta(x) \subset B'_\varepsilon(x), \quad B'_1(x) \subset B_\varepsilon(x).$$

für alle  $\varepsilon > 0$ .  $\square$

**Satz B4.6.** Seien  $\|\cdot\|$  und  $\|\cdot\|'$  zwei Normen auf einem Vektorraum  $V$ . Die beiden induzierten Metriken auf  $V$  sind genau dann äquivalent, wenn es  $\lambda \geq 1$  so gibt, dass für alle  $v \in V$  gilt

$$\frac{1}{\lambda} \|v\| \leq \|v\|' \leq \lambda \|v\|.$$

*Beweisskizze.* Gibt es ein solches  $\lambda$ , können wir in der Definition von Äquivalenz einfach  $\delta = \delta' = \varepsilon/\lambda$  wählen.

Umgekehrt, falls die Metriken äquivalent sind, heißt das insbesondere, es gibt  $\varepsilon$  so, dass  $B'_\varepsilon(0) \subset B_1(0)$  und  $B_\varepsilon(0) \subset B'_1(0)$ . Wir setzen  $\lambda := 1/\varepsilon$ . Aus Homogenität folgt die gewünschte Ungleichung.  $\square$

**Lemma B4.7.** Auf  $\mathbb{R}^n$  sind alle  $\ell^p$ -Normen ( $1 \leq p \leq \infty$ ) in diesem Sinne äquivalent.

*Beweis.* Wegen Transitivität reicht es aus zu zeigen, dass für jedes  $p$  die  $\ell^p$ -Norm und die  $\ell^\infty$ -Norm äquivalent sind. In der Tat behaupten wir,

$$\|v\|_\infty \leq \|v\|_p \leq \sqrt[n]{n} \|v\|_\infty.$$

Sei nämlich  $s := \|v\|_\infty = \sup |v_i|$ . Dann gilt  $0 \leq |v_i| \leq s$  für  $i = 1, 2, \dots, n$  mit Gleichheit auf der rechten Seite mindestens einmal. Das heißt,

$$s^p \leq \sum_{i=1}^n |v_i|^p \leq ns^p,$$

wovon die  $p$ -te Wurzel die behauptete Ungleichung ist.  $\square$

*Bemerkung B4.8.* Später sehen wir, alle Normen auf einem endlichdimensionalen Vektorraum sind äquivalent.

### B5. Beschränkte Räume

**Definition B5.1.** Der Durchmesser eines nichtleeren metrischen Raumes  $(X, d)$  ist

$$\text{diam } X := \sup_{x, y \in X} d(x, y) \in [0, +\infty].$$

Wir vereinbaren,  $\text{diam } \emptyset := 0$  (und nicht etwa  $-\infty$ ). Ein Raum  $X$  heißt *beschränkt*, falls  $\text{diam } X < \infty$ .

*Bemerkung B5.2.* Natürlich wenden wir diese Definitionen auch auf Teilmengen  $Y \subset X$  an, wobei Eigenschaften des Unterraumes  $(Y, d)$  gemeint sind. Jede Teilmenge eines beschränkten Raumes ist beschränkt.

*Beispiele B5.3.* Jeder diskrete Raum hat Durchmesser 1 und ist deshalb beschränkt. In der Standardmetrik auf  $\mathbb{R}$  sind  $\mathbb{N}$  und  $\mathbb{Q}$  unbeschränkt; ein Intervall mit Endpunkten  $a, b \in \mathbb{R}$  ist beschränkt mit Durchmesser  $|b - a|$ , egal, ob es offen oder halboffen oder kompakt ist.

**Lemma B5.4.** Sei  $(X, d)$  ein metrischer Raum ( $X \neq \emptyset$ ) und sei  $T \subset X$ . Folgende Aussagen sind äquivalent:

1.  $T$  ist beschränkt,

2. es gibt einen Ball  $B_R(x)$  in  $X$  mit  $T \subset B_R(x)$ ,

3. zu jedem  $x \in X$  gibt es  $R > 0$  mit  $T \subset B_R(x)$ .

*Beweis.* Wir dürfen annehmen,  $T \neq \emptyset$ . (Für  $T = \emptyset$  gelten offensichtlich alle drei Aussagen.)

(1  $\implies$  2): Ist  $T$  beschränkt, dann gilt  $T \subset B_R(t)$  für jedes  $t \in T$  und jedes  $R > \text{diam } T$ .

(2  $\implies$  3): Ist  $T \subset B_r(x)$  und ist  $y \in X$ , dann gilt

$$T \subset B_r(x) \subset B_R(y) \quad \text{mit} \quad R = r + d(x, y).$$

(3  $\implies$  1): Ist  $T \subset B_R(x)$ , dann gilt

$$\text{diam } T \leq \text{diam } B_R(x) \leq 2R. \quad \square$$

**Korollar B5.5.** Sei  $(X, d)$  ein metrischer Raum und seien  $A, A' \subset X$  beschränkt. Dann ist auch  $A \cup A'$  beschränkt. (Es folgt mit Induktion, dass jede endliche Vereinigung beschränkter Teilmengen auch beschränkt ist.)

*Beweis.* Sei  $x \in X$ . Aus dem Lemma gibt es  $R, R'$  so, dass  $A \subset B_R(x)$  und  $A' \subset B_{R'}(x)$ . Dann gilt  $A \cup A' \subset B_r(x)$  mit  $r := \max\{R, R'\}$ .  $\square$

---

Ende der Vorlesung 2008 November 10

---

*Beispiel B5.6.* Auf  $\mathbb{R}$  definieren wir die Metrik  $d(x, y) := |\arctan y - \arctan x|$ . Mit dieser Metrik ist  $\mathbb{R}$  beschränkt:

$$\text{diam}(\mathbb{R}, d) = \lim_{x \rightarrow \infty} \arctan x - \lim_{x \rightarrow -\infty} \arctan x = \pi.$$

Die Metrik ist aber äquivalent zur Standardmetrik (warum?), was heißt, Beschränktheit ist nicht topologisch. Bezüglich dieser Metrik ist die Folge  $a_n := n$  eine Cauchyfolge, die nicht konvergiert. Das heißt,  $(\mathbb{R}, d)$  ist nicht vollständig; Vollständigkeit ist auch nicht topologisch.

**Definition B5.7.** Sei  $A$  eine Menge und  $(X, d)$  ein metrischer Raum. Eine Abbildung  $f: A \rightarrow X$  heißt dann *beschränkt*, falls das Bild  $f(A)$  beschränkt ist. Den Raum aller beschränkten Abbildungen  $f: A \rightarrow X$  bezeichnen wir als  $\mathcal{B}(A, X)$ . Für  $f, g \in \mathcal{B}(A, X)$  definieren wir

$$d_\infty(f, g) := \sup_{a \in A} d(f(a), g(a)).$$

**Lemma B5.8.** Auf  $\mathcal{B}(A, X)$  ist  $d_\infty$  eine Metrik.

*Beweis.* Endlichkeit: Weil  $f$  und  $g$  beschränkt sind, gilt

$$d_\infty(f, g) \leq \text{diam}(f(A) \cup g(A)) < \infty.$$

Definitheit: Ist  $d_\infty(f, g) = 0$ , dann gilt  $f(a) = g(a)$  für alle  $a$ , d.h.,  $f = g$ . Dreiecksungleichung: Seien  $f, g, h \in \mathcal{B}(A, X)$ . Zu jedem  $\varepsilon > 0$  gibt es  $a_0 \in A$  so, dass

$$\begin{aligned} d_\infty(f, h) &= \sup d(f(a), h(a)) \leq d(f(a_0), h(a_0)) + \varepsilon \\ &\leq d(f(a_0), g(a_0)) + d(g(a_0), h(a_0)) + \varepsilon \\ &\leq d_\infty(f, g) + d_\infty(g, h) + \varepsilon \end{aligned}$$

Weil  $\varepsilon > 0$  beliebig war, gilt  $d_\infty(f, h) \leq d_\infty(f, g) + d_\infty(g, h)$ .  $\square$

**Definition B5.9.** Sei  $(f_n)$  eine Folge in  $\mathcal{B}(A, X)$  und sei  $f \in \mathcal{B}(A, X)$ . Falls  $f_n \rightarrow f$  (bezüglich der Metrik  $d_\infty$ ), sagen wir, die Abbildungen  $f_n$  konvergieren *gleichmäßig* gegen  $f$ .

**B6. Kompaktheit**

a. Totalbeschränkte und folgenkompakte Räume

**Definition B6.1.** Eine Überdeckung eines Raumes  $X$  ist ein System  $\{T_\alpha : \alpha \in A\}$  von Teilmengen  $T_\alpha \subset X$ , deren Vereinigung der ganze Raum  $X = \bigcup_\alpha T_\alpha$  ist. Die Überdeckung heißt *endlich*, falls die Indexmenge  $A$  endlich ist. Die Überdeckung heißt *offen*, falls jedes  $T_\alpha$  offen in  $X$  ist. Eine *Teilüberdeckung* ist durch eine Teilmenge  $B \subset A$  gegeben, falls  $\{T_\alpha : \alpha \in B\}$  den ganzen Raum  $X$  noch überdeckt.

**Beispiel B6.2.** Das Intervall  $[0, 1]$  hat folgende Überdeckung mit  $A = \mathbb{N}$ :

$$T_n := \begin{cases} \{0\}, & n = 0, \\ [2^{-n}, 2^{1-n}], & n > 0. \end{cases}$$

Diese Überdeckung ist weder endlich noch offen; sie hat keine Teilüberdeckungen (außer sich selbst).

**Beispiel B6.3.** Die offene Überdeckung  $\{(-n, n) : n \in \mathbb{N}^+\}$  von  $\mathbb{R}$  hat keine endlichen Teilüberdeckungen. Jede unendliche Teilmenge  $B \subset \mathbb{N}^+$  gibt aber eine Teilüberdeckung.

**Definition B6.4.** Sei  $(X, d)$  ein metrischer Raum und sei  $\varepsilon > 0$ . Eine Teilmenge  $A \subset X$  heißt  $\varepsilon$ -Netz, falls  $\{B_\varepsilon(a) : a \in A\}$  eine Überdeckung von  $X$  ist, d.h., falls es zu jedem  $x \in X$  ein  $a \in A$  so gibt, dass  $d(x, a) < \varepsilon$ .

**Beispiele B6.5.** In einem diskreten Raum  $X$  ist  $X$  das einzige 1-Netz; jedes  $\{x\}$  ist aber ein 2-Netz. In  $\mathbb{R}$  (mit der Standardmetrik) ist  $\mathbb{Q}$  für jedes  $\varepsilon > 0$  ein  $\varepsilon$ -Netz;  $\mathbb{Z}$  ist ein 1-Netz.

**Definition B6.6.** Ein metrischer Raum  $(X, d)$  heißt *totalbeschränkt*, falls es zu jedem  $\varepsilon > 0$  ein endliches  $\varepsilon$ -Netz gibt.

**Bemerkung B6.7.** Ein Raum ist genau dann totalbeschränkt, wenn es zu jedem  $\varepsilon > 0$  eine endliche Überdeckung durch Teilmengen von Durchmesser höchstens  $\varepsilon$  gibt. (Eine Richtung folgt, weil  $\text{diam } B_\varepsilon(x) \leq 2\varepsilon$ ; die andere Richtung folgt, weil jede Teilmenge  $T$  mit  $\text{diam } T \leq \varepsilon$  in einer  $2\varepsilon$ -Ball enthalten ist.)

**Lemma B6.8.** Jeder totalbeschränkte Raum ist beschränkt.

**Beweis.** Ein totalbeschränkter Raum ist eine endliche Vereinigung (beschränkter) 1-Bälle, aus Korollar B5.5 ist er deshalb beschränkt.  $\square$

**Beispiel B6.9.** Ein diskreter Raum ist genau dann totalbeschränkt, wenn er endlich ist.

**Satz B6.10.** Ein metrischer Raum  $(X, d)$  ist genau dann totalbeschränkt, wenn jede Folge in  $X$  eine Teilfolge besitzt, die eine Cauchyfolge ist.

**Definition B6.11.** [Vgl. I.D7.3.] Ein metrischer (oder topologischer) Raum heißt *folgenkompakt*, falls jede Folge eine konvergente Teilfolge hat.

**Korollar B6.12.** [Vgl. I.D7.4.] Ein metrischer Raum ist genau dann folgenkompakt, wenn er vollständig und totalbeschränkt ist.

**Beweis.** (Aufgabe.)  $\square$

**Beweis des Satzes.** Falls  $X$  nicht totalbeschränkt ist, können wir  $r > 0$  so wählen, dass es kein endliches  $r$ -Netz gibt. Wir wählen ein beliebiges  $a_0 \in X$  und definieren rekursiv eine Folge  $(a_n)$ , indem wir  $a_{n+1}$  beliebig in

$$X \setminus \bigcup_{j=0}^n B_r(a_j)$$

wählen. (Weil  $\{a_j : j = 0, \dots, n\}$  kein  $r$ -Netz ist, ist diese Menge nicht leer.) Die Folge  $(a_n)$  hat die Eigenschaft, dass  $d(a_n, a_m) \geq r$  für alle  $n, m \in \mathbb{N}$ . Weil jede Teilfolge dieselbe Eigenschaft hat, ist sie insbesondere keine Cauchyfolge.

Umgekehrt, sei  $(a_n)$  eine Folge im totalbeschränkten Raum  $(X, d)$ . Wir wollen unendliche Teilmengen

$$\mathbb{N} = J_0 \supset J_1 \supset J_2 \supset \dots$$

rekursiv so definieren, dass  $\text{diam}\{a_n : n \in J_k\} \leq 1/k$ . Um  $J_k$  zu definieren wählen wir eine endliche Überdeckung von  $X$  durch Teilmengen  $T$  mit Durchmesser höchstens  $1/k$ . Weil  $J_{k-1}$  unendlich ist, die Überdeckung aber endlich, gibt es (mindestens) ein  $T =: T_k$  so, dass

$$J_k := \{n \in J_{k-1} : a_n \in T_k\}$$

unendlich ist. Jetzt definieren wir rekursiv eine monotone Folge  $(n_k)$  in  $\mathbb{N}$ , indem wir  $n_0$  in  $\mathbb{N}$  und  $n_k > n_{k-1}$  in  $J_k$  beliebig wählen. Wir behaupten, die Teilfolge  $b_k := a_{n_k}$  ist eine Cauchyfolge. Für  $i, j \geq k > 1/\varepsilon$  gilt nämlich  $n_i, n_j \in J_k$  und deshalb  $b_i, b_j \in T_k$ . Das bedeutet,

$$d(b_i, b_j) \leq \text{diam } T_k \leq 1/k < \varepsilon. \quad \square$$

---

Ende der Vorlesung 2008 November 13

---

b. Kompaktheit durch Überdeckungen

**Definition B6.13.** Ein metrischer (oder topologischer) Raum heißt *kompakt*, falls jede offene Überdeckung eine endliche Teilüberdeckung hat.

**Beispiel B6.14.** Die Überdeckung  $\{(-n, n)\}$  zeigt, dass  $\mathbb{R}$  nicht kompakt ist.

**Bemerkung B6.15.** Kompaktheit und Folgenkompaktheit sind topologische Begriffe. Es gibt topologische Räume, die kompakt aber nicht folgenkompakt sind. Es gibt auch Räume, die folgenkompakt aber nicht kompakt sind. Wenn wir aber nur metrische Räume betrachten sind Kompaktheit und Folgenkompaktheit äquivalent. Um dies zu beweisen, benutzen wir vier Lemmata.

**Lemma B6.16.** Ein kompakter metrischer Raum ist totalbeschränkt.

*Beweis.* Sei  $(X, d)$  ein metrischer Raum. Für jedes  $\varepsilon > 0$  ist  $\{B_\varepsilon(x) : x \in X\}$  eine offene Überdeckung. Eine Teilüberdeckung ist genau ein  $\varepsilon$ -Netz. Wenn  $X$  kompakt ist, gibt es eine endliche Teilüberdeckung, d.h. ein endliches  $\varepsilon$ -Netz.  $\square$

**Lemma B6.17.** *Ein kompakter metrischer Raum ist vollständig.*

*Beweis.* Sei  $(X, d)$  nichtvollständig. D.h., es gibt eine Cauchyfolge  $(a_k)$ , die nicht konvergiert. Zu jedem  $n \in \mathbb{N}^+$  gibt es (aus Lemma B2.7) ein  $x_n \in X$  so, dass  $d(a_k, x_n) < 1/n$  für fast alle  $k$ . Wir setzen  $U_n := \{x \in X : d(x, x_n) > 1/n\}$  und bemerken, dass jedes  $U_n$  offen ist und nur endlich viele Folgenglieder  $a_k$  enthält.

Wir behaupten,  $\{U_n\}$  überdeckt  $X$ . Sonst gäbe es  $x \notin \bigcup U_n$ . Das hieße für jedes  $n$ , dass  $d(x, x_n) \leq 1/n$  und deshalb  $d(x, a_k) < 2/n$  für fast alle  $k$ . Dann wäre  $x$  aber ein Grenzwert von  $(a_k)$ .

Wir haben deshalb eine offene Überdeckung  $\{U_n\}$  von  $X$ . Weil jedes  $U_n$  nur endlich viele  $a_k$  enthält, gibt es keine endliche Teilüberdeckung. Das heißt,  $X$  ist nicht kompakt.  $\square$

**Definition B6.18.** Sei  $\{U_\alpha : \alpha \in A\}$  eine offene Überdeckung eines metrischen Raumes  $(X, d)$ . Ein  $\delta > 0$  heißt dann *Lebesguezahl* für  $\{U_\alpha\}$ , wenn es zu jeder Teilmenge  $C \subset X$  mit  $\text{diam } C < \delta$  ein  $\alpha \in A$  so gibt, dass  $C \subset U_\alpha$ .

*Beispiele B6.19.*

- Jedes  $\delta > 0$  ist eine Lebesguezahl für die offene Überdeckung  $\{(-n, n)\}$  von  $\mathbb{R}$ . (Weil  $\mathbb{R}$  archimedisch ist.)
- Die Überdeckung  $\{(2n, 2n + 3) : n \in \mathbb{Z}\}$  von  $\mathbb{R}$  hat Lebesguezahl  $\delta = 1$ . (Weil jedes  $C \subset \mathbb{R}$  mit  $\text{diam } C = d$  in einem Intervall  $[x, x + d]$  enthalten ist.)
- Die Überdeckung  $\{(2^n, 3 \cdot 2^n) : n \in \mathbb{Z}\}$  von  $(0, \infty)$  hat keine Lebesguezahl  $\delta > 0$ . (Weil  $C := (0, \varepsilon)$  auch für kleines  $\varepsilon$  in keinem Intervall  $(2^n, 3 \cdot 2^n)$  enthalten ist.)

**Lemma B6.20.** *Jede offene Überdeckung eines folgenkompakten metrischen Raumes hat eine Lebesguezahl.*

*Beweis.* Sei  $\{U_\alpha : \alpha \in A\}$  eine offene Überdeckung des folgenkompakten Raumes  $(X, d)$ . Wir nehmen an, es gibt keine Lebesguezahl  $\delta > 0$ . Das heißt, zu jedem  $n \in \mathbb{N}^+$  gibt es  $C_n \subset X$  mit  $\text{diam } C_n < 1/n$  so, dass  $C_n$  in kein  $U_\alpha$  enthalten ist. Wir wählen  $a_n \in C_n$ . Weil  $X$  folgenkompakt ist, hat  $(a_n)$  eine konvergente Teilfolge  $(a_{n_k})$ ; wir nennen den Grenzwert  $x \in X$ . Weil  $\{U_\alpha\}$  eine Überdeckung ist, gibt es  $\alpha \in A$  mit  $x \in U_\alpha$ . Weil  $U_\alpha$  offen ist, gibt es  $\varepsilon > 0$  mit  $B_\varepsilon(x) \subset U_\alpha$ . Weil  $a_{n_k} \rightarrow x$  gilt  $a_{n_k} \in B_{\varepsilon/2}(x)$  für fast alle  $k$ . Für  $n > 2/\varepsilon$  gilt aber  $\text{diam } C_n < \varepsilon/2$ . Mit  $a_n \in C_n$  heißt das,  $C_{n_k} \subset B_\varepsilon(x) \subset U_\alpha$  für fast alle  $k$ . Widerspruch (zur Wahl von  $C_n$ )!  $\square$

**Korollar B6.21.** *Ein metrischer Raum ist genau dann kompakt, wenn er folgenkompakt ist.*

*Beweis.* Sei  $(X, d)$  folgenkompakt und sei  $\{U_\alpha\}$  eine offene Überdeckung. Aus Lemma B6.20 hat sie eine Lebesguezahl  $\delta > 0$ . Weil  $(X, d)$  totalbeschränkt ist (Korollar B6.12) gibt es zu  $\varepsilon := \delta/3$  ein endliches  $\varepsilon$ -Netz  $\{x_i : i = 1, \dots, N\}$ , d.h. eine endliche Überdeckung durch die Bälle  $B_\varepsilon(x_i)$ . Weil

$\text{diam } B_\varepsilon(x_i) \leq 2\delta/3 < \delta$ , ist jeder Ball in einem  $U_\alpha$  enthalten, sagen wir  $U_{\alpha_i}$ . Dann ist  $\{U_{\alpha_i} : i = 1, \dots, N\}$  eine endliche Teilüberdeckung, weil  $\bigcup U_{\alpha_i} \supset \bigcup B_\varepsilon(x_i) = X$ .

Umgekehrt, sei  $(X, d)$  kompakt. Dann ist  $X$  aus Lemma B6.16 totalbeschränkt und aus Lemma B6.17 vollständig, mit Korollar B6.12 also folgenkompakt.  $\square$

### c. Abgeschlossene und kompakte Teilmengen

Im letzten Semester zeigten wir (I.D7.4, Satz von Bolzano–Weierstraß), dass eine Teilmenge  $K \subset \mathbb{R}$  genau dann folgenkompakt ist, wenn  $K$  abgeschlossen und beschränkt ist. (Wenn man hier „folgenkompakt“ mit „kompakt“ ersetzt, heißt es dann der Satz von Heine–Borel. Weil wir alles in einer anderen Reihenfolge machen, sehen wir diesen Satz als solche nicht.) Nun haben wir gezeigt, ein metrischer Raum  $K$  ist genau dann (folgen)kompakt, wenn  $K$  vollständig und totalbeschränkt ist. Jetzt wollen wir die Zusammenhänge näher untersuchen.

**Lemma B6.22.** [Vgl. I.D7.2.] *Sei  $(X, d)$  ein metrischer Raum. Eine Teilmenge  $A \subset X$  ist genau dann abgeschlossen in  $X$ , wenn für jede konvergente Folge  $a_n \rightarrow x \in X$  mit  $a_n \in A$  gilt  $x \in A$ .*

*Beweis.* (Aufgabe.)  $\square$

**Korollar B6.23.** *Eine abgeschlossene Teilmenge eines vollständigen metrischen Raumes ist (mit der induzierten Metrik) vollständig.*

*Beweis.* Sei  $(X, d)$  vollständig, sei  $A \subset X$  abgeschlossen und sei  $(a_n)$  eine Cauchyfolge in  $A$ . Als Cauchyfolge in  $X$  hat  $(a_n)$  einen Grenzwert  $x \in X$ . Aus dem Lemma liegt aber  $x \in A$ .  $\square$

**Satz B6.24.** *Sei  $(X, d)$  ein metrischer Raum und sei  $(A, d)$  ein vollständiger Unterraum. Dann ist  $A \subset X$  abgeschlossen.*

*Beweis.* Sei  $(a_n)$  eine Folge in  $A$ , die in  $X$  gegen  $x \in X$  konvergiert. Wir zeigen, dass  $x \in A$ ; dann folgt aus dem letzten Lemma, dass  $A$  abgeschlossen ist. Als konvergente Folge ist  $(a_n)$  aber eine Cauchyfolge in  $X$ , d.h. auch in  $A$ , weil  $d(a_n, a_m)$  unverändert ist. Weil  $(A, d)$  vollständig ist, hat  $a_n$  in  $A$  einen Grenzwert  $a$ . Weil Grenzwerte eindeutig sind (Lemma B2.2), gilt  $x = a \in A$ .  $\square$

**Korollar B6.25.** *Eine Teilmenge eines kompakten metrischen Raumes ist genau dann kompakt, wenn sie abgeschlossen ist.*

*Beweis.* Ist  $T \subset K$  kompakt, dann ist  $T$  vollständig und (aus dem Satz) abgeschlossen. Umgekehrt gilt folgendes: weil  $K$  totalbeschränkt ist, ist jede Teilmenge auch totalbeschränkt; weil  $K$  vollständig ist, ist jede abgeschlossene Teilmenge vollständig.  $\square$

**Korollar B6.26.** *Jede kompakte Teilmenge eines metrischen Raumes ist abgeschlossen und beschränkt.*

*Beweis.* Ein kompaktes  $K \subset (X, d)$  ist vollständig und totalbeschränkt, deshalb abgeschlossen (aus Satz B6.24) und beschränkt (aus Lemma B6.8).  $\square$

Ende der Vorlesung 2008 November 17

Umgekehrt wissen wir, dass abgeschlossene Teilmengen  $A \subset X$  nur dann alle vollständig sind, wenn  $X$  vollständig ist. Wir wissen auch, dass nicht jeder beschränkte Raum totalbeschränkt ist. Im euklidischen Raum sieht es aber besser aus, einfach weil wir jeden Würfel in beliebig kleine Stücke zerschneiden können.

**Lemma B6.27.** *Eine beschränkte Teilmenge im euklidischen Raum  $\mathbb{E}^n$  ist totalbeschränkt.*

*Beweis.* Sei  $T \subset \mathbb{E}^n$  beschränkt. Dann gibt es  $R > 0$  mit

$$T \subset B_R(0) \subset [-R, R]^n.$$

Zu  $\varepsilon > 0$  wählen wir  $k \geq 2\sqrt{n}R/\varepsilon$  und zerschneiden den Würfel in  $k^n$  kleine Würfel mit Seitenlänge höchstens  $\varepsilon/\sqrt{n}$  und deshalb Durchmesser höchstens  $\varepsilon$ . □

*Bemerkung B6.28.* Die Kompaktheit des Würfels folgt auch aus dem sehr allgemeinen Satz von Tichonow: jedes (auch unendliche) kartesische Produkt kompakter Räume ist (mit der sogenannten Produkttopologie) selbst kompakt.

**Korollar B6.29.** *Eine Teilmenge im euklidischen Raum ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.* □

*Bemerkung B6.30.* Dieser Satz gilt auch, wenn wir die euklidische Metrik mit einer anderen  $\ell^p$ -Metrik ersetzen. Auch wenn Beschränktheit und Vollständigkeit keine topologischen Eigenschaften sind, bleiben sie bei äquivalenten Normen unverändert.

**Lemma B6.31.** *Eine Folge im  $\mathbb{E}^n$  konvergiert genau dann, wenn sie komponentenweise konvergiert. Genauer gesagt, sei  $(a_k)$  eine Folge im  $\mathbb{R}^n$  mit  $a_k = (a_{k,1}, \dots, a_{k,n})$ . Dann gilt*

$$\lim_{k \rightarrow \infty} a_k = x = (x_1, \dots, x_n) \iff \forall j \lim_{k \rightarrow \infty} a_{k,j} = x_j.$$

*Beweis.* (Aufgabe.) □

## B7. Stetigkeit

### a. Definition

**Definition B7.1.** [Vgl. I.D2.1.] Seien  $(X, d)$  und  $(Y, d')$  metrische Räume und sei  $p \in X$ . Eine Funktion  $f: X \rightarrow Y$  heißt *stetig in  $p$* , falls

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad f(B_\delta(p)) \subset B'_\varepsilon(f(p)),$$

das heißt, falls zu jedem  $\varepsilon > 0$  es ein  $\delta > 0$  gibt, so dass

$$d(x, p) < \delta \implies d'(f(x), f(p)) < \varepsilon.$$

Die Abbildung  $f$  heißt *stetig*, falls sie in jedem Punkt  $p \in X$  stetig ist.

**Satz B7.2.** [Vgl. I.D2.5.] *Eine Abbildung  $f: X \rightarrow Y$  ist genau dann stetig, wenn für jede offene Menge  $U \subset Y$  das Urbild  $f^{-1}(U)$  offen in  $X$  ist.*

*Beweis.* Sei  $f$  stetig und sei  $U \subset Y$  offen. Wir müssen zeigen, dass jedes  $x \in f^{-1}(U)$  ein innerer Punkt ist. Sei  $y := f(x) \in U$ . Weil  $U$  offen ist, gibt es  $\varepsilon > 0$  so, dass  $B'_\varepsilon(y) \subset U$ . Weil  $f$  stetig ist, gibt es zu diesem  $\varepsilon$  ein  $\delta > 0$  so, dass  $f(B_\delta(x)) \subset B'_\varepsilon(y)$ . Das bedeutet

$$B_\delta(x) \subset f^{-1}(B'_\varepsilon(y)) \subset f^{-1}(U).$$

Umgekehrt, seien die Urbilder aller offenen Mengen offen. Wir müssen zeigen, dass  $f$  in jedem Punkt  $x \in X$  stetig ist. Sei  $y := f(x) \in Y$ . Für jedes  $\varepsilon > 0$  ist  $f^{-1}(B'_\varepsilon(y))$  offen und enthält  $x$ . Deshalb gibt es  $\delta > 0$  mit  $B_\delta(x) \subset f^{-1}(B'_\varepsilon(y))$ , d.h.,  $f(B_\delta(x)) \subset B'_\varepsilon(y)$ . □

*Bemerkung B7.3.* Für eine Abbildung zwischen topologischen Räumen ist diese Eigenschaft die Definition von Stetigkeit.

*Beispiele B7.4.*

- (a) Die identische Abbildung  $\text{id}_X: X \rightarrow X$  ist auf jedem Raum  $(X, d)$  stetig.
- (b) Jede konstante Abbildung ist stetig.
- (c) Ist  $(X, d)$  diskret (und  $(Y, d')$  beliebig), dann ist jede Abbildung  $f: X \rightarrow Y$  stetig.
- (d) Sind  $f: X \rightarrow Y$  und  $g: Y \rightarrow Z$  stetig, dann ist  $g \circ f$  auch stetig. (Weil  $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}U)$ .)
- (e) Eine Abbildung

$$f = (f_1, \dots, f_m): X \rightarrow \mathbb{R}^m$$

ist genau dann stetig, wenn jede Komponentenfunktion  $f_j: X \rightarrow \mathbb{R}$  stetig ist.

**Korollar B7.5.** *Eine Abbildung  $f: X \rightarrow Y$  ist genau dann stetig, wenn für jede abgeschlossene Menge  $A \subset Y$  das Urbild  $f^{-1}(A)$  abgeschlossen in  $X$  ist.*

*Beweis.* (Aufgabe.) □

**Satz B7.6.** [Vgl. I.D2.7.] *Sei  $X = \bigcup_{\alpha \in A} U_\alpha$  eine offene Überdeckung des Raumes  $X$  und sei  $f: X \rightarrow Y$  eine Abbildung. Falls für jedes  $\alpha \in A$  die Einschränkung  $f|_{U_\alpha}: U_\alpha \rightarrow Y$  stetig ist, dann ist  $f$  stetig.*

*Beweis.* Sei  $V \subset Y$  offen. Wir wollen zeigen, dass  $f^{-1}(V)$  offen ist. Zu jedem  $p \in f^{-1}(V)$  gibt es  $\alpha \in A$  mit  $p \in U_\alpha$ . Weil  $f|_{U_\alpha}$  stetig ist, ist  $f|_{U_\alpha}^{-1}(V) = f^{-1}(V) \cap U_\alpha$  offen. Deshalb gibt es  $\varepsilon > 0$  mit  $B_\varepsilon(p) \subset f^{-1}(V) \cap U_\alpha \subset f^{-1}(V)$ . □

### b. Grenzwerte

**Definition B7.7.** [Vgl. I.D2.1.] Sei  $(X, d)$  ein metrischer Raum. Ein Punkt  $p \in X$  heißt *Häufungspunkt* (von  $X$ ), falls  $\{p\}$  nicht offen in  $X$  ist.

**Beispiel B7.8.** Ein diskreter Raum hat keine Häufungspunkte. Jeder Punkt in  $\mathbb{E}^n$  ist Häufungspunkt.

**Lemma B7.9.** Ein Punkt  $p \in (X, d)$  ist genau dann ein Häufungspunkt, falls es eine Folge  $(x_n)$  in  $X \setminus \{p\}$  mit  $x_n \rightarrow p$  gibt.

*Beweis.* Falls  $p$  ein Häufungspunkt ist, gibt es zu jedem  $n \in \mathbb{N}^+$  einen Punkt  $x_n \in B_{1/n}(p) \setminus \{p\}$ ; die Folge  $(x_n)$  konvergiert dann gegen  $p$ . Umgekehrt, falls  $x_n \rightarrow p$ , dann enthält jedes offene  $U \ni p$  sogar fast alle  $x_n$ .  $\square$

**Definition B7.10.** [Vgl. I.D3.3.] Seien  $(X, d)$  und  $(Y, d')$  metrische Räume, sei  $f: X \rightarrow Y$  eine Abbildung und sei  $p \in X$  ein Häufungspunkt. Wir sagen,  $y \in Y$  ist ein Grenzwert von  $f(x)$  für  $x \rightarrow p$ , geschrieben

$$\lim_{x \rightarrow p} f(x) = y \quad \text{oder} \quad f(x) \rightarrow y \text{ für } x \rightarrow p,$$

falls es zu jedem offenen  $V \subset Y$  mit  $y \in V$  ein offenes  $U \subset X$  mit  $p \in U$  so gibt, dass  $f(U \setminus \{p\}) \subset V$ .

**Bemerkung B7.11.** Der Grenzwert  $\lim_{x \rightarrow p} f(x)$  hat nichts mit dem Wert  $f(p)$  zu tun, nur mit Werten in naheliegenden Punkten. Wir könnten auch mit einer Funktion  $f: X \setminus \{p\} \rightarrow Y$  anfangen.

**Bemerkung B7.12.** Falls ein Grenzwert existiert, ist er eindeutig. (Warum?) Hier ist es wichtig, dass  $p$  ein Häufungspunkt ist. (Sonst erfüllt jedes  $y \in Y$  die Bedingungen mit  $U = \{p\}$ .)

**Lemma B7.13.** [Vgl. I.D3.8.] Sei  $f: X \rightarrow Y$  eine Abbildung und sei  $p \in X$  ein Häufungspunkt. Es gilt  $\lim_{x \rightarrow p} f(x) = y \in Y$  genau dann, wenn für jede Folge  $(x_n)$  in  $X \setminus \{p\}$  mit  $x_n \rightarrow p$  gilt  $f(x_n) \rightarrow y$ .

*Beweis.* Wir nehmen zunächst an,  $f(x) \rightarrow y$  für  $x \rightarrow p$ . Sei  $(x_n)$  eine Folge in  $X \setminus \{p\}$  mit  $x_n \rightarrow p$ . Zu jedem  $V \ni y$  offen in  $Y$  gibt es  $U \ni p$  offen in  $X$ , so dass  $f(U \setminus \{p\}) \subset V$ . Weil  $x_n \rightarrow p$ , enthält  $U$  fast alle  $x_n$ . Das bedeutet,  $f(x_n) \in V$  für fast alle  $n$ . Weil  $V$  beliebig war, konvergiert  $f(x_n) \rightarrow y$ .

Umgekehrt, sei  $y$  kein Grenzwert von  $f(x)$  für  $x \rightarrow p$ . D.h., es gibt  $V \ni y$  offen in  $Y$ , so dass zu jedem  $U \ni p$  offen in  $X$  es ein  $x \in U \setminus \{p\}$  mit  $f(x) \notin V$  existiert. Zu jedem  $n \in \mathbb{N}^+$  können wir insbesondere  $U = B_{1/n}(p)$  benutzen, um einen Punkt  $x_n$  zu definieren. Die Folge  $(x_n)$  in  $X \setminus \{p\}$  konvergiert offensichtlich gegen  $p$ . Weil aber  $f(x_n) \notin V$ , konvergiert  $f(x_n)$  nicht gegen  $y$ .  $\square$

**Bemerkung B7.14.** Dieses Lemma gilt nicht in jedem topologischen Raum. Man kann sagen, es ist die archimedische Eigenschaft der reellen Zahlen (insbesondere  $1/n \rightarrow 0$ ), die uns in einem metrischen Raum erlaubt, Konvergenz in abzählbar vielen Schritten (d.h. mittels einer Folge) zu betrachten.

---

Ende der Vorlesung 2008 November 20

**Korollar B7.15.** Eine Abbildung  $f: X \rightarrow Y$  ist genau dann stetig in  $p \in X$ , wenn für jede Folge  $(x_n)$  aus  $X$  mit  $x_n \rightarrow p$  auch  $f(x_n) \rightarrow f(p)$  gilt.

*Beweis.* (Aufgabe.)  $\square$

c. Rechenregeln für reelle Funktionen

Aus den Rechenregeln für reelle Folgen (I.B5.2) folgt:

**Korollar B7.16 (Rechenregeln).** [Vgl. I.D3.12.] Sei  $(X, d)$  ein metrischer Raum, seien  $f, g: X \rightarrow \mathbb{R}$  Abbildungen und sei  $p \in (X, d)$  ein Häufungspunkt, in dem Grenzwerte von  $f$  und  $g$  existieren. Dann gelten

$$\begin{aligned} \lim_{x \rightarrow p} (f \pm g)(x) &= \lim_{x \rightarrow p} f(x) \pm \lim_{x \rightarrow p} g(x), \\ \lim_{x \rightarrow p} (f \cdot g)(x) &= \lim_{x \rightarrow p} f(x) \cdot \lim_{x \rightarrow p} g(x), \end{aligned}$$

und, falls  $g \neq 0$  in einer Umgebung von  $p$ , auch

$$\lim_{x \rightarrow p} (f/g)(x) = \lim_{x \rightarrow p} f(x) / \lim_{x \rightarrow p} g(x). \quad \square$$

**Korollar B7.17.** Seien  $f, g: X \rightarrow \mathbb{R}$  stetig in  $p \in X$ . Dann sind auch  $f \pm g$ ,  $f \cdot g$  und (falls  $g(p) \neq 0$ )  $f/g$  stetig in  $p$ .  $\square$

**Bemerkung B7.18.** Alternativ wissen wir (Beispiel B7.4e), dass  $(f, g): X \rightarrow \mathbb{R}^2$  stetig ist. Die arithmetische Verknüpfungen (Addition usw.) auf  $\mathbb{R}$  sind auch als Abbildungen  $\mathbb{R}^2 \rightarrow \mathbb{R}$  stetig. Deshalb ist  $f + g = + \circ (f, g)$  als Verknüpfung zweier stetigen Abbildungen selbst stetig (Beispiel B7.4d).

d. Nochmal die erweiterten reellen Zahlen

**Definition B7.19.** Sei  $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$  die erweiterten reellen Zahlen (wie im letzten Semester, I.D4). Wir erweitern den Arcus Tangens, indem wir  $\arctan(\pm\infty) := \pm\pi/2$  setzen. Dann ist  $\arctan: \overline{\mathbb{R}} \rightarrow [-\pi/2, \pi/2]$  eine Bijektion. Auf  $\overline{\mathbb{R}}$  definieren wir die Metrik

$$d_a(x, y) := |\arctan x - \arctan y|.$$

**Bemerkung B7.20.** Auf  $\mathbb{R}$  ist  $d_a$  äquivalent zur Standardmetrik;  $\overline{\mathbb{R}}$  ist die Vervollständigung von  $(\mathbb{R}, d_a)$ . Der Raum  $(\overline{\mathbb{R}}, d_a)$  ist kompakt (und sogar isometrisch zu  $[-\pi/2, \pi/2]$ ).

**Bemerkung B7.21.** Anstelle von  $\arctan$  könnten wir natürlich eine beliebige monotone, beschränkte Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  benutzen. Die Metrik  $d_f(x, y) := |f(x) - f(y)|$  wäre zu  $d_a$  äquivalent. Es gibt aber auf  $\mathbb{R}$  keine zu  $d_a$  äquivalente Metrik, die auf der Teilmenge  $\mathbb{R} \subset \overline{\mathbb{R}}$  die Standardmetrik induziert.

**Bemerkung B7.22.** Die unendlichen Grenzwerte und Grenzwerte im Unendlichen, die wir im letzten Semester definierten, sind ganz normale Grenzwerte bezüglich der Metrik  $d_a$ . Das heißt, wir betrachten eine Funktion  $f: \overline{\mathbb{R}} \rightarrow \overline{\mathbb{R}}$ . Wenn wir  $p = \lim_{x \rightarrow a} f(x)$  schreiben, heißt das einfach Konvergenz im metrischen Raum  $(\overline{\mathbb{R}}, d_a)$  – auch im Falle  $a = \pm\infty$  bzw.  $p = \pm\infty$ .

**Bemerkung B7.23.** Konvergenz von Folgen ist auch ein Spezialfall metrischer Konvergenz. Sei  $\overline{\mathbb{N}} := \mathbb{N} \cup \{+\infty\} \subset \overline{\mathbb{R}}$ . Auf  $\overline{\mathbb{N}}$  ist diese Metrik  $d_a$  äquivalent zur Standardmetrik auf  $\mathbb{N} \subset \mathbb{R}$  und deshalb auch zur diskreten Metrik. Damit ist jede Abbildung  $f: (\overline{\mathbb{N}}, d_a) \rightarrow (X, d)$  stetig. Die Folge  $(f(n))$  konvergiert genau dann gegen  $x \in X$ , wenn die durch  $f(\infty) := x$  erweiterte Abbildung  $f: \overline{\mathbb{N}} \rightarrow X$  stetig ist.



**B8. Stetige Funktionen auf kompakten Räumen***b. Gleichmäßige Stetigkeit**a. Erhaltung von Kompaktheit*

Seien  $(X, d)$  und  $(Y, d')$  metrische Räume und sei  $f : X \rightarrow Y$  stetig. Wir wissen, Stetigkeit ist dadurch charakterisiert, dass die Urbilder  $f^{-1}(T)$  offener (bzw. abgeschlossener) Teilmengen  $T \subset Y$  selbst offen (bzw. abgeschlossen) sind. Hingegen können wir nichts über das Bild  $f(T)$  einer offenen bzw. abgeschlossenen Teilmenge  $T \subset X$  sagen. Welche Eigenschaften bleiben unter stetigen Abbildungen erhalten? Kompaktheit ist die Wichtigste!

**Satz B8.1.** [Vgl. I.D7.6.] *Seien  $K$  und  $Y$  metrische Räume mit  $K$  kompakt und sei  $f : K \rightarrow Y$  eine stetige Abbildung. Dann ist  $f(K) \subset Y$  kompakt.*

*Beweis mit Überdeckungen.* Es sei  $\{U_\alpha\}$  eine offene Überdeckung von  $f(K)$ . Die Urbilder  $V_\alpha := f^{-1}(U_\alpha)$  überdecken  $K$  und sind (wegen Stetigkeit) offen in  $K$ . Weil  $K$  kompakt ist, gibt es eine endliche Teilüberdeckung  $\{V_{\alpha_i} : i = 1, \dots, N\}$  von  $K$ . Dann ist  $\{U_{\alpha_i} : i = 1, \dots, N\}$  eine endliche Teilüberdeckung von  $f(K)$ .  $\square$

Weil ein metrischer Raum genau dann kompakt ist, wenn er folgenkompakt ist, hätten wir auch Folgenkompaktheit benutzen können.

*Beweis mit Folgen.* Sei  $(y_n)$  eine Folge in  $f(K)$ . Zu jedem  $y_n$  wählen wir  $x_n \in K$  mit  $f(x_n) = y_n$ . Weil  $K$  kompakt ist, hat die Folge  $(x_n)$  eine konvergente Teilfolge  $x_{n_k} \rightarrow x \in K$ . Weil  $f$  stetig ist, konvergiert auch die Teilfolge

$$y_{n_k} = f(x_{n_k}) \rightarrow f(x) \in f(K). \quad \square$$

*Bemerkung B8.2.* Die beiden Beweise haben nichts mit der genauen Metrik zu tun. Für eine stetige Abbildung zwischen topologischen Räumen haben wir zwei verschiedene Sätze bewiesen:

- das stetige Bild eines kompakten Raumes ist kompakt,
- das stetige Bild eines folgenkompakten Raumes ist folgenkompakt.

Die beiden Aussagen sind natürlich für metrische Räume äquivalent.

**Korollar B8.3.** *Seien  $X$  und  $Y$  metrische Räume und sei  $f : X \rightarrow Y$  stetig. Das Bild  $f(K)$  einer kompakten Teilmenge  $K \subset X$  ist kompakt und damit beschränkt und abgeschlossen in  $Y$ .*  $\square$

**Korollar B8.4 (Satz vom Maximum).** [Vgl. I.D7.7.] *Sei  $f : K \rightarrow \mathbb{R}$  eine stetige Funktion auf einem nichtleeren kompakten Raum  $K$ . Dann nimmt  $f$  auf  $K$  sein Minimum und Maximum an.*  $\square$

**Definition B8.5.** [Vgl. I.D7.8.] Seien  $(X, d)$  und  $(Y, d')$  metrische Räume. Eine Abbildung  $f : X \rightarrow Y$  heißt dann *gleichmäßig stetig*, falls

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x_1 \in X \quad \forall x_2 \in X \\ d(x_1, x_2) < \delta \implies d'(f(x_1), f(x_2)) < \varepsilon.$$

*Bemerkung B8.6.* [Vgl. I.D7.9.] Dass  $f$  stetig ist, heißt dasselbe – nur mit „ $\exists \delta > 0$ “ und „ $\forall x \in D$ “ getauscht. Da darf  $\delta$  von  $x$  abhängen. Hier muss  $\delta$  unabhängig von  $x$  (oder „gleichmäßig“ für alle  $x$ ) gewählt werden können. (Natürlich hängt  $\delta$  immer noch von  $\varepsilon$  ab.)

*Bemerkung B8.7.* Gleichmäßige Stetigkeit ist kein topologischer Begriff; sie hängt von der Metrik ab.

Folgender Satz hat im Wesentlichen denselben Beweis wie im letzten Semester.

**Satz B8.8.** [Vgl. I.D7.12.] *Seien  $K$  und  $Y$  metrische Räume mit  $K$  kompakt. Dann ist jede stetige Abbildung  $f : K \rightarrow Y$  gleichmäßig stetig.*

*Beweis.* (Aufgabe.)  $\square$

Wichtige Beispiele gleichmäßig stetiger Abbildungen sind die Lipschitzabbildungen.

**Definition B8.9.** Seien  $(X, d)$  und  $(Y, d')$  metrische Räume. Eine Abbildung  $f : X \rightarrow Y$  heißt *Lipschitzabbildung*, falls es  $\lambda > 0$  so gibt, dass für alle  $x_1, x_2 \in X$  gilt

$$d'(f(x_1), f(x_2)) \leq \lambda d(x_1, x_2).$$

Die Konstante  $\lambda$  heißt *Lipschitzkonstante* von  $f$  und wir sagen auch,  $f$  ist eine  $\lambda$ -Lipschitzabbildung.

**Satz B8.10.** *Jede Lipschitzabbildung ist gleichmäßig stetig.*

*Beweis.* Ist  $\lambda$  eine Lipschitzkonstante für  $f$ , können wir in der Definition gleichmäßiger Stetigkeit  $\delta := \varepsilon/\lambda$  setzen.  $\square$

**Definition B8.11.** Sei  $(X, d)$  ein metrischer Raum. Eine Lipschitzabbildung  $X \rightarrow X$  mit Lipschitzkonstante  $\lambda < 1$  heißt *Kontraktion*.

**Definition B8.12.** Sei  $f : X \rightarrow X$  eine Abbildung. Ein Punkt  $x \in X$  mit  $f(x) = x$  heißt *Fixpunkt* von  $f$ .

**Satz B8.13 (Fixpunktsatz von Banach).** *Sei  $(X, d)$  ein nichtleerer vollständiger metrischer Raum und sei  $f : X \rightarrow X$  eine Kontraktion. Dann hat  $f$  genau einen Fixpunkt.*

*Beweis.* Sei  $\lambda < 1$  eine Lipschitzkonstante für  $f$ . Zunächst zeigen wir,  $f$  hat höchstens einen Fixpunkt. Aus  $f(x) = x$  und  $f(y) = y$  folgt

$$d(x, y) = d(f(x), f(y)) \leq \lambda d(x, y).$$

Damit folgt  $d(x, y) = 0$ , also  $x = y$ .

Nun sei  $a_0 \in X$ . Wir definieren rekursiv  $a_{n+1} := f(a_n)$  und behaupten,  $(a_n)$  ist eine Cauchyfolge. Es gilt

$$d(a_j, a_{j+1}) \leq \lambda d(a_{j-1}, a_j) \leq \dots \leq \lambda^j d(a_0, a_1).$$

Damit gilt für  $k \geq j$

$$d(a_j, a_k) \leq \sum_{i=j}^{k-1} \lambda^i d(a_0, a_1) \leq \sum_{i=j}^{\infty} \lambda^i d(a_0, a_1) = \frac{\lambda^j}{1-\lambda} d(a_0, a_1),$$

und  $(a_n)$  ist Cauchyfolge.

Ende der Vorlesung 2008 November 24

Weil  $X$  vollständig ist, hat  $(a_n)$  einen Grenzwert  $x \in X$ . Wir zeigen,  $x$  ist ein Fixpunkt von  $f$ . Für jedes  $k$  gilt

$$\begin{aligned} d(x, f(x)) &\leq d(x, a_{k+1}) + d(a_{k+1}, f(x)) \\ &= d(x, a_{k+1}) + d(f(a_k), f(x)) \\ &\leq d(x, a_{k+1}) + \lambda d(x, a_k). \end{aligned}$$

Weil Letzteres gegen Null konvergiert, gilt  $d(f(x), x) = 0$ , also  $f(x) = x$ .  $\square$

c. Gleichmäßige Konvergenz

Seien  $A$  eine Menge und  $(X, d)$  ein metrischer Raum. Wir haben auf der Menge  $\mathcal{B}(A, X)$  aller beschränkten Abbildungen  $f: A \rightarrow X$  die Metrik  $d_\infty$  definiert; Konvergenz bezüglich  $d_\infty$  heißt gleichmäßige Konvergenz.

**Satz B8.14.** Falls  $(X, d)$  vollständig ist, ist  $(\mathcal{B}(A, X), d_\infty)$  auch vollständig.

*Beweis.* Sei  $(f_n)$  eine Cauchyfolge in  $\mathcal{B}(A, X)$ . Für jedes  $a \in A$  ist  $(f_n(a))$  deswegen eine Cauchyfolge in  $X$ , weil

$$d(f_n(a), f_m(a)) \leq d_\infty(f_n, f_m).$$

Weil  $X$  vollständig ist, hat  $f_n(a)$  einen Grenzwert, den wir  $f(a)$  nennen. Die Folge  $(f_n)$  konvergiert punktweise gegen die damit definierte Abbildung  $f: A \rightarrow X$ . Wir zeigen, die Konvergenz ist gleichmäßig. Dazu sei  $\varepsilon > 0$ . Es gibt  $N \in \mathbb{N}$  so, dass für alle  $m, n \geq N$  gilt  $d_\infty(f_n, f_m) < \varepsilon/2$ , d.h. für alle  $a \in A$  gilt  $d(f_n(a), f_m(a)) < \varepsilon/2$ . Aus Lemma B3.13 (über den Rand eines Balls) gilt deshalb  $d(f_n(a), f(a)) \leq \varepsilon/2 < \varepsilon$  für alle  $a \in A$  und alle  $n \geq N$ .  $\square$

Sei nun  $K$  ein kompakter metrischer Raum. Die Menge  $C(K, X)$  aller stetigen Abbildungen  $K \rightarrow X$  ist dann eine Teilmenge  $C(K, X) \subset \mathcal{B}(K, X)$ .

**Satz B8.15.** Bezüglich  $d_\infty$  ist  $C(K, X) \subset \mathcal{B}(K, X)$  abgeschlossen. Das heißt, falls  $(f_n)$  eine Folge von stetigen Abbildungen  $K \rightarrow X$  ist, die gleichmäßig gegen  $f: K \rightarrow X$  konvergiert, dann ist auch  $f$  stetig.

*Beispiel B8.16.* Gleichmäßige Konvergenz ist hier wichtig. Die stetigen Abbildungen  $f_k: [0, 1] \rightarrow [0, 1]$  mit  $f_k(x) = x^k$  aus Beispiel A3.2 konvergieren punktweise gegen einen nicht-stetigen Limes  $f: [0, 1] \rightarrow [0, 1]$ .

*Beweis.* Die Äquivalenz der beiden Aussagen folgt aus der Charakterisierung abgeschlossener Mengen mittels Folgen (Lemma B6.22).

Seien nun  $f_n$  stetig mit  $f_n \rightarrow f$  gleichmäßig und sei  $p \in K$ . Um zu zeigen, dass  $f$  in  $p$  stetig ist, sei  $(x_j)$  eine Folge in  $K$ , die gegen  $p$  konvergiert. Wir müssen zeigen,  $f(x_j) \rightarrow f(p)$ . Dazu sei  $\varepsilon > 0$ . Weil  $f_n \rightarrow f$  gleichmäßig, gibt es  $N \in \mathbb{N}$  so, dass für alle  $x \in K$  und alle  $n \geq N$  gilt

$$d(f_n(x), f(x)) < \varepsilon/3.$$

Weil  $f_N$  in  $p$  stetig ist und  $x_j \rightarrow p$  konvergiert, gibt es  $J \in \mathbb{N}$  so, dass für alle  $j \geq J$  gilt

$$d(f_N(x_j), f_N(p)) < \varepsilon/3.$$

Damit gilt für jedes  $j \geq J$

$$\begin{aligned} d(f(x_j), f(p)) &\leq d(f(x_j), f_N(x_j)) + d(f_N(x_j), f_N(p)) + d(f_N(p), f(p)) \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon. \end{aligned} \quad \square$$

*Bemerkung B8.17.* Alternativ kann man mit einer unstetigen Abbildung  $f: K \rightarrow X$  anfangen. Wenn  $f$  in  $p$  unstetig ist, gibt es  $\varepsilon > 0$  so, dass  $f^{-1}(B_\varepsilon(f(p)))$  keinen Ball um  $p$  enthält. Dann kann man zeigen, der  $\varepsilon/3$ -Ball um  $f$  besteht nur aus Abbildungen, die ebenso unstetig in  $p$  sind. (Der Fall, wo  $f$  in  $p$  einen Sprung hat, ist besonders anschaulich.) Damit ist die Menge aller unstetigen Abbildungen offen.

*Bemerkung B8.18.* Kompaktheit ist hier nicht so wichtig; man kann gleichmäßige Konvergenz auch für unbeschränkte Abbildungen definieren; ein gleichmäßiger Limes stetiger Abbildungen ist immer noch stetig.

*Bemerkung B8.19.* [Vgl. I.F1.9.] Man kann sagen, der Satz erlaubt uns (im Fall gleichmäßiger Konvergenz), zwei Limites zu tauschen:

$$\lim_{n \rightarrow \infty} \lim_{x \rightarrow p} f_n(x) = \lim_{n \rightarrow \infty} f_n(p) = f(p) = \lim_{x \rightarrow p} f(x) = \lim_{x \rightarrow p} \lim_{n \rightarrow \infty} f_n(x).$$

Hier gilt natürlich die dritte Gleichung nur deshalb, weil  $f$  stetig ist. (Alle anderen gelten schon, wenn  $f_n \rightarrow f$  punktweise.)

Hinter den allermeisten wichtigen Begriffen der Analysis (u.a. Ableitung und Integral) stecken Grenzwerte. Das heisst, wenn man z.B. zwei Integrale oder Ableitungen tauschen will, ist die zugrunde liegende Frage die, ob zwei Limites getauscht werden dürfen. Gleichmäßigkeit spielt immer wieder eine wichtige Rolle.

**B9. Zusammenhang**

Um den Zwischenwertsatz auf beliebige metrische Räume zu erweitern, müssen wir ein Analogon zu Intervallen finden. Das sind zusammenhängende Räume. Wir werden sehen, Zusammenhang ist die zweite wichtige Eigenschaft, die unter stetigen Abbildungen erhalten bleibt.

**Definition B9.1.** Ein Raum  $X$  heißt dann *zusammenhängend*, falls  $\emptyset$  und  $X$  die einzigen Teilmengen sind, die zugleich offen und abgeschlossen sind.

**Beispiel B9.2.** Offensichtlich ist jeder Raum mit höchstens einem Punkt zusammenhängend. Sonst sind diskrete Räume (insbes. endliche metrische Räume) nie zusammenhängend.

**Beispiel B9.3.** Die rationalen Zahlen  $\mathbb{Q}$  sind nicht zusammenhängend, weil z.B.

$$\{x \in \mathbb{Q} : x^2 < 2\} = \{x \in \mathbb{Q} : x^2 \leq 2\}$$

zugleich offen und abgeschlossen in  $\mathbb{Q}$  ist.

**Satz B9.4.** Sei  $(X, d)$  ein metrischer Raum und sei  $T \subset X$  eine Teilmenge. Dann ist  $(T, d)$  mit der induzierten Metrik genau dann zusammenhängend, wenn folgendes gilt:

Sind  $\hat{U}, \hat{V} \subset X$  disjunkte, offene Mengen mit  $T \subset \hat{U} \cup \hat{V}$ , dann gilt entweder  $T \subset \hat{U}$  oder  $T \subset \hat{V}$ .

**Lemma B9.5.** Sei  $T \subset (X, d)$ . Eine Teilmenge  $U \subset T$  ist genau dann offen in  $T$ , wenn es ein offenes  $\hat{U} \subset X$  gibt mit  $U = T \cap \hat{U}$ .

*Beweis.* Deses Lemma wurde schon in der Übung bewiesen. Hier nochmal ein Beweis zum Lesen.

Für  $t \in T$  sei  $B_\varepsilon(t)$  der  $\varepsilon$ -Ball um  $t$  in  $X$ . Dann ist es klar, dass  $T \cap B_\varepsilon(t)$  der  $\varepsilon$ -Ball um  $t$  in  $T$  ist.

Nun sei  $U \subset T$  offen. Zu jedem  $t \in U$  gibt es  $\varepsilon(t) > 0$  so, dass  $T \cap B_{\varepsilon(t)}(t) \subset U$ . Wir definieren

$$\hat{U} := \bigcup_{t \in U} B_{\varepsilon(t)}(t),$$

sodass  $U = T \cap \hat{U}$ . Als Vereinigung offener Mengen ist  $\hat{U}$  offen in  $X$ .

Umgekehrt, sei  $\hat{U} \subset X$  offen und sei  $U := T \cap \hat{U}$ . Zu jedem  $t \in U \subset \hat{U}$  gibt es  $\varepsilon > 0$  so, dass  $B_\varepsilon(t) \subset \hat{U}$  und damit  $T \cap B_\varepsilon(t) \subset U$ .  $\square$

*Beweis des Satzes.* Sei  $(T, d)$  zusammenhängend und seien  $\hat{U}, \hat{V} \subset X$  wie im Satz. Wir setzen  $U := T \cap \hat{U}$  und  $V := T \cap \hat{V}$ . Dann sind  $U, V$  disjunkte offene Teilmengen in  $T$ , deren Vereinigung  $T$  ist. Das heißt,  $V = T \setminus U$ , und damit sind  $U, V$  zugleich abgeschlossen. Weil  $T$  zusammenhängend ist, gilt dann entweder  $U = \emptyset, V = T$  oder  $U = T, V = \emptyset$ .

Umgekehrt wollen wir zeigen,  $T$  ist zusammenhängend. Dazu sei  $U$  zugleich offen und abgeschlossen in  $T$ , d.h.,  $V := T \setminus U$  ist auch offen. Wenn wir einfach aus dem Lemma beliebige offene Mengen  $\hat{U}, \hat{V} \subset X$  mit  $U = T \cap \hat{U}$  und  $V = T \cap \hat{V}$  wählen, dann gilt natürlich  $T \subset \hat{U} \cup \hat{V}$  aber  $\hat{U}$  und  $\hat{V}$  sind nicht unbedingt disjunkt. Deshalb wiederholen wir die Konstruktion aus dem Lemma, halbieren aber die Radien:

$$\hat{U} := \bigcup_{u \in U} B_{\varepsilon(u)/2}(u), \quad \hat{V} := \bigcup_{v \in V} B_{\varepsilon(v)/2}(v).$$

Dann behaupten wir,  $\hat{U}$  und  $\hat{V}$  sind disjunkt. Aus dieser Behauptung und der Voraussetzung folgt, dass entweder  $T \subset \hat{U}$  oder  $T \subset \hat{V}$ , d.h., dass entweder  $U = T$  oder  $V = T$  (und damit  $U = \emptyset$ ).

Um die Behauptung zu beweisen, sei  $w \in \hat{U} \cap \hat{V}$ . Dann gibt es  $u \in U \subset T$  und  $v \in V \subset T$  mit

$$w \in B_{\varepsilon(u)/2}(u) \cap B_{\varepsilon(v)/2}(v).$$

Sei etwa  $\varepsilon(u) \leq \varepsilon(v)$ . Dann gilt

$$d(u, v) \leq d(u, w) + d(w, v) < \frac{\varepsilon(u)}{2} + \frac{\varepsilon(v)}{2} \leq \varepsilon(u).$$

D.h.,  $v \in B_{\varepsilon(u)}(u)$ . Es gilt aber  $B_{\varepsilon(u)}(u) \cap T \subset U$ . Weil  $v \in T$  gilt damit  $v \in U$  – im Widerspruch zu  $v \in V = T \setminus U$ .  $\square$

**Bemerkung B9.6.** Obwohl Zusammenhang natürlich ein topologischer Begriff ist, gilt dieser Satz nicht in jedem topologischen Raum  $X$ .

---

Ende der Vorlesung 2008 November 27

---

Wie Kompaktheit bleibt auch Zusammenhang unter stetigen Abbildungen erhalten.

**Satz B9.7.** Sei  $f : X \rightarrow Y$  eine stetige Abbildung. Ist  $X$  zusammenhängend, dann ist auch  $f(X)$  zusammenhängend.

*Beweis.* Seien  $U, V$  disjunkt und offen in  $Y$  mit  $f(X) \subset U \cup V$ . Wir wollen zeigen, entweder  $f(X) \subset U$  oder  $f(X) \subset V$ . Weil  $f$  stetig ist, sind  $f^{-1}(U)$  und  $f^{-1}(V)$  offen in  $X$ . Die sind aber disjunkt und überdecken  $X$ . Weil  $X$  zusammenhängend ist, gilt dann entweder  $f^{-1}(U) = \emptyset$  oder  $f^{-1}(V) = \emptyset$ .  $\square$

Um diesen Satz als Erweiterung des Zwischenwertsatzes zu sehen, zeigen wir, dass die zusammenhängenden Teilmengen von  $\mathbb{R}$  genau die Intervalle sind. Wir erinnern uns, ein Intervall in  $\mathbb{R}$  ist ein Teilmenge  $I$  mit folgender Eigenschaft (I.D5.3):

Sind  $x, y \in I$  und ist  $x < p < y$ , dann ist  $p \in I$ .

**Satz B9.8.** Eine Teilmenge  $T \subset \mathbb{R}$  ist genau dann zusammenhängend, wenn sie ein Intervall ist.

*Beweis.* Sei  $T \subset \mathbb{R}$  zusammenhängend. Seien  $x < p < y$  mit  $x, y \in T$ . Wäre  $p \notin T$ , so wäre

$$T \subset (-\infty, p) \cup (p, +\infty),$$

also (weil  $T$  zusammenhängend) entweder  $T \subset (-\infty, p)$  oder  $T \subset (p, +\infty)$ . Das Erste gilt aber nicht, weil  $y \in T$ , und das Zweite gilt nicht, weil  $x \in T$ .

Umgekehrt, sei  $T \subset \mathbb{R}$  ein Intervall. Seien  $U, V \subset \mathbb{R}$  disjunkt und offen mit  $T \subset U \cup V$ . Wir nehmen an, es gibt  $x \in T \cap U$  und  $y \in T \cap V$ , und folgern daraus einen Widerspruch. Ohne Einschränkung sei  $x < y$  und damit  $[x, y] \subset T$ . Wir definieren  $X := U \cap [x, y]$ ,  $Y := V \cap [x, y] = [x, y] \setminus X$  und  $p := \sup X \in [x, y]$ . Es gilt entweder  $p \in X$  oder  $p \in Y$ ; wir zeigen, dass beide zum Widerspruch führen.

Ist  $p \in X$ , dann ist  $p \neq y$ ; weil  $X$  offen in  $[x, y]$  ist, gibt es dann  $\varepsilon > 0$  mit  $[p, p + \varepsilon) \subset X$ , insbesondere mit  $p + \varepsilon/2 \in X$ . Das heißt,  $p$  ist keine obere Schranke für  $X$ .

Ist hingegen  $p \in Y$ , dann ist  $p \neq x$ ; weil  $Y$  offen in  $[x, y]$  ist, gibt es  $\varepsilon > 0$  mit  $(p - \varepsilon, p] \subset Y$ , d.h.,  $(p - \varepsilon, p]$  ist disjunkt von  $X$ . Weil  $p$  eine obere Schranke für  $X$  ist, ist auch  $(p, y]$  disjunkt von  $X$ . D.h.,  $X \subset [x, p - \varepsilon]$  und damit ist  $p - \varepsilon$  eine kleinere obere Schranke für  $X$ .  $\square$

**Korollar B9.9.** [Vgl. I.D6.1.] Sei  $X$  zusammenhängend und sei  $f : X \rightarrow \mathbb{R}$  stetig. Dann ist  $f(X)$  ein Intervall. Das heißt,  $f$  nimmt jeden Wert an, der zwischen zwei Funktionswerten liegt.

*Beweis.* Mit Satz B9.8 ist dies der Spezialfall  $Y = \mathbb{R}$  vom Satz B9.7.  $\square$

*Bemerkung B9.10.* Die großen Sätze vom letzten Semester über reelle Funktionen (I.D7.7; I.D7.12; Zwischenwertsatz I.D6.1; usw.) verstehen wir jetzt als Folgerungen zweier topologischer Eigenschaften eines Intervalls  $[a, b]$ : Kompaktheit und Zusammenhang. Mit den beiden eher abstrakten Definitionen konnten wir ganz schnell beweisen, dass beide Eigenschaften unter stetigen Abbildungen erhalten bleiben (B8.1, B9.7). Eigentlich steckt mehr Arbeit darin, zu zeigen, dass  $[a, b]$  die beiden Eigenschaften hat (B6.29, B9.8).

**Definition B9.11.** Sei  $X$  ein Raum, seien  $x, y \in X$  und sei  $[a, b] \subset \mathbb{R}$  ein kompaktes Intervall. Eine stetige Abbildung  $\gamma: [a, b] \rightarrow X$  mit  $\gamma(a) = x$  und  $\gamma(b) = y$  heißt ein *Weg* in  $X$  von  $x$  nach  $y$ . Das Bild  $\gamma([a, b])$  nennen wir auch die *Spur* des Weges.

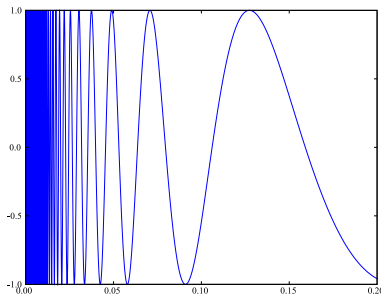
**Korollar B9.12.** Die Spur  $\gamma([a, b])$  eines Weges  $\gamma$  ist zusammenhängend.

*Beweis.* Dies ist der Spezialfall  $X = \mathbb{R}$  vom Satz B9.7.  $\square$

**Definition B9.13.** Ein Raum  $X$  heißt *wegzusammenhängend*, falls es zu je zwei Punkten  $x, y \in X$  einen Weg in  $X$  von  $x$  nach  $y$  gibt.

*Beispiel B9.14.* Es gibt zusammenhängende Räume, die nicht wegzusammenhängend sind. Das gebräuchlichste Beispiel, die sogenannte „Sinuskurve der Topologen“, ist

$$S := \{(x, \sin 1/x) : x \in (0, 1]\} \cup \{(0, y) : y \in [-1, 1]\} \subset \mathbb{E}^2.$$



**Satz B9.15.** Ein wegzusammenhängender Raum ist zusammenhängend.

*Beweis.* Sei  $X$  ein nichtzusammenhängender Raum. Dann gibt es nichtleere, disjunkte, offene Teilmengen  $U, V \subset X$  mit  $U \cup V = X$ . Jede zusammenhängende Teilmenge  $T \subset X$  ist entweder in  $U$  oder in  $V$  enthalten. Mit Korollar B9.12 gibt es deshalb keinen Weg von einem Punkt  $u \in U$  nach einem Punkt  $v \in V$ . Damit ist  $X$  nicht wegzusammenhängend.  $\square$

C. DIE KOMPLEXEN ZAHLEN

C1. Definition

Sei  $K$  ein Körper. Jede lineare Gleichung  $bx = c$  (mit  $0 \neq b \in K$ ) hat die eindeutige Lösung  $x = c/b$ . Hingegen hat nicht jede quadratische Gleichung  $ax^2 + bx = c$  eine Lösung. In den rationalen Zahlen  $\mathbb{Q}$  z.B. hat  $x^2 = 2$  keine Lösung. Letztes Semester behoben wir diese Problem, in dem wir die reellen Zahlen  $\mathbb{R}$  als Vervollständigung von  $\mathbb{Q}$  konstruierten. In  $\mathbb{R}$  hat  $x^2 = c$  genau dann eine Lösung  $x = \pm\sqrt{c}$ , wenn  $c \geq 0$ .

Wollen wir alle quadratische Gleichungen lösen können, müssen wir die reellen Zahlen ergänzen und einen größeren Körper  $\mathbb{C}$  konstruieren. Dabei wird  $\mathbb{C}$  kein angeordneter Körper.

*Bemerkung C1.1.* Jeder Körper  $K \supset \mathbb{R}$  ist ein reeller Vektorraum. Die Vektoraddition ist die Körperaddition und die Skalarmultiplikation ist ein Spezialfall der Körpermultiplikation.

Sei nun  $K \supset \mathbb{R}$  ein Körper, der eine Lösung  $\mathbf{i}$  zur Gleichung  $x^2 = -1$  enthält. Hier ist  $\mathbf{i}$  einfach ein Name für ein neues Element  $\mathbf{i} \in K \setminus \mathbb{R}$  mit der Eigenschaft  $\mathbf{i}^2 = -1$ . Als Vektorraum enthält  $K$  die lineare Hülle

$$\mathbb{C} := \text{span}(1, \mathbf{i}) := \{x + y\mathbf{i} : x, y \in \mathbb{R}\} \subset K$$

als 2-dimensionalen Unterraum. Um  $\mathbb{C}$  als Körper zu sehen, müssen wir noch die Multiplikation erklären. Wollen wir tatsächlich, dass  $\mathbf{i}^2 = -1$ , gibt es keine Wahl: aus den Körperaxiomen (insbes. dem Distributivgesetz) muss gelten

$$(x + y\mathbf{i})(a + b\mathbf{i}) = xa + y\mathbf{a}\mathbf{i} + x\mathbf{b}\mathbf{i} + y\mathbf{b}\mathbf{i}^2 = (xa - yb) + (ya + xb)\mathbf{i}.$$

Es ist dann einfach zu verifizieren, dass  $\mathbb{C}$  mit dieser Multiplikation (und der Vektoraddition) tatsächlich ein Körper ist. Das heißt, wir haben einen Körper  $\mathbb{C} \supset \mathbb{R}$  gefunden, in dem  $x^2 = -1$  eine Lösung (genauer gesagt, die zwei Lösungen  $\pm\mathbf{i}$ ) hat.

**Definition C1.2.** Die  $\mathbb{R}$ -lineare Abbildung  $z = x + y\mathbf{i} \mapsto \bar{z} := x - y\mathbf{i}$  heißt (komplexe) Konjugation und ist ein Körperautomorphismus. Das heißt, neben  $\bar{z} + \bar{w} = \overline{z + w}$  gelten auch  $\overline{\bar{z}} = z$  und  $\overline{z/w} = \bar{z}/\bar{w}$ . Wir definieren den *Realteil* und den *Imaginärteil* einer komplexen Zahl:

$$\begin{aligned} \text{Re } z &:= \frac{1}{2}(z + \bar{z}) \in \mathbb{R}, & \text{Re}(x + y\mathbf{i}) &= x, \\ \text{Im } z &:= \frac{1}{2\mathbf{i}}(z - \bar{z}) \in \mathbb{R}, & \text{Im}(x + y\mathbf{i}) &= y. \end{aligned}$$

Damit gilt  $z = (\text{Re } z) + (\text{Im } z)\mathbf{i}$ . Eine komplexe Zahl  $z \in \mathbb{C}$  ist natürlich genau dann reell, wenn  $\text{Im } z = 0$ ; wir sagen,  $z$  ist (*rein*) *imaginär*, falls  $\text{Re } z = 0$ . Die Zahl  $\mathbf{i} = \sqrt{-1}$  heißt *imaginäre Einheit*.

---

Ende der Vorlesung 2008 Dezember 1

*Bemerkung C1.3.* Über mehrere Jahrhunderte wurden komplexen Zahlen benutzt – aber nur als Hilfsmittel, reellen Lösungen zu finden. Man glaubte, die imaginären Zahlen – auch

wenn nützlich – seien nur Hirngespinnste, die reellen Zahlen hingegen existieren in der realen Welt.

Als Beispiel hat jede kubische Gleichung mit reellen Koeffizienten mindestens eine reelle Lösung; es gibt sogar eine (allerdings komplizierte) Formel dazu, die 1545 von Cardano veröffentlicht wurde. Mit einer linearen Substitution kann man die Gleichung in die reduzierte Form  $x^3 + 3bx = 2c$  bringen. Definieren wir die *Diskriminante*  $D := b^3 + c^2$ , ist

$$\sqrt[3]{c + \sqrt{D}} + \sqrt[3]{c - \sqrt{D}}$$

eine reelle Lösung. Für  $D \geq 0$  ist alles reell; für  $D < 0$  hingegen arbeiten wir vorübergehend mit komplexen Zahlen, auch wenn die Endergebnis wieder reell ist.

Spätestens seit dem Versuch (etwa vor 100 Jahren), die reellen Zahlen rigoros zu definieren, verstehen Mathematiker das alles ganz anders. Schon die reellen Zahlen sind „imaginär“ im Sinne, dass sie ein menschliches Hirngespinnst sind. Hingegen sind die komplexen Zahlen auch „reell“ sind im Sinne, dass sie eine sehr vernünftige mathematische Struktur bilden. Sie spielen z.B. bei der von Eugene Wigner benannten „unverschämten Effektivität der Mathematik in den Naturwissenschaften“ eine entscheidende Rolle.

Wir haben  $\mathbb{C}$  gebaut, damit jede reelle Zahl eine Wurzel hat. Es stellt sich heraus, dass sogar jede komplexe Zahl  $z \in \mathbb{C}$  eine Wurzel  $w \in \mathbb{C}$  hat:

$$w = \sqrt{1/2(|z| + \text{Re } z)} + \sqrt{1/2(|z| - \text{Re } z)}\mathbf{i}.$$

Mit der aus der Schule bekannten Lösungsformel hat dann jede quadratische Gleichung eine explizite Lösung in  $\mathbb{C}$ .

Der Fundamentalsatz der Algebra, den wir später beweisen, sagt, dass jedes Polynom  $P(z)$  mit komplexen Koeffizienten eine Nullstelle  $P(z) = 0$  in  $\mathbb{C}$  hat. (Es folgt, dass die Anzahl der Nullstellen gleich dem Grad des Polynoms ist, wenn diese mit einer angemessenen Vielfachheit gezählt werden.)

C2. Metrik und Topologie

Die natürliche Basis  $\{1, \mathbf{i}\}$  gibt uns eine Identifizierung

$$\mathbb{C} \cong \mathbb{R}^2, \quad x + y\mathbf{i} \longleftrightarrow (x, y), \quad z \longleftrightarrow (\text{Re } z, \text{Im } z).$$

Es ist oft sehr hilfreich, wenn man sich komplexe Zahlen bildlich als Punkte in dieser sogenannten *Gauß'schen Zahlenebene* vorstellt.

**Definition C2.1.** Auf  $\mathbb{C}$  benutzen wir die euklidische Norm auf  $\mathbb{R}^2$ , die sich auch komplex ausdrücken lässt:

$$\|x + y\mathbf{i}\| = |z| := \sqrt{z\bar{z}} = \sqrt{x^2 + y^2} = \|(x, y)\|_2.$$

Wir nennen  $|z|$  den *Betrag* von  $z$ .

*Bemerkung C2.2.* Natürlich gilt die Dreiecksungleichung  $|z + w| \leq |z| + |w|$ . Ähnlich zum Betrag auf  $\mathbb{R}$  gilt aber auch  $|zw| = |z||w|$  und (für  $w \neq 0$ )  $|\bar{z}/\bar{w}| = |z|/|w|$ .

**Bemerkung C2.3.** Als metrischer Raum ist  $\mathbb{C}$  deshalb isometrisch zu  $\mathbb{E}^2$ . Damit sind alle topologischen Fragen (Konvergenz, Stetigkeit, Kompaktheit usw.) geklärt. Auch für  $\mathbb{C}^n$  gilt  $\mathbb{C}^n \cong \mathbb{E}^{2n}$ .

**Beispiele C2.4.** Es gelten z.B.:

1. Eine Folge  $(z_n)$  in  $\mathbb{C}$  konvergiert genau dann gegen  $w \in \mathbb{C}$ , wenn  $\operatorname{Re} z_n \rightarrow \operatorname{Re} w$  und  $\operatorname{Im} z_n \rightarrow \operatorname{Im} w$ .
2. Eine komplexe Funktion  $f: X \rightarrow \mathbb{C}$  ist genau dann stetig, wenn  $\operatorname{Re} f$  und  $\operatorname{Im} f$  stetige Funktionen  $X \rightarrow \mathbb{R}$  sind.
3. Der Raum  $\mathbb{C}^n$  ist vollständig und wegzusammenhängend.
4. Eine Teilmenge  $T \subset \mathbb{C}^n$  ist genau dann kompakt, wenn sie beschränkt und abgeschlossen in  $\mathbb{C}^n$  ist.

**Beispiel C2.5.** Komplexe Multiplikation  $(z, w) \mapsto zw$  ist eine stetige Abbildung  $\mathbb{C}^2 \rightarrow \mathbb{C}$ . (Weil  $(x, y, a, b) \mapsto (xa - yb, xb + ya)$  eine stetige Abbildung  $\mathbb{R}^4 \rightarrow \mathbb{R}^2$  ist.) Es folgt, dass die Rechenregeln für reelle Folgen (I.B5.2) auch für komplexe Folgen gelten, insbes., dass

$$\lim(z_n w_n) = (\lim z_n)(\lim w_n),$$

falls die beiden Grenzwerte auf der rechten Seite existieren.

**Bemerkung C2.6.** Hingegen machen viele Sachen, die wir in  $\mathbb{R}$  bzw. für reelle Funktionen definiert haben, in  $\mathbb{C}$  bzw. für komplexe Funktionen deswegen keinen Sinn mehr, weil  $\mathbb{C}$  nicht angeordnet ist:

- Eine Teilmenge in  $\mathbb{C}$  hat kein Supremum bzw. Infimum.
- Eine Abbildung  $f: \mathbb{C} \rightarrow Y$  hat keine einseitigen Grenzwerte.
- Monotonie macht keinen Sinn für komplexe Funktionen  $\mathbb{C} \rightarrow \mathbb{C}$  (bzw.  $\mathbb{R} \rightarrow \mathbb{C}$  bzw.  $\mathbb{C} \rightarrow \mathbb{R}$ ).

**Bemerkung C2.7.** Es gibt in  $\mathbb{C}$  keine Intervalle und deshalb keinen Zwischenwertsatz. (Um komplexe Nullstellen topologisch zu finden, braucht man den komplizierteren Begriff eines *einfach zusammenhängenden* Raumes.)

**Bemerkung C2.8.** In der komplexen Analysis lernt man die erweiterten komplexen Zahlen (die Riemann'sche Sphäre)  $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$  kennen. Hier gibt es nur das eine unendliche Element  $\infty$  und die Rechenregeln sind ein bisschen anders als in  $\overline{\mathbb{R}}$ . In diesem Semester betrachten wir  $\overline{\mathbb{C}}$  nicht weiter.

**Bemerkung C2.9.** Später werden wir Ableitungen für Funktionen  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  definieren. Die komplexe Ableitung einer Funktion  $\mathbb{C} \rightarrow \mathbb{C}$  ist aber etwas Strengeres, die wir hier nicht betrachten.

### C3. Komplexe Reihen

#### a. Reihen in Banachräumen

Bezeichne  $V$  in diesem Abschnitt einen *Banachraum*, d.h. ein vollständiger normierter Vektorraum. Im letzten Semester (I.C) betrachteten wir Reihen in  $\mathbb{R}$ . Die meisten Ergebnisse können wir leicht auf Reihen in  $V$  übertragen.

**Definition C3.1.** [Vgl. I.C1.1.] Sei  $(a_n)_{n \geq 1}$  eine Folge in  $V$ . Wir definieren dazu die Folge  $(s_n)$  der *Partialsommen*:

$$s_n := \sum_{k=1}^n a_k := a_1 + a_2 + \dots + a_n.$$

Diese neue Folge nennen wir die zu  $(a_n)$  gehörige (*unendliche*) *Reihe*. Falls sie konvergiert, nennen wir den Grenzwert die *Reihensumme*:

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n a_k =: \sum_{k=1}^{\infty} a_k =: \sum a_k =: a_1 + a_2 + \dots.$$

Wir benutzen  $\sum a_k$  als Name sowohl für die Reihe selbst als auch für deren Grenzwert (falls dieser existiert).

**Satz C3.2 (Rechenregeln).** [Vgl. I.C1.4.] Seien  $\sum b_n$  und  $\sum c_n$  konvergente Reihen in  $V$ , und sei  $a \in \mathbb{R}$ . Die Reihe  $\sum(ab_n + c_n)$  konvergiert auch. Es gilt

$$\sum_{n=1}^{\infty} (ab_n + c_n) = a \sum_{n=1}^{\infty} b_n + \sum_{n=1}^{\infty} c_n.$$

Weil  $V$  vollständig ist, konvergiert eine Folge genau dann, wenn sie eine Cauchyfolge ist. Das heißt:

**Satz C3.3 (Cauchy Kriterium).** [Vgl. I.C1.7.] Die Reihe  $\sum a_k$  in  $V$  konvergiert genau dann, wenn es zu jedem  $\varepsilon > 0$  ein  $N \in \mathbb{N}$  gibt, so dass für alle  $n \geq m \geq N$  gilt

$$\left\| \sum_{k=m}^n a_k \right\| < \varepsilon.$$

**Korollar C3.4.** [Vgl. I.C1.8.] Ist  $(\|a_k\|)$  keine Nullfolge, dann konvergiert  $\sum a_k$  nicht.

**Definition C3.5.** [Vgl. I.C2.3.] Eine Reihe  $\sum a_k$  in  $V$  heißt *absolut konvergent*, wenn die Reihe  $\sum \|a_k\|$  in  $\mathbb{R}$  konvergent ist.

**Lemma C3.6.** [Vgl. I.C2.7.] Eine absolut konvergente Reihe in  $V$  ist konvergent und es gilt

$$\left\| \sum_{k=1}^{\infty} a_k \right\| \leq \sum_{k=1}^{\infty} \|a_k\|.$$

**Bemerkung C3.7.** In jedem *unvollständigen* normierten Vektorraum gibt es hingegen eine absolut konvergente Reihe, die *nicht* konvergiert.

**Satz C3.8 (Wurzelkriterium).** [Vgl. I.C2.10.] Sei  $\sum a_k$  eine Reihe in  $V$ . Falls es  $q < 1$  gibt, so dass  $\sqrt[k]{\|a_k\|} \leq q$  für fast alle  $k \in \mathbb{N}$ , dann ist die Reihe  $\sum a_k$  absolut konvergent. Falls hingegen  $\sqrt[k]{\|a_k\|} \geq 1$  für unendlich viele  $k$ , dann ist  $\sum a_k$  divergent.

**Bemerkung C3.9.** Man kann den Satz auch wie folgt formulieren: Sei  $a := \overline{\lim} \sqrt[k]{\|a_k\|}$ . Falls  $a < 1$ , konvergiert  $\sum a_k$  absolut; falls  $a > 1$  divergiert  $\sum a_k$ .

**Bemerkung C3.10.** Der Satz hat fast nichts mit  $V$  zu tun. Alle Aussagen beziehen sich auf die reelle Folge  $\sum \|a_k\|$  mit nur einer Ausnahme. Für  $a > 1$  schließen wir nicht nur, dass  $\sum a_k$  nicht absolut konvergiert, sondern (aus Korollar C3.4), dass  $\sum a_k$  divergiert.

**Korollar C3.11 (Quotientenkriterium).** [Vgl. I.C2.12.] Sei  $\sum a_k$  eine Reihe in  $V$ , wobei  $a_k \neq 0$  für fast alle  $k$ . Falls es  $q < 1$  gibt, so dass  $\|a_{k+1}\|/\|a_k\| \leq q$  für fast alle  $k$ , dann ist  $\sum a_k$  absolut konvergent. Falls hingegen  $\|a_{k+1}\|/\|a_k\| \geq 1$  für fast alle  $k$ , dann ist  $\sum a_k$  divergent.

b. Komplexe Potenzreihen

Weil  $\mathbb{C} \cong \mathbb{E}^2$  ein Banachraum ist, gelten alle obigen Sätze für komplexe Reihen.

**Beispiel C3.12.** [Vgl. I.C1.10.] Sei  $z \in \mathbb{C}$ . Wir betrachten die geometrische Reihe  $\sum z^n$ . Für  $|z| \geq 1$  ist sie divergent; für  $|z| < 1$  konvergiert sie absolut:

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}.$$

---

Ende der Vorlesung 2008 December 4

---

**Definition C3.13.** [Vgl. I.F1.] Sei  $(a_n)$  eine Folge in  $\mathbb{C}$ . Für jedes  $z \in \mathbb{C}$  betrachten wir die Potenzreihe

$$\sum_{k=0}^{\infty} a_k z^k.$$

Ist  $T \subset \mathbb{C}$  die Menge aller  $z$ , für die die Reihe konvergiert, dann definiert die Potenzreihe eine Funktion  $f: T \rightarrow \mathbb{C}$  durch  $f(z) := \sum a_k z^k$ .

**Definition C3.14.** Sei  $\sum a_k z^k$  eine Potenzreihe. Wir setzen  $a := \overline{\lim} \sqrt[k]{|a_k|} \in [0, +\infty]$ . Wir definieren wie folgt den Konvergenzradius  $R \in [0, +\infty]$  der Potenzreihe:

$$R := \begin{cases} +\infty, & a = 0, \\ 1/a, & 0 < a < +\infty, \\ 0, & a = +\infty. \end{cases}$$

**Beispiele C3.15.** Aus der Regel von l'Hôpital (I.E3) gilt

$$\lim_{k \rightarrow +\infty} \frac{\log k}{k} = 0.$$

Für jedes  $b \in \mathbb{R}$  gilt daher  $\lim \sqrt[k]{k^b} = e^{0b} = 1$ . Hingegen wissen wir schon (I.F1.8), dass  $\lim \sqrt[k]{k!} = +\infty$ . Deshalb gelten z.B.:

- $\sum z^k$  hat Konvergenzradius  $R = 1$ ,
- $\sum k^b z^k$  hat ebenfalls  $R = 1$ ,
- $\sum z^k/k!$  hat  $R = +\infty$ ,
- $\sum k! z^k$  hat  $R = 0$ .

**Satz C3.16.** Sei  $R$  der Konvergenzradius der Potenzreihe  $\sum a_k z^k$ . Für  $|z| < R$  (d.h., für  $z \in B_R(0)$ ) ist die Reihe absolut konvergent. Für  $|z| > R$  ist sie divergent.

**Beweis.** Wegen  $\sqrt[k]{|a_k z^k|} = z \sqrt[k]{|a_k|}$  folgt dieser Satz direkt aus dem Wurzelkriterium. □

**Bemerkung C3.17.** Man kann auch Potenzreihen um ein Zentrum  $z_0 \in \mathbb{C}$  betrachten. Falls  $f(z) = \sum a_k z^k$  Konvergenzradius  $R$  hat, dann ist

$$g(z) := \sum a_k (z - z_0)^k = f(z - z_0)$$

absolut konvergent für  $z \in B_R(z_0)$  und divergent für  $|z - z_0| > R$ .

**Bemerkung C3.18.** Sei  $f(z) = \sum_0^\infty a_k z^k$  eine Potenzreihe mit Konvergenzradius  $R$ . Die Partialsummen  $\sum_0^n a_k z^k$  sind Polynome und haben Ableitungen  $\sum_0^n k a_k z^{k-1}$ . (Dies wissen wir im Falle  $z, a_k \in \mathbb{R}$ . Es gilt aber auch im komplexen Falle, sobald man komplexe Ableitungen kennt.) Wir definieren

$$g(z) := \sum_1^\infty a_k k z^{k-1} = \sum_0^\infty (k+1) a_{k+1} z^k.$$

Weil  $\sqrt[k]{k} \rightarrow 1$ , hat auch diese Potenzreihe Konvergenzradius  $R$ . Auf  $B_R(0)$  ist  $g$  die Ableitung von  $f$ . Wir haben einen Spezialfall (etwa  $R = \infty$  und natürlich alles reell) im letzten Semester bewiesen (I.F1.11). Der Beweis im allgemeinen Fall ist sehr ähnlich. Es folgt natürlich, dass  $f$  auf  $B_R(0)$  stetig ist; einen einfacheren Beweis der Stetigkeit wird in der Übung gemacht.

**Bemerkung C3.19.** Sei  $T$  die Menge aller  $z \in \mathbb{C}$ , für die  $f(z) = \sum a_k z^k$  konvergiert. Dann gilt  $B_R(0) \subset T \subset \overline{B_R(0)}$  aber viel mehr können wir nicht sagen, weil Satz C3.16 im Falle  $|z| = R$  gar nichts sagt. Obwohl  $f$  auf  $B_R(0)$  stetig ist, kann es sein, dass  $f$  auf  $T$  unstetig ist. (Der Abel'sche Grenzwertsatz sagt, im reellen Fall ist  $f$  stetig. Das heißt, falls  $a_k \in \mathbb{R}$ , dann ist  $f$  stetig auf  $T \cap \mathbb{R} \subset [-R, R]$ .)

**Beispiel C3.20.** Nehmen wir z.B.  $a_k = 1/k!$ ; die Potenzreihe  $\sum z^k/k!$  hat  $R = 1$ . Für  $z = 1$  ist die Reihe divergent; für  $z = -1$  ist sie konvergent (aber nicht absolut). Der Abel'sche Grenzwertsatz ist hier anwendbar und garantiert (vgl. I.F6.18), dass

$$1 - 1/2 + 1/3 - \dots = \log 2.$$

c. Elementare komplexe Funktionen

Im letzten Semester konstruierten wir die Exponentialfunktion  $\exp: \mathbb{R} \rightarrow \mathbb{R}$  und die trigonometrischen Funktionen  $\sin, \cos: \mathbb{R} \rightarrow \mathbb{R}$  durch Potenzreihen mit unendlichem Konvergenzradius. Diese Potenzreihen können wir auch komplex betrachten.

**Definition C3.21.** Wir definieren die komplexe Exponentialfunktion  $\exp: \mathbb{C} \rightarrow \mathbb{C}$  und die komplexen trigonometrischen

Funktionen  $\cos, \sin: \mathbb{C} \rightarrow \mathbb{C}$  durch

$$\begin{aligned} \exp z &:= \sum_{k=0}^{\infty} \frac{z^k}{k!} = 1 + z + \frac{z^2}{2} + \frac{z^3}{3!} + \frac{z^4}{4!} + \dots, \\ \cos z &:= \sum_{j=0}^{\infty} (-1)^j \frac{z^{2j}}{(2j)!} = 1 - \frac{z^2}{2} + \frac{z^4}{4!} - \dots, \\ \sin z &:= \sum_{j=0}^{\infty} (-1)^j \frac{z^{2j+1}}{(2j+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots \end{aligned}$$

Aus diesen drei, definieren wir die weiteren Funktionen  $\cosh, \sinh, \tanh, \tan$  (usw.) genau wie im reellen Fall:

$$\begin{aligned} \cosh z &:= \frac{1}{2}(\exp z + \exp(-z)), & \tanh z &:= \sinh z / \cosh z, \\ \sinh z &:= \frac{1}{2}(\exp z - \exp(-z)), & \tan z &:= \sin z / \cos z. \end{aligned}$$

**Satz C3.22 (Euler'sche Identität).** Für  $z \in \mathbb{C}$  gilt

$$\exp iz = \cos z + i \sin z.$$

*Beweis.* Wir bekommen eine Teilfolge der Partialsummen (die ebenfalls gegen  $\exp iz$  konvergiert), indem wir Klammern in die Reihe setzen (I.C1.5):

$$\begin{aligned} \exp iz &:= \sum_{k=0}^{\infty} \frac{i^k z^k}{k!} = (1 + iz) + \left( \frac{i^2 z^2}{2} + \frac{i^3 z^3}{3!} \right) + \dots \\ &= \sum_{j=0}^{\infty} i^{2j} \left( \frac{z^{2j}}{(2j)!} + i \frac{z^{2j+1}}{(2j+1)!} \right). \end{aligned}$$

Weil die Reihen für  $\cos z$  und  $\sin z$  beide konvergieren, können wir die Rechenregeln anwenden, um eine lineare Kombination zu summieren:

$$\begin{aligned} \cos z + i \sin z &= \sum_{j=0}^{\infty} (-1)^j \frac{z^{2j}}{(2j)!} + i \sum_{j=0}^{\infty} (-1)^j \frac{z^{2j+1}}{(2j+1)!} \\ &= \sum_{j=0}^{\infty} (-1)^j \left( \frac{z^{2j}}{(2j)!} + i \frac{z^{2j+1}}{(2j+1)!} \right). \quad \square \end{aligned}$$

*Bemerkung C3.23.* Weil auch im komplexen Fall  $\cos$  gerade ist und  $\sin$  ungerade, gilt

$$\exp(-iz) = \cos z - i \sin z.$$

Damit gelten auch

$$\cos z = \cosh iz, \quad \sin z = i \sinh iz.$$

Meistens betrachtet man die Euler'sche Identität für  $z = \theta \in \mathbb{R}$ . Dann gelten

$$\operatorname{Re} \exp(i\theta) = \cos \theta, \quad \operatorname{Im} \exp(i\theta) = \sin \theta$$

und mit  $|\exp i\theta| = 1$  liegt  $\exp i\theta$  auf dem Einheitskreis in der Gauß'schen Zahlenebene.

**Satz C3.24 (Cauchyproduktformel).** Seien  $\sum a_k$  und  $\sum b_k$  zwei absolut konvergente Reihen. Definieren wir

$$c_k := \sum_{i+j=k} a_i b_j = \sum_{i=0}^k a_i b_{k-i} = a_0 b_k + \dots + a_k b_0,$$

dann konvergiert  $\sum c_k$  mit Reihensumme  $(\sum a_k)(\sum b_k)$ .

*Bemerkung C3.25.* Allgemeiner könnten wir die sogenannte *Doppelreihe*

$$\sum_{i,j \in \mathbb{N}} a_i b_j$$

betrachten. Weil sie absolut konvergent ist, konvergiert sie in jeder Umordnung, d.h. für jede Abzählung von  $\mathbb{N} \times \mathbb{N}$ . Hier betrachten wir nur die diagonale Abzählung (vgl. I.G1.7), die in der Definition von  $c_k$  eingebaut ist.

*Beweis.* Wir betrachten die Partialsummen

$$s_n := \sum_0^n a_k, \quad t_n := \sum_0^n b_k, \quad u_n := \sum_0^n c_k.$$

Wir müssen zeigen,  $\lim s_n t_n - u_n = 0$ . Dazu betrachten wir ein Quadrat und ein Dreieck in  $\mathbb{N}^2$ :

$$\square_n := \{(i, j) \in \mathbb{N}^2 : i, j \leq n\} \supset \triangle_n := \{(i, j) \in \mathbb{N}^2 : i + j \leq n\}.$$

Dann gelten

$$s_n t_n = \sum_{(i,j) \in \square_n} a_i b_j, \quad u_n = \sum_{(i,j) \in \triangle_n} a_i b_j.$$

Damit gilt

$$\begin{aligned} |s_n t_n - u_n| &= \left| \sum_{\square_n \setminus \triangle_n} a_i b_j \right| \leq \sum_{\square_n \setminus \triangle_n} |a_i| |b_j| \\ &= \sum_{\square_n} |a_i| |b_j| - \sum_{\triangle_n} |a_i| |b_j| \\ &\leq \sum_{\square_n} |a_i| |b_j| - \sum_{\square_{\lfloor n/2 \rfloor}} |a_i| |b_j|, \end{aligned}$$

wo im letzten Schritt wir  $\square_{\lfloor n/2 \rfloor} \subset \triangle_n$  benutzen.

Wir wissen aber auch, dass  $\sum a_k$  und  $\sum b_k$  absolut konvergieren. Seien

$$A_n := \sum_0^n |a_k|, \quad B_n := \sum_0^n |b_k|$$

die Partialsummen. Es gilt

$$A_n B_n = \sum_{\square_n} |a_i| |b_j|.$$

Weil  $(A_n)$  und  $(B_n)$  konvergieren, konvergiert auch die Produktfolge  $(A_n B_n)$  und sie ist deshalb Cauchyfolge. Es folgt,

$$\sum_{\square_n} |a_i| |b_j| - \sum_{\square_{\lfloor n/2 \rfloor}} |a_i| |b_j| = A_n B_n - A_{\lfloor n/2 \rfloor} B_{\lfloor n/2 \rfloor} \rightarrow 0.$$

Damit gilt  $\lim s_n t_n - u_n = 0$ , wie gewünscht. □



Ende der Vorlesung 2008 Dezember 9

**Korollar C3.26.** Für alle  $z, w \in \mathbb{C}$  gilt

$$\exp(z + w) = \exp z \exp w.$$

*Beweis.* Seien  $a_k := z^k/k!$  und  $b_k := w^k/k!$ . Wir wenden die Cauchyproduktformel an. Es gilt aus dem binomischen Lehrsatz,

$$c_k := \sum_{i+j=k} a_i b_j = \frac{1}{k!} \sum_{i=0}^k \binom{k}{i} z^i w^{k-i} = \frac{1}{k!} (z + w)^k.$$

Damit gilt

$$\exp(z + w) = \sum c_k = \left( \sum a_k \right) \left( \sum b_k \right) = \exp z \exp w. \quad \square$$

*Bemerkung C3.27.* Im letzten Semester haben wir diese Gleichung (für  $z, w \in \mathbb{R}$ ) aus der Differentialgleichung  $\exp' = \exp$  bewiesen. (Eigentlich geht dies auch komplex, sobald man komplexe Differentiation kennt.) Der Beweis hier aus der Potenzreihe ist unabhängig.

**Definition C3.28.** Für  $0 < a \in \mathbb{R}$  und  $z \in \mathbb{C}$  definieren wir  $a^z := \exp(z \log a)$ . Insbesondere gilt  $e^z = \exp z$ .

*Bemerkung C3.29.* Diese Definition ist deshalb vernünftig, weil mit Korollar C3.26 die normalen Rechenregeln für Potenzen gelten:

$$a^{z+w} = a^z a^w, \quad (ab)^z = a^z b^z, \quad (a^x)^z = a^{xz}$$

für alle  $a, b, x \in \mathbb{R}$  und  $z \in \mathbb{C}$ . Es gilt auch  $\overline{a^z} = a^{\overline{z}}$ .

**Definition C3.30.** Betrachten wir nochmal die Euler'sche Identität. Sei

$$z := e^{\rho+i\theta} = e^\rho e^{i\theta} = e^\rho (\cos \theta + i \sin \theta) =: r(\cos \theta + i \sin \theta)$$

für  $\rho, \theta \in \mathbb{R}$ . Diese Darstellung der komplexen Zahl  $z$  heißt *Polarform*. Der Radius  $r$  ist der Betrag  $|z| = r = e^\rho \geq 0$ . Der Winkel  $\theta$  heißt *Argument* von  $z$ . Natürlich gilt  $re^{i\theta} = re^{i(\theta+2\pi)}$ . Das heißt, die Polarform ist nicht eindeutig. Um das Argument einer komplexen Zahl eindeutig zu definieren, wählen wir immer  $\theta = \arg z \in (-\pi, \pi]$ . Dann ist aber  $\arg: \mathbb{C} \setminus \{0\} \rightarrow \mathbb{R}$  entlang der negativen reellen Achse unstetig.

*Bemerkung C3.31.* In Polarform gelten

$$\operatorname{Re} re^{i\theta} = r \cos \theta, \quad \operatorname{Im} re^{i\theta} = r \sin \theta, \quad \overline{re^{i\theta}} = re^{-i\theta}.$$

*Bemerkung C3.32.* Sonderfälle der Euler'schen Identität liefern schöne Formeln für bestimmten komplexen Zahlen auf dem Einheitskreis:

$$e^{2\pi i} = 1, \quad e^{\pi i} = -1, \quad e^{\pi i/2} = i.$$

Es wird gesagt,  $e^{\pi i} + 1 = 0$  sei die schönste Formel in der Mathematik, weil sie die fünf wichtigsten Zahlen verbindet.

*Bemerkung C3.33.* Jetzt können wir das Additionstheorem (I.F3.8) für  $\cos$  bzw.  $\sin$  neu herleiten als Realteil bzw. Imaginärteil der Gleichung

$$\begin{aligned} \cos(x + y) + i \sin(x + y) &= e^{i(x+y)} = e^{ix} e^{iy} = (\cos x + i \sin x)(\cos y + i \sin y) \\ &= (\cos x \cos y - \sin x \sin y) + i(\sin x \cos y + \sin y \cos x) \end{aligned}$$

für  $x, y \in \mathbb{R}$ . (Die Formeln gelten auch für nichtreelle  $x, y$ .)

Im reellen Fall wurde Logarithmus als Umkehrfunktion der injektiven Exponentialfunktion definiert. Die komplexe Exponentialfunktion ist nicht mehr injektiv ( $e^0 = 1 = e^{2\pi i}$ ).

**Definition C3.34.** Der komplexe Logarithmus  $\mathbb{C} \setminus \{0\} \rightarrow \mathbb{C}$  wird durch

$$\log z := \log |z| + i \arg z$$

definiert; diese Funktion ist auf der negativen reellen Achse unstetig, sonst stetig.

*Bemerkung C3.35.* Es gilt  $e^{\log z} = z$ . Umgekehrt gilt  $\log e^z = z$  nur dann, wenn  $z$  nah an der reellen Achse liegt. Im Allgemeinen gilt  $\log e^z = z + 2\pi i n$  mit  $n \in \mathbb{Z}$  so gewählt, dass  $\operatorname{Im} z + 2\pi n \in (-\pi, \pi]$ . (D.h.,  $n = -\lfloor \operatorname{Im} z / 2\pi + 1/2 \rfloor$ .)

**Definition C3.36.** Sei  $n \in \mathbb{N}^+$  und sei  $\omega := e^{2\pi i/n}$ . Dann gilt  $\omega^n = 1$ . Für jedes  $k \in \mathbb{Z}$  gilt  $(\omega^k)^n = 1$ . Die  $n$  Nullstellen von  $z^n = 1$ , die *n-ten Einheitswurzeln*, sind  $\omega^k = e^{2\pi i k/n}$  für  $k = 0, 1, \dots, n-1$ .

*Bemerkung C3.37.* Sei  $0 \neq z \in \mathbb{C}$ . Setzen wir  $w := e^{(\log z)/n}$ , dann gilt  $w^n = z$ . Die anderen  $n$ -ten Wurzeln aus  $z$  sind  $w$  mal die Einheitswurzeln:  $\omega^k w$ .

#### C4. Polynome

Bezeichne  $\mathbb{K}$  in diesem Abschnitt entweder den Körper  $\mathbb{R}$  oder den Körper  $\mathbb{C}$ .

**Definition C4.1.** Ein *Polynom* (über  $\mathbb{K}$ ) ist eine Abbildung  $P: \mathbb{K} \rightarrow \mathbb{K}$ , die sich entweder als  $P(z) \equiv 0$  oder wie folgt darstellen lässt:

$$P(z) = \sum_{j=0}^n c_j z^j,$$

wobei  $n \in \mathbb{N}$  und die *Koeffizienten*  $c_j$  Zahlen aus  $\mathbb{K}$  sind mit  $c_n \neq 0$ . Hier heißt  $n \in \mathbb{N}$  der *Grad*  $\deg(P)$  von  $P$ ; wir definieren  $\deg(0) = -\infty$ . Eine Zahl  $z \in \mathbb{K}$  mit  $P(z) = 0$  heißt *Nullstelle* von  $P$ .

*Bemerkung C4.2.* Summe und Produkt zweier Polynome sind Polynome. Die Vereinbarung  $\deg 0 = -\infty$  ist sinnvoll: Nur deshalb gelten die Regeln

$$\deg(fg) = \deg f + \deg g, \quad \deg(f + g) \leq \max(\deg f, \deg g)$$

auch dann, wenn  $f = 0$  bzw.  $g = 0$ .

**Lemma C4.3.** Sei  $P(z) = \sum_0^n c_j z^j$  ein Polynom mit  $c_n = 1$  und sei  $C := \sum |c_j|$ . Dann hat  $P$  keine Nullstellen außerhalb  $B_C(0)$ .

*Beweis.* Sei  $|z| \geq C$ . Für  $j < n$  gilt  $|z|^j \leq |z|^{n-1}$  weil  $|z| \geq 1$ . Deshalb gilt nach der Dreiecksungleichung

$$|P(z) - z^n| \leq \sum_{j=0}^{n-1} |c_j| |z|^j < C |z|^{n-1}.$$

Daher gilt

$$|P(z)| > |z|^n - C |z|^{n-1} = (|z| - C) |z|^{n-1} \geq 0. \quad \square$$

**Korollar C4.4.** Die Darstellung eines Polynoms ist eindeutig. (Damit ist der Grad wohldefiniert).

*Beweis.* Die Differenz zweier unterschiedlicher Darstellungen ergäbe eine Darstellung  $\sum c_j z^j$  (mit  $c_n \neq 0$ ) für das Nullpolynom, was dem Lemma widerspricht, sobald man durch  $c_n$  teilt.  $\square$

---

Ende der Vorlesung 2008 Dezember 14

**Bemerkung C4.5.** Tatsächlich ist ein Polynom von Grad  $n$  schon durch seine Werte  $P(z_j)$  in je  $n + 1$  Punkten  $z_j \in \mathbb{K}$  bestimmt.

**Bemerkung C4.6.** Wenn man in der abstrakten Algebra Polynome über einem beliebigen Ring betrachtet, stimmt das Korollar nicht mehr. Deswegen muss man zwischen *Polynomfunktionen* (unsere Polynome) und *Polynomen* (etwa unsere Darstellungen) unterscheiden.

Wir haben  $\mathbb{C}$  konstruiert, um Wurzeln ziehen zu können, und haben gesehen, dass jedes quadratische Polynom über  $\mathbb{C}$  eine Nullstelle hat. Tatsächlich sagt der Fundamentalsatz der Algebra, jedes komplexe Polynom hat eine Nullstelle.

**Lemma C4.7.** Sei  $P$  ein Polynom über  $\mathbb{C}$  und sei  $w \in \mathbb{C}$  mit  $P(w) \neq 0$ . Dann gibt es ein  $z \in \mathbb{C}$  so, dass  $|P(z)| < |P(w)|$ .

*Beweis.* Wir dürfen annehmen,  $w = 0$ , indem wir  $P(z)$  mit  $Q(z) := P(z + w)$  ersetzen. Nun sei  $a := P(0) \neq 0$ . Es gibt  $0 \neq b \in \mathbb{C}$  und  $k \in \mathbb{N}^+$  so, dass wir  $P$  in der Form

$$P(z) = a + bz^k(1 + c_1z + \dots + c_nz^n)$$

(mit  $c_1, \dots, c_n \in \mathbb{C}$ ) schreiben können. Für die stetige reelle Funktion  $g(r) := 1 - \sum_1^n |c_j| r^j$  gilt  $g(0) > 0$ . Dann gibt es  $\varepsilon > 0$  so, dass für  $|r| < \varepsilon$  gilt  $g(r) > 0$ .

Sei nun  $w \in \mathbb{C}$  eine  $k$ -te Wurzel aus  $-a/b$  und wähle  $\lambda \in (0, 1) \subset \mathbb{R}$  klein genug, dass  $z := \lambda w$  den Betrag  $r := |z| = \lambda |w| < \varepsilon$  hat. Dann gilt

$$a + bz^k = a + b\lambda^k w^k = a(1 - \lambda^k)$$

und damit

$$\begin{aligned} |P(z)| &\leq |a + bz^k| + |b|r^k \sum_1^n |c_j| r^j \\ &= |a| - \lambda^k |a| + |b|r^k \sum_1^n |c_j| r^j \\ &= |a| + |b|r^k \left(-1 + \sum_1^n |c_j| r^j\right) \\ &= |a| - |b|r^k g(r) < |a|. \quad \square \end{aligned}$$

**Satz C4.8 (Fundamentalsatz der Algebra).** Jedes nichtkonstante Polynom über  $\mathbb{C}$  hat eine Nullstelle.

*Beweis.* Sei  $P$  ein Polynom. Wir setzen  $a := |P(0)|$  und dürfen annehmen,  $a > 0$  (weil sonst 0 eine Nullstelle ist). Weil  $|P|: \mathbb{C} \rightarrow \mathbb{R}$  stetig ist, ist das Urbild  $A := \{z \in \mathbb{C} : |P(z)| \leq a\}$  von  $[0, a]$  kompakt. Auf  $A$  nimmt die stetige Funktion  $|P|$  deshalb ihre Minimum an: es gibt  $w \in A$  mit  $|P(w)| \leq |P(z)|$  für alle  $z \in A$ . Aus der Definition von  $A$  ist  $w$  dann ein globales Minimum für  $|P|$  auf  $\mathbb{C}$ . Aus dem Lemma folgt  $P(w) = 0$ .  $\square$

Aus der Schule kennen wir den Divisionsalgorithmus für Polynome:

**Satz C4.9.** Seien  $F$  und  $G$  Polynome über  $\mathbb{K}$  mit  $\deg G > 0$ . Dann gibt es eindeutig bestimmte Polynome  $Q$  (Quotient) und  $R$  (Rest) über  $\mathbb{K}$  mit

$$F = QG + R \quad \text{und} \quad \deg R < \deg G.$$

**Korollar C4.10.** Ist  $P \neq 0$  ein Polynom über  $\mathbb{K}$  und ist  $a \in \mathbb{K}$  eine Nullstelle, dann gibt es ein eindeutig bestimmtes Polynom  $Q$  mit  $P(z) = (z - a)Q(z)$ .

**Korollar C4.11.** Jedes komplexe Polynom von Grad  $n \geq 0$  hat  $n$  Nullstellen im Sinne, dass es  $a_1, \dots, a_n \in \mathbb{C}$  und  $c \in \mathbb{C}$  gibt mit

$$P(z) = c \prod_{j=1}^n (z - a_j).$$

*Beweis.* Wir benutzen Induktion über  $n$ . Für  $n = 0$  ist  $P \equiv c$  eine Konstante. Für  $n > 0$  hat  $P$  aus dem Fundamentalsatz eine Nullstelle  $a_n$ . Mit dem letzten Korollar können wir  $P(z) = (z - a_n)Q(z)$  schreiben, wobei  $\deg Q = n - 1$ . Aus der Induktionsvoraussetzung folgt dann

$$Q(z) = c \prod_{j=1}^{n-1} (z - a_j). \quad \square$$

**Bemerkung C4.12.** Sei  $P(z) = \sum c_j z^j$  ein komplexes Polynom. Sei  $\bar{P}$  das konjugierte Polynom  $\bar{P}(z) = \sum \bar{c}_j z^j$ . Dann gilt  $\overline{P(z)} = \bar{P}(\bar{z})$ . Ist  $a$  eine Nullstelle von  $P$ , dann ist  $\bar{a}$  eine Nullstelle von  $\bar{P}$ .

Im Falle, dass die Koeffizienten reell sind ( $c_j \in \mathbb{R}$ ), gilt  $\bar{P} = P$ . Für jede nichtreelle Nullstelle  $a \in \mathbb{C}$  ist  $\bar{a}$  auch eine Nullstelle. Das heißt  $(z - a)(z - \bar{a}) = z^2 - 2 \operatorname{Re} a + |a|^2$  teilt  $P$ . Sind  $a_j, \bar{a}_j$  die nichtreellen und  $b_k$  die reellen Nullstellen, dann gilt

$$P(z) = c \prod (z - b_k) \prod (z^2 - 2 \operatorname{Re} a_j + |a_j|^2).$$

**D. VEKTORWERTIGE FUNKTIONEN AUF EINEM INTERVALL**

In diesem Abschnitt sei  $W$  ein Banachraum, d.h. ein vollständiger normierter Vektorraum. Fast alle Sätze über reelle Reihen (aus I.C) konnten wir leicht auf Reihen in  $W$  erweitern (im Abschnitt C3).

Ähnlich der Reihensumme sind auch Ableitungen bzw. Integrale als Grenzwerte von linearen Operationen (Differenzenquotienten bzw. Riemann'schen Summen) definiert. Damit können wir deren Definitionen leicht auf vektorwertige Funktionen übertragen.

**D1. Ableitungen vektorwertiger Funktionen**

Sei  $I \subset \mathbb{R}$  ein Intervall und sei  $f : I \rightarrow W$  eine Abbildung, eine vektorwertige Funktion auf  $I$ .

*Bemerkung D1.1.* Falls  $f : I \rightarrow W$  stetig ist, ist  $f$  ein Weg oder eine (parametrisierte) Kurve im Raum  $W$ .

Wir können die Ableitung von  $f$  in  $p \in I$  genau so definieren, wie im Falle  $W = \mathbb{R}$ :

**Definition D1.2.** [Vgl. I.E1.1.] Die Abbildung  $f : I \rightarrow W$  heißt differenzierbar in  $p \in I$ , falls der Limes

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

existiert. Dann heißt  $f'(p)$  die Ableitung von  $f$  in  $p$ .

*Beispiel D1.3.* [Vgl. I.E1.2.] Für  $v, w \in W$  heißt  $f(t) := tv + w$  eine affin lineare Funktion  $\mathbb{R} \rightarrow W$ . Sie hat in jedem  $t \in \mathbb{R}$  die Ableitung  $f'(t) = v$ .

*Bemerkung D1.4.* [Vgl. I.E1.5.] Falls  $f : I \rightarrow W$  differenzierbar in  $p$  ist, dann ist  $f$  in  $p$  stetig, weil für  $x \rightarrow p$  gilt  $f(x) - f(p) \rightarrow f'(p) \cdot 0$ .

*Bemerkung D1.5.* Die Ableitung  $f'(t) \in W$  heißt auch Tangentenvektor der parametrisierten Kurve  $f$  im Punkt  $f(t)$ . Wenn man den Parameter  $t$  als Zeit betrachtet, ist  $\dot{f}(t) = f'(t)$  die Geschwindigkeit der Kurve. Die Norm  $\|\dot{f}(t)\|$  nennt man auch Schnelligkeit der Kurve.

**Lemma D1.6.** [Vgl. I.E1.17.] Eine Funktion  $f : I \rightarrow W$  ist genau dann differenzierbar in  $p \in I$ , wenn es eine Funktion  $f_1 : I \rightarrow W$  so gibt, dass  $f(x) = f(p) + (x-p)f_1(x)$  und  $f_1$  in  $p$  stetig ist. Es gilt dann  $f'(p) = f_1(p)$ .

*Beweis.* Für  $x \neq p$  ist  $f_1(x)$  nichts anders als der Differenzenquotient  $\frac{f(x)-f(p)}{x-p}$ . Diese Funktion hat genau dann eine stetige Erweiterung auf  $x = p$ , wenn die Ableitung  $f'(p)$  existiert. □

**Satz D1.7.** [Vgl. I.E1.9.] Seien  $f, g : I \rightarrow W$  und  $h : I \rightarrow \mathbb{R}$  differenzierbar in  $p \in I$ . Dann sind auch  $f \pm g$  und  $hf$  in  $p$  differenzierbar und es gelten die Rechenregeln

$$(f \pm g)'(p) = f'(p) \pm g'(p), \quad (hf)'(p) = h'(p)f(p) + h(p)f'(p).$$

*Beweis.* Weil die Regeln für  $f \pm g$  noch einfacher sind, betrachten wir nur die Produktregel. Wir benutzen die Funktionen  $f_1 : I \rightarrow W$  und  $h_1 : I \rightarrow \mathbb{R}$  aus dem Lemma. Aus

$$f(x) = f(p) + (x-p)f_1(x), \quad h(x) = h(p) + (x-p)h_1(x)$$

folgt

$$(hf)(x) = (hf)(p) + (x-p)s_1(x)$$

mit

$$s_1(x) := h(p)f_1(x) + h_1(x)f(p) + (x-p)f_1(x)h_1(x).$$

Die Funktion  $s_1$  ist in  $p$  stetig mit Wert  $s_1(p) = h(p)f_1(p) + h_1(p)f(p)$ . Aus dem Lemma (nochmal) folgt die Produktregel. □

**Satz D1.8.** [Vgl. I.E1.19.] Seien  $I, J \subset \mathbb{R}$  Intervalle und seien  $f : I \rightarrow W$  und  $g : J \rightarrow I$  differenzierbare Abbildungen. Dann gilt die Kettenregel

$$(f \circ g)'(p) = f'(g(p))g'(p).$$

*Beweis.* (Aufgabe.) □

**D2. Endlich-dimensionale Banachräume**

Nun sei  $W$  endlich-dimensional,  $\dim W = n$ . Jede Basis  $\{w_1, \dots, w_n\}$  für  $W$  gibt uns durch

$$(a_1, \dots, a_n) \mapsto a_1w_1 + \dots + a_nw_n = \sum a_iw_i$$

einen Isomorphismus  $\mathbb{R}^n \rightarrow W$ .

Wir wissen schon, alle  $\ell^p$ -Normen auf  $\mathbb{R}^n$  sind äquivalent (Lemma B4.7). Darüber hinaus gilt:

**Lemma D2.1.** Alle Normen auf einem endlich-dimensionalen Raum  $W$  sind äquivalent.

*Beweis.* Wir dürfen annehmen,  $W = \mathbb{R}^n$ . Sei  $\{e_i\}$  die kanonische Basis für  $\mathbb{R}^n$  und sei  $\|\cdot\|$  eine beliebige Norm. Es reicht aus zu zeigen, dass diese Norm und die  $\ell^1$ -Norm äquivalent sind.

Mit  $M := \max_i \|e_i\|$  gilt aus der Dreiecksungleichung

$$\|v\| = \|(v_1, \dots, v_n)\| \leq \sum |v_i| \|e_i\| \leq M \sum |v_i| = M\|v\|_1.$$

Es folgt, dass

$$\| \|v\| - \|w\| \| \leq M \|v - w\|_1.$$

Das heißt, die Funktion  $\mathbb{R}^n \rightarrow \mathbb{R}, v \mapsto \|v\|$  ist Lipschitz bezüglich der  $\ell^1$ -Metrik, insbesondere stetig bezüglich der Standardtopologie auf  $\mathbb{R}^n$ .

Damit hat  $\|\cdot\|$  auf der kompakten  $\ell^1$ -Einheitssphäre

$$S_1 := \{v \in \mathbb{R}^n : \|v\|_1 = 1\}$$

ein Minimum  $m$ . Weil  $\|v\| > 0$  für  $v \neq 0$  (insbes. für  $v \in S_1$ ), gilt  $m > 0$ . Es folgt, dass  $\|v\| \geq m\|v\|_1$  für alle  $v \in \mathbb{R}^n$ . Mit

$$m\|v\|_1 \leq \|v\| \leq M\|v\|_1$$

sind  $\|\cdot\|$  und  $\|\cdot\|_1$  äquivalent. □

**Korollar D2.2.** *Ein endlich-dimensionaler Vektorraum  $W$  ist mit jeder Norm vollständig, d.h. ein Banachraum.*  $\square$

**Korollar D2.3.** *Sei  $W$  ein endlich-dimensionaler Vektorraum und  $f: I \rightarrow W$  eine Abbildung. Die Differenzierbarkeit von  $f$  und die Ableitungen  $f'(p)$  sind unabhängig von der Norm auf  $W$ .*  $\square$

**Bemerkung D2.4.** Das heißt, die (Vektor-)Geschwindigkeit  $\gamma'(t)$  einer Kurve  $\gamma$  im endlich-dimensionalen Raum  $W$  ist unabhängig von der Norm  $\|\cdot\|$  auf  $W$  hängt natürlich deren Norm  $\|\gamma'(t)\|$  – die Schnelligkeit von  $\gamma$  – hängt natürlich von der ausgewählten Norm ab.

**Satz D2.5.** *Sei  $f: I \rightarrow \mathbb{R}^n$  eine Abbildung,  $f = (f_1, \dots, f_n)$ . Die Abbildung  $f$  ist genau dann in  $p$  differenzierbar, wenn jedes  $f_i: I \rightarrow \mathbb{R}$  in  $p$  differenzierbar ist. Es gilt*

$$f'(p) = (f'_1(p), \dots, f'_n(p)).$$

*Beweis.* Die Ableitung ist per Definition ein Limes. Mit Lemma B6.31 können wir diesen Limes komponentenweise betrachten.  $\square$

Später werden wir Ableitungen auch für Funktionen auf einem Banachraum  $V$  betrachten. Wegen folgenden Satzes ist es wichtig, dass  $V$  endlich-dimensional ist.

**Satz D2.6.** *Seien  $V$  und  $W$  Banachräume und sei  $L: V \rightarrow W$  eine lineare Abbildung. Falls  $V$  endlich-dimensional ist, ist  $L$  eine Lipschitzabbildung (und damit stetig).*

*Beweis.* Falls  $\dim V = n$ , dürfen wir annehmen,  $V = \mathbb{R}^n$  mit der  $\ell^1$ -Norm. Sei  $w_i := L(e_i) \in W$  für  $i = 1, \dots, n$  und sei  $\lambda := \max \|w_i\|$ . Für  $v = (v_1, \dots, v_n)$  gilt dann

$$\|L(v)\| = \left\| \sum v_i w_i \right\| \leq \sum |v_i| \|w_i\| \leq \lambda \sum |v_i| = \lambda \|v\|_1.$$

Damit ist  $\lambda$  eine Lipschitzkonstante für  $L$ .  $\square$

---

Ende der Vorlesung 2009 Januar 5

---

### D3. Mehr über Ableitungen

**Definition D3.1.** Sei  $f: I \rightarrow W$  differenzierbar. Genau wie im reellen Fall heißt  $f$  stetig differenzierbar, falls  $f': I \rightarrow W$  stetig ist.

**Bemerkung D3.2.** Für Ableitungen reeller Funktionen – auch wenn diese unstetig sind – hatten wir einen Zwischenwertsatz (I.F4.1). Dieser macht keinen Sinn mehr für vektorwertige Funktionen, weil  $f'(I)$  eine Teilmenge von  $W$  (und nicht von  $\mathbb{R}$ ) ist.

**Beispiel D3.3.** Der Mittelwertsatz (I.E2.4) hingegen macht immer noch Sinn für vektorwertige Funktionen: wir können die Frage stellen, ob es  $c \in [a, b]$  gibt mit  $f'(c) = \frac{f(b)-f(a)}{b-a}$ . Die Antwort ist aber „nein“, der Mittelwertsatz stimmt einfach nicht mehr. Sei z.B.  $f(t) := (\cos t, \sin t) \in \mathbb{R}^2$  für  $t \in [0, 2\pi]$ . Es gilt  $f(2\pi) = f(0)$ ; weil  $f'(t) = (-\sin t, \cos t)$ , gibt es aber kein  $t$  mit  $f'(t) = (0, 0)$ .

Folgender Satz ist im Falle  $W = \mathbb{R}$  ein unmittelbares Korollar des Mittelwertsatzes.

**Satz D3.4.** *Sei  $f: [a, b] \rightarrow W$  stetig und in jedem inneren Punkt differenzierbar. Setze  $M := \sup\{\|f'(t)\| : t \in (a, b)\}$ . Dann gilt  $\|f(b) - f(a)\| \leq M(b - a)$ .*

*Beweis.* Im Falle  $M = \infty$  gibt es nichts zu beweisen. Sonst sei  $c \in (a, b)$  und sei  $M < C \in \mathbb{R}$ . Wir setzen

$$S := \{t \in [c, b] : \|f(t) - f(c)\| \leq C(t - c)\}$$

und  $s := \sup S$ . Weil  $S$  nicht leer ist ( $c \in S$ ), gilt  $s \in [c, b]$ . Weil  $f$  stetig ist, ist  $S$  abgeschlossen und damit gilt  $s \in S$ . Wir behaupten,  $s = b$ . Dann gilt  $\|f(b) - f(c)\| \leq C(b - c)$  und zwar für jedes  $C > M$ , also  $\|f(b) - f(c)\| \leq M(b - c)$ . Weil dies für jedes  $c \in (a, b)$  gilt und  $f$  stetig ist, gilt

$$\|f(b) - f(a)\| \leq M(b - a).$$

Wir beweisen die Behauptung indirekt. Falls  $s < b$ , sei  $t \in (s, b)$ . Aus der Dreiecksungleichung gilt

$$\|f(t) - f(c)\| \leq \|f(t) - f(s)\| + C(s - c).$$

Aus der Differenzierbarkeit von  $f$  in  $s$  folgt, dass

$$\frac{\|f(t) - f(s)\|}{t - s} \rightarrow \|f'(s)\| < C$$

für  $t \rightarrow s$ . Deshalb gibt es  $\delta > 0$  so, dass für  $s < t < s + \delta$  gilt

$$\|f(t) - f(s)\| \leq C(t - s)$$

und damit

$$\|f(t) - f(c)\| \leq C(t - c).$$

Das heißt,  $t \in S$ , was die Definition von  $s$  widerspricht.  $\square$

**Korollar D3.5.** [Vgl. I.E2.8.] *Seien  $f, g: I \rightarrow W$  differenzierbar mit  $f' \equiv g'$ . Dann gibt es  $w \in W$  mit  $f(x) - g(x) \equiv w$ .*

*Beweis.* Sei  $h = f - g$ . Es gilt  $h' \equiv 0$ . Für alle  $a < b \in I$  sagt der Satz,  $\|h(b) - h(a)\| \leq 0$ , d.h.,  $h(b) = h(a)$ .  $\square$

**Definition D3.6.** Die höheren Ableitungen einer vektorwertigen Funktion  $f: I \rightarrow W$  definieren wir genau wie im reellen Fall:

$$f^{(k+1)} := (f^{(k)})';$$

Taylorpolynome bzw. Taylorreihen auch so:

$$\sum_k f^{(k)}(p) \frac{(x - p)^k}{k!}.$$

**D4. Integrale vektorwertiger Funktionen**

Seien  $[a, b] \subset \mathbb{R}$  ein kompaktes Intervall,  $X$  ein metrischer Raum und  $f: [a, b] \rightarrow X$  eine Funktion. Treppen- bzw. Regelfunktionen definiert man genau wie im Falle  $X = \mathbb{R}$ .

**Definition D4.1.** Die Funktion  $f: [a, b] \rightarrow X$  heißt *Treppenfunktion*, falls sie stückweise konstant ist, d.h. konstant auf jedem offenen Teilintervall  $(z_{j-1}, z_j)$  einer Treppenzерlegung  $Z$ . Sie heißt *Regelfunktion*, falls alle einseitigen Grenzwerte  $f(p^\pm)$  existieren.

**Satz D4.2.** [Vgl. A3.5.] *Sei  $X$  vollständig. Eine Funktion  $f: [a, b] \rightarrow X$  ist genau dann eine Regelfunktion, wenn es eine Folge  $(\varphi_k)$  von Treppenfunktionen gibt, die gegen  $f$  gleichmäßig konvergiert.*

*Beweisidee.* Der Beweis des Satzes – und die Beweise der zwei Lemmata – gelten ungeändert, wenn man einfach alle Begriffe der Form  $|f(x) - f(y)|$  mit  $d(f(x), f(y))$  ersetzt. Die Vollständigkeit wird (nur) am Ende des zweiten Lemmas benutzt, um einen einseitigen Grenzwert zu finden.  $\square$

Obwohl wir diesen Satz allgemein für vollständige metrische Räume bewiesen haben, haben wir nur an dem Fall Interesse, wo  $X = W$  ein Banachraum ist. Nur in einem Vektorraum können wir Riemann'sche Summen als lineare Kombinationen definieren.

**Definition D4.3.** [Vgl. A4.1.] Sei  $f: [a, b] \rightarrow W$  eine Funktion und  $Z$  eine markierte Zerlegung von  $[a, b]$ . Die *Riemann'sche Summe* von  $f$  bezüglich  $Z$  ist

$$S(f, Z) := \sum_{j=1}^k \Delta z_j f(\xi_j) = \sum_{j=1}^k (z_j - z_{j-1}) f(\xi_j).$$

Sei  $\varphi: [a, b] \rightarrow W$  eine Treppenfunktion. Das (*Treppen-*)*Integral* von  $\varphi$  ist

$$\int_a^b \varphi := S(\varphi, Z),$$

wo  $Z$  eine beliebige Treppenzерlegung für  $\varphi$  ist.

*Bemerkung D4.4.* Dass das Treppnenintegral wohldefiniert ist, folgt aus einem Analogon zu Lemma A4.2, das wir nicht explizit aufschreiben.

Folgendes Lemma gilt genau wie im reellen Fall:

**Lemma D4.5.** [Vgl. A4.4.] *Es gelten Linearität und Intervalladditivität:*

$$\begin{aligned} \int_a^b (f + cg) &= \int_a^b f + c \int_a^b g, \\ \int_a^b f &= \int_a^p f + \int_p^b f. \end{aligned}$$

**Lemma D4.6.** *Es gilt*

$$\left\| \int_a^b f \right\| \leq \int_a^b \|f\|.$$

*Beweis.* Sei  $Z$  eine markierte Treppenzерlegung für  $f$ . Dann ist  $Z$  auch für die Treppenfunktion  $\|f\|: [a, b] \rightarrow \mathbb{R}$  eine Treppenzерlegung. Es gilt

$$\begin{aligned} \left\| \int_a^b f \right\| &= \left\| S(f, Z) \right\| = \left\| \sum \Delta z_j f(\xi_j) \right\| \\ &\leq \sum \|\Delta z_j f(\xi_j)\| = \sum \Delta z_j \|f(\xi_j)\| \\ &= S(\|f\|, Z) = \int_a^b \|f\|, \end{aligned}$$

wo die Ungleichung in der Mitte die Dreiecksungleichung für  $W$  ist.  $\square$

*Bemerkung D4.7.* Für reelle Integrale folgte diese Ungleichung aus der Monotonie ( $f \geq 0 \implies \int f \geq 0$ ). Für vektorwertige Integrale macht Monotonie natürlich keinen Sinn.

**Lemma D4.8.** [Vgl. A4.8.] *Sei  $f: [a, b] \rightarrow W$  eine Regelfunktion, und  $(\varphi_n)$  eine Folge von Treppenfunktionen, die gleichmäßig gegen  $f$  konvergiert. Dann konvergiert in  $W$  die Folge  $(\int_a^b \varphi_n)$ . Der Grenzwert ist unabhängig von der gewählten Folge  $(\varphi_n)$ ; wir nennen ihn das (Regel-)Integral  $\int_a^b f$ .*

*Beweisskizze.* Der Beweis, dass  $(\int_a^b \varphi_n)$  eine Cauchyfolge ist, geht genau wie im reellen Fall. Weil  $W$  vollständig ist, konvergiert diese Cauchyfolge.  $\square$

*Bemerkung D4.9.* Aus den obigen Eigenschaften des Treppnenintegrals – Linearität, Intervalladditivität und die Ungleichung  $|\int f| \leq \int |f|$  – folgen unmittelbar dieselben für das Regelintegral.

*Bemerkung D4.10.* Uneigentliche Integrale definiert man wie im reellen Fall als Grenzwerte eigentlicher Integrale. Man kann auch Riemann- bzw. Lebesgue- bzw. Gauge-Integrale für vektorwertige Funktionen definieren. Wie im reellen Fall ist jedes Regelintegral auch ein Riemann-Integral:

$$\int_a^b f = \lim_{\Delta(Z) \rightarrow 0} S(f, Z).$$

*Bemerkung D4.11.* [Vgl. A4.11.] Man kann natürlich den Mittelwert definieren

$$\int_a^b f := \frac{\int_a^b f}{\int_a^b 1} = \frac{1}{b-a} \int_a^b f,$$

der Mittelwertsatz der Integralrechnung (A4.12) gilt aber nicht für vektorwertige Funktionen. (Beispiel D3.3 kann man hier wieder verwenden.) Man hat aber immer noch eine wichtige Folgerung: Falls  $\|f\| \leq \varepsilon$  auf  $[a, b]$  gilt, dann gilt

$$\left\| \int_a^b f \right\| \leq \int_a^b \|f\| \leq \varepsilon.$$

Im Fall eines endlich-dimensionalen Banachraums  $W$  kann man natürlich komponentenweise integrieren:

**Satz D4.12.** *Eine Funktion*

$$f = (f_1, \dots, f_n): [a, b] \rightarrow \mathbb{R}^n$$

ist genau dann eine Regelfunktion, wenn jedes  $f_i$  eine Regelfunktion ist. Dann gilt

$$\int_a^b f = \left( \int_a^b f_1, \dots, \int_a^b f_n \right).$$

*Beweis.* (Aufgabe.) □

### D5. Der Hauptsatz

Unbestimmte Integrale, Stammfunktionen und der Hauptsatz der Differential- und Integralrechnung kann man auch leicht auf vektorwertige Funktionen übertragen.

**Definition D5.1.** [Vgl. A5.3.] Sei  $f: [a, b] \rightarrow W$  eine Regelfunktion und sei  $c \in [a, b]$ . Die Funktion  $F: [a, b] \rightarrow W$ , definiert durch

$$F(p) := \int_c^p f(x) dx,$$

heißt ein *unbestimmtes Integral* von  $f$ .

**Satz D5.2.** [Vgl. A5.4.] Sei  $f: [a, b] \rightarrow W$  eine Regelfunktion. Ein unbestimmtes Integral  $F$  von  $f$  ist stetig und hat in jedem  $p \in [a, b]$  einseitige Ableitungen  $F'_\pm(p) = f(p^\pm)$ . Falls  $f$  in  $p$  stetig ist, folgt, dass  $F$  in  $p$  differenzierbar ist mit  $F'(p) = f(p)$ .

*Beweisskizze.* Wie im reellen Fall gilt

$$F'_+(p) = \lim_{h \rightarrow 0^+} \int_p^{p+h} f.$$

Weil  $f$  eine Regelfunktion ist, existiert  $f(p^+)$ . D.h.,

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in (p, p + \delta) \quad \|f(p^+) - f(x)\| < \varepsilon.$$

Für  $h < \delta$  gilt dann aus Bemerkung D4.11, dass

$$\left\| f(p^+) - \int_p^{p+h} f \right\| < \varepsilon.$$

Es folgt, dass

$$f(p^+) = \lim_{h \rightarrow 0^+} \int_p^{p+h} f = F'_+(p).$$

Ähnliches gilt für die linksseitigen Ableitungen. □

**Definition D5.3.** [Vgl. A5.5.] Sei  $f: [a, b] \rightarrow W$  stetig. Eine stetig differenzierbare Funktion  $\Phi: [a, b] \rightarrow W$  mit  $\Phi' = f$  heißt eine *Stammfunktion* für  $f$ .

**Satz D5.4.** [Vgl. A5.7.] Sei  $f: [a, b] \rightarrow W$  stetig. Für jedes  $c \in [a, b]$  ist das unbestimmte Integral  $F(x) := \int_c^x f$  eine Stammfunktion für  $f$ . Sei nun  $\Phi$  eine beliebige Stammfunktion für  $f$ . Für  $p, q \in [a, b]$  gilt

$$\int_p^q f = \Phi(q) - \Phi(p) =: \Phi \Big|_p^q.$$

*Beweis.* Weil  $f$  stetig ist, sagt Satz D5.2, dass  $F' \equiv f$ . D.h.,  $F$  ist eine Stammfunktion. Aus Korollar D3.5 hat jede Stammfunktion die Form  $\Phi = F + w$  für  $w \in W$ . Aus Intervalladditivität folgt, dass

$$\Phi \Big|_p^q = F \Big|_p^q = \int_p^q f. \quad \square$$

### D6. Schlusswort

In diesem Abschnitt haben wir Ableitungen und Integrale für Funktionen mit Werten in einem Banachraum  $W$  betrachtet. Wie wir gesehen haben, ist die Theorie weitgehend unabhängig von  $W$  und kommt dem Falle  $W = \mathbb{R}$  fast gleich.

Wenn wir nur an endlich-dimensionalen Räumen  $W$  Interesse hätten, könnten wir einfach sagen: Man wählt eine Basis,  $W \cong \mathbb{R}^n$ , und alles wird komponentenweise gemacht.

Hingegen ziehen wir (auch im Falle  $\dim W < \infty$ ) einen anderen Gesichtspunkt vor. Die Wahl einer Basis ist unwichtig. Ableitungen und Integrale sind lineare Verfahren. Die Definitionen, Ergebnisse und Beweise kann man direkt mit Vektoren betrachten.

**E. ABLEITUNGEN IN MEHREREN DIMENSIONEN**

In diesem Abschnitt, seien  $V$  und  $W$  Banachräume mit  $\dim V = m < \infty$ , sei  $G \subset V$  offen und sei  $f: G \rightarrow W$  eine Abbildung.

Die Ableitung von  $f$  in einem Punkt  $p \in G$  ist jetzt kein Vektor in  $W$  sondern eine lineare Abbildung  $L: V \rightarrow W$ . Weil dies ein bisschen kompliziert wird, fangen wir mit etwas einfacherem an.

**E1. Richtungsableitungen**

Sei  $I \subset \mathbb{R}$  ein Intervall und sei  $\gamma: I \rightarrow G \subset V$  eine Kurve in  $G$ . Dann ist  $f \circ \gamma$  eine Abbildung  $I \rightarrow W$  und wir können ihre Ableitungen (im Sinne von §D1) untersuchen.

**Definition E1.1.** Sei  $s \in I$  und sei  $p := \gamma(s) \in G$ . Die Ableitung von  $f$  in  $p$  entlang der Kurve  $\gamma$  ist

$$(f \circ \gamma)'(s) = \left. \frac{d}{dt} \right|_{t=s} f(\gamma(t)) = \lim_{h \rightarrow 0} \frac{f(\gamma(s+h)) - f(\gamma(s))}{h} \in W.$$

Der wichtigste Fall ist der, wo  $\gamma$  eine Gerade durch  $p$  in Richtung  $v \in V$  ist, d.h.,  $\gamma(t) := p + tv$ .

**Definition E1.2.** Seien  $p \in G$  und  $v \in V$ . Die Ableitung entlang der Geraden  $\gamma(t) := p + tv$  heißt *Richtungsableitung* von  $f$  im Punkt  $p$  in Richtung  $v$ :

$$\partial_v f(p) := (f \circ \gamma)'(0) = \left. \frac{d}{dt} \right|_{t=0} f(p + tv) = \lim_{t \rightarrow 0} \frac{f(p + tv) - f(p)}{t}.$$

**Beispiel E1.3.** Sei  $f: V \rightarrow W$  linear. Für alle  $p, v \in V$  gilt dann  $\partial_v f(p) = f(v)$ .

Weiter sei  $\gamma: I \rightarrow G$  eine differenzierbare Kurve mit  $\gamma(s) = p$ . Die Ableitung von  $f$  in  $p$  entlang  $\gamma$  können wir wie folgt berechnen, indem wir zunächst die Linearität und danach die Stetigkeit von  $f$  benutzen:

$$\begin{aligned} (f \circ \gamma)'(s) &= \lim_{h \rightarrow 0} \frac{f(\gamma(s+h)) - f(\gamma(s))}{h} \\ &= \lim_{h \rightarrow 0} f\left(\frac{1}{h}(\gamma(s+h) - \gamma(s))\right) \\ &= f\left(\lim_{h \rightarrow 0} \frac{1}{h}(\gamma(s+h) - \gamma(s))\right) \\ &= f(\gamma'(s)). \end{aligned}$$

**Definition E1.4.** Eine Abbildung  $f: V \rightarrow W$  heißt *homogen*, falls  $f(\lambda v) = \lambda f(v)$  für alle  $\lambda \in \mathbb{R}$  und  $v \in V$ .

**Beispiel E1.5.** Sei  $S := \{v : \|v\| = 1\}$  die Einheitskugel bezüglich der Norm auf  $V$ . Jede Funktion  $f: S \rightarrow W$  mit  $f(-x) = -f(x)$  (auch z.B. wenn nirgendwo stetig) hat eine eindeutige Erweiterung zu einer homogenen Abbildung  $V \rightarrow W$ . Diese ist in 0 stetig (aber möglicherweise in keinem anderen Punkt).

**Beispiel E1.6.** Auf  $\mathbb{R}^2$  ist die Funktion

$$f = r \cos 3\theta = \frac{x^3 - 3xy^2}{x^2 + y^2}$$

homogen. Hier existiert  $\partial_v f(p)$  für alle  $p$  und  $v$ .

**Lemma E1.7.** Sei  $f: V \rightarrow W$  homogen. Für jedes  $v \in V$  gilt dann  $\partial_v f(0) = f(v)$ .

*Beweis.* Es gilt  $f(tv) = tf(v)$ , insbes.  $f(0) = 0$ . Damit gilt

$$\partial_v f(0) = \lim_{t \rightarrow 0} \frac{1}{t}(f(tv) - f(0)) = \lim_{t \rightarrow 0} \frac{1}{t}tf(v) = f(v). \quad \square$$

**Bemerkung E1.8.** Entlang anderen Kurven  $\gamma$  durch 0 muss die Ableitung von einem homogenen  $f$  nicht existieren.

**Beispiel E1.9.** Sei  $P := \{(x, y) \in \mathbb{R}^2 : y = x^2 \neq 0\}$  und sei  $f = \chi_P: \mathbb{R}^2 \rightarrow \mathbb{R}$  die charakteristische Funktion von  $P$ , d.h.,  $f(p) = 1$  für  $p \in P$  und  $f(p) = 0$  für  $p \notin P$ . Nun sei  $p := (0, 0)$ . Auf jeder Gerade durch  $p$  ist  $f$  lokal konstant um  $p$ . Deswegen gilt  $\partial_v f(p) = 0$  für alle  $v \in \mathbb{R}^2$ . Hingegen ist  $f$  in  $p$  unstetig. Entlang der Kurve  $\gamma(t) := (t, t^2)$  hat  $f$  in  $p$  keine Ableitung.

Diese Beispiele dienen um zu zeigen, dass auch die Existenz aller Richtungsableitungen in einem Punkt nicht besonders aussagekräftig ist. Für eine differenzierbare Funktion (definiert unten) sind aber die Richtungsableitungen nützlich, um andere Ableitungen zu berechnen. Im Falle  $V = \mathbb{R}^m$  sind die Richtungsableitungen in den Richtungen  $e_i$  aus der Standardbasis von besonderem Interesse.

**Definition E1.10.** Sei  $G \subset \mathbb{R}^n$  offen, sei  $f: G \rightarrow W$  eine Abbildung und sei  $p \in G$ . Die *partiellen Ableitungen* von  $f$  in  $p$  sind

$$\frac{\partial}{\partial x_i} f(p) = \partial_i f(p) := \partial_{e_i} f(p), \quad i = 1, \dots, n.$$

Es gilt also

$$\partial_i f(p) = \left. \frac{d}{dt} \right|_{t=0} f(p_1, \dots, p_i + t, \dots, p_n).$$

**Bemerkung E1.11.** Um die partielle Ableitung  $\partial_i f(p)$  zu berechnen, nimmt man die Ableitung von  $f$  bezüglich der  $i$ -ten Variablen  $x_i$  (im Sinne von §IE1 bzw. von §D1); dabei betrachtet man alle anderen Variablen  $x_j$  ( $j \neq i$ ) als Konstanten  $x_j = p_j$ .

---

Ende der Vorlesung 2009 Januar 12

---

**E2. Ableitungen als lineare Abbildungen**

Sei  $G \subset V$  offen und sei  $f: G \rightarrow W$  eine Funktion mehrerer Variabler. Wie können wir sinnvoll eine Ableitung von  $f$  definieren? Ein Differenzenquotient macht keinen Sinn, weil  $h = x - p$  ein Vektor wäre: wir können nicht durch einen Vektor teilen.

Die Ableitung einer Funktion  $g: I \rightarrow W$  versteht man normalerweise als die (Änderungs-)Geschwindigkeit der Werte von  $g$ . Dies macht auch wenig Sinn für  $f: G \rightarrow W$ . Wir sollten uns eher die Ableitung als „beste lineare Approximation“ vorstellen.

Man kann Lemma D1.6 leicht wie folgt umschreiben:

**Lemma E2.1.** Eine Funktion  $g: I \rightarrow W$  hat genau dann eine Ableitung  $g'(p)$  in  $p \in I$ , wenn

$$g(x) - g(p) = g'(p)(x - p) + R(x), \quad 0 = \lim_{x \rightarrow p} \frac{R(x)}{|x - p|}.$$

Hier ist  $R$  natürlich das Restglied des ersten Taylorpolynoms von  $g$  und dieses Kriterium ist der Fall  $n = 1$  der Taylorapproximation (Satz I.F6.1).

Jede lineare Abbildung  $L: \mathbb{R} \rightarrow W$  hat die Form  $L(t) = wt$  für einen konstanten Vektor  $w \in W$ . Das letzte Lemma können wir wie folgt verstehen:  $g'(p)(x - p)$  ist eine lineare Funktion von  $x - p$ , die beste lineare Approximation zur Differenz  $g(x) - g(p)$ .

In dieser Form können wir die Ableitung auch für Funktionen mehrerer Variabler definieren:

**Definition E2.2.** Sei  $f: G \rightarrow W$  eine Abbildung und sei  $p \in G$ . Falls es eine lineare Abbildung  $L: V \rightarrow W$  so gibt, dass

$$0 = \lim_{x \rightarrow p} \frac{R(x)}{\|x - p\|} \quad \text{für das durch} \quad f(x) - f(p) - L(x - p) =: R(x)$$

definierte Restglied  $R: G \rightarrow W$ , dann sagen wir,  $f$  ist in  $p$  differenzierbar und  $D_p f := L$  ist die Ableitung von  $f$  in  $p$ .

*Bemerkung E2.3.* Äquivalent ist folgende Umschreibung: Differenzierbarkeit in  $x$  heißt,

$$f(x + h) = f(x) + D_x f(h) + \|h\| r(h),$$

wobei  $r(h) \rightarrow 0$  für  $h \rightarrow 0$ . (Hier ist  $r$  zunächst auf einer punktierten Umgebung von  $h = 0$  definiert.)

*Bemerkung E2.4.* Dass die Ableitung  $D_p f$  eindeutig bestimmt ist (wenn sie existiert), folgt aus dem nächsten Lemma.

**Lemma E2.5.** Sei  $f$  in  $p$  differenzierbar und sei  $v \in V$ . Dann existiert für jedes  $v \in V$  die Richtungsableitung  $\partial_v f(p)$  und es gilt

$$D_p f(v) = \partial_v f(p).$$

*Beweis.* Es gilt

$$\begin{aligned} \frac{f(p + tv) - f(p)}{t} &= \frac{f(p) + D_p f(tv) + R(p + tv) - f(p)}{t} \\ &= D_p f(v) + \frac{R(p + tv)}{t} \\ &= D_p f(v) \pm \frac{R(p + tv)}{\|tv\|} \|v\| \end{aligned}$$

und damit gilt im Limes  $t \rightarrow 0$   $\partial_v f(p) = D_p f(v) + 0$ . □

*Beispiel E2.6.* Die Funktion  $f$  aus Beispiel E1.6 ist im Ursprung nicht differenzierbar, weil die Richtungsableitungen  $\partial_v f(0) = f(v)$  nicht linear in  $v$  sind.

**Korollar E2.7.** Im Falle  $V = \mathbb{R}^n$  ist die Ableitung  $D_p f$  (wenn sie existiert) durch die partiellen Ableitungen  $\partial_i f(p)$  bestimmt: für  $v = (v_1, \dots, v_n)$  gilt

$$D_p f(v) = \sum_{i=1}^n \partial_i f(p) v_i.$$

*Beweis.* Es gilt  $v = \sum v_i e_i$  und damit aus der Linearität von  $D_p f$  gilt

$$D_p f(v) = \sum v_i D_p f(e_i) = \sum v_i \partial_{e_i} f(p). \quad \square$$

*Bemerkung E2.8.* Weitere Namen für die partiellen Ableitungen sind z.B.

$$f_{x_i} := \frac{\partial}{\partial x_i} f := \partial_{e_i} f.$$

Nennt man die Koordinaten auf  $\mathbb{R}^3$  z.B.  $x, y, z$ , dann schreibt man auch etwa

$$f_y := \frac{\partial}{\partial y} f := \partial_y f := \partial_{e_2} f.$$

**Lemma E2.9.** Ist  $f: G \rightarrow W$  in  $p \in G$  differenzierbar, dann ist  $f$  in  $p$  stetig.

*Beweis.* Wir haben

$$f(x) = f(p) + D_p f(x - p) + R(x).$$

Hier ist  $R$  sicherlich in  $p$  stetig. ( $R(x) \rightarrow 0$  für  $x \rightarrow p$ , sogar schneller als  $\|x - p\|$ .) Weil  $V$  endlich-dimensional ist, ist auch die lineare Abbildung  $D_p f$  stetig. □

**Lemma E2.10.** Sei  $W = \mathbb{R}^n$  und sei

$$f = (f_1, \dots, f_n): G \rightarrow \mathbb{R}^n.$$

Diese Abbildung ist genau dann in  $p \in G$  differenzierbar, wenn jedes  $f_i$  in  $p$  differenzierbar ist. In diesem Fall gilt

$$D_p f = (D_p f_1, \dots, D_p f_n): V \rightarrow \mathbb{R}^n.$$

*Beweis.* (Aufgabe.) □

*Bemerkung E2.11.* Ähnliches gilt für  $W = W_1 \times W_2$  (z.B. mit der Norm  $\|(w_1, w_2)\| = \|w_1\| + \|w_2\|$ ) und  $f = (f_1, f_2)$ . Auch dann haben wir  $D_p f = (D_p f_1, D_p f_2)$ .

### E3. Rechenregeln

**Satz E3.1.** Seien  $f, g: G \rightarrow W$  differenzierbar in  $p$  und sei  $\lambda \in \mathbb{R}$ . Dann ist auch  $f + \lambda g$  differenzierbar in  $p$  und es gilt  $D_p(f + \lambda g) = D_p f + \lambda D_p g$ .

*Beweis.* (Aufgabe.) □

**Satz E3.2 (Kettenregel).** Seien  $U, V, W$  Banachräume ( $U, V$  endlich-dimensional). Seien  $G \subset U$  und  $H \subset V$  offen. Seien  $f: G \rightarrow V$  und  $g: H \rightarrow W$  Abbildungen mit  $f(G) \subset H$ , damit die Komposition  $g \circ f: G \rightarrow W$  existiert. Seien  $x \in G$  und  $y := f(x) \in H$ . Ist  $f$  bzw.  $g$  differenzierbar in  $x$  bzw.  $y$ , dann ist  $g \circ f$  in  $x$  differenzierbar, mit

$$D_x(g \circ f) = D_y g \circ D_x f.$$



**Bemerkung E3.3.** Das heißt, die Ableitung der Komposition ist die Komposition der Ableitungen. Dies sollte nicht überraschend sein, wenn man die Ableitungen als die besten linearen Approximationen betrachtet.

*Beweis.* Mit Bemerkung E2.3 haben wir

$$\begin{aligned} f(x+h) &= f(x) + D_x f(h) + \|h\|r(h), \\ g(y+k) &= g(y) + D_y g(k) + \|k\|s(k). \end{aligned}$$

Wir definieren

$$k(h) := f(x+h) - y = D_x f(h) + \|h\|r(h).$$

Damit gilt

$$\begin{aligned} g(f(x+h)) &= g(y) + D_y g(k(h)) + \|k(h)\|s(k(h)) \\ &= g(f(x)) + D_y g(D_x f(h)) + \\ &\quad \|h\|D_y g(r(h)) + \|k(h)\|s(k(h)) \\ &= (g \circ f)(x) + (D_y g \circ D_x f)(h) + \|h\|u(h), \end{aligned}$$

wobei  $u(h) := D_y g(r(h)) + \|k(h)\|s(k(h))/\|h\|$ . Wir müssen zeigen,  $u(h) \rightarrow 0$  für  $h \rightarrow 0$ .

Der erste Summand ist einfach: weil  $D_y g$  stetig und linear ist, gilt

$$\lim_{h \rightarrow 0} D_y g(r(h)) = D_y g\left(\lim_{h \rightarrow 0} r(h)\right) = D_y g(0) = 0.$$

Für den zweiten, sei  $M := \|D_x f\|$  die Operatornorm von  $D_x f$ . Dann gilt

$$\|k(h)\| \leq \|D_x f(h)\| + \|h\|\|r(h)\| \leq \|h\|(M + \|r(h)\|).$$

Insbesondere gilt  $k(h) \rightarrow 0$  für  $h \rightarrow 0$  und damit  $s(k(h)) \rightarrow 0$ . Weil  $r(h) \rightarrow 0$  gilt  $r(h) \leq 1$  für hinreichend kleines  $h$  und damit  $\|k(h)\| \leq (M+1)\|h\|$ . Deshalb haben wir

$$\lim_{h \rightarrow 0} \frac{\|k(h)\|}{\|h\|} \|s(k(h))\| \leq \lim_{h \rightarrow 0} (M+1) \|s(k(h))\| = 0. \quad \square$$

**Definition E3.4.** Seien  $V, V'$  und  $W$  Vektorräume. Eine Abbildung  $\mu: V \times V' \rightarrow W$  heißt dann *bilinear*, falls sie in jedem Argument linear ist, d.h., falls

$$\mu(\lambda v + w, v') = \lambda \mu(v, v') + \mu(w, v')$$

und

$$\mu(v, \lambda v' + w') = \lambda \mu(v, v') + \mu(v, w')$$

für alle  $v, w \in V, v', w' \in V'$  und  $\lambda \in \mathbb{R}$ .

**Beispiele E3.5.** Folgende Abbildungen sind bilinear:

1. das Produkt  $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, (x, y) \mapsto xy$ ,
2. jedes Skalarprodukt  $V \times V \rightarrow \mathbb{R}, (v, w) \mapsto \langle v, w \rangle$ ,
3. die Skalarmultiplikation  $\mathbb{R} \times V \rightarrow V, (\lambda, v) \mapsto \lambda v$ ,
4.  $L(V, W) \times V \rightarrow W, (f, v) \mapsto f(v)$ .

**Lemma E3.6.** [Vgl. D2.6.] Seien  $V$  und  $V'$  endlich-dimensionale Banachräume und sei  $W$  ein Banachraum. Zu jedem bilinearen  $\mu: V \times V' \rightarrow W$  gibt es  $C > 0$  so, dass

$$\|\mu(v, v')\| \leq C \|v\| \|v'\|$$

für alle  $v \in V, v' \in V'$ .

*Beweis.* Wir dürfen annehmen,  $V = \mathbb{R}^m, V' = \mathbb{R}^n$ , jeweils mit der  $\ell^1$ -Norm. Sei  $w_{ij} := \mu(e_i, e_j)$  für  $i = 1, \dots, m$  und  $j = 1, \dots, n$  und sei  $C := \max_{i,j} \|w_{ij}\|$ . Für  $v = (v_1, \dots, v_m)$  und  $v' = (v'_1, \dots, v'_n)$  gilt dann

$$\begin{aligned} \|\mu(v, v')\| &= \left\| \sum_{i,j} v_i v'_j w_{ij} \right\| \leq \sum_{i,j} |v_i| |v'_j| \|w_{ij}\| \\ &\leq C \sum_i |v_i| \sum_j |v'_j| = C \|v\|_1 \|v'\|_1. \quad \square \end{aligned}$$

**Satz E3.7 (Allgemeine Produktregel).** Seien  $V$  und  $V'$  endlich-dimensionale Banachräume, sei  $W$  ein Banachraum und sei  $\mu: V \times V' \rightarrow W$  bilinear. Dann ist  $\mu$  differenzierbar und es gilt

$$D_{(p,p')}\mu(v, v') = \mu(p, v') + \mu(v, p').$$

**Bemerkung E3.8.** Allgemeiner: eine *multilineare* Abbildung  $\mu: V_1 \times \dots \times V_k \rightarrow W$  ist differenzierbar mit

$$D_p \mu(v) = \sum_{i=1}^k \mu(p_1, \dots, p_{i-1}, v_i, p_{i+1}, \dots, p_k).$$

*Beweis.* (Aufgabe.) □

---

Ende der Vorlesung 2009 Januar 15

---

In konkreten Fällen benutzt man die allgemeine Produktregel (z.B. für eine der bilinearen Funktionen aus E3.5 zusammen mit der Kettenregel. Zu E3.5.3 haben wir:

**Korollar E3.9.** Seien  $V$  und  $W$  endlich-dimensionale Banachräume, sei  $G \subset V$  offen und seien  $f: G \rightarrow \mathbb{R}$  und  $g: G \rightarrow W$  differenzierbare Abbildungen. Dann ist  $fg: G \rightarrow W$  auch differenzierbar; in  $p \in G$  gilt

$$D_p(fg)(v) = f(p) D_p g(v) + D_p f(v) g(p).$$

*Beweis.* Wir schreiben  $fg = \mu \circ h$ , wobei  $\mu: \mathbb{R} \times W \rightarrow W$  die Skalarmultiplikation ist und  $h := (f, g): G \rightarrow \mathbb{R} \times W$ . Für  $p \in G$  seien  $t := f(p)$  und  $q := g(p)$ , sodass  $h(p) = (t, q)$ . Nach Bemerkung E2.11 gilt  $D_p h = (D_p f, D_p g)$ ; nach der allgemeinen Produktregel gilt

$$D_{(t,q)}\mu(a, w) = \mu(t, w) + \mu(a, q) = tw + aq.$$

Mit der Kettenregel gilt

$$D_p(fg) = D_p(\mu \circ h) = D_{(t,q)}\mu \circ D_p h.$$

Das heißt dann,

$$\begin{aligned} D_p(fg)(v) &= D_{(t,q)}\mu(D_p f(v), D_p g(v)) \\ &= t D_p g(v) + D_p f(v) g \\ &= f(p) D_p g(v) + D_p f(v) g(p). \quad \square \end{aligned}$$

*Beispiel E3.10.* Sei  $V$  ein endlich-dimensionaler Vektorraum mit Skalarprodukt  $\langle \cdot, \cdot \rangle$ . (Ein Skalarprodukt heißt auch inneres Produkt.) Die Norm  $\|v\| := \sqrt{\langle v, v \rangle}$  ist eine Abbildung  $r : V \rightarrow \mathbb{R}$ ,  $r(v) := \|v\|$ . In 0 ist sie nicht differenzierbar. (Warum?) Wir behaupten, sie ist in jedem Punkt  $p \neq 0$  differenzierbar mit  $D_p r(v) = \langle v, p \rangle / \|p\|$ . Definieren wir

$$\delta : v \mapsto (v, v), \quad \mu : (v, w) \mapsto \langle v, w \rangle, \quad \sigma : x \mapsto \sqrt{x},$$

gilt dann

$$r = \sigma \circ \mu \circ \delta : V \rightarrow V \times V \rightarrow I \rightarrow \mathbb{R},$$

wobei  $I := [0, \infty) \subset \mathbb{R}$ . Weil  $\delta$  linear ist und  $\mu$  bilinear, gelten

$$\begin{aligned} D_p \delta(v) &= \delta(v) = (v, v), \\ D_{(p,p)} \mu(v, v) &= \mu(p, v) + \mu(v, p) = 2 \langle v, p \rangle, \\ D_x \sigma(h) &= \frac{1}{2\sqrt{x}} h. \end{aligned}$$

Deshalb gilt nach der Kettenregel

$$\begin{aligned} D_p r(v) &= D_{\langle p,p \rangle} \sigma \left( D_{(p,p)} \mu (D_p \delta(v)) \right) \\ &= \frac{1}{2\sqrt{\langle p, p \rangle}} 2 \langle v, p \rangle = \frac{\langle v, p \rangle}{\|p\|}. \end{aligned}$$

*Bemerkung E3.11.* Eine Norm  $\|\cdot\|$ , die nicht aus einem Skalarprodukt kommt, ist natürlich stetig, muss aber nicht differenzierbar sein. Die  $\ell^1$ -Norm ist z.B. in  $e_i$  nicht differenzierbar.

Ein wichtiger Fall der Kettenregel sagt uns, dass Ableitungen entlang einer differenzierbaren Kurve einfach Richtungsableitungen sind.

**Lemma E3.12.** Seien  $\gamma : I \rightarrow G$  und  $f : G \rightarrow W$  differenzierbar und sei  $t \in I$ . Die Ableitung von  $f$  in  $p := \gamma(t)$  entlang der Kurve  $\gamma$  ist

$$(f \circ \gamma)'(t) = \partial_{\gamma'(t)} f(p).$$

*Beweis.* Es gilt

$$\begin{aligned} (f \circ \gamma)'(t) &= D_t(f \circ \gamma)(1) = D_p f \circ D_t \gamma(1) \\ &= D_p f(\gamma'(t)) = \partial_{\gamma'(t)} f(p). \quad \square \end{aligned}$$

*Bemerkung E3.13.* Ist  $f$  in  $p$  differenzierbar, hängt nach diesem Lemma die Ableitung entlang einer Kurve  $\gamma$  nur von dessen Tangentenvektor ab. Ist  $f$  hingegen in  $p$  nicht differenzierbar, kann es sein, dass die Ableitungen entlang verschiedenen Kurven mit demselben Tangentenvektor unterschiedlich sind bzw. teilweise nicht existieren. [Vgl. E1.9.]

#### E4. Existenz der Ableitung

Wollen wir zeigen, eine Abbildung  $f : G \rightarrow W$  ist in  $p \in G$  differenzierbar, können wir bisher nur die Definition „per Hand“ anwenden.

Wir kennen Beispiele, wo alle Richtungsableitungen existieren,  $D_p f$  aber nicht (Beispiel E2.6). Auch wenn die verschiedenen Richtungsableitungen  $\partial_v f(p)$  in  $p$  linear in  $v$  sind (z.B. wenn alle Null sind wie im Beispiel E1.9) muss  $D_p f$  nicht existieren, muss  $f$  nicht mal in  $p$  stetig sein.

Es ist deshalb schön, ein Kriterium für Differenzierbarkeit zu finden, das nur von den Richtungs- bzw. partiellen Ableitungen abhängt, deren Stetigkeit aber verlangt. Dabei nehmen wir an, dass nicht nur  $V$  sondern auch  $W$  endlich-dimensional ist.

**Satz E4.1.** Sei  $U \subset \mathbb{R}^m$  offen und sei  $f : U \rightarrow \mathbb{R}$  eine Funktion. Existieren in jedem Punkt  $x \in U$  die  $m$  partiellen Ableitungen  $\partial_i f(x)$ , betrachten wir diese als Funktionen  $\partial_i f : U \rightarrow \mathbb{R}$ . Sind diese auf  $U$  stetig, dann ist  $f$  auf  $U$  stetig differenzierbar.

*Beweis.* Sei  $x = (x_1, \dots, x_m) \in U$ . Aus Lemma E2.5 wissen wir, welche lineare Abbildung die Ableitung  $D_x f$  sein muss, falls sie existiert. Für  $h = \sum h_i e_i \in \mathbb{R}^m$  gälte nämlich

$$D_x f(h) = \sum h_i D_x f(e_i) = \sum h_i \partial_i f(x).$$

Diese lineare Abbildung  $h \mapsto \sum h_i \partial_i f(x)$  nennen wir  $L$ . Wir müssen nur zeigen, das Restglied  $f(x+h) - f(x) - Lh$  wächst wie in der Definition von Differenzierbarkeit) sublinear.

Weil  $U$  offen ist, gibt es  $r > 0$  so, dass  $B_r(x) \subset U$ . Für  $\|h\| < r$  definieren wir wie folgt  $m+1$  Punkte  $y_0, \dots, y_m$ :

$$y_0 := x, \quad y_1 := y_0 + h_1 e_1, \quad \dots, \quad y_k := y_{k-1} + h_k e_k, \quad \dots, \quad y_m = x+h.$$

Wir haben

$$f(x+h) - f(x) = \sum_{k=1}^m f(y_k) - f(y_{k-1}).$$

Die Punkte  $y_k$  (und die geraden Strecken dazwischen) liegen in  $B_r(x)$  und deshalb in  $U$ .

Auf jede Strecke wenden wir den Mittelwertsatz auf  $g(t) := f(y_k + t e_k)$  an. Hier liegt  $t$  zwischen 0 und  $h_k$ ; um die Notation zu vereinfachen, nehmen wir an,  $h_k > 0$ . (Für  $h_k = 0$  gilt  $y_k = y_{k-1}$  und wir setzen  $\xi_k = 0$ .) Aus dem Mittelwertsatz gibt es  $\xi_k \in [0, h_k]$  so, dass  $g(h_k) - g(0) = h_k g'(\xi_k)$ . Das heißt,

$$f(y_k) - f(y_{k-1}) = h_k \partial_k f(z_k),$$

wobei wir  $z_k := y_{k-1} + \xi_k e_k$  setzen. Dies bedeutet,

$$f(x+h) - f(x) = \sum_{k=1}^m h_k \partial_k f(z_k)$$

und damit

$$\begin{aligned} |f(x+h) - f(x) - Lh| &= \left| \sum_{k=1}^m h_k (\partial_k f(z_k) - \partial_k f(x_k)) \right| \\ &\leq \|h\| \sum_{k=1}^m |\partial_k f(z_k) - \partial_k f(x_k)|. \end{aligned}$$

Für  $h \rightarrow 0$  konvergiert  $z_k \rightarrow x$  (weil  $\|z_k - x\| \leq \|h\|$ ). Weil jedes  $\partial_k f$  stetig ist, konvergiert jeder Summand – und deshalb auch die Summe – gegen 0.  $\square$

Ende der Vorlesung 2009 Januar 19

**Korollar E4.2.** Seien  $V$  und  $W$  endlich-dimensionale Banachräume und sei  $G \subset V$  offen. Ferner sei  $f: G \rightarrow W$  eine Abbildung und sei  $\{v_1, \dots, v_m\}$  eine Basis für  $V$ . Existieren in jedem Punkt  $p \in G$  die  $m$  Richtungsableitungen  $\partial_{v_i} f(p)$  und sind diese stetige Funktionen von  $p$ , dann ist  $f$  auf  $G$  stetig differenzierbar.

*Beweis.* Die Basis  $\{v_i\}$  gibt uns einen Isomorphismus  $\mathbb{R}^m \rightarrow V$ , wodurch die Richtungsableitungen  $\partial_{v_i}$  den partiellen Ableitungen  $\partial_i$  entsprechen. Um jeden Punkt  $p \in G$  gibt es einen offenen Ball  $U \subset G$ .

Ist  $n := \dim W$ , dann sind auch  $W$  und  $\mathbb{R}^n$  isomorph und wir können  $f = (f_1, \dots, f_n)$  komponentenweise betrachten. Wir wenden den Satz auf jede Komponente an.  $\square$

### E5. Ableitungen in Koordinaten

*Bemerkung E5.1.* Aus der linearen Algebra wissen wir, eine lineare Abbildung  $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$  wird (bezüglich der Standardbasen  $\{e_1, \dots, e_m\}$  für  $\mathbb{R}^m$  und  $\{e'_1, \dots, e'_n\}$  für  $\mathbb{R}^n$ ) durch eine  $(m \times n)$ -Matrix  $(L_{ij})$  dargestellt. Die Einträge  $L_{ij}$  werden durch  $L(e_i) = \sum_j L_{ij} e'_j$  definiert. Für jeden Vektor  $v = \sum_i v_i e_i \in V$  gilt dann

$$L\left(\sum_{i=1}^m v_i e_i\right) = \sum_{i=1}^m v_i L(e_i) = \sum_{i=1}^m v_i \sum_{j=1}^n L_{ij} e'_j.$$

**Definition E5.2.** Sei  $G \subset \mathbb{R}^m$  offen und  $f: G \rightarrow \mathbb{R}^n$  differenzierbar. Für  $p \in G$  ist die Ableitung  $D_p f$  eine lineare Abbildung  $\mathbb{R}^m \rightarrow \mathbb{R}^n$ . Deren Darstellungsmatrix  $J_p f$  heißt die *Jacobimatrix* von  $f$  in  $p$ .

**Lemma E5.3.** Die Einträge der Jacobimatrix  $J_p f$  sind die partiellen Ableitungen  $\partial_i f_j(p)$  der Komponenten von  $f$ . D.h., es gilt

$$Jf = (\partial_i f_j) = \begin{pmatrix} \partial_1 f_1 & \cdots & \partial_m f_1 \\ \vdots & & \vdots \\ \partial_1 f_n & \cdots & \partial_m f_n \end{pmatrix}.$$

*Beweis.* Es gilt  $D_p f(e_i) = \sum_j D_p f_j(e_i) e'_j = \sum_j (\partial_i f_j(p)) e'_j$ .  $\square$

*Bemerkung E5.4.* Die  $i$ -te Spalte von  $J_p f$  ist die partielle Ableitung  $\partial_i f(p) = D_p f(e_i)$ . Die  $j$ -te Reihe ist  $J_p f_j$ , die Darstellungsmatrix von  $D_p f_j$ . (Diese nennen wir später auch Gradient von  $f_j$ .)

Folgendes Lemma kennen wir aus der linearen Algebra.

**Lemma E5.5.** Die Darstellungsmatrix der Komposition zweier linearer Abbildungen ist das Matrixprodukt der jeweiligen Darstellungsmatrizen. Das heißt, für  $N = M \circ L$  gilt  $N_{ik} = \sum_j L_{ij} M_{jk}$ .  $\square$

*Bemerkung E5.6.* Die Kettenregel sagt, die Ableitung einer Komposition ist die Komposition der Ableitungen. Das heißt, die Jacobimatrix der Komposition ist das Matrixprodukt der

Jacobimatrizen. Seien  $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$  und  $g: \mathbb{R}^l \rightarrow \mathbb{R}^m$  Abbildungen und sei  $h = f \circ g: \mathbb{R}^l \rightarrow \mathbb{R}^n$  deren Komposition. (Natürlich können wir auch Funktionen mit offenen Teilmengen als Definitionsbereichen betrachten.) Sei  $x \in \mathbb{R}^l$ . Wir setzen  $y := g(x) \in \mathbb{R}^m$  und  $z := f(y) \in \mathbb{R}^n$ . Dann sagt die Kettenregel,  $D_x h = D_y f \circ D_x g$ . Für die Jacobimatrizen haben wir deshalb  $J_x h = J_y f \cdot J_x g$ . Das heißt,  $\partial_i h_k = \sum_j \partial_j f_k \partial_i g_j$  oder

$$\frac{\partial h_k}{\partial x_i} = \sum_{j=1}^m \frac{\partial f_k}{\partial y_j} \frac{\partial g_j}{\partial x_i}.$$

Wegen  $y_j = g_j(x)$  und  $z_k = f_k(y) = h_k(x)$  schreibt man auch gern

$$\frac{\partial z_k}{\partial x_i} = \sum_{j=1}^m \frac{\partial z_k}{\partial y_j} \frac{\partial y_j}{\partial x_i}.$$

### E6. Höhere Ableitungen

Seien weiter  $V$  und  $W$  Banachräume mit  $\dim V = m < \infty$  und sei  $G \subset V$  offen. Bezeichne  $L(V, W)$  den Vektorraum aller linearen Abbildungen  $V \rightarrow W$ . Mit der Operatornorm ist  $L(V, W)$  auch ein Banachraum.

Ist eine Abbildung  $f: G \rightarrow W$  auf  $G$  differenzierbar, dann ist  $Df: p \mapsto D_p f$  eine Abbildung  $G \rightarrow L(V, W)$ . Falls  $Df$  stetig ist, heißt  $f$  *stetig differenzierbar*. Dann kann man fragen, ob  $Df$  sogar in  $p \in G$  differenzierbar ist. Wenn ja, dann ist  $D_p(Df)$  eine lineare Abbildung  $V \rightarrow L(V, W)$ . (Wir könnten schreiben  $D_p(Df) \in L(V, L(V, W))$ .) Das heißt, für  $v \in V$  ist  $D_p(Df)(v) = \partial_v(Df)(p) \in L(V, W)$  eine lineare Abbildung  $V \rightarrow W$ . Ist weiter  $v' \in V$ , dann haben wir

$$(D_p(Df)(v))(v') = (\partial_v(Df)(p))(v') \in W.$$

Diese Notation und die Rollen der beiden Vektoren  $v$  und  $v'$  bleiben dabei ein bisschen unklar.

Es hilft, ein Lemma aus der linearen Algebra zu benutzen.

**Definition E6.1.** Für  $k \in \mathbb{N}$  sei  $L^k(V, W)$  der Vektorraum aller  $k$ -linearen Abbildungen

$$f: V^k = V \times \dots \times V \rightarrow W,$$

d.h. Abbildungen, die in jedem der  $k$  Argumente linear sind. Es gelten  $L^0(V, W) \cong W$  und  $L^1(V, W) = L(V, W)$ .

**Lemma E6.2.** Für jedes  $k \in \mathbb{N}$  gibt es einen natürlichen Isomorphismus

$$\varphi_k: L(V, L^k(V, W)) \rightarrow L^{k+1}(V, W)$$

definiert durch

$$(\varphi_k(\mu))(v, v_1, v_2, \dots, v_k) := (\mu(v))(v_1, \dots, v_k) \in W.$$

*Beweis.* Dass  $\varphi_k(\mu)$  eine Abbildung  $V^{k+1} \rightarrow W$  ist, ist offensichtlich. Dass diese Abbildung multilinear ist, ist eine unkomplizierte Aufgabe.  $\square$

Jetzt können wir die höheren Ableitungen  $D^k f$  rekursiv so definieren, dass  $D_p^k f \in L^k(V, W)$ .

**Definition E6.3.** Sei  $f: G \rightarrow W$  eine Abbildung. Für  $k = 0$  definieren wir  $D_p^0 f = f(p) \in L^0(V, W) \cong W$ , d.h.,  $D^0 f = f$ , und wir sagen, jede Abbildung  $f$  ist 0-mal differenzierbar.

Für  $k \in \mathbb{N}^+$  nehmen wir an,  $f$  ist  $(k-1)$ -mal differenzierbar. Falls  $D^{k-1} f: G \rightarrow L^{k-1}(V, W)$  in  $p \in G$  differenzierbar ist, sagen wir,  $f$  ist  $k$ -mal in  $p$  differenzierbar. Wir definieren die  $k$ -te Ableitung von  $f$  in  $p$  als

$$D_p^k f = \varphi_{k-1}(D_p(D^{k-1} f)) \in L^k(V, W).$$

Falls  $D_p^k f$  in jedem  $p \in G$  existiert, sagen wir,  $f$  ist  $k$ -mal differenzierbar (auf  $G$ ). Wir definieren

$$D^k f: G \rightarrow L^k(V, W), \quad p \mapsto D_p^k f.$$

*Bemerkung E6.4.* Damit ist die erste Ableitung die Ableitung:  $D_p^1 f = D_p f$ . Die zweite Ableitung in  $p$  ist eine bilineare Abbildung  $D_p^2 f: V \times V \rightarrow W$ .

Wir werden gleich sehen, dass auch höhere Ableitungen sich mittels Richtungsableitungen ausdrücken lassen. Zunächst berechnen wir aber einmal eine zweite Ableitung direkt aus der Definition.

*Beispiel E6.5.* Sei  $V$  ein endlich-dimensionaler Vektorraum mit Skalarprodukt  $\langle \cdot, \cdot \rangle$ . Es gibt einen natürlichen Isomorphismus

$$\varphi: V \rightarrow L(V, \mathbb{R}), \quad p \mapsto \langle p, \cdot \rangle.$$

Das heißt,  $\varphi(p)(v) := \langle p, v \rangle$ .

Sei  $r(v) := \|v\| = \sqrt{\langle v, v \rangle}$  die Norm auf  $V$ . In Beispiel E3.10 haben wir die Ableitung  $D_p r(v) = \langle v, p \rangle / \|p\|$  berechnet. Das heißt,

$$D_p r = \frac{1}{r(p)} \varphi(p) \in L(V, \mathbb{R}).$$

Damit ist  $D_r$  das Produkt der beiden Funktionen  $1/r$  und  $\varphi$ . Nach der Kettenregel gilt

$$D_p \frac{1}{r} = -\frac{1}{r(p)^2} D_p r = -\frac{1}{r(p)^3} \varphi(p).$$

Weil  $\varphi$  linear ist, gilt natürlich  $D_p \varphi = \varphi$ . Korollar E3.9 erklärt die Produktregel für Skalarmultiplikation; danach ist

$$\begin{aligned} D_p(D_r)(v) &= D_p\left(\frac{1}{r}\varphi\right)(v) = \frac{1}{r(p)}\varphi(v) - \frac{1}{r(p)^3}(\varphi(p)(v))\varphi(p) \\ &= \frac{\varphi(v)}{\|p\|} - \frac{\langle p, v \rangle \varphi(p)}{\|p\|^2}. \end{aligned}$$

Nach der Definition der zweiten Ableitung gilt dann

$$D_p^2 r(v, w) = (D_p(D_r)(v))(w) = \frac{\langle v, w \rangle \langle p, p \rangle - \langle p, v \rangle \langle p, w \rangle}{\|p\|^3}.$$

---

Ende der Vorlesung 2009 Januar 22

Mithilfe des folgenden Lemmas können wir höhere Ableitungen leichter berechnen.

**Lemma E6.6.** Die Funktion  $f: G \rightarrow W$  sei  $k$ -mal differenzierbar in  $p \in G$ . (Dies setzt voraus, dass  $f$  auf  $G$   $(k-1)$ -mal differenzierbar ist.) Für alle  $v_1, \dots, v_k \in V$  gilt dann

$$D_p^k f(v_1, \dots, v_k) = \partial_{v_1} \cdots \partial_{v_k} f(p).$$

Insbesondere existiert die  $k$ -fache Richtungsableitung auf der rechten Seite.

*Beweis.* Wir benutzen Induktion über  $k$ . Der Fall  $k = 0$  ist die Definition  $D_p^0 f = f(p)$ . (Und der Fall  $k = 1$  ist Lemma E2.5.)

Für den Induktionsschritt definieren wir die lineare Auswertungsabbildung

$$g: L^{k-1}(V, W) \rightarrow W, \quad \mu \mapsto \mu(v_2, \dots, v_k).$$

Wir betrachten die Komposition

$$g \circ D^{k-1} f: G \rightarrow W, \quad p \mapsto D_p^{k-1} f(v_2, \dots, v_k).$$

Nach Lemma E2.5 und der Induktionsvoraussetzung gilt dann

$$D_p(g \circ D^{k-1} f)(v_1) = \partial_{v_1}(D_p^{k-1} f(v_2, \dots, v_k)) = \partial_{v_1} \partial_{v_2} \cdots \partial_{v_k} f.$$

Andererseits gilt aus der Kettenregel

$$\begin{aligned} D_p(g \circ D^{k-1} f)(v_1) &= g(D_p(D^{k-1} f)(v_1)) \\ &= (D_p^{k-1} f(v_1))(v_2, \dots, v_k) \\ &= D_p^k f(v_1, \dots, v_k). \end{aligned}$$

Ein Vergleich dieser beiden Gleichungen schließt den Beweis ab. □

*Bemerkung E6.7.* Sei  $G$  offen in  $V = \mathbb{R}^m$ . Die  $k$ -te Ableitung einer Funktion  $f: G \rightarrow W$  lässt sich mittels  $k$ -facher partieller Ableitungen berechnen. Seien  $k$  Vektoren  $v_1, \dots, v_k \in \mathbb{R}^m$  gegeben, wobei

$$v_i = (v_{i,1}, \dots, v_{i,m}) = \sum_{j=1}^m v_{i,j} e_j \quad \text{für } i = 1, \dots, k.$$

Ist  $f$  in  $p$   $k$ -mal differenzierbar, dann gilt

$$\begin{aligned} D_p^k f(v_1, \dots, v_k) &= \sum_{j_1, \dots, j_k=1}^m v_{1,j_1} \cdots v_{k,j_k} D_p^k f(e_{j_1}, \dots, e_{j_k}) \\ &= \sum_{j_1, \dots, j_k=1}^m v_{1,j_1} \cdots v_{k,j_k} \partial_{j_1} \cdots \partial_{j_k} f(p). \end{aligned}$$

**Definition E6.8.** Ist  $f: G \rightarrow W$   $k$ -mal differenzierbar und ist  $D^k f: G \rightarrow L^k(V, W)$  stetig, dann sagen wir  $f$  ist (auf  $G$ )  $k$ -mal stetig differenzierbar. Wir schreiben  $f \in C^k(G, W)$ . Falls  $f \in C^k$  für alle  $k \in \mathbb{N}$ , sagen wir,  $f$  ist beliebig oft differenzierbar,  $f \in C^\infty(G, W)$ .

**Satz E6.9.** Seien  $V$  und  $W$  endlich-dimensionalen Banachräume und sei  $G \subset V$  offen. Ferner sei  $f: G \rightarrow W$  eine Abbildung und sei  $\{v_1, \dots, v_m\}$  eine Basis für  $V$ . Existieren in jedem Punkt  $p \in G$  alle  $k$ -fachen Richtungsableitungen

$$\partial_{v_1} \cdots \partial_{v_k} f(p)$$

und sind diese stetige Funktionen von  $p$ , dann ist  $f$  auf  $G$   $k$ -mal stetig differenzierbar.

*Beweisskizze.* Wir dürfen annehmen,  $V = \mathbb{R}^m$  und  $W = \mathbb{R}^n$ ; wir benutzen die Standardbasen  $\{e_i\}$  und  $\{e'_j\}$ . Seien  $f_j$  die Komponenten von  $f = \sum_j f_j e'_j$ . Dann sind die  $(k-1)$ -fachen partiellen Ableitungen

$$\partial_{i_1} \cdots \partial_{i_{k-1}} f_j$$

die Komponenten von  $D^{k-1}f$  bzgl. einer natürlichen Basis für  $L^{k-1}(\mathbb{R}^m, \mathbb{R}^n)$ . Existieren deren (einfache) partielle Ableitungen, ist  $D^{k-1}f$  nach Korollar E4.2 stetig differenzierbar.  $\square$

*Bemerkung E6.10.* Es gilt

$$\begin{aligned} D_p^2 f(u, v) &= \partial_u \partial_v f(p) = \lim_{s \rightarrow 0} \frac{\partial_v f(p + su) - \partial_v f(p)}{s} \\ &= \lim_{s \rightarrow 0} \frac{1}{s} \lim_{t \rightarrow 0} \left( \frac{f(p + su + tv) - f(p + su)}{t} - \frac{f(p + tv) - f(p)}{t} \right) \\ &= \lim_{s \rightarrow 0} \lim_{t \rightarrow 0} \frac{1}{st} (f(p + su + tv) - f(p + su) - f(p + tv) + f(p)). \end{aligned}$$

**Definition E6.11.** Wir nennen

$$\Delta_p^2 f(h, k) := f(p + h + k) - f(p + h) - f(p + k) + f(p)$$

die *Differenz zweiter Ordnung* von  $f$ . Es gilt  $\Delta_p^2 f(h, k) = \Delta_p^2 f(k, h)$ .

*Bemerkung E6.12.* Für „kleine“ Vektoren  $h, k$  gilt  $\Delta_p^2 f(h, k) \approx D_p^2 f(h, k)$ . Genauer gesagt können wir die letzte Bemerkung wie folgt schreiben:

$$D_p^2 f(u, v) = \partial_u \partial_v f(p) = \lim_{s \rightarrow 0} \lim_{t \rightarrow 0} \frac{\Delta_p^2 f(su, tv)}{st}.$$

Entsprechend gilt

$$D_p^2 f(v, u) = \partial_v \partial_u f(p) = \lim_{t \rightarrow 0} \lim_{s \rightarrow 0} \frac{\Delta_p^2 f(su, tv)}{st}.$$

Natürlich darf man nicht immer zwei Limes tauschen. Viele wichtige Sätze der Analysis geben Voraussetzungen, unter denen man bestimmte Limes tauschen darf. Hier reicht es (nach einem H.A. Schwarz bzw. A. Clairaut zugeordneten Satz) aus, dass  $f$  in  $p$  zweimal differenzierbar ist.

*Beispiel E6.13.* Die existenz der zweifachen Richtungsableitungen ist nicht ausreichend. Sei  $p := (0, 0) \in \mathbb{R}^2$  und sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  durch  $f(p) := 0$  und

$$f(x, y) := xy \frac{x^2 - y^2}{x^2 + y^2}$$

für  $(x, y) \neq p$  definiert. Dann gilt  $\partial_x \partial_y f(p) \neq \partial_y \partial_x f(p)$ . (Beweis als Aufgabe. Wegen  $f(x, y) = -f(y, x)$  ist es klar, dass  $\partial_x \partial_y f(x, y) = -\partial_y \partial_x f(y, x)$ .)

**Satz E6.14.** Sei  $f: G \rightarrow W$  differenzierbar und in  $p \in G$  zweimal differenzierbar. Dann gilt

$$\partial_u \partial_v f(p) = \lim_{t \rightarrow 0} \frac{\Delta_p^2 f(tu, tv)}{t^2} = \partial_v \partial_u f(p).$$

für alle  $u, v \in V$ .

*Beweis.* Weil der Ausdruck in der Mitte symmetrisch in  $u, v$  ist, reicht es aus, die erste Gleichung zu beweisen. Um der Einfachheit willen, nehmen wir an,  $\dim W < \infty$ . (Für den allgemeinen Fall, siehe z.B. Amman–Escher VII.5.2.) Weil wir dann alles komponentenweise in  $W$  betrachten können, dürfen wir  $W = \mathbb{R}$  annehmen.

Wir wissen,  $f$  ist in  $p$  zweimal differenzierbar. Das heißt per Definition, dass  $Df: G \rightarrow W$  existiert und in  $p$  differenzierbar ist, d.h., dass

$$R(h) := D_{p+h}f - D_p f - D_p(Df)(h) \in L(V, \mathbb{R})$$

sublinear wächst ( $\lim_{h \rightarrow 0} R(h)/\|h\| = 0$ ). Sei  $\varepsilon > 0$  gegeben. Dass der Limes 0 ist heißt, es gibt  $\delta > 0$  so, dass  $\|R(h)\| < \varepsilon\|h\|$  für  $\|h\| < \delta$ . (Hier ist  $\|R\|$  die Operatornorm.) Das heißt, für  $\|h\| < \delta$  und alle  $v \in V$  gilt

$$\begin{aligned} |D_{p+h}f(v) - D_p f(v) - D_p^2 f(h, v)| &= |R(h)(v)| \leq \|R(h)\| \|v\| \\ &\leq \varepsilon \|h\| \|v\|. \end{aligned}$$

Diese Ungleichung können wir auch wie folgt schreiben:

$$|\partial_v f(p + h) - \partial_v f(p) - \partial_h \partial_v f(p)| \leq \varepsilon \|h\| \|v\|, \quad (*)$$

(Ist nun  $h = su$  mit  $|s| < \delta/\|u\|$ , dann folgt

$$\left| \lim_{t \rightarrow 0} \frac{\Delta_p^2 f(su, tv)}{st} - \partial_u \partial_v f(p) \right| \leq \varepsilon \|u\| \|v\|,$$

was eine genauere Version der ersten Gleichung in Bemerkung E6.12 ist.)

Nun sei  $g(t) := f(p + su + tv) - f(p + tv)$ . (Natürlich macht diese Definition nur dann Sinn, wenn  $s$  und  $t$  hinreichend klein sind, damit die Punkte im Definitionsbereich  $G$  von  $f$  liegen. Weil wir sowieso am Ende den Limes  $s = t \rightarrow 0$  betrachten, schreiben wir diese Schranken nicht explizit.) Es gelten  $\Delta_p^2 f(su, tv) = g(t) - g(0)$  und  $g'(t) = \partial_v f(p + su + tv) - \partial_v f(p + tv)$ . Nach dem Mittelwertsatz gibt es zu jedem  $t$  ein  $\tau$  (mit  $0 < \tau/t < 1$ ) so, dass

$$\begin{aligned} \frac{1}{t} \Delta_p^2 f(su, tv) &= \frac{1}{t} (g(t) - g(0)) = g'(\tau) \\ &= \partial_v f(p + su + \tau v) - \partial_v f(p + \tau v) \\ &= (\partial_v f(p + su + \tau v) - \partial_v f(p)) \\ &\quad - (\partial_v f(p + \tau v) - \partial_v f(p)). \end{aligned}$$

Jetzt benutzen wir (\*) zweimal mit  $h = su + \tau v$  bzw.  $h = \tau v$ . Es gilt

$$\begin{aligned} &|g'(\tau) - s \partial_u \partial_v f(p)| \\ &\leq |\partial_v f(p + su + \tau v) - \partial_v f(p) - \partial_{su+\tau v} \partial_v f(p)| \\ &\quad + |\partial_v f(p + \tau v) - \partial_v f(p) - \partial_{\tau v} \partial_v f(p)| \\ &\leq \varepsilon \|v\| (\|su + \tau v\| + \|\tau v\|). \end{aligned}$$

Damit gilt (natürlich unter der Voraussetzung, dass  $\|\tau v\| < \delta$  und  $\|su + \tau v\| < \delta$ )

$$\left| \frac{\Delta_p^2 f(su, tv)}{st} - \partial_u \partial_v f(p) \right| \leq \frac{\varepsilon \|v\|}{|s|} (\|su + \tau v\| + \|\tau v\|).$$

Für  $s = t$  (mit  $t$  hinreichend klein) gilt dann

$$\left| \frac{\Delta_p^2 f(tu, tv)}{t^2} - \partial_u \partial_v f(p) \right| \leq \varepsilon \|v\| \left( \|u + \frac{\tau}{t} v\| + \left\| \frac{\tau}{t} v \right\| \right) \leq \varepsilon \|v\| (\|u\| + 2\|v\|).$$

Weil  $\varepsilon$  beliebig war, konvergiert die linke Seite wie gewünscht gegen 0.  $\square$

Ende der Vorlesung 2008 Januar 26

**Bemerkung E6.15.** Häufig sieht man eine schwächere Version des Satzes, die die Stetigkeit von  $D^2 f$  voraussetzt.

**Korollar E6.16.** Die Abbildung  $f: G \rightarrow W$  sei  $k$ -mal in  $p \in G$  differenzierbar. Für jede Permutation

$$\sigma: (1, \dots, k) \mapsto (\sigma 1, \dots, \sigma k)$$

und für alle  $v_1, \dots, v_k \in V$  gilt

$$\begin{aligned} \partial_{v_1} \cdots \partial_{v_k} f(p) &= D_p^k f(v_1, \dots, v_k) \\ &= D_p^k f(v_{\sigma 1}, \dots, v_{\sigma k}) = \partial_{v_{\sigma 1}} \cdots \partial_{v_{\sigma k}} f(p). \end{aligned}$$

Das heißt, die  $k$ -lineare Abbildung  $D_p^k f$  ist symmetrisch.

**Beweisidee.** Jede Permutation von  $(1, \dots, k)$  ist eine Komposition von (höchstens  $\binom{k}{2}$ ) Transpositionen.  $\square$

**Beispiel E6.17.** Sei  $U \subset \mathbb{R}^m$  offen und seien  $f_1, \dots, f_m: U \rightarrow \mathbb{R}$  differenzierbare Funktionen. Man kann sich fragen, ob es eine Funktion  $g: U \rightarrow \mathbb{R}$  so gibt, dass  $\partial_i g = f_i$  für alle  $i = 1, \dots, m$ . Der letzte Satz gibt uns eine notwendige (aber nicht hinreichende) Bedingung dafür, dass dieses System partieller Differentialgleichungen eine Lösung  $g$  hat: falls  $g$  existiert, dann gilt

$$\partial_i f_j = \partial_i \partial_j g = \partial_j \partial_i g = \partial_j f_i$$

für alle  $i, j$ .

**Definition E6.18.** Sei  $V$  ein Vektorraum und seien  $p, q \in V$ . Die Strecke  $\overline{pq}$  zwischen  $p$  und  $q$  ist

$$\overline{pq} := \{(1-t)p + tq : t \in [0, 1]\} \subset V.$$

Das heißt,  $\overline{pq}$  ist die Spur des geraden Weges  $\gamma: [0, 1] \rightarrow V$  mit  $\gamma(t) := (1-t)p + tq$ . Eine Teilmenge  $T \subset V$  heißt *konvex*, falls

$$p, q \in T \implies \overline{pq} \subset T.$$

**Bemerkung E6.19.** Sei  $f: G \rightarrow W$  und sei  $\overline{pq} \subset G$ . Die Werte von  $f$  auf dieser Strecke  $\overline{pq}$  können wir als Funktion von  $t$  betrachten:

$$g(t) := f(\gamma(t)) = f((1-t)p + tq).$$

Es gilt  $\gamma'(t) \equiv q - p$  und damit  $g'(t) = \partial_{q-p} f(\gamma(t))$  und per Induktion weiter

$$\begin{aligned} g^{(k)}(t) &= \partial_{q-p} \cdots \partial_{q-p} f(\gamma(t)) \\ &= \partial_{q-p}^k f(\gamma(t)) = D_{\gamma(t)}^k f(q-p, \dots, q-p). \end{aligned}$$

Auf dieser Weise können wir den Schrankensatz D3.4 auf  $f$  anwenden.

**Korollar E6.20.** Sei  $f: G \rightarrow W$  differenzierbar und sei  $\overline{pq} \subset G$ . Setze  $M := \sup\{\|D_x f\| : x \in \overline{pq}\}$ . Dann gilt

$$\|f(q) - f(p)\| \leq M \|q - p\|.$$

*Beweis.* (Folgt direkt aus Satz D3.4; Details als Aufgabe.)  $\square$

**Definition E6.21.** Sei  $f: G \rightarrow W$  eine  $n$ -mal differenzierbare Abbildung. Das  $n$ -te Taylorpolynom von  $f$  an der Entwicklungsstelle  $p \in G$  ist

$$T_n(x) := T_{n,p}^f(x) := \sum_{k=0}^n \frac{\partial_{x-p}^k f(p)}{k!}.$$

**Satz E6.22.** Sei  $f: G \rightarrow W$   $n$ -mal differenzierbar. Das  $n$ -te Taylorpolynom ist eine gute Approximation für  $f$ , im Sinne, dass für das Restglied  $R(x) := f(x) - T_n(x)$  gilt

$$\lim_{x \rightarrow p} \frac{\|R(x)\|}{\|x - p\|^n} = 0.$$

**Bemerkung E6.23.** Der Fall  $n = 0$  bzw.  $n = 1$  ist einfach die Definition von Stetigkeit bzw. Differenzierbarkeit. Würden wir wie oben alles nur eindimensional entlang der Strecke  $\overline{xp}$  betrachten, könnten wir daraus schließen, dass entlang jeder Geraden  $R(x)/\|x - p\|^n \rightarrow 0$ . Dies ist natürlich eine deutlich schwächere Aussage. (Für  $n = 1$  vgl. Differenzierbarkeit mit der Existenz aller Richtungsableitungen.)

*Beweis.* Wir benutzen Induktion über  $n$  und haben schon den Induktionsbeginn erledigt. Für den Induktionsschritt, sei  $R(x) = f(x) - T_{n,p}^f(x)$  das Restglied zu einem  $n$ -mal differenzierbaren  $f$ . Betrachten wir (für  $k > 0$ ) nun die Abbildung

$$g: x \mapsto \partial_{x-p}^k f(p) = D_p^k f(x-p, \dots, x-p)$$

als Komposition von  $x \mapsto (x-p, \dots, x-p)$  mit dem multilinearen  $D_p^k f$ , ergibt sich nach der allgemeinen Produktregel E3.8 und dem Korollar E6.16:

$$\begin{aligned} \partial_v g(x) &= D_x g(v) = k D_p^k f(x-p, \dots, x-p, v) \\ &= k \left( D_p^{k-1} (Df)(x-p, \dots, x-p) \right)(v) \\ &= k \left( \partial_{x-p}^{k-1} (Df)(p) \right)(v). \end{aligned}$$

Damit gilt

$$\partial_v R(x) = \partial_v f(x) - \sum_{k=1}^n \frac{1}{(k-1)!} \left( \partial_{x-p}^{k-1} (Df)(p) \right)(v)$$

und deshalb

$$\begin{aligned} D_x R &= D_x f - \sum_{k=1}^n \frac{1}{(k-1)!} \partial_{x-p}^{k-1} (Df)(p) \\ &= D_x f - \sum_{j=0}^{n-1} \frac{1}{j!} \partial_{x-p}^j (Df)(p) = D_x f - T_{n-1,p}^{Df}(x), \end{aligned}$$

d.h.,  $D_x R$  ist das Restglied für das Taylorpolynom von  $Df$ . Weil  $Df$  eine  $(n-1)$ -mal differenzierbare Abbildung ist, folgt

aus der Induktionsvoraussetzung, dass  $D_x R / \|x - p\|^{n-1} \rightarrow 0$ . Das heißt, zu jedem  $\varepsilon > 0$  gibt es  $\delta > 0$  so, dass

$$0 < \|x - p\| < \delta \implies \|D_x R\| < \varepsilon \|x - p\|^{n-1}.$$

Damit gilt für  $0 < \|x - p\| < \delta$ , dass

$$M_x := \sup_{y \in \bar{x}p} \|D_y R\| < \varepsilon \|x - p\|^{n-1}.$$

Nach Korollar E6.20 und wegen  $R(p) = 0$  gilt dann

$$\|R(x)\| = \|R(x) - R(p)\| \leq M_x \|x - p\| < \varepsilon \|x - p\|^n.$$

Weil  $\varepsilon > 0$  beliebig war, heißt das wie gewünscht, dass  $\|R(x)\| / \|x - p\|^n \rightarrow 0$  für  $x \rightarrow p$ .  $\square$

### E7. Klassische Vektoranalysis

In diesem Abschnitt betrachten wir den Raum  $V = \mathbb{R}^n$  mit dem euklidischen Skalarprodukt  $\langle x, y \rangle = x \cdot y = \sum_i x_i y_i$ . Das Skalarprodukt erlaubt uns, die klassischen Differentialoperatoren  $\text{div}$ ,  $\text{grad}$  und (für  $n = 3$ )  $\text{rot}$  zu definieren. Die bekannten eigenschaften dieser Operatoren erwähnen wir hier, ohne die (meistens sehr einfache) Beweise aufzuschreiben.

**Definition E7.1.** Sei  $G \subset \mathbb{R}^n$  offen. Aus der Physik kennt man die folgenden Namen: eine reelle Funktion  $f: G \rightarrow \mathbb{R}$  heißt *skalares Feld* auf  $G$ ; eine Abbildung  $X: G \rightarrow \mathbb{R}^n$  heißt *Vektorfeld* auf  $G$ .

#### a. Gradient

**Definition E7.2.** Sei  $f$  ein differenzierbares skalares Feld auf  $G \subset \mathbb{R}^n$ . Der *Gradient* von  $f$  ist das folgende Vektorfeld  $\nabla f = \text{grad } f$  auf  $G$ :

$$\text{grad}_p f := \nabla f(p) := (\partial_1 f(p), \dots, \partial_n f(p)).$$

*Bemerkung E7.3.* Das Symbol  $\nabla$  heißt „Nabla“. Als Gedächtnisstütze kann man  $\nabla = (\partial_1, \dots, \partial_n)$  schreiben. (Dieser Gedächtnisstütze eine genaue Bedeutung zu geben – in einem Vektorraum von Differentialoperatoren – würde den Rahmen dieser Vorlesung sprengen.)

*Bemerkung E7.4.* Der Gradient  $\nabla f = \text{grad } f$  wird durch

$$\langle v, \text{grad}_p f \rangle = \langle v, \nabla f(p) \rangle = \partial_v f(p) = D_p f(v) \text{ für alle } v \in V$$

charaktisiert.

*Bemerkung E7.5.* Der Richtungsableitungoperator  $\partial_v$  wird oft (besonders in der Physik) als  $\langle v, \nabla \rangle$  oder  $v \cdot \nabla$  geschrieben:

$$\langle v, \nabla \rangle f := \langle v, \nabla f \rangle = \partial_v f = v \cdot (\nabla f) = (v \cdot \nabla) f.$$

*Bemerkung E7.6.* Der Gradient  $\nabla f$  gibt die Richtung und die Größe des stärksten Wachstums von  $f$  an. Für  $\|v\| = 1$  ist

$$\partial_v f = \langle v, \nabla f \rangle = \|\nabla f\| \cos \varphi,$$

wobei  $\varphi$  der Winkel zwischen  $v$  und  $\nabla f$  ist. Insbesondere ist  $\nabla f$  senkrecht zu den Niveaumengen von  $f$ :  $f$  ist genau dann konstant entlang einer Kurve  $\gamma(t)$  in  $G$ , wenn  $\nabla f \perp \dot{\gamma}(t)$  für alle  $t$ .

*Bemerkung E7.7.* In der Physik ist z.B. das (statische) elektrische Feld der Gradient des elektrischen Potentials. (Allgemeiner ist der Gradient einer jenen potentiellen Energie ein Kraftfeld.)

*Bemerkung E7.8.* Ein skalares Feld  $f$  mit  $\text{grad } f = 0$  ist lokal konstant.

*Beispiel E7.9.* Sei  $v \in \mathbb{R}^n$ . Das skalare Feld  $f(x) := \langle v, x \rangle$  hat konstanten Gradienten  $\text{grad } f \equiv v$ .

#### b. Divergenz

**Definition E7.10.** Sei  $X = (X_1, \dots, X_n)$  ein differenzierbares Vektorfeld auf  $G \subset \mathbb{R}^n$ . Die *Divergenz* von  $X$  ist das folgende skalare Feld  $\langle \nabla, X \rangle = \nabla \cdot X = \text{div } X$  auf  $G$ :

$$\text{div}_p X := \langle \nabla, X \rangle(p) := (\nabla \cdot X)(p) := \sum_{i=1}^n \partial_i X_i(p).$$

*Bemerkung E7.11.* Die *Spur* einer  $(n \times n)$ -Matrix  $A = (A_{ij})$  ist die Summe der Diagonalelemente:  $\text{tr } A := \sum_i A_{ii}$ . Damit gilt  $\text{div}_p X = \text{tr } J_p X$ .

*Bemerkung E7.12.* Die Divergenz eines Vektorfeldes  $X$  kann physikalisch als „Quellendichte“ von  $X$  interpretiert werden. Die Divergenz des Schwerefeldes z.B. ist die (Massen-) Dichte; die Divergenz des elektrischen Feldes ist die Ladungsdichte. Ein Vektorfeld  $X$  mit  $\text{div } X = 0$  heißt *divergenzfrei* oder *quellenfrei*; das Magnetfeld ist quellenfrei. Weitere Beispiele sind Strömungsfelder von inkompressiblen Flüssigkeiten.

*Bemerkung E7.13.* Das Skalarprodukt  $\langle v, w \rangle = \langle w, v \rangle$  ist symmetrisch. Die Schreibweisen  $\langle v, \nabla \rangle = v \cdot \nabla$  bzw.  $\langle \nabla, X \rangle = \nabla \cdot X$  für Richtungsableitungen bzw. Divergenz darf man hingegen nicht verwechseln.

*Beispiel E7.14.* Das Vektorfeld  $X(x) = x$  hat konstante Divergenz  $\text{div } X \equiv n$ .

#### c. Rotation

Auf  $\mathbb{R}^3$  wird das Vektorprodukt  $v \times w \in \mathbb{R}^3$  durch

$$v \times w = (v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_2 w_1)$$

definiert. Dazu verwandt ist die Rotation eines Vektorfeldes.

**Definition E7.15.** Sei  $X = (X_1, X_2, X_3)$  ein differenzierbares Vektorfeld auf  $G \subset \mathbb{R}^3$ . Die *Rotation* von  $X$  ist das folgende Vektorfeld  $\nabla \times X = \text{rot } X = \text{curl } X$  auf  $G$ :

$$\text{rot}_p X := \text{curl}_p X := (\nabla \times X)(p)$$

$$:= (\partial_2 X_3 - \partial_3 X_2, \partial_3 X_1 - \partial_1 X_3, \partial_1 X_2 - \partial_2 X_1)(p).$$

**Bemerkung E7.16.** Sei  $\{e_1, e_2, e_3\}$  die Standardbasis für  $\mathbb{R}^3$ . Als Gedächtnisstütze kann man schreiben:

$$v \times w = \det \begin{pmatrix} e_1 & v_1 & w_1 \\ e_2 & v_2 & w_2 \\ e_3 & v_3 & w_3 \end{pmatrix}, \quad \nabla \times X = \det \begin{pmatrix} e_1 & \partial_1 & X_1 \\ e_2 & \partial_2 & X_2 \\ e_3 & \partial_3 & X_3 \end{pmatrix}.$$

**Bemerkung E7.17.** Es gibt vier sinnvolle Möglichkeiten, Operatoren aus  $\{\text{div}, \text{grad}, \text{rot}\}$  zu kombinieren:  $\text{div}(\text{rot } X)$  und  $\text{div}(\text{grad } f)$  sind skalare Felder;  $\text{rot}(\text{rot } X)$  und  $\text{rot}(\text{grad } f)$  sind Vektorfelder. Den Fall  $\text{rot}(\text{rot } X)$  betrachten wir nicht weiter.

**Satz E7.18.** Seien  $f$  und  $X$  zweimal differenzierbare Felder auf  $G \subset \mathbb{R}^3$ . Es gelten

$$\begin{aligned} 0 &= \text{rot}(\text{grad } f) = \nabla \times (\nabla f), \\ 0 &= \text{div}(\text{rot } X) = \langle \nabla, \nabla \times X \rangle. \end{aligned}$$

*Beweis.* (Aufgabe nach dem Satz von Schwarz.) □

**Bemerkung E7.19.** Um diesen Satz sinnvoll auf höhere Dimensionen zu erweitern, muss man die sogenannte äußere Ableitung von Differentialformen benutzen.

**Bemerkung E7.20.** Beispiel E6.17 können wir jetzt wie folgt verstehen: Sei  $X$  ein Vektorfeld auf  $G \subset \mathbb{R}^3$ . Wann gibt es ein skalares Feld  $g$  mit  $\nabla g = X$ ? Eine notwendige Bedingung dafür ist, dass  $\nabla \times X = 0$ .

**Bemerkung E7.21.** Die Rotation  $\nabla \times X$  heißt auch *Wirbeldichte*. Ein Vektorfeld  $X$  mit  $\nabla \times X = 0$  heißt *wirbelfrei*. Für jedes  $f$  ist der Gradient  $\nabla f$  wirbelfrei. Das Gesetz der elektromagnetischen Induktion sagt, die Wirbeldichte des elektrischen Feldes ist die zeitliche Ableitung des Magnetfeldes.

**Beispiel E7.22.** Sei  $v \in \mathbb{R}^3$ . Das Vektorfeld  $X(x) = v \times x$  beschreibt eine Rotation um die Achse  $\overline{0v}$ . Es gilt  $\text{rot } X \equiv 2v$ .

#### d. Laplaceoperator

**Definition E7.23.** Sei  $f$  ein zweimal differenzierbares skalares Feld auf  $G \subset \mathbb{R}^n$ . Der *Laplaceoperator*  $\Delta = \nabla \cdot \nabla$  liefert ein skalares Feld  $\Delta f$  auf  $G$ :

$$\Delta f := \text{div}(\text{grad } f) = \langle \nabla, \nabla f \rangle = \nabla \cdot (\nabla f) = \sum_{i=1}^n \partial_i^2 f.$$

**Bemerkung E7.24.** Die symmetrische Matrix  $Hf = (\partial_i \partial_j f)$  heißt *Hessematrix* von  $f$ . Es gilt  $\Delta f = \text{tr}(Hf)$ .

**Bemerkung E7.25.** Der Laplaceoperator ist sehr wichtig in der Physik. Ist  $\Phi$  ein Potential, dann ist  $\nabla \Phi$  ein konservatives (d.h. wirbelfreies) Kraftfeld und  $\Delta \Phi$  eine Dichte. Der Laplaceoperator taucht auch z.B. in den partiellen Differentialgleichungen für Wärmeleitung und Wellenausbreitung auf.

**Definition E7.26.** Eine *harmonische Funktion* auf  $G \subset \mathbb{R}^n$  ist ein skalares Feld  $f$  mit  $\Delta f = 0$ .

**Beispiel E7.27.** Jede affin lineare Funktion  $f(x) = \langle x, v \rangle + b$  ist harmonisch. Eine quadratische Funktion  $f$  hat  $\Delta f \equiv c$ , z.B. für  $f(x) = \|x\|^2 = \sum x_i^2$  gilt  $\Delta f \equiv 2n$ .

**Beispiel E7.28.** In der Elektrostatik ist das elektrische Potential harmonisch in jedem Gebiet, wo die Ladungsdichte verschwindet.

**Bemerkung E7.29.** Die Operatoren Gradient, Divergenz, Rotation und Laplaceoperator sind linear:

$$\begin{aligned} \nabla(f + \lambda g) &= \nabla f + \lambda \nabla g, \\ \langle \nabla, X + \lambda Y \rangle &= \langle \nabla, X \rangle + \lambda \langle \nabla, Y \rangle \\ \nabla \times (X + \lambda Y) &= \nabla \times X + \lambda \nabla \times Y, \\ \Delta(f + \lambda g) &= \Delta f + \lambda \Delta g, \end{aligned}$$

### E8. Umkehrsatz

In der Mathematik geht es oft darum, nach den Lösungen einer Gleichung (bzw. eines Gleichungssystems) zu suchen. Schreiben wir die Gleichungen als  $f(x) = y$ , dann sind die Folgenden vielleicht die ersten Fragen: ob es zu jedem  $y$  ein  $x$  gibt mit  $f(x) = y$ , ob dieses  $x$  eindeutig ist und ob kleine Änderungen in den Parametern  $y$  zu kleinen Änderungen in der Lösung  $x$  führen. Das heißt, wir möchten wissen, ob  $f$  eine Umkehrabbildung  $f^{-1}$  hat und ob diese stetig (und vielleicht auch differenzierbar) ist.

Es geht hier nur um die (theoretische) Existenzfrage und nicht z.B. darum, eine schöne Formel für  $f^{-1}$  zu finden. Schon bei den meisten elementaren Funktionen musste man neue Notationen für die Umkehrfunktionen erfinden ( $\sqrt[n]{\cdot}$ ,  $\log$ ,  $\arcsin$  usw.).

**Beispiel E8.1.** Der Fall einer linearen Gleichung  $y = Ax$  mit  $A \in L(\mathbb{R}^m, \mathbb{R}^n)$  kann man mit der linearen Algebra betrachten. Es gibt zu jedem  $y$  eine eindeutige Lösung  $x$  (d.h.,  $A$  ist invertierbar) genau dann, wenn  $n = m$  und  $\det A \neq 0$ . In diesem Fall ist  $A^{-1}$  auch linear, insbesondere stetig und sogar differenzierbar.

**Bemerkung E8.2.** Jede differenzierbare Abbildung  $f$  kann man lokal gut durch ihre Ableitung  $D_p f$  approximieren. Der Umkehrsatz sagt, falls  $D_p f$  invertierbar ist, dann ist auch  $f$  bei  $p$  lokal invertierbar. Wir betrachten die Gleichung  $y = f(x)$  etwa als Störung der affin linearen Gleichung  $y - f(p) = D_p f(x - p)$ . Falls Letzteres invertierbar ist, bleibt diese Eigenschaft unter der Störung erhalten.

**Bemerkung E8.3.** Die Invertierbarkeit von  $D_p f$  ist natürlich keine notwendige Bedingung. Auch wenn z.B. die Ableitung von  $x \mapsto x^3$  in 0 verschwindet, ist  $x \mapsto x^3$  sogar global invertierbar.

**Definition E8.4.** Seien  $U \subset \mathbb{R}^m$  und  $V \subset \mathbb{R}^n$  offen und sei  $f: U \rightarrow V$  eine Bijektion. Mengentheoretisch existiert die Umkehrabbildung  $f^{-1}: V \rightarrow U$ . Falls  $f$  und  $f^{-1}$  stetig sind, heißt  $f$  *Homöomorphismus* zwischen  $U$  und  $V$ . Falls  $f$  und  $f^{-1}$  stetig differenzierbar sind, heißt  $f$  *Diffeomorphismus*.

**Beispiele E8.5.**

- $x \mapsto x^3$  ist ein Homöomorphismus  $\mathbb{R} \rightarrow \mathbb{R}$ , ist aber kein Diffeomorphismus, weil  $\sqrt[3]{x}$  in 0 nicht differenzierbar ist.



- ist  $f$  streng monoton auf einem Intervall  $I \subset \mathbb{R}$ , dann ist  $f$  ein Homöomorphismus  $I \rightarrow f(I)$ .
- $x \mapsto \tan x$  ist ein Diffeomorphismus  $(-\pi/2, \pi/2) \rightarrow \mathbb{R}$ .
- ist  $f' \neq 0$  auf einem Intervall  $I \subset \mathbb{R}$ , dann ist  $f$  ein Diffeomorphismus  $I \rightarrow f(I)$ .

Ende der Vorlesung 2009 Februar 2

**Lemma E8.6.** Seien  $U \subset \mathbb{R}^m$  und  $V \subset \mathbb{R}^n$  offen und sei  $f: U \rightarrow V$  eine Bijektion mit Umkehrabbildung  $g: V \rightarrow U$ . Sei  $p \in U$  so, dass  $f$  in  $p$  differenzierbar ist und  $g$  in  $q := f(p) \in V$  differenzierbar ist. Dann gelten  $m = n$  und  $D_qg = (D_p f)^{-1}$ , insbesondere ist  $D_p f$  invertierbar.

*Beweis.* (Aufgabe nach der Kettenregel.) □

*Bemerkung E8.7.* Dieses Lemma sagt insbesondere, Dimension wird unter Diffeomorphismen erhalten. Dasselbe gilt unter Homöomorphismen, der Beweis dazu ist aber nicht einfach und braucht Techniken aus der algebraischen Topologie. Für je zwei offene Mengen  $U \subset \mathbb{R}^m$  und  $V \subset \mathbb{R}^n$  gibt es aber eine Bijektion  $U \rightarrow V$ : die Mengen sind gleich groß (gleichmächtig). Für  $m \neq n$  sind aber solche Bijektionen nie stetig; sie sind Stoffe eher für die Mengentheorie als für die Analysis.

**Definition E8.8.** Ein Automorphismus des Vektorraumes  $\mathbb{R}^n$  ist eine invertierbare lineare Abbildung  $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Die Menge aller Automorphismen heißt  $GL(\mathbb{R}^n) \subset L(\mathbb{R}^n, \mathbb{R}^n)$ . Sei  $\varphi: GL(\mathbb{R}^n) \rightarrow GL(\mathbb{R}^n)$  die Inversenabbildung  $A \mapsto A^{-1}$ .

**Lemma E8.9.** Für  $A \in GL(\mathbb{R}^n)$  und  $v \in \mathbb{R}^n$  gilt

$$\frac{\|v\|}{\|A^{-1}\|} \leq \|A(v)\| \leq \|A\| \|v\|.$$

*Beweis.* Es gilt  $\|v\| = \|A^{-1}(Av)\| \leq \|A^{-1}\| \|Av\|$ . □

Aus der linearen Algebra wissen wir:

**Lemma E8.10.** In  $L(\mathbb{R}^n, \mathbb{R}^n)$  ist  $GL(\mathbb{R}^n)$  offen. Die Abbildung  $\varphi$  ist stetig.

*Bemerkung E8.11.* Bezüglich der Standardbasis wird  $A \in L(\mathbb{R}^n, \mathbb{R}^n)$  durch eine  $(n \times n)$ -Matrix  $(A_{ij})$  dargestellt. Es gilt genau dann  $A \in GL(\mathbb{R}^n)$ , wenn  $\det(A_{ij}) \neq 0$ .

*Bemerkung E8.12.* Tatsächlich ist  $\varphi$  sogar differenzierbar mit  $D_A\varphi(B) = -A^{-1}BA^{-1}$ . Nehmen wir die Differenzierbarkeit an, können wir leicht diese Formel bestätigen: Wir betrachten

$$\alpha: A \mapsto A\varphi(A) = AA^{-1} = I.$$

Dann gilt

$$0 = D_A\alpha(B) = B\varphi(A) + AD_A\varphi(B) = BA^{-1} + AD_A\varphi(B).$$

**Lemma E8.13.** Seien  $U, V \subset \mathbb{R}^n$  offen und sei  $f: U \rightarrow V$  ein stetig differenzierbarer Homöomorphismus. Falls  $D_p f$  invertierbar ist für jedes  $p \in U$ , dann ist  $f$  ein Diffeomorphismus.

*Beweis.* Weil  $f$  ein Homöomorphismus ist, ist die Umkehrabbildung  $g: V \rightarrow U$  stetig. Wir behaupten,  $g$  ist differenzierbar, d.h., für jedes  $p \in U$  existiert  $D_{f(p)}g$ . Nach Lemma E8.6 können wir dann die Ableitung als  $Dg = \varphi \circ Df \circ g$  schreiben. Die Stetigkeit von  $Dg$  folgt aus der Stetigkeit von  $g$ , von  $Df$  und von  $\varphi$  (Lemma E8.10).

Um die Behauptung zu beweisen, sei  $p \in U$ . Wir wissen,  $D_p f$  ist invertierbar und wollen zeigen,  $g$  ist in  $f(p)$  differenzierbar. Ohne Einschränkung dürfen wir  $p = 0$  annehmen, indem wir  $f$  mit  $x \mapsto f(p+x)$  ersetzen (Translation im Definitionsbereich). Weiter dürfen wir  $f(0) = 0$  annehmen, indem wir  $f$  mit  $x \mapsto f(x) - f(0)$  ersetzen (Translation in der Zielmenge). Schließlich dürfen wir  $D_0 f = \text{id}$  annehmen, indem wir  $f$  mit  $(D_0 f)^{-1} \circ f$  ersetzen (lineare Abbildung in der Zielmenge). Unter diesen Annahmen gilt dann auch  $g(0) = 0$ . Zudem gilt  $D_0 g = \text{id}$ , falls  $D_0 g$  existiert, was wir jetzt zeigen wollen.

Mit der Definition von Differenzierbarkeit im Auge setzen wir

$$f(h) - h =: r(h), \quad g(k) - k =: s(k)$$

für hinreichend kleine  $h, k \in \mathbb{R}^n$ . Für  $k = f(h)$  gilt dann

$$s(k) = g(k) - k = h - f(h) = -r(h).$$

Wir wissen, dass  $r(h)/\|h\| \rightarrow 0$  für  $h \rightarrow 0$  und müssen zeigen, dass  $s(k)/\|k\| \rightarrow 0$  für  $k \rightarrow 0$ .

Es gibt  $\varepsilon > 0$  so, dass für  $\|h\| < \varepsilon$  gilt  $\|r(h)\| \leq \|h\|/2$ . Weil  $g$  stetig ist, gibt es zu diesem  $\varepsilon$  ein  $\delta > 0$  so, dass  $\|g(k)\| < \varepsilon$  für  $\|k\| < \delta$ . Damit gilt  $\|s(k)\| \leq \frac{1}{2}\|g(k)\|$ . Wegen  $g(k) = k + s(k)$  gilt dann nach der Dreiecksungleichung  $\|g(k)\| \leq 2\|k\|$ . Sei nun  $h := g(k)$ . Es gelten  $\|s(k)\| = \|r(h)\|$  und  $\|h\| \leq 2\|k\|$ . Für  $k \rightarrow 0$  konvergiert auch  $h \rightarrow 0$ . In diesem Limes gilt dann

$$\frac{s(k)}{2\|k\|} \leq \frac{r(h)}{\|h\|} \rightarrow 0,$$

wie gewünscht. □

*Beispiel E8.14.* Wir werden den Fixpunktsatz B8.13 von Banach benutzen, um den Umkehrsatz zu beweisen. Wir fangen mit einem einfacheren Beispiel an, wie man eine Gleichung mit dem Fixpunktsatz lösen kann. Sei  $A$  eine lineare Abbildung  $A \in L(\mathbb{R}^n, \mathbb{R}^n)$  mit Operatornorm  $\|A\| < 1$ . Wir behaupten,  $\text{id} - A$  ist invertierbar, d.h., zu jedem  $y \in \mathbb{R}^n$  gibt es ein eindeutiges  $x \in \mathbb{R}^n$  mit

$$(\text{id} - A)(x) = x - Ax = y.$$

Sei  $F_y: \mathbb{R}^n \rightarrow \mathbb{R}^n$  die Abbildung  $x \mapsto Ax + y$ . Sie ist so gewählt, dass ein Fixpunkt von  $F_y$  eine Lösung zur Gleichung ist:

$$x = F_y(x) = Ax + y \iff (\text{id} - A)x = x - Ax = y.$$

Wir sehen,  $F_y$  ist eine Kontraktion, weil

$$\|F_y(x) - F_y(x')\| = \|Ax - Ax'\| \leq \|A\| \|x - x'\|$$

mit  $\|A\| < 1$ . Nach dem Fixpunktsatz hat dann  $F_y$  einen eindeutigen Fixpunkt  $x = Ax + y$ . Das heißt,  $\text{id} - A$  ist invertierbar.

Der Fixpunktsatz gibt uns darüber hinaus eine rekursiv definierte Folge  $(x_n)$ , die gegen den Fixpunkt konvergiert. Mit  $x_0 = 0$  und  $x_{n+1} = F_y(x_n) = Ax_n + y$  gilt

$$x_n = \sum_{i=0}^{n-1} A^i y \quad \rightarrow \quad x = \sum_{i=0}^{\infty} A^i y.$$

Das heißt, wir haben eine Formel für die Umkehrabbildung:  $(\text{id} - A)^{-1} = \sum_{i=0}^{\infty} A^i$  als Reihensumme im Banachraum  $L(V, V)$ .

**Satz E8.15 (Umkehrsatz).** Sei  $U \subset \mathbb{R}^n$  offen und sei  $f: U \rightarrow \mathbb{R}^n$  stetig differenzierbar. Sei  $p \in U$  so, dass  $D_p f \in \text{GL}(\mathbb{R}^n)$  invertierbar ist. Dann ist  $f$  „lokal bei  $p$  ein Diffeomorphismus“, das heißt, es gibt eine offene Menge  $V \subset U$  mit  $V \ni p$  so, dass  $f|_V: V \rightarrow f(V)$  ein Diffeomorphismus ist.

*Beweis.* Wie im Lemma dürfen wir ohne Einschränkung annehmen, dass

$$p = 0, \quad f(0) = 0, \quad D_0 f = \text{id},$$

indem wir  $f$  mit

$$x \mapsto (D_p f)^{-1}(f(p + x) - f(p))$$

ersetzen.

Weil  $Df$  stetig ist und  $\text{GL}(\mathbb{R}^n)$  offen ist, dürfen wir annehmen, dass  $D_q f$  invertierbar ist für alle  $q \in U$ . (Wenn nicht, ersetzen wir  $U$  mit einer kleineren offenen Umgebung von  $p$ .)

Für gegebenes  $y \in \mathbb{R}^n$  suchen wir nach einer Lösung der Gleichung  $f(x) = y$ . Um die Lösung  $x$  als Fixpunkt zu sehen, betrachten wir die Abbildung

$$F_y: U \rightarrow \mathbb{R}^n, \quad F_y(x) := x - f(x) + y.$$

Wir wollen zeigen – für jedes hinreichend kleine  $y$  –, dass  $F_y$  eine Kontraktion  $K \rightarrow K$  ist, wobei  $K \subset U$  kompakt ist.

Wir bemerken,  $D_x F_y = \text{id} - D_x f$ . Weil  $Df$  stetig ist, ist

$$x \mapsto \|D_x F_y\| = \|\text{id} - D_x f\|$$

eine stetige Abbildung  $U \rightarrow \mathbb{R}$  mit  $0 \mapsto 0$ . Das heißt, es gibt  $r > 0$  so, dass  $\|D_x F_y\| < 1/2$  für  $x \in K$ , wobei

$$K := \{x \in \mathbb{R}^n : \|x\| \leq 2r\}$$

eine (konvexe und kompakte) Teilmenge von  $U$  ist. Aus dem Schrankensatz E6.20 gilt dann für  $x, x' \in K$

$$\|F_y(x) - F_y(x')\| \leq \sup_{p \in K} \|D_p F_y\| \|x - x'\| \leq \frac{1}{2} \|x - x'\|. \quad (*)$$

(Insbesondere mit  $x' = 0$  und  $y = 0$  gilt  $\|F_0(x)\| \leq 1/2 \|x\|$ .) Für großes  $y$  ist natürlich  $F_y$  keine Abbildung  $K \rightarrow K$ . Für  $\|y\| < r$  (und  $x \in K$ ) gilt allerdings

$$\|F_y(x)\| \leq \|F_0(x)\| + \|y\| \leq \frac{1}{2} \|x\| + \|y\| < 2r,$$

d.h.,  $F_y: K \rightarrow K$ . Diese Abbildung ist dann eine Kontraktion (mit Lipschitzkonstante  $1/2$ ) und hat deshalb einen eindeutigen Fixpunkt  $x \in K$ .

Das heißt, zu jedem  $y \in B_r(0)$  gibt es genau ein  $x \in K$  mit  $f(x) = y$ . Wir setzen  $V := \{x \in K : f(x) \in B_r(0)\}$  und sehen, dass  $f|_V: V \rightarrow B_r(0)$  eine Bijektion ist.

Sei  $g := (f|_V)^{-1}$ . Nach (\*) und der Dreiecksungleichung gilt

$$\begin{aligned} \|x - x'\| - \|f(x) - f(x')\| &\leq \|x - x' - f(x) + f(x')\| \\ &= \|F_0(x) - F_0(x')\| \leq \frac{1}{2} \|x - x'\|. \end{aligned}$$

Damit folgt, dass  $g$  eine Lipschitzabbildung ist mit Lipschitzkonstante 2:

$$\|g(y) - g(y')\| \leq 2 \|f(g(y)) - f(g(y'))\| = 2 \|y - y'\|.$$

Insbesondere ist  $g$  stetig. Damit ist  $f|_V$  ein Homöomorphismus. Weil  $D_x f$  invertierbar ist für jedes  $x \in V \subset U$ , ist nach dem Lemma dieser Homöomorphismus sogar ein Diffeomorphismus.  $\square$

---

Ende der Vorlesung 2009 Februar 5

---

### E9. Implizite Funktionen und Mannigfaltigkeiten

Der Umkehrsatz untersucht die Lösbarkeit eines nichtlinearen Gleichungssystems in dem Fall, wo die Anzahl der Gleichungen mit der Anzahl der Variablen übereinstimmt und damit eine eindeutige Lösung zu erwarten ist. Wenn wir mehr Variablen als Gleichungen haben, erwarten wir mehrere Lösungen. Ist  $F$  eine Abbildung  $\mathbb{R}^{n+k} \rightarrow \mathbb{R}^k$ , dann erwarten wir – für ein festes  $c \in \mathbb{R}^k$  – eine  $n$ -dimensionale Familie von Lösungen zu  $F(z) = c$ . Wenn wir  $z \in \mathbb{R}^{n+k} = \mathbb{R}^n \times \mathbb{R}^k$  als  $z = (x, y)$  schreiben (mit  $x \in \mathbb{R}^n$  und  $y \in \mathbb{R}^k$ ), dann können wir hoffen, diese Familie wird (mindestens lokal) durch  $x$  parametrisiert mit  $y = g(x)$ . Das heißt, wir suchen nach einer Auflösung der Gleichung  $F(x, y) = c$  nach  $y$  als Funktion von  $x$ . Wir sagen auch, die Gleichung  $F(x, y) = c$  definiert implizit die Funktion  $y = g(x)$ .

Hier betrachten wir den festen Wert  $c \in \mathbb{R}^k$  als Konstante. Indem wir  $F$  mit  $F - c$  ersetzen, können und werden wir durchgehend  $c = 0$  annehmen.

*Beispiel E9.1.* Eine lineare Abbildung  $F: \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^k$  hat immer die Form

$$F(x, y) = F(x, 0) + F(0, y) =: F'(x) + F''(y),$$

wobei  $F': \mathbb{R}^n \rightarrow \mathbb{R}^k$  und  $F'': \mathbb{R}^k \rightarrow \mathbb{R}^k$  linear sind. Falls  $F''$  invertierbar ist, hat die Gleichung  $F(x, y) = 0$  eine eindeutige Auflösung  $y = g(x)$ : es gilt nämlich

$$y = g(x) := -(F'')^{-1}(F'(x)).$$

*Bemerkung E9.2.* Sei  $F: \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^k$  eine differenzierbare Abbildung. Wir schreiben

$$D'_{(p,q)} F(v) := D_{(p,q)} F(v, 0), \quad D''_{(p,q)} F(w) := D_{(p,q)} F(0, w).$$

Damit ist  $D'_{p,q} F: \mathbb{R}^k \rightarrow \mathbb{R}^k$  linear.

*Beispiel E9.3.* Sei  $n = k$  und sei  $F(x, y) := -x + f(y)$  für eine stetig differenzierbare Abbildung  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Die Gleichung  $F(x, y) = 0$  heißt dann  $x = f(y)$ . Der Umkehrsatz sagt, es gibt eine lokale Auflösung der Form  $y = g(x)$  (wobei natürlich  $g = f^{-1}$ ), falls  $D_q f$  invertierbar ist. Wie sieht  $D_{p,q} F$  aus? Wir haben  $D_{(p,q)} F(v, w) = -v + D_q f(w)$ . Das heißt, die Ableitung von  $F$  bezüglich der  $y$ -Variablen ist  $D''_{(p,q)} F = D_q f$ ; diese muss invertierbar sein, um den Umkehrsatz anzuwenden.

*Beispiel E9.4.* Sei  $F: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  durch  $F(x, y) = x^2 + y^2 - 1$  definiert. Die Lösungen der Gleichung  $F(x, y) = 0$  sind natürlich genau die Punkte des Einheitskreises  $K$ . Lokal können wir  $y = \pm \sqrt{1 - x^2} =: g_{\pm}(x)$  schreiben: Bei einem Punkt  $(p, q) \in K$  mit  $q > 0$  (bzw.  $q < 0$ ) gilt lokal  $y = g_+(x)$  (bzw.  $y = g_-(x)$ ). Bei den beiden Punkten  $(\pm 1, 0)$  lässt sich  $x^2 + y^2 - 1 = 0$  nicht in der Form  $y = g(x)$  auflösen. (Da gilt  $\partial_y F = 0$ , d.h.  $D'' f = 0$ ). Im Allgemeinen gilt  $D_{(p,q)} f(v, w) = 2pv + 2qw$ .

*Bemerkung E9.5.* Sei  $y = g(x)$  eine lokale Auflösung der Gleichung  $F(x, y) = 0$ . Geometrisch können wir sagen, die Niveaumenge  $\{(x, y) : F = 0\}$  stimmt lokal mit dem Graphen  $\{(x, g(x))\}$  von  $g$  überein. Dies ist nicht der Fall bei einem Punkt wo die Niveaumenge „vertikal“ ist.

*Beispiel E9.6.* Sei  $f(x, y) = x^2 - y^2$ . Die Niveaumenge  $\{f = 0\}$  besteht aus zwei Geraden  $y = \pm x$ . Bei  $(0, 0)$  lässt sie sich nicht als einen Graphen darstellen. Die Ableitung von  $f$  verschwindet in  $(0, 0)$ .

**Satz E9.7 (Satz über implizite Funktionen).** Sei  $W$  eine offene Teilmenge in  $\mathbb{R}^n \times \mathbb{R}^k$  und sei  $F: W \rightarrow \mathbb{R}^k$  stetig differenzierbar. Sei  $(p, q) \in W$  mit  $F(p, q) = 0$ . Falls die lineare Abbildung

$$D''_{(p,q)} F: \mathbb{R}^k \rightarrow \mathbb{R}^k$$

invertierbar ist, dann lässt sich  $F(x, y) = 0$  lokal bei  $(p, q)$  nach einer stetig differenzierbaren Abbildung  $y = g(x)$  auflösen. Das heißt, es gibt Umgebungen  $U \subset \mathbb{R}^n$  von  $p$  und  $V \subset \mathbb{R}^k$  von  $q$  mit  $U \times V \subset W$  und eine stetig differenzierbare Abbildung  $g: U \rightarrow V$  so, dass

$$F(x, y) = 0 \text{ mit } x \in U, y \in V \iff y = g(x) \text{ mit } x \in U.$$

Ferner ist  $D''_{(x,g(x))} F$  invertierbar für jedes  $x \in U$  und es gilt

$$D_x g = -(D''_{(x,g(x))} F)^{-1} \circ D'_{(x,g(x))} F.$$

*Beweis.* Wir wollen den Umkehrsatz auf die stetig differenzierbare Abbildung

$$f: W \rightarrow \mathbb{R}^n \times \mathbb{R}^k, \quad f(x, y) = (x, F(x, y))$$

anwenden. Es gilt  $D_{(x,y)} f(v, w) = (v, D_{(x,y)} F(v, w))$  und damit

$$\begin{aligned} D_{(x,y)} f(v, w) = 0 &\iff v = 0 \text{ und } D_{(x,y)} F(0, w) = 0 \\ &\iff v = 0 \text{ und } D''_{(x,y)} F(w) = 0. \end{aligned}$$

Daraus folgt, dass  $D_{(x,y)} f$  genau dann invertierbar ist, wenn  $D''_{(x,y)} F$  invertierbar ist. (Hier gilt die obige Gleichung nur für  $w = 0$ .)

Insbesondere ist  $D_{(p,q)} f$  invertierbar. Nach dem Umkehrsatz gibt es dann eine Umgebung  $G \subset W$  von  $(p, q)$  so, dass  $f$  eine stetig differenzierbare Umkehrabbildung

$$(x, y) \mapsto (\varphi(x, y), \psi(x, y))$$

auf  $f(G)$  hat. Wegen

$$(x, y) = f(\varphi(x, y), \psi(x, y)) = (\varphi(x, y), F(\varphi(x, y), \psi(x, y)))$$

gilt  $\varphi(x, y) = x$ . Für  $(x, y) \in G$  gilt deshalb

$$\begin{aligned} F(x, y) = 0 &\iff f(x, y) = (x, 0) \\ &\iff (x, y) = (x, \psi(x, 0)) \iff y = \psi(x, 0). \end{aligned}$$

Wir definieren  $\tilde{U} := \{x : (x, 0) \in G\} \subset \mathbb{R}^n$  als Schnitt von  $G$  und  $g: \tilde{U} \rightarrow \mathbb{R}^k$  durch  $g(x) := \varphi(x, 0)$ . Dann ist  $g$  eine Auflösung der Gleichung nach  $y$ .

---

Ende der Vorlesung 2009 Februar 9

Wir müssen nur ggf. eine kleinere Umgebung  $U \subset \tilde{U}$  wählen, falls  $\tilde{U} \times g(\tilde{U}) \not\subset G$ . Dazu wählen wir zunächst  $\varepsilon > 0$  hinreichend klein, dass  $B_\varepsilon(p) \times B_\varepsilon(q) \subset G$ ; wir definieren  $V := B_\varepsilon(q)$ . Weil  $g$  stetig ist, gibt es  $\delta \in (0, \varepsilon]$  so, dass mit  $U := B_\delta(p)$  gilt  $g(U) \subset V$ . Es gilt auch  $U \times V \subset G \subset W$  und wir haben die gewünschte Auflösung.

Um die letzte Aussage zu beweisen, sei  $h: U \rightarrow G$  die Abbildung  $x \mapsto (x, g(x))$ . Nach dem Umkehrsatz ist  $f$  ein Diffeomorphismus auf  $G$ , insbesondere ist  $D_{h(x)} f$  invertierbar für jedes  $x \in U$ . Wie wir oben zeigten, heißt das,  $D''_{h(x)} F$  ist invertierbar. Es gilt  $0 \equiv F \circ h$  und damit

$$\begin{aligned} 0 = D_x(F \circ h)(v) &= D_{h(x)} F(D_x h)(v) = D_{h(x)} F(v, D_x g(v)) \\ &= D'_{h(x)} F(v) + D''_{h(x)} F(D_x g(v)). \end{aligned}$$

Weil  $D''_{h(x)} F$  invertierbar ist, folgt die gewünschte Formel für  $D_x g$ .  $\square$

*Bemerkung E9.8.* Für  $k = 1$  haben wir  $F: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ . Die zu untersuchende Ableitung  $D''_{(p,q)} F$  ist einfach die letzte partielle Ableitung von  $F$ , die man lieber  $\partial_{n+1} F(p, q) = \partial_y F(p, q)$  nennt. Ist diese ungleich Null, können wir lokal  $y = g(x)$  schreiben.

*Beispiel E9.9.* Sei  $P = t^2 - 2at + b$  ein quadratisches Polynom. Sind die Nullstellen differenzierbare Funktionen von  $(a, b)$ ? Wir betrachten  $P$  als Abbildung

$$P: \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}, \quad P(a, b, t) = t^2 - 2at + b.$$

Der Satz über implizite Funktionen sagt, wir können  $P = 0$  lokal in der Form  $t = t(a, b)$  auflösen, falls

$$0 \neq \partial_3 P(a, b, t) = \frac{\partial P}{\partial t}(a, b, t) = 2t - 2a,$$

d.h., falls  $t \neq a$ . Aus der Schule wissen wir,

$$P(a, b, t) = t^2 - 2at + b = (t - t_+)(t - t_-),$$

wobei

$$t_{\pm} := t_{\pm}(a, b) := a \pm \sqrt{a^2 - b}.$$

Wir bemerken,

$$t = a \iff a^2 = b \iff t_+ = t_- \iff P = (t - a)^2.$$

Das heißt, der Satz über implizite Funktionen scheitert genau dann, wenn  $P$  eine doppelte Nullstelle hat.

Die Menge  $\{(a, b, t) : t^2 - 2at + b = 0\}$  ist eine Fläche in  $\mathbb{R}^2 \times \mathbb{R}$ . Über dem Gebiet  $\{(a, b) : a^2 > b\}$  in der Ebene  $\mathbb{R}^2$  hat die Fläche zwei Schichten. Jede Schicht ist der Graph einer differenzierbaren Funktion  $t_{\pm}$ . Über der Parabel  $\{a^2 = b\}$  wird die Fläche gefaltet; da ist die Tangentenebene vertikal.

*Beispiel E9.10.* Allgemeiner betrachten wir ein Polynom beliebigen Grades  $n$  als Abbildung

$$P: \mathbb{R}^{n+1} \times \mathbb{R} \rightarrow \mathbb{R}, \quad P(a, t) = P_a(t) = \sum_{i=0}^n a_i t^i.$$

Sei  $(a, c)$  eine Nullstelle von  $P$ , d.h., das Polynom  $P_a$  hat  $c$  als Nullstelle:  $P_a(c) = 0$ . Wir können den Satz über implizite Funktionen genau dann anwenden, wenn  $\partial_t P_a(c) \neq 0$ . Was ist aber diese partielle Ableitung? Nach Satz C4.9 können wir  $P_a(t) = (t - c)Q(t)$  schreiben, wobei  $Q$  ein Polynom ist. Dann gilt  $\partial_t P_a(c) = 0Q'(c) + 1Q(c) = Q(c)$ . Das heißt, der Satz über implizite Funktionen scheitert genau dann, wenn  $c$  eine Nullstelle von  $Q$  ist, d.h., wenn  $c$  eine doppelte Nullstelle von  $P_a$  ist. Daraus schließen wir, die Nullstellen eines Polynoms sind differenzierbare Funktionen der Koeffizienten, solange jede Nullstelle eine einfache Nullstelle ist.

Eine Mannigfaltigkeit ist ein metrischer Raum, der lokal dem euklidischen Raum  $\mathbb{R}^n$  gleicht. Man kann Mannigfaltigkeiten abstrakt definieren (vgl. Vorlesungen zu Differentialtopologie bzw. -geometrie). Man zeigt, jede Mannigfaltigkeit kann in einen höher-dimensionalen euklidischen Raum  $\mathbb{R}^{n+k}$  eingebettet werden. Hier betrachten wir nur solche Untermannigfaltigkeiten; lokal gleichen diese dem Untervektorraum  $\mathbb{R}^n \subset \mathbb{R}^{n+k}$ .

**Definition E9.11.** Eine Teilmenge  $M \subset \mathbb{R}^{n+k}$  heißt *n-dimensionale Untermannigfaltigkeit*, falls es zu jedem Punkt  $p \in M$  folgendes gibt:

- eine offene Umgebung  $U \subset \mathbb{R}^{n+k}$  von  $p$ ,
- eine offene Menge  $V \subset \mathbb{R}^{n+k}$  und
- einen Diffeomorphismus  $F: U \rightarrow V$  so, dass

$$F(U \cap M) = V \cap (\mathbb{R}^n \times \{0\}).$$

Wir sagen auch,  $M$  hat *Kodimension*  $k$  in  $\mathbb{R}^{n+k}$ . Der Diffeomorphismus  $F$  heißt *Karte* für  $M$  bei  $p$ .

*Beispiel E9.12.* Sei  $G \subset \mathbb{R}^n$  offen und sei  $g: G \rightarrow \mathbb{R}^k$  stetig differenzierbar. Der Graph

$$\Gamma := \{(x, g(x)) : x \in G\} \subset G \times \mathbb{R}^k$$

ist dann eine  $n$ -dimensionale Untermannigfaltigkeit: unabhängig von  $p \in \Gamma$  können wir  $U := V := G \times \mathbb{R}^k$  und

$$F(x, y) := (x, y - g(x))$$

setzen. Auch die Umkehrabbildung  $F^{-1}(x, y) = (x, y + g(x))$  ist offensichtlich stetig differenzierbar.

*Beispiel E9.13.* Jede offene Teilmenge  $U \subset \mathbb{R}^n$  ist eine  $n$ -dimensionale Untermannigfaltigkeit von  $\mathbb{R}^n$ .

**Definition E9.14.** Sei  $U \subset \mathbb{R}^{n+k}$  offen und sei  $f: U \rightarrow \mathbb{R}^k$  stetig differenzierbar. Ein Punkt  $x \in U$  heißt *kritischer Punkt*, falls  $D_x f: \mathbb{R}^{n+k} \rightarrow \mathbb{R}^k$  nicht surjektiv ist. Ein Punkt  $y \in \mathbb{R}^k$  heißt *kritischer Wert*, falls es einen kritischen Punkt  $x \in U$  gibt mit  $y = f(x)$ ; sonst heißt  $y \in \mathbb{R}^k$  *regulärer Wert*.

*Beispiele E9.15.*

- Für eine konstante Abbildung  $f(x) \equiv c \in \mathbb{R}^k$  ( $k > 0$ ) ist jeder Punkt  $x \in U$  kritisch. Jeder Wert  $y \neq c$  ist regulär, während  $c$  ein kritischer Wert ist.
- Für  $f: \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ ,  $f(x) := \|x\|^2 = \langle x, x \rangle$  ist  $x = 0$  der einzige kritische Punkt. Damit ist  $y = 0$  der einzige kritische Wert.
- Ein Wert, der nicht angenommen wird ( $y \notin f(U)$ ), ist regulär.
- Für  $k = 1$  ist  $x$  genau dann ein regulärer Punkt, wenn  $\text{grad}_x f \neq 0$ .

Der Satz über implizite Funktionen sagt, die Niveaumenge in einem regulären Wert ist eine Untermannigfaltigkeit. Dazu brauchen wir ein Lemma aus der linearen Algebra.

**Lemma E9.16.** Sei  $F: \mathbb{R}^{n+k} \rightarrow \mathbb{R}^k$  eine lineare und surjektive Abbildung. Nach einer Permutation (Umordnung) der Koordinaten in  $\mathbb{R}^{n+k}$  können wir  $F(x, y) = F'(x) + F''(y)$  schreiben, wobei  $F'': \mathbb{R}^k \rightarrow \mathbb{R}^k$  bijektiv ist.

**Satz E9.17.** Eine Teilmenge  $M \subset \mathbb{R}^{n+k}$  ist genau dann eine  $n$ -dimensionale Untermannigfaltigkeit, wenn es zu jedem Punkt  $p \in M$  eine offene Umgebung  $U \subset \mathbb{R}^{n+k}$  und eine stetig differenzierbare Abbildung  $h: U \rightarrow \mathbb{R}^k$  so gibt, dass  $M \cap U = h^{-1}(c)$ , wobei  $c \in \mathbb{R}^k$  ein regulärer Wert von  $h$  ist.

*Beweis.* „ $\implies$ “: Sei  $p \in M$  und sei  $F: U \rightarrow \mathbb{R}^{n+k}$  eine Karte für  $M$  bei  $p$ . Wir schreiben  $F = (f, h)$ , wobei  $h: U \rightarrow \mathbb{R}^k$ . Dann gilt  $M \cap U = h^{-1}(0)$ . Weil  $F$  ein Diffeomorphismus ist, ist  $D_x F$  in jedem Punkt  $x$  surjektiv, daher auch  $D_x h$ .

„ $\impliedby$ “: Sei  $p \in M$ . Weil  $p$  ein regulärer Punkt von  $h$  ist, können wir die Koordinaten nach dem Lemma so umordnen, dass  $D'_p h$  surjektiv ist. Damit können wir den Satz über implizite Funktionen anwenden: lokal ist die Niveaumenge  $h^{-1}(c)$  der Graph einer Funktion  $g: U' \rightarrow \mathbb{R}^k$ . Nach Beispiel E9.12 ist dieser Graph eine Untermannigfaltigkeit.  $\square$

**Korollar E9.18.** Sei  $c$  ein regulärer Wert von  $h: U \rightarrow \mathbb{R}^k$ . Dann ist die Niveaumenge  $h^{-1}(c) \subset U \subset \mathbb{R}^{n+k}$  eine Untermannigfaltigkeit.

*Beweis.* Zu jedem  $p \in M$  benutzen wir im Satz die gegebenen  $U$  und  $h$ .  $\square$

*Beispiel E9.19.* Die Einheitssphäre  $\mathbb{S}^n \subset \mathbb{R}^{n+1}$  ist eine Untermannigfaltigkeit.