Analysis 2

Andreas Kriegl

23. November 2004

Dies ist das Skriptum zu meiner gleichnamigen Vorlesung im Studienjahr 2003/2004. Darin wird die Integration von Funktionen in einer 1-dimensionalen Variable sowie die Differentiation von Funktionen in mehreren Variablen behandelt.

Ich setzte dabei die konsequente Verwendung von Räumen von Funktionen aus der Analysis 1 fort, da meiner Meinung nach nur dadurch der vollständige Sachverhalt einiger zentraler Sätze der Analysis verstanden werden kann. Dies führt dazu Banach-Räume zu betrachten, denn die auftretenden Funktionenräume sind, abgesehen von jenen linearer Abbildungen, unendlich dimensional.

Mein Hauptmotiv dafür alle Begriffe der mehrdimensionalen Analysis in Banach-Räumen vorzustellen und erst in der Folge auf endlich dimensionalen Vektorräumen und insbesonders dem $\mathbb{R}^n$ zu diskutieren besteht allerdings darin, daß ich die geometrische Natur der Konzepte betonen will und nicht vorzeitig die für Rechenzwecke unvermeidlichen, jedoch geometrisch nicht a priori vorhandenen Koordinaten verwenden will. Die LeserIn kann also die auftretenden Banach-Räume E, F etc. in weiten Teilen des Skriptums als endlich dimensionale Vektorräume auffassen, die allerdings nicht mit einer vorgegebenen Basis versehen sind. Es ist mir klar, daß dadurch höhere Abstraktionsansprüche an die LeserIn gestellt werden, und möglicherweise auch die vermeintliche Sicherheit, die konkretes Rechnen erzeugt leidet. Aber einerseits ist Abstraktion – jedenfalls in meinen Augen – auch das Erfolgsrezept der Mathematik – man denke nur an den Gauß'schen Algorithmus für ein lineares Gleichungssystem im Vergleich zur Interpretation als lineare Gleichung $a(x) = y$ für Vektoren – und andererseits ist Rechnen mit Abbildungen (nach entsprechender Übung) leichter als sich durch den Dschungel der Koordinaten zu wursteln, wo es oft genialer Intuition bedarf allgemeine Resultate zu erkennen (wie z.B. eine Formel für die Ableitung des Invertierens von Matrizen).

Ich habe mich bemüht, als Gegengewicht zu diesem abstrakteren Zugang durch Visualisierung der geometrischen Konzepte in Form zahlreicher Graphiken und Animationen auch einen sehr konkrete Blickwinkel zu bieten. Links auf (gif-)Animationen und solche auf interaktive (java-)Animationen sind durch nebenstehende Symbole gekennzeichnet:

Ich habe diejenigen Teile, die über diese Vorlesung hinausgehen, und für jene gedacht sind, die keine Angst haben zeitweilig ein wenig den Boden unter den Füßen zu verlieren und in höhere Sphären aufzusteigen, durch linkseitiges Symbol eines Hängegleiters gekennzeichnet. Insbesonders habe ich mich bemüht, an den adäquaten Stellen einige Ausblicke auf andere mathematische Gebiete zu geben, die eine natürliche Fortsetzung der hier vorgestellten Konzepte bilden.

EXAM

Weitere Resultate oder Beweise, die zwar für die Vorlesung relevant sind, aber auf Grund ihrer Komplexität nicht zur Prüfung kommen, habe ich mit linksseitigem Symbol gekennzeichnet.

Sicherlich wird die aufmerksame LeserIn (Tipp-)Fehler finden können. Ich möchte folglich wie immer die Bitte aussprechen, mir diese mitzuteilen (geteiltes Leid ist halbes Leid). Ich werde diese in der redigierten Version natürlich berücksichtigen.

Andreas Kriegl, Wien im April 2004

Zahlreiche Korrekturen verdanke ich Helge Krüger. Auf Anregung von Roland Steinbauer habe ich bei der Definition (5.5.3) eines Banach-Raums den Zusammenhang zwischen Norm und Metrik expliziter gemacht und aud Empfehlung von Michael Grosser habe ich die Definition (6.1.1) der Differenzierbarkeit in einzelne Teile zerlegt.

Andreas Kriegl, Wien im Juli 2004

Sehr umfangreiche Listen mit Korrekturen haben mir dankenswerter Weise Annegret Burtscher und Philipp John sowie Florian Gach zur Verfügung gestellt.

Andreas Kriegl, Wien im Oktober 2004

Inhaltsverzeichnis

5 Integralrechnung **1**

5.1 Bestimmte Integrale . . . 1

5.2 Unbestimmte Integrale . . . 16

5.3 Uneigentliche Integrale . . . 30

5.4 Banach-Räume . . . 36

5.5 Analysis von Kurven . . . 50

6 Differentiation im mehrdimensionalen **64**

6.1 Differenzierbarkeit . . . 64

6.2 Lösung von Gleichungen . . . 85

6.3 Höhere Ableitungen . . . 114

6.4 Lokale Extrema . . . 127

6.5 Stammfunktionen & Kurven-Integrale . . . 138

6.6 Komplexe Analysis . . . 148

Bibliographie **156**

Liste der Symbole **157**

Index **161**

5 Integralrechnung

5.1 Bestimmte Integrale

5.1.1 Definition. Integral.
Wir wollen nun allgemeine Flächen berechnen. Für Rechtecke ist das (mehr oder minder nach Definition) das Produkt aus Länge und Breite. Für den Kreis haben wir das in (2.1.2) durch Approximation mit ein- und umschriebenen Vielecken erreicht. Ganz ähnlich gehen wir nun vor, um die durch eine Funktion $f : [a, b] \to \mathbb{R}$ und die x-Achse begrenzte Fläche zu bestimmen.

Dazu benötigen wir Zerlegungen Z des Intervalls $[a, b]$ in endlich viele Teilintervalle $I_1, \dots, I_N$, die wir durch die endliche Folge der Teilungspunkte $a = t_0 \le t_1 \le \dots \le t_{N_1} \le t_N = b$ vermöge $I_1 := [t_0, t_1]$,…,$I_i := [t_{i-1}, t_i]$,…,$I_N := [t_{N-1}, t_N]$ beschreiben können. Wir werden (logisch nicht völlig korrekt) zwischen der Interpretation der Zerlegung Z als $Z = \{I_1, \dots, I_N\}$ bzw. als $Z = \{t_0, \dots, t_N\}$ je nach Bedarf hin- und herwechseln.

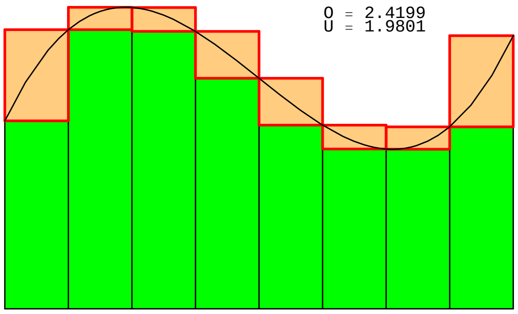
Es sei $f : \mathbb{R} \supseteq I \to \mathbb{R}$ eine beschränkte Abbildung auf einem kompakten Intervall I. Es sei $Z = \{I_1, \dots, I_N\}$ eine endliche Zerlegung von $I = [a, b]$ in Teilintervalle $I_1, \dots, I_N$. Als Obersumme von f bzgl. der Zerlegung Z bezeichnet man

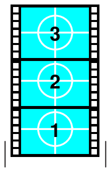
$$O(f, Z) := \sum_{i=1}^{N} |I_i| \sup(f(I_i)) = \sum_{J \in Z} |J| \sup(f(J))$$

und als Untersumme

$$U(f, Z) := \sum_{i=1}^{N} |I_i| \inf(f(I_i)) = \sum_{J \in Z} |J| \inf(f(J)),$$

wobei $|J|$ die Länge des Intervalls J bezeichnet.





Weiters sei

$$O(f) := \inf\{O(f, Z) : Z \text{ ist Zerlegung von } I\}$$
$$U(f) := \sup\{U(f, Z) : Z \text{ ist Zerlegung von } I\}$$

Idee von $O(f)$ ist, daß wir die gesuchte Fläche durch Obersummen $O(f, Z)$ beliebig genau von oben approximieren können, sie also $O(f)$ sein sollte. Und ganz analog sollte sie von unten durch Untersummen $U(f, Z)$ approximierbar sein, also gleich $U(f)$ sein.

Es ist $O(f) \ge U(f)$, denn seien Z_1 und Z_2 zwei Zerlegungen und $Z := Z_1 \cup Z_2$ die gemeinsame Verfeinerung bestehend aus allen in Z_1 oder Z_2 vorkommenden Teilungspunkten. Dann ist $U(f, Z_1) \le U(f, Z) \le O(f, Z) \le O(f, Z_2)$ und somit $U(f) = \sup\{U(f, Z_1) : Z_1\} \le O(f, Z_2)$ und schließlich $U(f) \le \inf\{O(f, Z_2) : Z_2\} = O(f)$.

Die Abbildung f heißt nun Darboux-integrierbar (kurz D-integrierbar oder auch integrierbar), falls $O(f) = U(f)$, d.h.

$$\forall \varepsilon > 0\, \exists\, Z_1, Z_2 : O(f, Z_1) < U(f, Z_2) + \varepsilon,$$

und man schreibt

$$\int_I f = \int_a^b f = \int_a^b f(x)$$

für diesen Wert, und nennt dies das (BESTIMMTE) DARBOUX-INTEGRAL (oder kurz INTEGRAL) der Abbildung f von a bis b.

Es ist $O(f) = U(f)$ genau dann, wenn $\forall \varepsilon > 0 \;\; \exists Z_1, Z_2$ Zerlegungen : $O(f, Z_1) - U(f, Z_2) < \varepsilon$.

Beweis. ($\Rightarrow$) Sei $\varepsilon > 0$.

$$\begin{aligned} O(f) = \inf\{O(f,Z) : Z\} &\Rightarrow \;\exists Z_1 : O(f, Z_1) < O(f) + \frac{\varepsilon}{2} \\ U(f) = \sup\{U(f,Z) : Z\} &\Rightarrow \;\exists Z_2 : U(f, Z_2) > U(f) - \frac{\varepsilon}{2} \\ &\Rightarrow O(f, Z_1) - U(f, Z_2) < O(f) + \frac{\varepsilon}{2} - U(f) + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

($\Leftarrow$) Für $\varepsilon > 0$ ist nach Voraussetzung $0 \le O(f) - U(f) \le O(f, Z_1) - U(f, Z_2) < \varepsilon$ für gewisse Zerlegungen Z_1 und Z_2 und somit ist $O(f) = U(f)$. □

5.1.2 Lemma. Vieles ist integrierbar.
Jede monotone Funktion und jede stetige Funktion ist D-integrierbar.

Monotone Funktionen können in einzelnen Punkten a unstetig sein, denn es existiert $\lim_{x\to a-} f(x)$, da f monoton und durch $f(a)$ beschränkt ist, und ebenso existiert $\lim_{x\to a+} f(x) \ge f(a) \ge \lim_{x\to a-} f(x)$ (für monoton wachsende f), also kann schlimmstenfalls eine Sprungstelle vorliegen. Davon können aber nicht zu viele vorhanden sein, denn die Summe der Sprunghöhen von $f|_{[a,b]}$ ist höchstens $f(b)-f(a)$. Es können also nicht überabzählbar viele Sprungstellen vorhanden sein, denn andernfalls gäbe es ein $\varepsilon > 0$, s.d. überabzahlbar viele mindestens ε sind, und deren gesamte Sprunghöhe wäre somit $\ge \infty \cdot \varepsilon = \infty$, ein Widerspruch. Siehe in diesen Zusammenhang auch (5.1.16).

Beweis. (1) O.B.d.A. sei f monoton wachsend. Weiters sei eine ÄQUIDISTANTE ZERLEGUNG Z von $[a, b]$ durch $a = t_0 < t_1 < \cdots < t_N = b$ gegeben mit $t_i := a + i\frac{b-a}{N}$, also $Z = \{I_1, \ldots, I_N\}$ mit $I_i := [t_{i-1}, t_i]$. Dann ist $\sup f(I_i) = f(t_i)$ und $\inf f(I_i) = f(t_{i-1})$ also ist $O(f, Z) - U(f, Z) = \sum_{i=1}^N (f(t_i) - f(t_{i-1}))\,\frac{b-a}{N} = (f(b) - f(a))\,\frac{b-a}{N} \to 0$ für $N \to \infty$.

(2) Sei nun f stetig, dann ist f gleichmäßig stetig nach (3.3.5), also existiert zu $\varepsilon > 0$ ein $\delta > 0$ mit $|f(x) - f(y)| \le \varepsilon$ für $|x - y| \le \delta$. Sei nun Z eine äquidistante Zerlegung wie zuvor mit Schrittweite $|t_i - t_{i-1}| \le \delta$. Für $J \in Z$ ist dann $\sup f(J) - \inf f(J) = \sup\{f(x) - f(y) : x, y \in J\} \le \varepsilon$ und somit $O(f, Z) - U(f, Z) \le \sum_{j=1}^N \varepsilon\,|J_j| = \varepsilon\,(b - a)$. Also gilt $O(f) = U(f)$. □

5.1.3 Beispiele integrierbarer Funktionen.

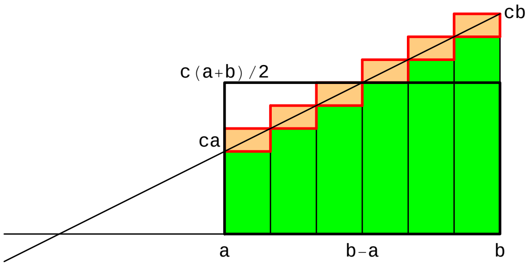
1. Für konstante Funktionen $f : x \mapsto c$ ist $U(f, Z) = (b - a)\,c = O(f, Z)$ und somit $\int_a^b f(x) = (a - b)\,c$. Beachte, daß für $c < 0$ das Integral $\int_a^b c < 0$ ist, also die so berechnete Fläche mit einem Vorzeichen versehen ist (welches die Orientierung der Fläche kodiert).

2. Für lineare Funktionen $f : x \mapsto c\,x$ mit $c > 0$ und Z gegeben durch $a = t_0 < \cdots < t_n = b$ ist

   $$U(f, Z) = \sum_{j=1}^n (t_j - t_{j-1})\,c\,t_{j-1} \quad \text{und} \quad O(f, Z) = \sum_{j=1}^n (t_j - t_{j-1})\,c\,t_j$$

   und somit

   $$\begin{aligned} \frac{1}{2}\Big(U(f, Z) + O(f, Z)\Big) &= \frac{1}{2}\sum_{j=1}^n (t_j - t_{j-1})\,c\,(t_j + t_{j-1}) = \frac{c}{2}\sum_{j=1}^n (t_j^2 - t_{j-1}^2) \\ &= \frac{c}{2}\,(b^2 - a^2) = \frac{c}{2}\,(a + b)\,(b - a) \end{aligned}$$

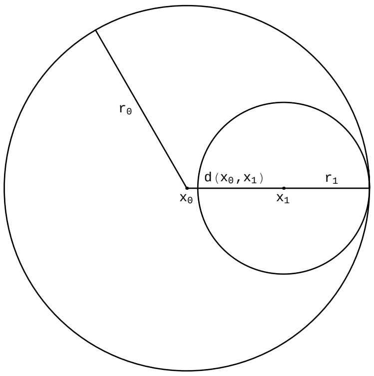


Dies ist $\int_a^b f(x)$, denn $U(f,Z)-O(f,Z)<\varepsilon$, wenn Z äquidistant mit hinreichend kleiner Schrittweite wie im Beweis von (5.1.2) gewählt wird.

Wir haben in (5.1.2) einerseits gesehen, daß alle stetigen Funktionen D-integrierbar sind und andererseits, daß auch unstetige Funktionen (nämlich alle monotonen) integrierbar sind. Wir wollen nun in (5.1.4) eine genaue Beschreibung der D-integrierbaren Funktionen geben, brauchen dazu allerdings einige Vorbereitungen kompakte Mengen betreffend.

5.1.4c Definition. Offene Teilmengen.
Eine Teilmenge W eines metrischen Raums X heißt OFFEN, wenn für jedes $x \in W$ ein Ball $U_\delta(x)$ mit $\delta > 0$ existiert mit $U_\delta(x) \subseteq W$.

Z.B. ist jeder offene Ball $U_r(x)$ offen, denn sei $x' \in U_r(x)$, d.h. $d(x',x) < r$ und $\delta := r - d(x,x') > 0$, dann ist $U_\delta(x') \subseteq U_r(x)$ nach Aufgabe (2.7), denn aus $x'' \in U_\delta(x')$ folgt $d(x'',x) \leq d(x'',x') + d(x',x) < \delta + d(x,x') = r$, d.h. $x'' \in U_r(x)$.



Es ist leicht einzusehen, daß eine Teilmenge $W \subseteq X$ genau dann offen ist, wenn das Komplement $X \setminus W$ abgeschlossen ist.

Eine Menge $\mathcal{W}$ von Mengen heißt ÜBERDECKUNG von X, wenn $X \subseteq \bigcup \mathcal{W}$ ist, d.h. jedes $x \in X$ in einer Menge $W \in \mathcal{W}$ enthalten ist.

Dabei ist die VEREINIGUNG $\bigcup \mathcal{W}$ einer Menge $\mathcal{W}$ von Mengen definiert durch (siehe (1.1.4)):

$$\bigcup \mathcal{W} := \Big\{x : \exists W \in \mathcal{W} : x \in W\Big\}.$$

Dies ist eine allgemeinere Schreibweise als

$$\bigcup_{i \in I} U_i = \{x : \exists i \in I : x \in U_i\} = \bigcup \{U_i : i \in I\}.$$

Vergleiche dies auch mit der analogen Schreibweise:

$$\sup f(I) = \sup\{f(t) : t \in I\} = \sup_{t \in I} f(t).$$

5.1.4a Überdeckungssatz von Lebesgue.
Es sei X ein kompakter metrischer Raum und $\mathcal{W}$ eine Überdeckung von X mit offenen Mengen. Dann existiert ein $\delta > 0$ (eine sogenannte Lebesgue-Zahl *der Überdeckung) s.d. jeder offene Ball mit Radius δ in einem $W \in \mathcal{W}$ enthalten ist.*

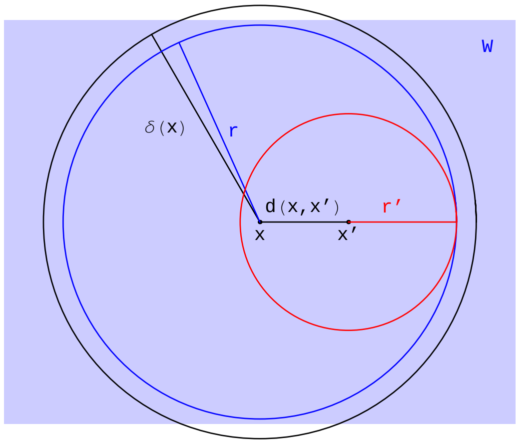
Man sagt in dieser Situation, daß $\mathcal{U} := \{U_\delta(x) : x \in X\}$ eine Verfeinerung der Überdeckung $\mathcal{W}$ ist, d.h. $\forall U \in \mathcal{U}\ \ \exists W \in \mathcal{W}$: $U \subseteq W$.

Beachte, daß dies zur Folge hat, daß jede Teilmenge $A \subseteq X$ mit Durchmesser $d(A) < \delta$ ganz in einem $W \in \mathcal{W}$ enthalten ist. Denn sei o.B.d.A. $x_0 \in A$, dann ist $A \subseteq U_\delta(x_0)$ und somit existiert ein $W \in \mathcal{W}$ mit $W \supseteq U_\delta(x_0) \supseteq A$.

Ein (notwendigerweise hinkender) Vergleich: Der Waldboden sei mit verschiedenen (offenen) Blättern bedeckt. Ist eine Erdbeere kleiner als die Lebesgue-Zahl der Überdeckung, so können wir sicher sein, daß wo immer sie am Boden liegt sie von mindestens einem Blatt vollständig bedeckt ist.

EXAM **Beweis.** Sei also $\mathcal{W}$ eine Überdeckung von X mit offenen Teilmengen. Wegen $X = \bigcup \mathcal{W}$ können wir für jedes $x \in X$ ein $W \in \mathcal{W}$ wählen mit $x \in W$. Da $W \in \mathcal{W}$ offen ist existiert ein $\delta > 0$ mit $U_\delta(x) \subseteq W$. Es sei $\delta(x) := \sup\{0 < \delta \leq 1 :\ \exists W \in \mathcal{W} : U_\delta(x) \subseteq W\}$. Beachte, daß dann für jedes $\delta < \delta(x)$ der Ball $U_\delta(x)$ in einer der Mengen $W \in \mathcal{W}$ enthalten ist. Wir wollen nun ein gemeinsam für alle Punkte verwendbares $\delta_0 > 0$ finden. Das sollte uns an den Beweis von (3.3.5) erinnern und wir versuchen folglich die Stetigkeit der Funktion $\delta : X \to \mathbb{R}$ zu zeigen.

Beh. : $|\delta(x) - \delta(x')| \leq d(x, x')$ für alle $x, x' \in X$:
Aus Symmetriegründen genügt es $\delta(x') \geq \delta(x) - d(x, x')$ zu zeigen. Sei dazu $r' < \delta(x) - d(x, x') \leq 1$ beliebig und somit $r := r' + d(x, x') < \delta(x)$. Also existiert ein $W \in \mathcal{W}$ mit $U_r(x) \subseteq W$ und nach Aufgabe (2.7) ist $U_{r'}(x') \subseteq U_r(x) \subseteq W$ (denn aus $x_1 \in U_{r'}(x')$, d.h. $d(x_1, x') < r'$ folgt $d(x_1, x) \leq d(x_1, x') + d(x', x) < r' + d(x, x') = r$). Damit ist aber $\delta(x') \geq r'$ für alle $r' < \delta(x) - d(x, x')$ und somit $\delta(x') \geq \delta(x) - d(x, x')$.



Wir haben damit gezeigt, daß die Funktion $\delta : X \to (0,1] \subseteq \mathbb{R}$ stetig ist und sie besitzt somit ein Minimum $\delta_0 > 0$ nach (3.3.6). Also ist jeder Ball mit Radius kleiner als $\delta_0 \leq \delta(x)$ um irgendeinen Punkt $x \in X$ ganz in einer Menge $W \in \mathcal{W}$ enthalten. □

Mittels Überdeckungssatz von Lebesgue können wir unter anderen einen einsichtigeren Beweis von (3.3.5) geben:

Folgerung.
Es sei $f : X \to Y$ stetig und X kompakt. Dann ist f gleichmäßig stetig.

Beweis. Da f stetig ist, existiert für jedes $x \in X$ ein (offener) Ball U_x um x, s.d. $d(f(x'), f(x)) < \varepsilon$ für alle $x' \in U_x$ gilt. Also ist $\mathcal{W} := \{U_x : x \in X\}$ eine offene Überdeckung von X. Nach (5.1.4a) existiert somit eine Lebesgue-Zahl $\delta > 0$ für diese Überdeckung. Sei nun $x_1, x_2 \in X$ mit $d(x_1, x_2) < \delta$. Dann hat $A := \{x_1, x_2\}$ Durchmesser $d(x_1, x_2) < \delta$ und somit existiert eine Menge $U_x \in \mathcal{W}$ mit $A \subseteq U_x$ und damit ist $d(f(x_i), f(x)) < \varepsilon$ für $i \in \{1, 2\}$, also $d(f(x_1), f(x_2)) < 2\varepsilon$. Dies zeigt die gleichmäßige Stetigkeit von f. □

5.1.4b Folgerung. Überdeckungssatz von Heine-Borel.
Ein metrischer Raum X ist genau dann kompakt, wenn jede seiner Überdeckungen mit offenen Mengen eine endliche Teilüberdeckung besitzt, d.h. aus $X = \bigcup \mathcal{W}$ und $\forall W \in \mathcal{W}$: $W \subseteq X$ ist offen, die Existenz einer endlichen Menge $\mathcal{W}_0 \subseteq \mathcal{W}$ mit $X = \bigcup \mathcal{W}_0$ folgt.

Dieser Satz zeigt, daß, obwohl kompakte Räume sehr viele (z.B. überabzählbar viele) Punkte besitzen können, man in vielen Situationen sich dennoch auf endlich viele Teilmengen beschränken kann. Er wird in der Topologie in allgemeineren Räumen als es die metrischen sind zur Definition von Kompaktheit erhoben – 'Old theorems never die, they turn into definitions!'

Beweis. ($\Rightarrow$) Es sei $\mathcal{W}$ eine Überdeckung von X mit offenen Mengen. Nach (5.1.4a) existiert eine Lebesgue-Zahl δ, d.h. $\{U_\delta(x) : x \in X\}$ ist eine Verfeinerung von $\mathcal{W}$.

Wir wollen nun zeigen, daß schon endlich viele dieser Mengen (und damit von $W \in \mathcal{W}$) ausreichen um X zu überdecken. Andernfalls wählen wir rekursiv eine Folgen $(x_n) \in X$ mit $x_n \notin \bigcup_{i=0}^{n-1} U_\delta(x_i)$ indem wir bei einem beliebigen $x_0 \in X$ starten. Da X kompakt ist, müßte diese Folge einen Häufungswert besitzen und damit eine Teilfolge die Cauchy ist. Wegen $d(x_n, x_i) \geq \delta$ für alle $i < n$ ist dies aber nicht möglich.

($\Leftarrow$) Sei (x_n) eine Folge in X. Angenommen kein $x \in X$ ist Häufungswert der Folge, d.h. es existiert ein offener Ball U_x um x der nur endlich viele Folgeglieder enthält. Andererseits ist $\{U_x : x \in X\}$ eine Überdeckung von X mit offenen Mengen, also existiert eine endlich Teilüberdeckung $\{U_1, \ldots, U_N\}$ und damit enthält $X = \bigcup_{i=1}^{N} U_i$ nur für endlich viele Folgeglieder, ein Widerspruch. □

5.1.4 Folgerung. Lebesgue'sches Integrabilitätskriterium.
Eine beschränkte Funktion $f : [a, b] \to \mathbb{R}$ ist genau dann D-integrierbar, wenn sie FAST ÜBERALL *stetig ist, d.h. die Menge der Punkte x in denen sie unstetig ist eine Lebesgue-Nullmenge ist.*

Dabei heißt eine Teilmenge $M \subseteq \mathbb{R}$ Lebesgue-Nullmenge, wenn zu jedem $\varepsilon > 0$ abzählbar viele Intervalle $(I_i)_{i\in\mathbb{N}}$ existieren, s.d. $M \subseteq \bigcup_{i\in\mathbb{N}} I_i$ und $\sum_{i=0}^\infty |I_i| < \varepsilon$ ist, d.h. M kann durch abzählbar viele Intervalle der Gesamtlänge kleiner als ε überdeckt werden.

Z.B. ist jede abzählbare Menge eine Nullmenge, denn sei $M = \{t_i : i \in \mathbb{N}\}$ und $I_i := [t_i - \frac{\varepsilon}{2^i}, t_i + \frac{\varepsilon}{2^i}]$. Dann ist $M \subseteq \bigcup_{i\in\mathbb{N}} I_i$ und $\sum_{i\in\mathbb{N}} |I_i| = \sum_{i=0}^\infty \frac{\varepsilon}{2^{i-1}} = 4\varepsilon$.

Der Begriff Nullmenge dient dazu Kleinheit (im maßtheoretischen Sinn) von (komplizierten) Mengen zu beschreiben.

Beweis. Die Idee des Beweises besteht darin, daß eine Funktion genau dann D-integrierbar ist, wenn $O(f,Z) - U(f,Z) = \sum_{I\in Z} \Omega(f|_I)\,|I|$ klein wird, und somit für jene Intervalle I für die $\Omega(f|_I)$ groß ist, die Länge $|I|$ klein ist. Dies kann auch so ausgedrückt werden, daß die Menge $\{x : 0 < \omega_f(x) = \inf\{\Omega(f|_{U_\delta(x)}) : \delta > 0\}\}$ 'Länge 0' besitzt.

($\Leftarrow$) Es sei $f : [a,b] \to \mathbb{R}$ beschränkt, also $\|f\|_\infty := \sup\{|f(x)| : x \in [a,b]\} < \infty$, und die Menge $\Delta(f)$ der Unstetigkeitspunkte von f eine Nullmenge. Sei $\varepsilon > 0$, dann existieren (o.B.d.A. offene!) Intervalle I_k mit $\Delta(f) \subseteq \bigcup_k I_k$ und $\sum_k |I_k| < \varepsilon$. Für jedes $x \notin \Delta(f)$ ist nach (3.2.9) $0 = \omega_x(f) = \inf\{\Omega(f|_{U_\delta(x)}) : \delta > 0\}$, also existiert ein $\delta_x > 0$, s.d. $\Omega(f|_{U_{\delta_x}(x)}) < \varepsilon$. Wegen der Kompaktheit von $[a,b]$ gibt es nach (5.1.4a) eine Lebesgue-Zahl $\delta = \frac{b-a}{N}$ zur Überdeckung $\{I_k : k \in \mathbb{N}\} \cup \{U_{\delta_x}(x) : x \notin \Delta(f)\}$, d.h. $t_i := a + i\,\frac{b-a}{N}$ mit $i \in \{0,\dots,N\}$ definiert eine Zerlegung Z von $[a,b]$, die dieser Überdeckung untergeordnet ist. Dann ist

$$O(f,Z) - U(f,Z) = \sum_{J\in Z} (\sup f|_J - \inf f|_J) \cdot |J|.$$

Jene Summanden, wo J in einem der I_k enthalten ist, tragen insgesamt höchstens $2\|f\|_\infty \sum_k |I_k| < 2\|f\|_\infty\,\varepsilon$ bei. Die übrigen Summanden tragen insgesamt höchstens $(b-a)\varepsilon$ bei. In Summe ist also $O(f,Z) - U(f,Z) < \Big(2\|f\|_\infty + (b-a)\Big)\,\varepsilon$, d.h. f ist D-integrierbar.

($\Rightarrow$) Nach (3.2.9) ist $\Delta(f) = \{x : \omega_x(f) > 0\} = \bigcup_n \Delta_{1/n}(f)$, wobei $\Delta_r(f) := \{x : \omega_f(x) = \inf\{\Omega(f|_{U_\delta(x)}) : \delta > 0\} \geq r\}$. Es genügt also zu zeigen, daß $\Delta_r(f)$ eine Nullmenge ist für jedes $r > 0$. Sei dazu $\varepsilon > 0$ und Z eine Zerlegung mit $O(f,Z) - U(f,Z) < r\,\frac{\varepsilon}{2}$. Sei Z_0 die Menge der Teilintervalle $I = [t_{i-1}, t_i]$ von Z für welche $(t_{i-1}, t_i) \cap \Delta_r(f) \neq \emptyset$. Für solche I sei x ein Punkt dieses Durchschnitts, dann ist $\Omega(f|_{U_\delta(x)}) \geq \omega_f(x) \geq r$ für alle $\delta > 0$ und o.B.d.A. $U_\delta(x) \subseteq I$ für ein $\delta > 0$, also $\Omega(f|_I) \geq \Omega(f|_{U_\delta(x)}) \geq r$. Dann ist

$$r \sum_{I\in Z_0} |I| \leq \sum_{I\in Z_0} \Omega(f|_I)\,|I| \leq \sum_{I\in Z} \Omega(f|_I)\,|I| = O(f,Z) - U(f,Z) < r\,\frac{\varepsilon}{2},$$

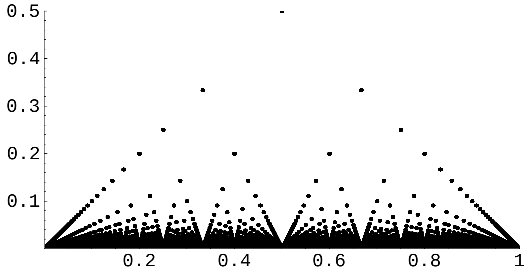
also $\sum_{I\in Z_0} |I| < \frac{\varepsilon}{2}$. Folglich ist $\Delta_r \setminus \{t_0,\dots,t_N\} \subseteq \bigcup_{I\in Z_0} I$ eine Nullmenge und damit auch $\Delta_r \subseteq (\Delta_r \setminus \{t_0,\dots,t_N\}) \cup \{t_0,\dots,t_N\}$ eine. □

5.1.6 Beispiele (nicht) integrierbarer Funktionen.
Die Dirichlet'sche Sprungfunktion $\chi_\mathbb{Q}$ ist überall unstetig, also auf $[0,1]$ nicht integrierbar.

Hingegen ist die Funktion $f : \mathbb{R} \to \mathbb{R}$ welche gegeben ist durch

$$f : x \mapsto \begin{cases} \frac{1}{q} & \text{für } x = \frac{p}{q} \in \mathbb{Q} \text{ mit teilerfreien } p \text{ und } q \\ 0 & \text{andernfalls,} \end{cases}$$

D-integrierbar über $[0,1]$, denn sie ist nur in den rationalen Punkten unstetig. Es ist leicht einzusehen, daß $\int_0^1 f(x) = 0$ ist.



5.1.8 Definition. Äußere Maß.
Für Teilmengen $X \subseteq \mathbb{R}$ definieren wir das sogenannte ÄUSSERE MASS von X als

$$|X| := \inf\Big\{\sum_{k=0}^{\infty} |I_k| : X \subseteq \bigcup_{k=0}^{\infty} I_k\Big\}.$$

Man schreibt für dieses Konzept der Maßtheorie (siehe z.B. [11, 2.4]) üblicherweise dafür auch $\mu^*(X)$. Beachte, daß X genau dann eine Lebesgue'sche Nullmenge ist, wenn ihr äußeres Maß $|X|$ gleich 0 ist. Man sieht leicht ein, daß $|\bigcup_{i=0}^{\infty} X_i| \leq \sum_{i=0}^{\infty} |X_i|$, $\forall \varepsilon > 0 \;\; \exists U$ offen mit $X \subseteq U$ und $|U| \leq |X| + \varepsilon$. Aber Achtung, selbst für disjunktes A und B kann $|A \cup B| < |A| + |B|$ sein.

Lemma. Existenz von Nicht-Nullmengen.
Für Intervalle $I := [a, b]$ ist das äußere Maß $|I|$ gleich der Länge $b - a$. Also ist kein Intervall $I := [a, b]$ mit $a < b$ eine 0-Menge.

Wir wollen also zeigen, daß die zwei Bedeutungen (als Länge und als äußeres Maß) die wir $|I|$ gegeben haben zum gleichen Ergebnis führen. Im Beweis verwenden wir die Notation $|I|$ nur im Sinne der Länge.

Beweis. Aus $I \subseteq \bigcup_{n\in\mathbb{N}} I_n$ folgt $|I| \leq \sum_{n=0}^{\infty} |I_n|$ und somit ist $|I|$ kleiner oder gleich dem äußeren Maß von I: O.B.d.A. seien die I_n offen. Dann existiert nach (5.1.4a) eine Lebesguezahl $\delta > 0$ und somit ist jedes Teilintervall einer äquidistanten Zerlegung Z mit Schrittweite $|Z| < \delta$ in einem der Intervalle I_n enthalten. Sei $Z_n := \{J \in Z : n = \min\{k : J \subseteq I_k\}\}$, also die Menge jener Interalle von Z die in I_n aber keinen I_k mit $k < n$ enthalten sind, dann ist $|I_n| \geq \sum_{J\in Z_n} |J|$, denn $I_n \supseteq \bigcup Z_n$ und die endlich vielen Intervalle aus Z_n sind nicht überlappend. Da Z offensichtlich die disjunkte Vereinigung der Z_n ist, ist $|I| = \sum_{J\in Z} |J| = \sum_{n=0}^{\infty} \sum_{J\in Z_n} |J| \leq \sum_{k=0}^{\infty} |I_k|$.

Wenn wir $I_0 := I$ und $I_n := \emptyset$ für $n > 0$ wählen, dann folgt, daß das äußere Maß von I kleiner oder gleich $\sum_n |I_n| = |I|$ ist. Also gilt Gleichheit. $\square$

5.1.4e Das Cantor'sche Diskontinuum.
Folgende Menge ist eine überabzählbare Nullmenge. Wir beginnen mit dem Intervall $I_0 := [0, 1]$ der Länge 1. Im ersten Schritt entfernen aus diesem das mittlere offene Drittel $(1/3, 2/3)$ und erhalten die 2 abgeschlossenen Intervalle $[0, 1/3]$ und $[2/3, 1]$, deren Vereinigung wir mit I_1 bezeichnen. Diese Vereinigung hat Gesamtlänge $2/3$. Im nächsten Schritt entfernen wir aus jedem dieser beiden Intervalle, die jeweiligen mittleren offenen Drittel $(1/9, 2/9)$ und $(7/9, 8/9)$ und erhalten 4 abgeschlossene Intervalle deren Vereinigung der Gesamtlänge $(2/3)^2$ wir mit I_2 bezeichnen. So fortfahren erhalten wir im k-ten Schritt eine Vereinigung I_k der Gesamtlänge $(2/3)^k$ von 2^k vielen abgeschlossenen Teilintervallen. Der Durchschnitt $I_\infty := \bigcap_{k\in\mathbb{N}} I_k$ ist dann eine abgeschlossene Teilmenge von $[0, 1]$, das sogenannte CANTOR'SCHE DISKONTINUUM. Dieses ist offensichtlich eine Nullmenge, denn $I_\infty \subseteq I_k$ und die Gesamtlänge dieser Vereinigung I_k endlich vieler Intervalle ist $(2/3)^k$, und $(2/3)^k \to 0$ für $k \to \infty$.

Es ist I_∞ kompakt, denn als Durchschnitt abgeschlossener Mengen I_k ist es abgeschlossen, und als Teilmenge von $[0, 1]$ beschränkt.

Es ist I_∞ überabzählbar. Um das einzusehen entwickeln wir alle $x \in (0, 1]$ im 3-er System in unendliche Ternärzahlen, d.h. als $x = \sum_{k=1}^{\infty} \frac{x_k}{3^k}$ mit $x_k \in \{0, 1, 2\}$ und nicht $x_k = 0$ für fast alle k. Das erste Intervall (und sein rechter Randpunkt), das wir entfernt haben besteht nur aus Zahlen mit $x_1 = 1$, die nächsten beiden (und deren jeweiliger rechte Endpunkt) nur aus Zahlen mit $x_1 \in \{0, 2\}$ und $x_2 = 1$, etc.. Wir haben also bei der Konstruktion von I_∞ nur Zahlen entfernt, bei deren Entwicklung im 3-er System mindestens ein 1-er auftritt. Alle Ternärzahlen, die nur 0'er und 2'er als Ziffern haben liegen also in I_∞. Davon gibt es mit gleichen Argument wie bei $\mathbb{R}$ (siehe (1.1.20)) aber überabzählbar viele.

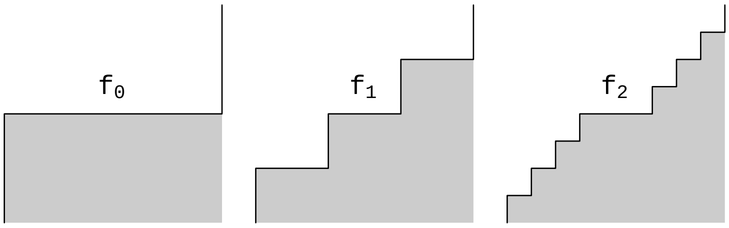
Umgekehrt, ist jede Ternärzahl die eine (möglicherweise abbrechende) Entwicklung ohne 1'er zuläßt in I_∞, denn die rechten Endpunkte der entfernten Intervalle haben zwar nicht-abbrechende Entwicklungen, die genau einen 1'er enthalten welcher von lauter 2'ern gefolgt wird, also auch eine abbrechende Entwicklung der Form, wo dieser 1'er durch 2 ersetzt ist und die nacholgenden 2'er durch 0.

Wenn wir nun die charakteristische Funktion $f := \chi_{[0,1]\setminus I_\infty}$ des Komplements von I_∞ betrachten, so ist diese auf der offenen Menge $[0,1] \setminus I_\infty$ konstant und somit stetig. Folglich ist die Menge der Unstetigkeitspunkte in I_∞ enthalten (ja sogar gleich, denn in der Nähe jedes Punktes aus I_∞ liegen auch Punkte in $[0,1] \setminus I_\infty$) und somit ist f fast überall stetig, also nach (5.1.4) D-integrierbar. Da in jedem echten Teilintervall von $[0,1]$ Punkte aus $[0,1] \setminus I_\infty$ liegen, ist jede Obersumme 1, also $\int_0^1 f = 1$.

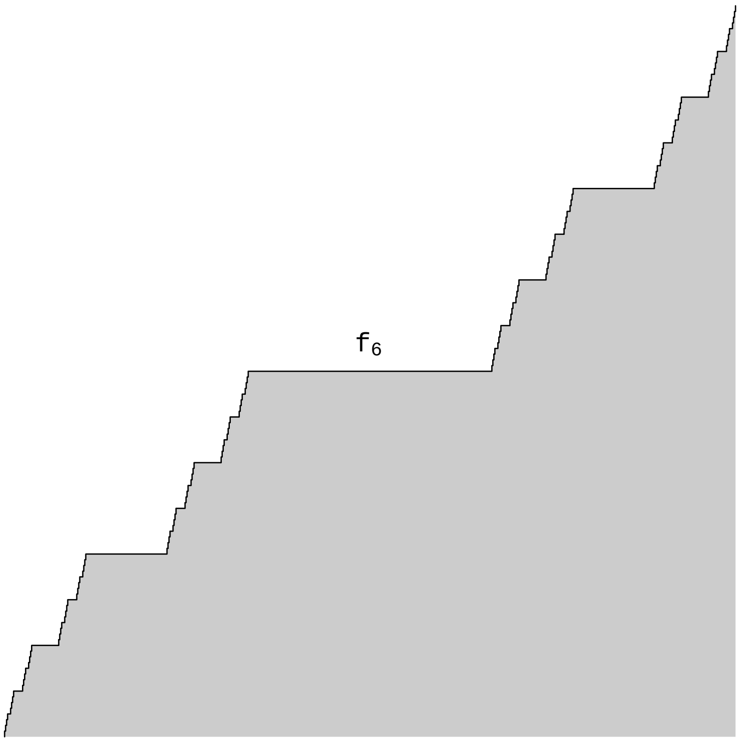
Wir wollen noch eine weitere pathologische Funktion $f : [0,1] \to [0,1]$ beschreiben, die von Vitali stammt und auch Teufelstreppe genannt wird. Dazu bezeichnen wir mit $f_n : [0,1] \to [0,1]$ jene Funktion, die x auf die Anzahl der Randpunkte von I_n die links von x liegen geteilt durch 2^{n+1} abbildet, d.h.

$$f_n(x) := \sum_{x \geq t \in \partial I_n} \frac{1}{2^{n+1}},$$

wobei wir mit ∂I_n die Menge der Randpunkte von I_n bezeichnen.

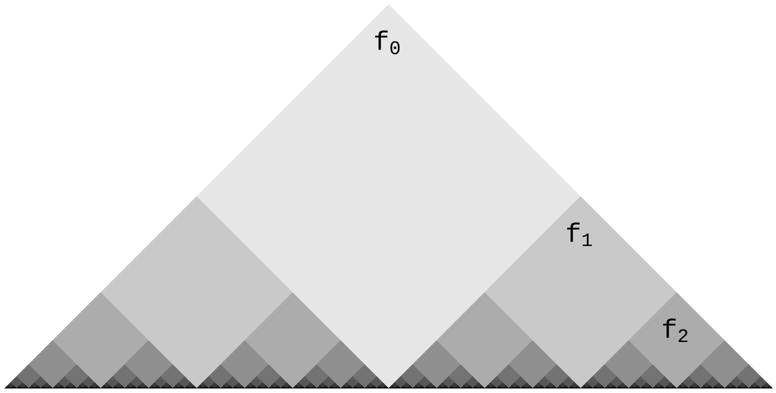


Es ist f_n monoton wachsend und auf den Teilintervallen von $[0,1] \setminus I_n$ konstant. Weiters konvergiert die Folge $(f_n)_n$ gleichmäßig, denn f_{n+1} unterscheidet sich von f_n (auf den Teilintervallen von I_n) um höchstens $1/2^{n+1}$. Somit ist die Grenzfunktion f_∞ nach (4.2.8) stetig und ebenfalls monoton wachsend. Da f_∞ auf jedem Teilintervall von $[0,1] \setminus I_n$ gleich f_n und somit konstant ist, ist $[0,1] = f_\infty([0,1]) = f_\infty(I_\infty)$, also das Bild der 0-Menge I_∞ weit entfernt davon eine 0-Menge zu sein. Dies zeigt gleichzeitig (nochmals), daß I_∞ gleichmächtig zu $[0,1]$, also überabzählbar ist. Beachte auch, daß f auf $[0,1] \setminus I_\infty$ differenzierbar mit Ableitung 0 ist, d.h. auf fast allen Punkten $x \in [0,1]$ existiert die Ableitung $f'(x)$ und ist 0.



5.1.4f Beispiel. Van der Waerden's Funktion.
Wir geben hier ein Beispiel einer stetigen Funktion die nirgends differenzierbar ist.

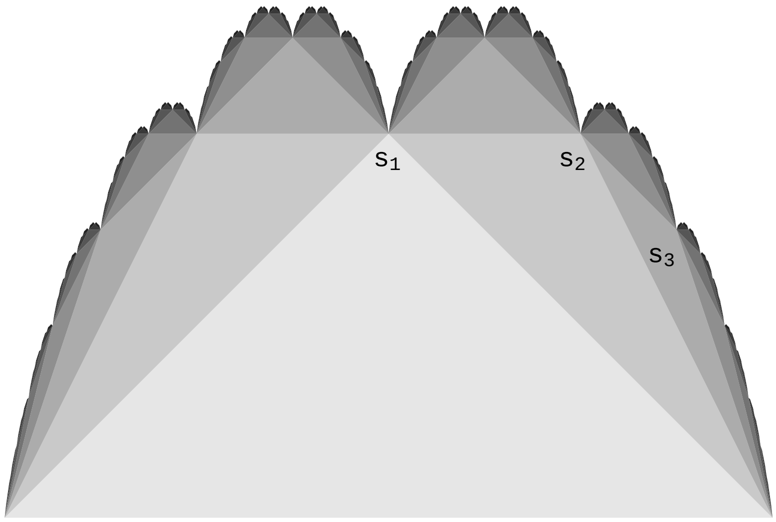
Es sei f_0 die Sägezahnkurve $f_0(x) := d(x, \mathbb{Z}) = \min\{|x-k| : k \in \mathbb{Z}\}$ und $f_n(x) := f_1(2^n x)/2^n$. Wegen $\|f_0\|_\infty = \frac{1}{2}$ ist $\|f_n\|_\infty = \frac{1}{2^{n+1}}$ und somit konvergiert die Reihe $\sum_k f_k$ gleichmäßig auf ganz $\mathbb{R}$.



Es sei $s_n := \sum_{k=0}^{n-1} f_k$ und $s_\infty := \sum_{k=0}^{\infty} f_k$ ihre Summe. Dann ist s_∞ (gleichmäßig) stetig nach (4.2.8). Es ist $f_0(k) = 0$ und $f_0(k+\frac{1}{2}) = \frac{1}{2}$ für alle $k \in \mathbb{Z}$. Somit ist $f_n(\frac{k}{2^n}) = 0$ und $f_n(\frac{2k+1}{2^{n+1}}) = \frac{1}{2^{n+1}}$ für alle $k \in \mathbb{Z}$. Somit ist $s_{n+1}\left(\frac{k}{2^n}\right) = s_n\left(\frac{k}{2^n}\right) + f_n\left(\frac{k}{2^n}\right) = s_n\left(\frac{k}{2^n}\right)$ und

$$s_{n+1}\left(\frac{2k+1}{2^{n+1}}\right) = \frac{1}{2}\left(s_n\left(\frac{k}{2^n}\right) + s_n\left(\frac{k+1}{2^n}\right)\right) + \frac{1}{2^{n+1}}.$$

Also ist $s_\infty(x) = s_n(x)$ für $x = \frac{k}{2^n}$ mit $k \in \mathbb{Z}$.



Es sei $x \in \mathbb{R}$. Wir wollen nun zeigen, daß die Ableitung von s_∞ bei x nicht existiert. Für jedes $n \in \mathbb{N}$ existiert ein eindeutig bestimmtes $k \in \mathbb{Z}$ mit $\alpha_n := \frac{k}{2^{n+1}} \le x < \frac{k+1}{2^{n+1}} =: \beta_n$. Für $x < \frac{\alpha_n+\beta_n}{2}$ ist $\alpha_{n+1} = \alpha_n$ und $\beta_{n+1} = \frac{\alpha_n+\beta_n}{2} = \alpha_n + \frac{1}{2^{n+2}}$. Für $x \ge \frac{\alpha_n+\beta_n}{2}$ ist $\beta_{n+1} = \beta_n$ und $\alpha_{n+1} = \frac{\alpha_n+\beta_n}{2} = \alpha_n + \frac{1}{2^{n+2}}$. Offensichtlich gilt $\alpha_n \nearrow x$ und $\beta_n \searrow x$. Wäre also s_∞ differenzierbar bei x, dann müßte

$$s'_\infty(x) = \lim_{n\to\infty} \frac{s_\infty(\beta_n) - s_\infty(\alpha_n)}{\beta_n - \alpha_n}$$

gelten. Im ersten Fall ist

$$\begin{aligned}
\frac{s_\infty(\beta_{n+1}) - s_\infty(\alpha_{n+1})}{\beta_{n+1} - \alpha_{n+1}} &= 2^{n+2}\big(s_\infty(\beta_{n+1}) - s_\infty(\alpha_n)\big) \\
&= 2^{n+2}\big(s_{n+1}(\beta_{n+1}) - s_n(\alpha_n)\big) \\
&= 2^{n+2}\left(\frac{1}{2}\Big(s_n(\alpha_n) + s_n(\beta_n)\Big) + \frac{1}{2^{n+2}} - s_n(\alpha_n)\right) \\
&= 2^{n+1}\big(s_n(\beta_n) - s_n(\alpha_n)\big) + 1 \\
&= \frac{s_n(\beta_n) - s_n(\alpha_n)}{\beta_n - \alpha_n} + 1
\end{aligned}$$

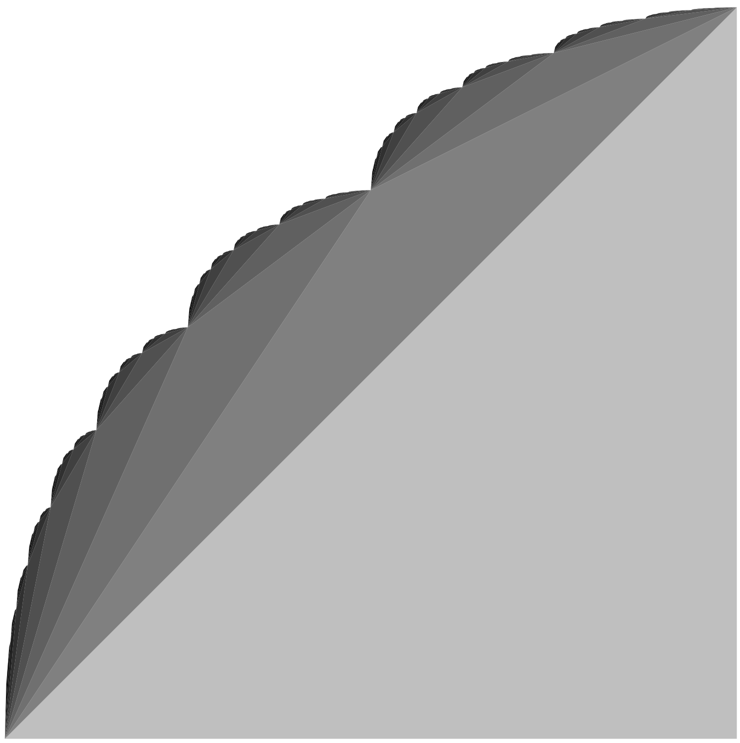
und im zweiten ist

$$\begin{aligned}
\frac{s_\infty(\beta_{n+1}) - s_\infty(\alpha_{n+1})}{\beta_{n+1} - \alpha_{n+1}} &= 2^{n+2}\big(s_\infty(\beta_n) - s_\infty(\alpha_{n+1})\big) \\
&= 2^{n+2}\big(s_n(\beta_n) - s_{n+1}(\alpha_{n+1})\big) \\
&= 2^{n+2}\left(s_n(\beta_n) - \frac{1}{2}\Big(s_n(\alpha_n) + s_n(\beta_n)\Big) - \frac{1}{2^{n+2}}\right) \\
&= 2^{n+1}\big(s_n(\beta_n) - s_n(\alpha_n)\big) - 1 \\
&= \frac{s_n(\beta_n) - s_n(\alpha_n)}{\beta_n - \alpha_n} - 1
\end{aligned}$$

Folglich kann obiger Grenzwert nicht existieren und s_∞ ist nirgends differenzierbar.

Beispiel.
Es folgt nun ein Beispiel einer stetigen, streng monoton wachsenden (und damit fast überall differenzierbaren) Funktion $[0,1] \to [0,1]$ *deren Ableitung über wo sie existiert gleich 0 ist.*

Dazu definieren wir für $0 < q < 1$ rekursiv eine monoton wachsende Folge strikt monotoner stückweise affiner stetiger Funktionen $h_n : [0,1] \to [0,1]$ durch $h_0(x) := x$, $h_{n+1}(\frac{k}{2^n}) = h_n(\frac{k}{2^n})$, $h_{n+1}(\frac{2k+1}{2^{n+1}}) = \frac{1-q}{2} h_n(\frac{k}{2^n}) + \frac{1+q}{2} h_n(\frac{k+1}{2^n})$ und dazwischen affin. Es sei $h(x) := \lim_{n\to\infty} h_n(x)$. Dann ist $h : [0,1] \to [0,1]$ monoton wachsend, $h(\frac{k}{2^n}) = h_n(\frac{k}{2^n})$.



Die Funktion h ist sogar streng monoton wachsend, denn zu $x < y$ existieren k und n mit $x < \frac{k}{2^n} < y$ und somit ist $h(x) \le h(\frac{k}{2^n}) = h_n(\frac{k}{2^n}) < h_n(y) \le h(y)$.

Die Funktion h ist stetig (und nach dem Satz von Dini – Aufgabe (5.6) – die Konvergenz gleichmäßig). Es sei $x \in [0,1]$. Wegen $h_n \le h$ ist h linksstetig. Für jedes n existiert ein k_n mit $\alpha_n := \frac{k_n}{2^n} \le x < \frac{k_n+1}{2^n} =: \beta_n$. Es genügt zu zeigen, daß $h(\beta_n) - h(\alpha_n) \to 0$ für $n \to \infty$. Im Falle $\alpha_{n+1} = \alpha_n$ (d.h. x liegt näher an α_n als an β_n) ist

$$\begin{aligned}h(\beta_{n+1})-h(\alpha_{n+1}) &= h_{n+1}(\beta_{n+1})-h_n(\alpha_n)\\ &= \frac{1-q}{2}h_n(\alpha_n)+\frac{1+q}{2}h_n(\beta_n)-h_n(\alpha_n)\\ &= \frac{1+q}{2}\Big(h_n(\beta_n)-h_n(\alpha_n)\Big)=\frac{1+q}{2}\Big(h(\beta_n)-h(\alpha_n)\Big).\end{aligned}$$

Analog ist $h(\beta_{n+1})-h(\alpha_{n+1})=\frac{1-q}{2}\Big(h(\beta_n)-h(\alpha_n)\Big)$ im Fall $\beta_{n+1}=\beta_n$. In jeden Fall erhalten wir rekursiv, daß

$$h(\beta_{n+1})-h(\alpha_{n+1})=\prod_{j=1}^{n}\frac{1+\varepsilon_j q}{2}\text{ mit }\varepsilon_j\in\{+1,-1\}.$$

Dieses Produkt konvergiert gegen 0 für $n\to\infty$, da die Faktoren vom Betrag kleiner als $\frac{1+q}{2}<1$ sind.

Da monotone Funktionen nach nachfolgenden Satz (5.1.16) von Lebesgue fast überall differenzierbar sind, existiert $h'(x)$ für fast alle x.

Wir zeigen nun, daß $h'(x)=0$ für alle x für die es existiert. Es ist

$$h'(x)=\lim_{n\to\infty}\frac{h(\beta_n)-h(\alpha_n)}{\beta_n-\alpha_n}=\lim_{n\to\infty}\frac{1}{\frac{1}{2^n}}\prod_{j=1}^{n}\frac{1+\varepsilon_j q}{2}=\lim_{n\to\infty}\prod_{j=1}^{n}(1+\varepsilon_j q).$$

Falls $h'(x)>0$ ist, so konvergiert $\sum_{j=1}^n\log(1+\varepsilon_j q)=\log(\prod_{j=1}^n(1+\varepsilon_j q))\to\log(h'(x))$ und somit $\log(1+\varepsilon_n q)\to 0$. Dies ist wegen $\varepsilon_n\in\{+1,-1\}$ ein Widerspruch.

Es kann die Menge der Punkte, in denen h nicht differenzierbar ist, nicht abzählbar sein, denn sonst wäre f nach einen allgemeineren Mittelwertsatz konstant, da $f'(x)=0$ für alle übrigen x ist.

Es sei nun $A:=\{x\in[0,1]:\ \exists\, h'(x)\}$. Dann ist nach obigen $[0,1]\setminus A$ eine Nullmenge. Wegen eines Satzes von Saks ist das äußere Maß von $|h(A)|\le\int_A|h'(x)|\,dx=0$, also $h(A)$ eine Nullmenge.

Da h streng monoton von $[0,1]$ auf $[0,1]$ ist, existiert die Umkehrfunktion $h^{-1}:[0,1]\to[0,1]$ und diese ist stetig und streng monoton, nach (3.4.3), also ebenfalls fast überall differenzierbar. Es ist $h^{-1}(h(A))=A$ eine Menge vom Maß 1. D.h. ein HOMÖOMORPHES BILD (d.h. Bild unter einer bijektiven stetigen Abbildung mit stetiger Inversen) einer 0-Menge muß keine 0-Menge mehr sein, kurz gesagt: Das stetige Bild von fast nichts kann fast alles sein. Für weitergehende Vergleiche zwischen topologischen Eigenschaften (wie Stetigkeit) und maßtheoretischen Eigenschaften von Funktionen siehe z.B. [19].

5.1.5 Folgerung. Operationen auf D-integrierbaren Funktionen.
Es sei f und g D-integrierbar. Dann sind auch $f+g$, $f\cdot g$, $\max(f,g):x\mapsto\max\{f(x),g(x)\}$ und $\min(f,g):x\mapsto\min\{f(x),g(x)\}$ integrierbar. Der Quotient f/g ist integrierbar, falls er wohldefiniert und beschränkt ist. Ist h stetig und beschränkt auf dem Bild von g, so ist auch $h\circ g$ integrierbar.

Es sei $g(x):=x$ für $x>0$ mit $g(0):=1$ und $f(x):=1/x$. Dann ist $f\circ g$ nicht beschränkt, aber f auf $g([0,1])$ stetig und g D-integrierbar.

Beachte, daß die Kurzschreibweise $\max(f,g)$ für $\max\circ(f,g)$ eine mißverstehbare ist, denn bei $\max(x,y)$ für $x,y\in\mathbb{R}$ war dies die größere der beiden Zahlen x und y, für Funktionen f und g (wo ja keine die größere zu sein braucht) ist $\max(f,g)$ nun die kleinste Funktion h, mit $h\ge f$ und $h\ge g$.

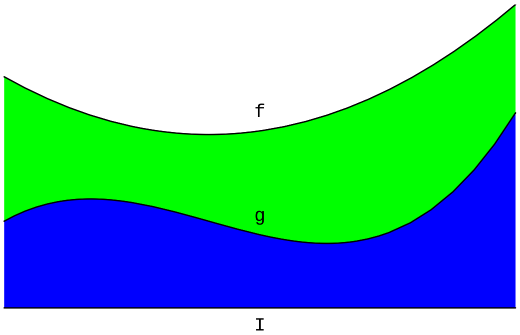
Beweis. Wegen (3.1.6) ist $\Delta(f+g)\subseteq\Delta(f)\cup\Delta(g)$, $\Delta(f\cdot g)\subseteq\Delta(f)\cup\Delta(g)$, $\Delta(\max(f,g))\subseteq\Delta(f)\cup\Delta(g)$, $\Delta(\min(f,g))\subseteq\Delta(f)\cup\Delta(g)$, $\Delta(f/g)\subseteq\Delta(f)\cup\Delta(g)$ (falls $g(x)\ne 0$ für alle x) und $\Delta(h\circ g)\subseteq\Delta(g)$ (falls $\Delta(h)=\emptyset$). Also sind nach (5.1.4) all diese Funktionen D-integrierbar sofern sie beschränkt sind. Letzteres ist für die erste 4 klar und für die übrigen vorausgesetzt. □

Beachte, daß diese Folgerung (ungleich der Kettenregel (4.1.14) und der Produktregel (4.1.15)) nur die Integrierbarkeit der Zusammensetzung $f\circ g$, der Summe $f+g$ und des Produkts $f\cdot g$, aber keine Formel für $\int_a^b f\circ g$ und $\int_a^b f\cdot g$ liefert. Für die Summe wollen wir ein solche nun dennoch zeigen.

5.1.7 Lemma. Integrieren ist linear am Raum $R(I,\mathbb{R})$ der integrierbaren Funktionen.
Die Menge der D-integrierbaren Funktionen $f : I \to \mathbb{R}$ auf einem Intervall I bildet einen Vektorraum $R(I,\mathbb{R})$ und Integrieren $\int_I : R(I,\mathbb{R}) \to \mathbb{R}$, $f \mapsto \int_I f$ ist linear, d.h. $\int_I (f + \lambda\, g) = \int_I f + \lambda \int_I g$ für alle $\lambda \in \mathbb{R}$ und alle D-integrierbare Funktionen $f, g : I \to \mathbb{R}$.

Insbesonders bedeutet dies, daß wir die (orientierte) Fläche A zwischen zwei Funktionen f und g über einem Intervall I wie folgt berechnen können:

$$A = \int_I f - \int_I g = \int_I (f - g).$$



Beweis. Nach (5.1.5) ist $R(I,\mathbb{R})$ ein linearer Teilraum des Vektorraums aller Abbildungen $f : I \to \mathbb{R}$ bezüglich der punktweise definierten Operationen $f + g : x \mapsto f(x) + g(x)$ und $\lambda \cdot f : x \mapsto \lambda \cdot f(x)$ (siehe lineare Algebra).

Nun zur Linearität von $\int_I : R(I,\mathbb{R}) \to \mathbb{R}$. Sei vorerst $\lambda \ge 0$. Dann ist $U(\lambda f, Z) = \sum_J \inf(\lambda\, f(J))\, |J| = \sum_J \lambda\, \inf(f(J))\, |J| = \lambda \sum_J \inf(f(J)) = \lambda U(f, Z)$ und somit $\int_I \lambda f = \sup_Z U(\lambda\, f, Z) = \sup_Z \lambda\, U(f, Z) = \lambda \sup_Z U(f, Z) = \lambda \int_I f$.

Für die Additivität verwenden wir

$$\sup f(J) + \sup g(J) \ge \sup(f + g)(J) \ge \inf(f + g)(J) \ge \inf f(J) + \inf g(J),$$

also ist

$$O(f, Z) + O(g, Z) \ge O(f + g, Z) \ge U(f + g, Z) \ge U(f, Z) + U(g, Z).$$

Sei nun $\varepsilon > 0$. Da f und g intergierbar sind existieren Zerlegungen Z_f und Z_g mit

$$\begin{aligned} \int_I f - \varepsilon \le U(f, Z_f) \le O(f, Z_f) \le \int_I f + \varepsilon \\ \int_I g - \varepsilon \le U(g, Z_g) \le O(g, Z_g) \le \int_I g + \varepsilon \end{aligned}$$

Für die gemeinsame Verfeinerung $Z = Z_f \cup Z_g$ von Z_f und Z_g anstelle von Z_f und Z_g gilt dies dann ebenfalls, und somit ist

$$\begin{aligned} \int_I f + \int_I g - 2\varepsilon \le U(f, Z) + U(g, Z) \le U(f + g, Z) \le \int_I (f + g) \\ \le O(f + g, Z) \le O(f, Z) + O(g, Z) \le \int_I f + \int_I g + 2\varepsilon \end{aligned}$$

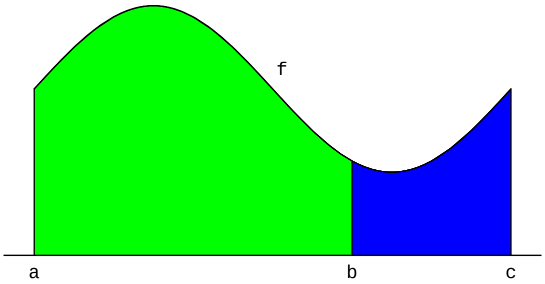
also gilt $\int_I (f + g) = \int_I f + \int_I g$.

Fehlt noch die Homogenität für $\lambda < 0$. Es genügt diese für $\lambda = -1$ zu zeigen, was aus $0 = \int_I 0 = \int_I (f + (-f)) = \int_I f + \int_I (-f)$ folgt. □

5.1.9 Lemma. Additivität des Integrals bzgl. der Grenzen.
Integrieren ist additiv in den Grenzen, d.h. $\int_a^c f = \int_a^b f + \int_b^c f$.

Diese Gleichung gilt vorerst nur für $a \leq b \leq c$. Damit sie auch für beliebige a, b, c gilt, muß für $a \leq b$ die Identität $0 = \int_a^a f = \int_a^b f + \int_b^a f$ erfüllt sein, d.h. wir definieren

$$\int_b^a f := -\int_a^b f \text{ für } b \geq a.$$



Beweis. Beachte, daß nach (5.1.4) aus der Integrierbarkeit von $f|_{[a,b]}$ und $f|_{[b,c]}$ jene von $f|_{[a,c]}$ folgt. Für beliebige Zerlegungen Z_- und Z_+ der Intervalle $[a, b]$ und $[b, c]$ ist $Z_- \cup Z_+$ eine Zerlegung von $[a, c]$ mit

$$U(f, Z_-) + U(f, Z_+) = U(f, Z_- \cup Z_+) \leq O(f, Z_- \cup Z_+) = O(f, Z_-) + O(f, Z_+).$$

Damit ist

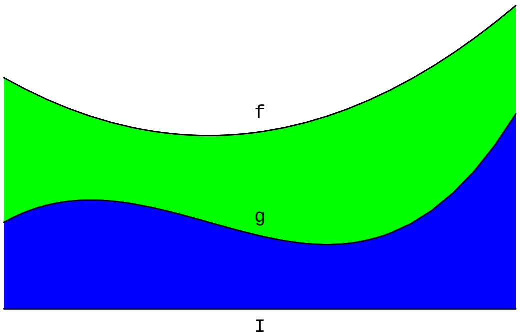
$$\begin{aligned} \int_a^b f + \int_b^c f &= \sup\{U(f, Z_-) : Z_-\} + \sup\{U(f, Z_+) : Z_+\} \\ &= \sup\{\underbrace{U(f, Z_-) + U(f, Z_+)}_{=U(f, Z_- \cup Z_+)} : Z_+, Z_-\} \leq \sup\{U(f, Z) : Z\} = \int_a^c f \end{aligned}$$

und ganz analog unter Verwendung der Obersummen folgt

$$\int_a^b f + \int_b^c f \geq \int_a^c f,$$

also gilt Gleichheit. □

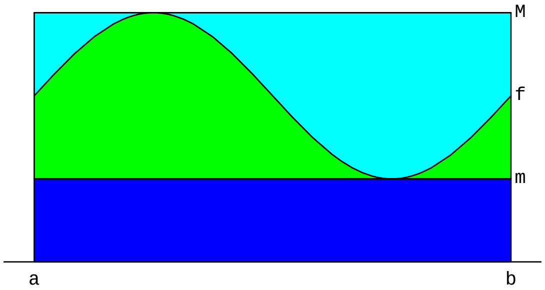
5.1.10 Lemma. Monotonie des Integrals.
Integrieren $\int_I : R(I, \mathbb{R}) \to \mathbb{R}$ ist monoton, d.h. aus $f \geq g$ (soll heißen $f(x) \geq g(x)$ für alle $x \in I$) folgt $\int_I f \geq \int_I g$.



Beachte dabei, daß für D-integrierbare f auch $|f| : x \mapsto |f(x)|$ D-integrierbar ist nach (5.1.5) und somit $|\int_I f| \leq \int_I |f|$ gilt.

Beweis. Wir betrachten $h := f - g \geq 0$. Dann ist $\int_I f - \int_I g = \int_I (f - g) = \int_I h \geq U(h, Z) \geq 0$. □

5.1.11 Mittelwertsatz der Integralrechnung.
Es sei $m \le f(x) \le M$ für alle $x \in [a,b]$ und f integrierbar. Dann ist $m\,(b-a) \le \int_a^b f \le M\,(b-a)$. Ist f sogar stetig, dann existiert ein $\xi \in [a,b]$ mit $\int_a^b f = f(\xi)\,(b-a)$.



Beweis. Dies folgt aus (5.1.10), da $\int_a^b m = (b-a)\,m$.

Für stetiges f liefert der Zwischenwertsatz die Existenz eines $\xi \in [a,b]$ mit $f(\xi) = \frac{1}{b-a}\int_a^b f$, da die rechte Seite zwischen $m := \inf(f)$ und $M := \sup(f)$ liegt. □

5.1.12 Proposition. Stetigkeit des Integrierens.
Es seien $f_n : [a,b] \to \mathbb{R}$ D-integrierbar und f_n konvergiere für $n \to \infty$ gegen f_∞ gleichmäßig. Dann ist auch f_∞ D-integrierbar und es gilt

$$\int_a^b \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_a^b f_n.$$

Beachte, daß dies einerseits besagt, daß $R(I,\mathbb{R})$ ein abgeschlossener Teilraum des Raumes $B(I,\mathbb{R})$ aller beschränkten Funktionen $I := [a,b] \to \mathbb{R}$ aus (4.2.8) ist (und somit nach (4.2.9) bzg. der Supremums-Metrik d_∞ vollständig ist) und andererseits, daß $\int_I : R(I,\mathbb{R}) \to \mathbb{R}$ stetig ist (nach (3.1.3)).

Beweis. Nach dem Lebesgue'schen Integrabilitätskriterium (5.1.4) ist f_∞ D-integrierbar, denn f_∞ ist nach (4.2.8) zumindestens dort stetig, wo es alle f_n sind, also ist $\Delta(f_\infty) \subseteq \bigcup_n \Delta(f_n)$ eine Lebesgue-0-Menge.

Wegen $|\int_a^b f - \int_a^b g| = |\int_a^b (f-g)| \le \int_a^b |f-g| \le |b-a|\,d_\infty(f,g)$ nach (5.1.10) und (5.1.11) ist das Integrieren $\int_I$ stetig (und sogar Lipschitz, siehe (5.5.9)), d.h. $\int_a^b f_n$ konvergiert gegen $\int_a^b f_\infty$ falls $d_\infty(f_n, f_\infty) \to 0$, also $f_n \to f_\infty$ gleichmäßig konvergiert. □

5.1.13 Bemerkung. Integration ist unstetig bzg. punktweiser Konvergenz.
Beachte, daß (5.1.12) für eine bloß punktweise konvergente Folge von Funktionen nicht mehr stimmt, wie das Beispiel $f_n := n\,f_1(n\,x)$ mit $f_1 = \chi_{[1,2]}$ (analog (4.2.7)) zeigt: Es konvergiert $f_n(x) \to 0$ für jedes x, also konvergiert $f_n \to 0$ punktweise. Hingegen ist $\int_0^2 f_n = n \cdot \frac{1}{n} = 1$.

Beachte weiters, daß die Untersumme $U(f,Z)$ gerade das Integral $\int_I f_Z$ der Treppenfunktion $f_Z := \sum_{I\in Z} \inf(f(I))\,\chi_I$ ist, wobei χ_I hier ausnahmsweise die charakteristische Funktion des Intervalls I vermindert um den rechten Randpunkt bezeichne. Wenn wir insbesonders äquidistante Zerlegungen Z_n mit 2^{n+1} mit den 2^{n+1} Teilungspunkten $t_j := a + j(b-a)2^{-n-1}$ betrachten, so ist die zugehörige Treppenfunktion $f_n := f_{Z_n}$ und es gilt $f_n \le f_{n+1}$. Für $x \in [a,b) \setminus \Delta(f)$ ist f stetig bei x, also existiert zu $\varepsilon > 0$ ein $\delta > 0$, sodaß $f(x') \ge f(x) - \varepsilon$ für alle x' mit $|x'-x| < \delta$. Für $(b-a)2^{-n-1} < \delta$ liegt x in einem $I \in Z_n$ (und ist nicht der rechte Randpunkt) mit $I \subseteq \{x' : |x'-x| < \delta\}$. Damit ist $\inf(f(I)) \ge f(x) - \varepsilon$, also $f_n(x) \ge f(x) - \varepsilon$. Somit konvergiert $f_n(x) \to f(x)$ für alle x mit Ausnahme der Nullmenge $\Delta(f) \cup \{b\}$, man sagt kurz: f_n konvergiert gegen f fast überall.

Es stellt sich somit die Frage, ob allgemeiner für D-integrierbare f_n mit $f_n \le f_{n+1}$ für alle n für welche f_n gegen eine D-integrierbare Funktion f_∞ fast überall konvergiert, auch $\int_I f_n \to \int_I f_\infty$ gilt. Indem wir $g_n := f_\infty - f_n$ betrachten, genügt es folgendes zu zeigen:

5.1.14 Theorem. Stetigkeit der Integration bzgl. monotoner pktw. Konvergenz fast überall.
Es sei $g_n \in R(I,\mathbb{R})$ mit $g_n \geq g_{n+1} \geq 0$ für alle $n \in \mathbb{N}$ und $g_n \to 0$ fast überall. Dann gilt $\int_I g_n \to 0$.

Beweis. Nach Voraussetzung ist die Vereinigung N aller $\Delta(f_n)$ und der Menge der Punkte $x \in I$ für welche $g_n(x)$ nicht gegen 0 konvergiert eine Nullmenge. Sei $\varepsilon > 0$, dann existieren also abzählbar viele (offene) Intervalle I_k mit $N \subseteq \bigcup_k I_k$ und $\sum_k |I_k| < \varepsilon$. Sei $x \in I \setminus N$. Dann ist $\lim_{n\to\infty} g_n(x) = 0$ und somit existiert ein $n = n_x$ mit $g_n(x) < \varepsilon$. Da g_n stetig ist, existiert ein offener Ball U_x um x mit $g_n|_{U_x} < \varepsilon$.

Zur offenen Überdeckung $\{I_k : k \in \mathbb{N}\} \cup \{U_x : x \notin N\}$ existiert nach (5.1.4b) eine endliche Teilüberdeckung $\{I_1, \dots, I_M\} \cup \{U_{x_1}, \dots, U_{x_K}\}$ mit $M, K \in \mathbb{N}$. Sei $n' := \max\{n_{x_1}, \dots, n_{x_N}\}$. Dann ist $g_n < \varepsilon$ auf $\bigcup_{i=1}^N U_{x_i}$ für alle $n \geq n'$. Für offene Intervalle $J = (a', b')$ bezeichnet wir mit $\int_J f$ das Integral $\int_{a'}^{b'} f$. Somit ist

$$\int_I g_n \leq \sum_{j=1}^{M} \int_{I_j} g_n + \int_{\bigcup_{j\leq K} U_{x_j}} g_n \leq \|g_n\|_\infty \sum_j |I_j| + (b-a)\,\varepsilon \leq \Big(\|g_1\|_\infty + (b-a)\Big)\,\varepsilon,$$

also $\int_I g_n \to 0$ für $n \to \infty$. □

Beachte, daß hier die D-Integrierbarkeit der Grenzfunktion vorausgesetzt werden muß, denn sei $\{t_n : n \in \mathbb{N}\}$ eine Abzählung von $\mathbb{Q}$, dann ist $f_n := \chi_{\{t_1,\dots,t_n\}}$ D-integrierbar und es gilt $f_n \leq f_{n+1}$ konvergiert punktweise gegen die nicht D-integrierbare Dirichlet-Funktion $\chi_{\mathbb{Q}}$.

Dieses Theorem ist einer der Ausgangspunkte der Lebesgue'schen Integrationstheorie, die dazu dient die Vervollständigung des Raums $R(I,\mathbb{R})$ der D-integrierbaren Funktionen bzgl der 1-Norm zu bestimmen, siehe z.B. [16, 4.13].

5.1.15 Proposition. Endliche Teilmengen von Vitali-Überdeckungen bedecken beinahe alles.
Es sei $X \subseteq \mathbb{R}$ mit äußeren Maß $|X| < \infty$ und $\mathcal{J}$ eine Vitali-Überdeckung von X. Dann existiert für jedes $\varepsilon > 0$ eine endliche Teilmenge $\mathcal{J}_0 \subseteq \mathcal{J}$ von paarweise disjunkten Intervallen mit $|X \setminus \bigcup \mathcal{J}_0| < \varepsilon$.

Dabei versteht man unter einer Vitali-Überdeckung einer Teilmenge $X \subseteq \mathbb{R}$ eine Menge $\mathcal{J}$ von Intervallen, s.d. $\forall \varepsilon > 0\ \ \forall x \in X\ \ \exists\, J \in \mathcal{J}$ mit $x \in J$ und $0 < |J| < \varepsilon$. D.h. jeder Punkt $x \in X$ ist in maßtheoretisch beliebig kleinen Mengen $J \in \mathcal{J}$ enthalten.

Beweis. Es sei $\varepsilon > 0$. Wegen $|X| < \infty$ existiert eine offene Menge $U \supseteq X$ (abzählbare disjunkte Vereinigung offener Mengen) mit $|U \setminus X| < \varepsilon$. O.B.d.A. gilt $I \subseteq U$ für alle $I \in \mathcal{I}$.

Wir gehen indirekt vor. Dann können wir rekursiv eine Folge I_n von disjunkten Intervallen in $\mathcal{I}$ finden mit $|I_{n+1}| \geq \frac{1}{2}\sup\{|I| : \mathcal{I} \ni I \text{ disjunkt zu } I_1, \dots, I_n\} > 0$, da $\mathcal{I}$ eine Vitali-Überdeckung ist. Wegen $I_i \subseteq U$ ist $\sum_{i=0}^\infty |I_i| \leq |U| \leq |U \setminus X| + |X| < \infty$. Es existiert somit ein N mit $\sum_{i=N+1}^\infty |I_i| < \frac{\varepsilon}{5}$. Sei $X_\varepsilon := X \setminus \bigcup_{i=0}^N I_i$. Für jedes Intervall $J = [a,b]$ sei $\widetilde{J} := [a - 2(b-a), b + 2(b-a)]$ also $|\widetilde{J}| = 5|J|$. Aus $I_m \cap I \neq \emptyset$ und $|I_m| \geq |I|/2$ folgt $I \subseteq \widetilde{I_m}$.

Es ist $X_\varepsilon \subseteq \bigcup_{n>N} \widetilde{I_n}$ und somit $|X_\varepsilon| \leq \bigcup_{n>N} |\widetilde{I_n}| \leq 5\sum_{n>N} |I_n| = 5\frac{\varepsilon}{5} = \varepsilon$: Für jedes $x \in X_\varepsilon$ existiert ein von $I_1, \dots, I_N$ disjunktes $I \in \mathcal{I}$ mit $x \in I$. Sei $m := \min\{n : I_n \cap I \neq \emptyset\}$ und somit $I \cap I_n = \emptyset$ für alle $n < m$. Dann ist $N < m < \infty$, denn aus $|I_n| \to 0$ folgt $\exists\, n > N$ mit $\frac{1}{2}|I| > |I_n| \geq \frac{1}{2}\sup\{|I| : \mathcal{I} \ni I \text{ disjunkt zu } I_1, \dots, I_{n-1}\} \Rightarrow\ \exists\, i < n$: $I \cap I_i \neq \emptyset$. Wegen $I_m \cap I \neq \emptyset$ und $|I_m| \geq |I|/2$ (da $I \cap I_n = \emptyset$ für alle $n < m$) ist $x \in I \subseteq \widetilde{I_m}$. □

5.1.16 Proposition. Differenzierbarkeit monotoner Funktionen.
Es sei $f : [a,b] \to \mathbb{R}$ monoton wachsend. Dann ist f höchstens abzählbar vielen Punkten unstetig und fast überall differenzierbar.

Beweis. Es ist $f(x') \leq f(x) \leq f(x'')$ für $x' < x < x''$, also existiert $f(x-) := \lim_{x'\to x-} f(x') \leq f(x) \leq \lim_{x''\to x+} f(x'') =: f(x+)$. Alle Unstetigkeitsstellen sind also Sprungstellen und davon kann es nur abzählbar viele geben, denn andernfalls gäbe es ein rationales $\delta > 0$ und (über)abzählbar viele

x mit $f(x_+) - f(x_-) > \delta$ und damit wäre $f(b) - f(a)$ mindestens so groß wie die Summe dieser Sprunghöhen $f(x_+) - f(x_-) > \delta$, also $+\infty$, ein Widerspruch.

Nun zur Differenzierbarkeitsaussage. Es sei dazu

$$d_+ f(x) := \overline{\lim_{h\to 0}} \frac{f(x+h) - f(x)}{h} \geq \underline{\lim_{h\to 0}} \frac{f(x+h) - f(x)}{h} =: d_- f(x) \geq 0.$$

Angenommen $|\{x : d_+ f(x) > d_- f(x)\}| > 0$. Dann existieren $a > b$ in $\mathbb{Q}$ mit $r := |A| > 0$, wobei $A := \{x : d_+ f(x) > a > b > d_- f(x)\}$. Sei $\varepsilon > 0$ beliebig, dann existiert eine abzählbare disjunkte Vereinigung U offener Intervalle mit $A \subseteq U$ und $|U| < |A| + \varepsilon$.

Wegen $d_- f(x) < b$ für $x \in A$ ist

$$\mathcal{I} := \Big\{ I : A \cap \partial I \neq \emptyset,\ I \subseteq U,\ 0 \leq \frac{f(\max I) - f(\min I)}{|I|} < b \Big\}$$

eine Vitali-Überdeckung von A. Nach (5.1.15) existiert ein endliches $\mathcal{I}_0 \subseteq \mathcal{I}$ paarweise disjunkter Intervalle mit $|A \setminus \bigcup \mathcal{I}_0| < \varepsilon$. Sei nun $A' := A \cap \bigcup_{I\in\mathcal{I}_0} (I \setminus \partial I)$, dann ist $A \setminus A' \subseteq A \setminus \bigcup_{I\in\mathcal{I}_0} (I \setminus \partial I)$ und somit $|A'| = |A| - |A \setminus A'| > r - \varepsilon$ und wegen $I \subseteq U$ for all $I \in \mathcal{I}$ ist $\sum_{I\in\mathcal{I}_0} |I| \leq |U|$.

Wegen $d_+ f(y) > a$ für $y \in A'$ ist

$$\Big\{ J : A' \cap \partial J \neq \emptyset,\ \ \exists I \in \mathcal{I}_0 : J \subseteq I,\ \frac{f(\max J) - f(\min J)}{|J|} > a \Big\}$$

eine Vitali-Überdeckung von A'. Nach (5.1.15) existiert ein endliches $\mathcal{J}_0 \subseteq \mathcal{J}$ paarweise disjunkter Intervalle mit und $|A' \setminus \bigcup \mathcal{J}_0| < \varepsilon$. Dann ist

$$\sum_{J\in\mathcal{J}_0} |J| = \mu^*\Big(\bigcup \mathcal{J}_0\Big) \geq |A'| - \mu^*\Big(A \setminus \bigcup \mathcal{J}_0\Big) > r - \varepsilon - \varepsilon = r - 2\varepsilon.$$

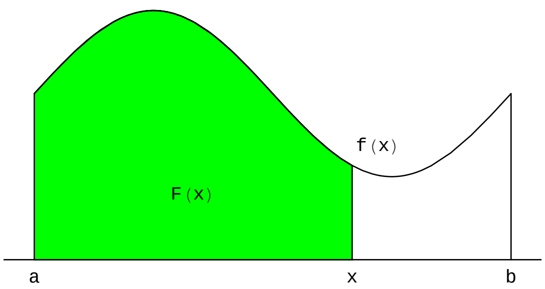
Somit ist

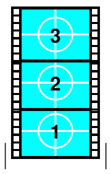
$$\begin{aligned}(r+\varepsilon)\, b > |U|\, b &\geq \sum_{I\in\mathcal{I}_0} |I|\, b > \sum_{I\in\mathcal{I}_0} f(\max I) - f(\min I) \geq \sum_{I\in\mathcal{I}_0} \sum_{I\supseteq J\in\mathcal{J}_0} f(\max J) - f(\min J) \\ &= \sum_{J\in\mathcal{J}_0} f(\max J) - f(\min J) > \sum_{J\in\mathcal{J}_0} |J|\, a > (r - 2\varepsilon)\, a\end{aligned}$$

für alle $\varepsilon > 0$ und wegen $r \neq 0$ ist $b \geq a$, ein Widerspruch. □

5.2 Unbestimmte Integrale

5.2.1 Definition. Unbestimmtes Integral.
Es sei $f : I \to \mathbb{R}$ D-integrierbar und $a \in I$. Nach dem Lebesgue'schen Integrabilitätskriterium (5.1.4) existiert dann für jedes $x \in I$ auch $F(x) := \int_a^x f(t)\, dt$. Diese Abbildung $F : I \to \mathbb{R}$ heißt (ein) UNBESTIMMTES INTEGRAL von f.





5.2.2 Hauptsatz der Differential und Integralrechnung.
Das unbestimmte Integral F jeder D-integrierbaren Funktion $f : [a, b] \to \mathbb{R}$ ist stetig.

Ist der Integrand f stetig bei $x \in [a, b]$, so ist F differenzierbar bei x mit Ableitung $F'(x) = f(x)$.

Umgekehrt sei $F : [a, b] \to \mathbb{R}$ stetig differenzierbar mit Ableitung $f := F'$. Dann ist f D-integrierbar und $\int_a^b f(x)\,dx = F(b) - F(a) =: F(x)\big|_{x=a}^{b}$.

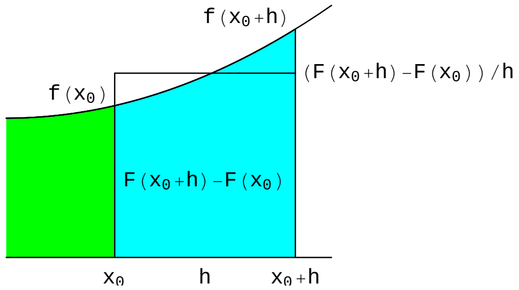
Jede Funktion F die $F' = f$ erfüllt heißt STAMMFUNKTION von f. Insbesonders ist $\int_a^b f(x)\,dx = F(x)\big|_{x=a}^{b}$ für jede Stammfunktion F von f und man schreibt auch $\int f$ für die (Familie der) Stammfunktion(en) von f und $\int f(t)\,dt$ für ihre(n) Wert(e) an der Stelle t.

Beweis. Es ist $F(x+h) - F(x) = \int_x^{x+h} f(t)\,dt$, also $|F(x+h) - F(x)| \le |h| \sup\{|f(t)| : t\}$ nach (5.1.11) und somit F (Lipschitz)-stetig, da f beschränkt ist.

Ist f zusätzlich stetig bei x, so existiert zu $\varepsilon > 0$ ein $\delta > 0$ s.d. $|f(t) - f(x)| \le \varepsilon$ für alle t mit $|t - x| \le \delta$ und damit ist

$$\begin{aligned}\Big|\frac{F(x+v) - F(x)}{v} - f(x)\Big| &= \Big|\frac{1}{v}\int_x^{x+v} f(t) - f(x)\,dt\Big| \\ &\le \Big|\frac{1}{v}\int_x^{x+v} |f(t) - f(x)|\,dt\Big| \le \varepsilon\end{aligned}$$

für $|v| \le \delta$. Also ist $F'(x) = \lim_{v\to 0} \frac{F(x+v)-F(x)}{v} = f(x)$.



Ist f also stetig, so ist $F : x \mapsto \int_a^x f(y)\,dy$ eine Stammfunktion von f.

Sei nun F stetig differenzierbar mit Ableitung $f := F'$. Somit ist F eine Stammfunktion von f und nach dem 1. Teil ist das unbestimmte Integral $F_a : x \mapsto \int_a^x f(t)\,dt$ ebenfalls eine Stammfunktion, also ist die Ableitung von $F - F_a$ gleich 0 und damit nach dem Spezialfall des Mittelwertsatzes (4.1.5) $F - F_a$ konstant, also $\int_a^b f(x)\,dx = F_a(b) = F_a(b) - F_a(a) = F(b) - F(a)$. □

Bemerkung. Interpretation des Hauptsatzes.
Der Hauptsatz besagt also, daß Integrieren und Differenzieren im wesentlichen, d.h. bis auf Addition einer Konstanten, invers zueinander sind.

Genauer: Es ist Differenzieren $d : f \mapsto f'$ auf der Menge der differenzierbaren Funktionen $I \to \mathbb{R}$ nicht injektiv, denn $f_1' = f_2' \Leftrightarrow f_1 - f_2$ ist konstant.

Jedoch ist Differenzieren d auf der Menge $C^1(I, \mathbb{R})$ der stetig differenzierbaren Funktionen $F : I \to \mathbb{R}$ surjektiv als Abbildung in die Menge $C(I, \mathbb{R})$ der stetigen Abbildungen $f : I \to \mathbb{R}$, denn zu jedem $f \in C(I, \mathbb{R})$ definiert $F : x \mapsto \int_a^x f$ eine Stammfunktion.

Somit definiert dieses unbestimmte Integral eine rechtsinverse Abbildung $i : C(I, \mathbb{R}) \to C^1(I, \mathbb{R})$ zum Differenzieren $d : C^1(I, \mathbb{R}) \to C(I, \mathbb{R})$, für die weiters gilt: $(i \circ d)(F) - F$ ist eine konstante Abbildung.

Wenn wir d auf den Teilvektorraum $C_a^1(I, \mathbb{R}) := \{F \in C^1(I, \mathbb{R}) : F(a) = 0\}$ einschränkten, dann sind Differenzieren

$$\begin{aligned} d : C_a^1(I, \mathbb{R}) &\to C(I, \mathbb{R}) \\ F &\mapsto F' \end{aligned}$$

und (unbestimmt) Integrieren

$$\begin{aligned} i : C(I,\mathbb{R}) &\to C^1_a(I,\mathbb{R}) \\ f &\mapsto (x \mapsto \int_a^x f) \end{aligned}$$

zueinander inverse Isomorphismen.

Bemerkung. Allgemeinere Formulierungen des Hauptsatzes und unbestimmter Integrale. Wir haben im Beweis von (5.2.2) wegen (5.1.4) gezeigt, daß unbestimmte Integrale F von D-integrierbare Funktion f fast überall differenzierbar sind, nämlich zumindest dort wo f stetig ist, und an diesen Stellen x die Ableitung $f(x)$ ist.

Auch umgekehrt können wir mehr ausssagen: Wenn eine Abbildung $F : I \to \mathbb{R}$ differenzierbar ist, und ihre Ableitung $f = F'$ D-integrierbar ist, so gilt $\int_a^b f = F(b) - F(a)$.
In der Tat sei Z eine Zerlegung mit Teilungsintervallen $I_i = [t_{i-1}, t_i]$. Dann ist nach dem Mittelwertsatz

$$\inf(F'(I_i)) \le \frac{F(t_i) - F(t_{i-1})}{t_i - t_{i-1}} \le \sup(F'(I_i))$$

und somit

$$U(f,Z) = \sum_{I\in Z} \inf(f(I))\,|I| \le \underbrace{\sum_{i=1}^{N} F(t_i) - F(t_{i-1})}_{=F(b)-F(a)} \le \sum_{I\in Z} \sup(f(I))\,|I| = O(f,Z),$$

also $|F(b) - F(a) - \int_a^b f| < \varepsilon$ für alle ε und somit $\int_a^b f = F(b) - F(a)$.

Ausgehend vom bestimmten Integral $\int_a^b f$ integrierbarer Funktionen, haben wir also unbestimmte Integrale $F_a : x \mapsto \int_a^x f$ betrachtet, und gezeigt, daß diese für stetiges f eine Stammfunktion zu f sind. Je zwei solche unbestimmte Integrale F_{a_1} und F_{a_0} unterscheiden sich nur um die Konstante $\int_{a_0}^{a_1} f$. Allgemein gilt nach (4.1.5), daß zwei beliebige Stammfunktionen F_0 und F_1 einer beliebigen Funktion f sich nur um eine Konstante unterscheiden können, denn $(F_1 - F_0)' = (F_1)' - (F_0)' = f - f = 0$, also ist $F_1 - F_0$ konstant. Wir schreiben auch in diesem allgemeinen Fall $\int f$ für die Familie der Stammfunktionen von f, und nennen auch dies UNBESTIMMTES INTEGRAL von f (obwohl es nicht immer durch D-Integration erhalten werden kann: Z.B. ist $F : x \mapsto \sqrt[3]{x^4}\sin(1/x)$ zur einer differenzierbaren Funktion $\mathbb{R} \to \mathbb{R}$ erweiterbar und die Ableitung $f(x) := F'(x) = 4\sqrt[3]{x}\sin(1/x)/3 - \cos(1/x)/\sqrt[3]{x^2}$ ist unbeschränkt bei 0). Der Hauptsatz der Analysis besagt dann unter anderen, daß das bestimmte Integral D-integrierbarer Funktionen f wie folgt berechnet werden kann:

$$\int_a^b f = \Big(\int f\Big)\Big|_a^b \quad \text{oder klassisch} \quad \int_a^b f(x)\,dx = \Big(\int f(x)\,dx\Big)\Big|_{x=a}^{b}.$$

Beachte jedoch, daß die klassische und weit verbreitete Schreibweise $\int f(x)\,dx$ für den Funktionswert $\big(\int f\big)(x)$ – oder kurz $\int f(x)$ – des/eines unbestimmten Integrals $\int f$ an der Stelle x mit Vorsicht zu genießen ist, denn x kommt hier in zwei wesentlich verschiedenen Bedeutungen vor: Einerseits bezeichnet es die (laufende) Integrationsvariable und andererseits die fixe Stelle, an welcher die resultierende Stammfunktion schlußendlich ausgewertet wird. Das entsprechende diskrete Pendant einer Summe wäre $\sum_{x=0}^{x} f_x$, was offensichtlicher Unsinn ist.

Es entspricht die obige kürzere aber ungebräuchlichere Schreibweise auch eher unserer Sprechweise, denn verbal formuliert besagt obiges ja, daß “das bestimmte Integral von a bis b einer Funktion f berechnet werden kann, indem man das/ein unbestimmte Integral der Funktion an den Stellen b und a auswertet und dann die Differenz nimmt” und dabei kommt die Variable x auch nicht vor.

Es wurde von Henstock und Kurzweil eine andere Integrationstheorie entwickelt mit der Eigenschaft, daß jede überall differenzierbare Funktion $F : I \to \mathbb{R}$ eine KURZWEIL-INTEGRIERBARE Ableitung besitzt und für das KURZWEIL-INTEGRAL $\int_a^b F' = F(b) - F(a)$ gilt, siehe [23].

5.2.3 Anwendungen des Hauptsatzes der Analysis.
Vermöge (5.2.2) können wir Formeln der Form $F' = f$ die wir für das Differenzieren erhalten haben äquivalent als Formeln der Form $\int f = F + C$ für (unbestimmte) Integrale interpretieren:

1. Insbesonders erhalten wir folgende Beispiele für Stammfunktionen bzw. unbestimmte Integrale:

$$\begin{aligned}
&\int x^a\,dx = \frac{x^{a+1}}{a+1} + C \text{ für alle } \mathbb{Q} \ni a \neq -1,\\
&\int \sin(x)\,dx = -\cos(x) + C,\\
&\int \cos(x)\,dx = \sin(x) + C \text{ wegen (4.1.3.6)},\\
&\int \frac{1}{\cos(x)^2}\,dx = \tan(x) + C \text{ wegen (4.1.21.2)},\\
&\int \frac{1}{\sin(x)^2}\,dx = -\cot(x) + C,\\
&\int \frac{1}{\sqrt{1-x^2}}\,dx = \arcsin(x) + C \text{ wegen (4.1.21.1)},\\
&\int \frac{1}{1+x^2}\,dx = \arctan(x) + C \text{ wegen (4.1.21.2)}.
\end{aligned}$$

2. Da Integrieren (fast) die Inverse zum Differenzieren ist und letzteres linear ist, gilt gleiches auch für das unbestimmte Integral von stetigen Funktionen f und g, denn mit $F(x) := \int_a^x f$ und $G(x) := \int_a^x g$ erhalten wir:

$$\int_a^x (f+\lambda g) = \int_a^x F' + \lambda G' = \int_a^x (F + \lambda\, G)' = (F + \lambda\, G)(x) = \int_a^x f + \lambda \int_a^x g,$$

also

$$\int (f + \lambda\, g) = \int f + \lambda \int g.$$

Dies folgt natürlich auch direkt aus der Linearität des bestimmten Integrals, die wir in (5.1.7) bewiesen haben.

Allgemeiner können wir diese Formel auch als Formel für Stammfunktionen von (nicht notwendigerweise D-integrierbarer) Funktionen interpretieren, denn dies ist (per Definition) äquivalent zur Aussage $(\int f + \lambda \int g)' = f + \lambda\, g$, welche unmittelbar aus (4.1.19), d.h. aus der Linearität des Differenzierens folgt: $(\int f + \lambda \int g)' = (\int f)' + \lambda\,(\int g)' = f + \lambda\, g$.

3. Aus der Leibniz'schen Produktregel (4.1.15) $(f \cdot g)' = f' \cdot g + f \cdot g'$ folgt durch Integration (für stetig differenzierbare f und g) die Formel für PARTIELLE INTEGRATION:

$$\int f\, g' = f\, g - \int f'\, g,$$

oder allgemeiner für differenzierbare f und g für welche $f'\, g$ eine Stammfunktion $\int f'\, g$ besitzt, denn dies ist äquivalent zur nach (4.1.15) gültigen Aussage $(f\, g - \int f'\, g)' = (f\, g)' - f'\, g = f'\, g + f\, g' - f'\, g = f\, g'$.
Z.B. erhalten wir durch partielle Integration

$$\int x\, \cos(x)\, dx = x\, \sin(x) - \int 1\, \sin(x)\, dx = x\, \sin(x) + \cos(x) + C.$$

Es kann uns im allgemeinen natürlich niemand garantieren, daß wir das zweite Integral $\int f'\, g$ eher bestimmen können als $\int f\, g'$.

4. Aus der Kettenregel (4.1.14) $(F \circ g)' = (F' \circ g) \cdot g'$ folgt durch Integrieren die Substitutionsformel für unbestimmte Integrale:

$$\int F'(g(t))\, g'(t)\, dt = F(g(t)) + C \quad \text{oder auch} \quad \int (f \circ g)\cdot g' = \Big(\int f\Big) \circ g.$$

Diese Formel gilt für differenzierbares F (bzw. f für welche eine Stammfunktion existiert) und differenzierbares g, denn dazu äquivalent ist die Kettenregel (4.1.14):

$$\frac{d}{dt} F(g(t)) + C = F'(g(t)) \cdot g'(t)$$

Z.B. erhalten wir durch Substitution:

$$\int \frac{2t}{t^2+1}\, dt = \int 2t\, \ln'(t^2+1)\, dt = \int \ln'(g(t))\, g'(t)\, dt = \ln(g(t)) + C = \ln(1+t^2) + C.$$

Die Substitutionsformel

$$\int (f \circ g)\cdot g' = \Big(\int f\Big) \circ g$$

kann in beide Richtungen verwendet werden, also als

$$\int f(g(x))\, g'(x)\, dx = \int f(y)\, dy\Big|_{y=g(x)}$$

oder – für invertierbares g – als

$$\int f(y)\, dy = \int f(g(x))\, g'(x)\, dx\Big|_{x=g^{-1}(y)}.$$

Probieren wir das am Beispiel $\int \frac{2x}{1+x^2}\, dx$. Bei diesem Integranden stört uns vor allem der komplizierte Ausdruck $1 + x^2$ im Nenner. Wenn wir diesen als Wert der Funktion $g : x \mapsto 1 + x^2$ auffassen, und dann den Zähler als $g'(x) = 2x$ identifizieren, und schließlich noch $f(y) := 1/y$ setzen, so ist wegen der ersten Variante der Substitutionsformel

$$\begin{aligned}\int \frac{2x}{1+x^2}\, dx &= \int f(g(x))\, g'(x)\, dx = \int f(y)\, dy\Big|_{y=g(x)}\\ &= \int \frac{1}{y}\, dy\Big|_{y=1+x^2} = \Big(\ln(y) + C\Big)\Big|_{y=1+x^2} = \ln(1+x^2) + C.\end{aligned}$$

Um die zweite Variante der Subsitutionsformel anwenden zu können setzen wir nun $f(y) := \frac{2y}{1+y^2}$ und erhalten somit für jedes invertierbare g

$$\int \frac{2y}{1+y^2}\, dy = \int f(y)\, dy = \int f(g(x))\, g'(x)\, dx\Big|_{x=g^{-1}(y)}.$$

Damit wir weiter rechnen können sollte $f(g(x)) = 2g(x)/(1 + g(x)^2)$ einfacher werden, also vor allem der Nenner $1 + g(x)^2$. Wir versuchen z.B. $1 + g(x)^2 = x$ zu erreichen, d.h. wir setzen $g(x) := \sqrt{x-1}$. Dann ist $f(g(x)) = \frac{2g(x)}{1+g(x)^2} = \frac{2\sqrt{x-1}}{x}$ und $g'(x) = \frac{1}{2\sqrt{x-1}}$ und somit

$$\begin{aligned}\int \frac{2y}{1+y^2}\, dy &= \int f(y)\, dy = \int f(g(x))\, g'(x)\, dx\Big|_{x=g^{-1}(y)}\\ &= \int \frac{2\sqrt{x-1}}{x}\, \frac{1}{2\sqrt{x-1}}\, dx\Big|_{x=1+y^2} = \int \frac{1}{x}\, dx\Big|_{x=1+y^2}\\ &= \Big(\ln(x) + C\Big)\Big|_{x=1+y^2} = \ln(1+y^2) + C.\end{aligned}$$

In beiden Fällen haben wir eine neue Variable vermöge $g(x) := 1 + x^2 = y$ (bzw. $g^{-1}(y) = 1 + y^2 = x$) eingeführt und die Substitutionsformel besagt dann, daß wir $g'(x)\,dx$ durch dy ersetzen müssen (bzw. dy durch $g'(x)\,dx$). Die Schreibweise $\frac{dy}{dx} = g'(x)$ (bzw. $(g^{-1})'(y) = \frac{dx}{dy}$) kann hier als Memotechnik dienen. Bevor das so entstandene Integral berechnet werden kann, muß es einen aber gelingen die ursprüngliche Variable völlig loszuwerden. Nach der Integration muß man dann noch unbedingt die entsprechende (Rück-)Substitution vornehmen.

Die Substitutionsformel ist also das Inverse zur Kettenregel und direkt aus dieser herleitbar. Die Anwendung der Substitutionsformel ist auch sehr einfach durchzuführen (abgesehen von der geschickten Wahl der Substitutionsvariable). Etwas aufwendiger war es hier den Zusammenhang zwischen Substitutionsformel und Anwendungsverfahren herzustellen, also letzteres zu begründen.

5. Die Substitutionsformel für bestimmte Integrale:

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

folgt aus dem vorigen Punkt, denn

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

mit einer Stammfunktion F von f.
Für das bestimmte Integral $\int_{-1}^1 \frac{2x}{1+x^2}\,dx$ erhalten wir somit

$$\int_{-1}^1 \frac{2x}{1+x^2}\,dx = \Big(\int \frac{2x}{1+x^2}\Big)\Big|_{x=-1}^1 = \ln(1+x^2)|_{x=-1}^1 = \ln(2) - \ln(2) = 0,$$

was wir allerdings auch ohne jegliche Rechnung hätten sehen können, denn $\mathrm{id} : x \mapsto x$ ist ungerade, also $x \mapsto 1 + x^2$ gerade und somit $f : x \mapsto \frac{2x}{1+x^2}$ ungerade, also mit Substitution $y = g(x) := -x$

$$\begin{aligned}\int_{-a}^0 f(x)\,dx &= \int_{g(a)}^{g(0)} f(-g(x))\,(-g'(x))\,dx = \int_{g(0)}^{g(a)} -f(g(x))\,g'(x)\,dx \\ &= \int_0^a -f(x)\,dx = -\int_0^a f(x)\,dx,\end{aligned}$$

also $\int_{-a}^a f = \int_{-a}^0 f + \int_0^a f = -\int_0^a f + \int_0^a f = 0$.
Eine weiteres Anwendungsbeispiel der Substitutionsformel ist:

$$\int_1^2 x\,\sqrt{1+x^2}\,dx = \int_2^5 \sqrt{g}\,\frac{1}{2}\,dg = \frac{g^{3/2}}{3}\Big|_{g=2}^5 = \frac{5\sqrt{5} - 2\sqrt{2}}{3}.$$

6. Wie geben nun auch eine geometrische Interpretationen der Substitutionsformel für bestimmte Integrale:

$$\int_{g(a)}^{g(b)} f(x)\,dx = \int f(x)\,dx\Big|_{x=g(a)}^{g(b)} = \int f(g(t))\,g'(t)\,dt\Big|_{t=a}^{b} = \int_a^b f(g(t))\,g'(t)\,dt$$

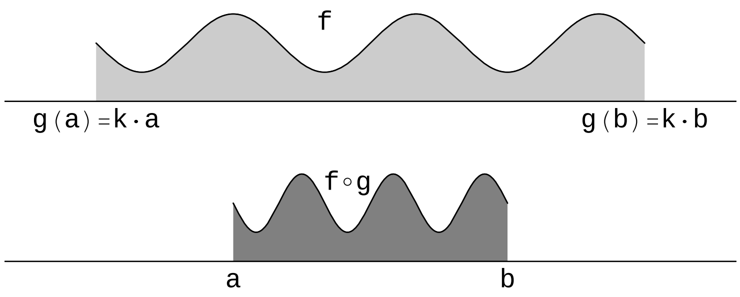
oder wenn wir $\alpha = g(a)$ und $\beta = g(b)$ setzen, also $a = g^{-1}(\alpha)$ und $b = g^{-1}(\beta)$, dann ist

$$\int_\alpha^\beta f(x)\,dx = \int_{g^{-1}(\alpha)}^{g^{-1}(\beta)} f(g(t))\,g'(t)\,dt.$$

Wenn insbesonders g linear (oder affin) ist, und somit g' konstant ist, also $g(x) = k\,x + d$ und $g'(x) = k$ gilt, dann ist

$$\int_{g(a)}^{g(b)} f = \int_a^b (f \circ g) \cdot g' = k \int_a^b (f \circ g)$$

was man auch folgender Zeichnung entnimmt:



Für allgemeines (nicht affines) g kompensiert der Faktor $g'(t)$ im rechtsstehenden Integral die Verzerrung, welche durch den Übergang von f zu $f \circ g$ an der Fläche verursacht wurde.

7. Das UNBESTIMMTE INTEGRAL DER UMKEHRFUNKTION ist gegeben durch:

$$\int f^{-1} = f^{-1} \cdot \mathrm{id} - \Big(\int f\Big) \circ f^{-1},$$

denn wegen der Substitutionsformel und partieller Integration ist

$$\begin{aligned}\Big(\int f^{-1}\Big) \circ f &= \int (f^{-1} \circ f) \cdot f' = \int \mathrm{id} \cdot f' = \mathrm{id} \cdot f - \int 1 \cdot f \\ \Rightarrow \int f^{-1} &= \Big(\int f^{-1}\Big) \circ f \circ f^{-1} = (\mathrm{id} \cdot f) \circ f^{-1} - \Big(\int f\Big) \circ f^{-1} = f^{-1} \cdot \mathrm{id} - \Big(\int f\Big) \circ f^{-1}.\end{aligned}$$

Z.B. ist

$$\begin{aligned}\int \arcsin(x)\, dx &= x \cdot \arcsin(x) - \Big(\int \sin\Big)(\arcsin(x)) \\ &= x \cdot \arcsin(x) + \cos(\arcsin(x)) + C \\ &= x \cdot \arcsin(x) \pm \sqrt{1 - \sin(\arcsin(x))^2} + C \\ &= x \cdot \arcsin(x) + \sqrt{1 - x^2} + C,\end{aligned}$$

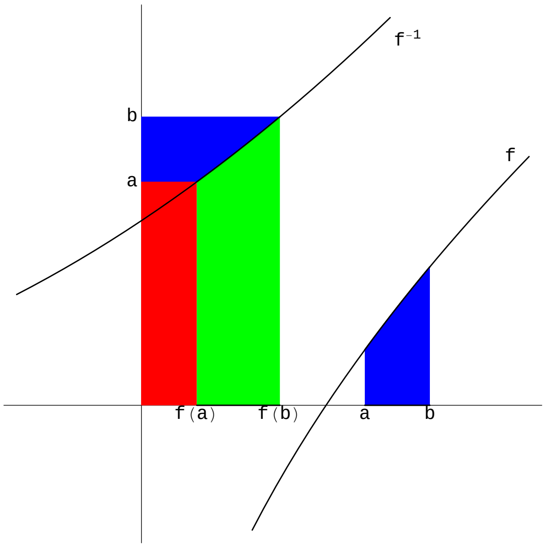
denn $\cos(y) \geq 0$ für $y = \arcsin(x) \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

8. Das BESTIMMTE INTEGRAL DER UMKEHRFUNKTION ist gegeben durch:

$$\int_{f(a)}^{f(b)} f^{-1}(y)\, dy = b\, f(b) - a\, f(a) - \int_a^b f(x)\, dx,$$

denn mittels Substitutionsformel und partieller Integration ist

$$\begin{aligned}\int_{f(a)}^{f(b)} f^{-1}(y)\, dy &= \int_a^b f^{-1}(f(x))\, f'(x)\, dx = \int_a^b x\, f'(x)\, dx = [x\, f(x)]_{x=a}^b - \int_a^b 1\, f(x)\, dx \\ &= b\, f(b) - a\, f(a) - \int_a^b f(x)\, dx.\end{aligned}$$



Beachte, daß wir uns zwar die korrekten Voraussetzungen für all diese Integrationsregeln überlegt haben, dies aber nicht wirklich notwendig ist, da die Aussage $\int f = F + C$ äquivalent zu $F' = f$ ist, also durch Differenzieren überprüft werden kann.

5.2.4 Definition. Logarithmus.
Wir haben in (3.4.14) die (Ableitung der) Logarithmusfunktion ln wie folgt erhalten:

$$\begin{aligned}
& a^2 := a \cdot a \\
\rightsquigarrow\quad & a^n \text{ für } n \in \mathbb{N} \text{ mittels Rekursion} \\
\rightsquigarrow\quad & a^k \text{ für } k \in \mathbb{Z} \text{ mittels } a^{n+m} = a^n \cdot a^m \text{ und somit } a^{-1} = \frac{1}{a} \\
\rightsquigarrow\quad & a^q \text{ für } q \in \mathbb{Q} \text{ durch } (a^n)^m = a^{nm} \text{ und somit } a^{1/n} = \sqrt[n]{a} \\
\rightsquigarrow\quad & a^x \text{ für } x \in \mathbb{R} \text{ durch stetige Erweiterung} \\
\rightsquigarrow\quad & e^x \text{ für } x \in \mathbb{R} \text{ durch } e := \lim_{n\to+\infty}(1+\frac{1}{n})^n = \sum_{n=1}^{\infty}\frac{1}{n!} \\
\rightsquigarrow\quad & \ln(x) \text{ für } x > 0 \text{ als Umkehrfunktion} \\
\rightsquigarrow\quad & \ln'(x) = \frac{1}{x} \text{ als Ableitung der Umkehrfunktion von } \mathit{exp} \text{ mit } \exp' = \exp\,.
\end{aligned}$$

Man könnte diesen Zugang auch umdrehen, d.h. den natürlichen Logarithmus ohne Verwendung der Exponentialfunktion definieren durch

$$\ln(x) := \int_1^x \frac{1}{y}\,dy.$$

Dann ist $\ln : \mathbb{R}^+ \to \mathbb{R}$ eine Stammfunktion zu $x \mapsto \frac{1}{x}$ also insbesonders streng monoton wachsend. Weiters ist $\ln(1) = \int_1^1 \frac{1}{y}\,dy = 0$ und

$$\ln(ab) = \int_1^{ab}\frac{1}{x}\,dx = \int_1^a \frac{1}{x}\,dx + \int_a^{ab}\frac{1}{x}\,dx = \ln(a) + \int_1^b \frac{1}{ay}\,a\,dy = \ln(a) + \ln(b).$$

Folglich ist $\ln(a^n) = n\,\ln(a)$ für alle $n \in \mathbb{N}$ und weiters $\ln(a^x) = x\,\ln(a)$ für $x \in \mathbb{Q}$. Denn sei $x = \frac{p}{q}$, dann ist $\ln(a) = \ln((a^{1/q})^q) = q\,\ln(a^{1/q})$, also $\ln(a^{1/q}) = \frac{1}{q}\,\ln(a)$ und somit $\ln(a^x) = x\,\ln(a)$ für $x > 0$. Für $x < 0$ ist $\ln(a^x) = \ln(1/a^{-x}) = -\ln(a^{-x}) = -(-x)\,\ln(a) = x\,\ln(a)$. Schließlich ist $\lim_{x\to+\infty}\ln(x) = +\infty$, denn $\int_1^n \frac{1}{x}\,dx \ge \sum_{k=2}^n \frac{1}{k} \to \infty$. Wegen $\ln(1/x) = -\ln(x)$ ist $\lim_{x\to 0+}\ln(x) = -\infty$.

Nach den Sätzen (3.4.3) und (7.5.4) über inverse Funktionen existiert somit die differenzierbare Umkehrfunktion $\exp : \mathbb{R} \to \mathbb{R}^+$ (die Exponentialfunktion) mit $\exp'(x) = \frac{1}{\ln'(\exp(x))} = \exp(x)$ (vgl. mit (4.1.21.3)) Weiters folgt aus den Aussagen für ln, daß $\exp(x+y) = \exp(x) \cdot \exp(y)$ und $\exp(0) = 1$. Es ist

$$\left(1+\frac{1}{n}\right)^n = \exp\left(\ln\left(\left(1+\frac{1}{n}\right)^n\right)\right) = \exp\left(n \ln\left(1+\frac{1}{n}\right)\right)$$

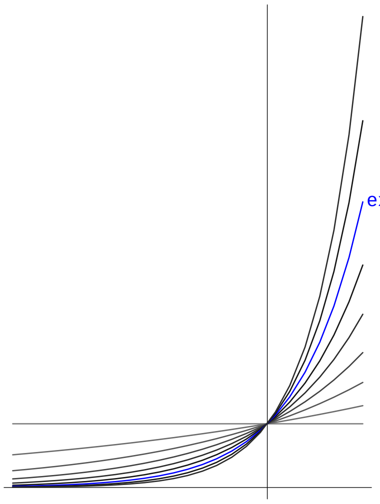
und somit

$$\begin{aligned} e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n &= \exp\left(\lim_{n\to\infty} n \ln\left(1+\frac{1}{n}\right)\right) \\ &= \exp\left(\lim_{h\to 0+}\frac{\ln(1+h)}{h}\right) = \exp(\ln'(1)) = \exp(1), \end{aligned}$$

d.h. $\exp(x) = e^x$ für alle $x \in \mathbb{Q}$.

Sei nun $a \in \mathbb{R}^+$. Dann setzen wir

$$a^x := \exp(x \ln(a)) \text{ für } x \in \mathbb{R}$$



Und für $a \neq 1$ heißt die Umkehrfunktion zu $x \mapsto a^x$ der Logarithmus $\log_a$ zur Basis a. Diese Definition von a^x stimmt für $x = \frac{p}{q} \in \mathbb{Q}$ mit der in der Analysis 1 gegebenen Definition $a^x := \sqrt[q]{a^p}$ überein, denn

$$(a^x)^q = \exp\Big(x \ln(a)\Big)^q = \exp\Big(q\, x \ln(a)\Big) = \exp\Big(p \ln(a)\Big) = \exp\Big(\ln(a)\Big)^p = a^p.$$

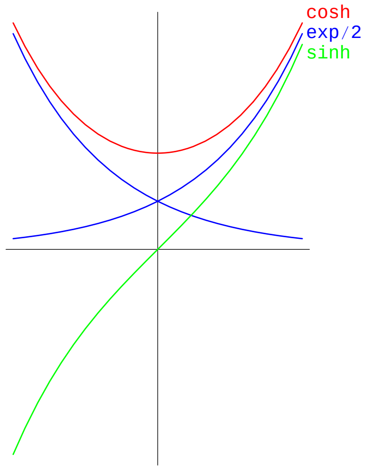
Beachte, daß $y = a^x = e^{x \ln(a)}$ die Gleichungen $x = \log_a(y)$ und $x \ln(a) = \ln(y)$ zur Folge hat, also ist

$$\ln(y) = \ln(a) \log_a(y),$$

d.h. Logarithmenfunktionen zu verschiedenen Basen sind zueinander proportional.

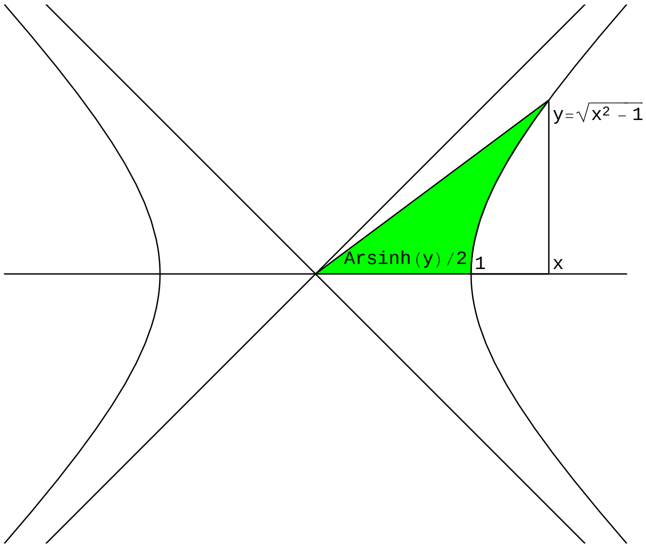
5.2.6 Definition. Hyperbolische Winkelfunktionen.
Es sei $f : \mathbb{R} \to \mathbb{R}$ eine Funktion. Dann ist $g : x \mapsto \frac{f(x)+f(-x)}{2}$ eine gerade und $u : x \mapsto \frac{f(x)-f(-x)}{2}$ eine ungerade Funktion mit $f = g + u$, die Zerlegung in geraden und ungeraden Teil.

Wenden wir dies insbesonders auf die Funktione $f := \exp$ an, so erhalten wir den Sinushyperbolicus $\sinh(x) := \frac{e^x - e^{-x}}{2}$, und den Cosinushyperbolicus $\cosh(x) := \frac{e^x + e^{-x}}{2}$. In Analogie zu den üblichen Winkelfunktionen definieren wir den Tangenshyperbolicus als $\tanh(x) := \frac{\sinh(x)}{\cosh(x)}$.



Man kann leicht folgende Aussagen beweisen:

1. $\sinh' = \cosh$ und $\cosh' = \sinh$, denn $\frac{d}{dx}\frac{e^x \pm e^{-x}}{2} = \frac{e^x \mp e^{-x}}{2}$.
2. $\cosh(x)^2 - \sinh(x)^2 = 1$, denn $\left(\frac{e^x \pm e^{-x}}{2}\right)^2 = \frac{e^{2x}+e^{-2x}-2}{4}$.
3. $\cosh(2x) - 1 = 2\sinh(x)^2$, denn $2\left(\frac{e^x - e^{-x}}{2}\right)^2 = \frac{e^{2x}+e^{-2x}-2}{2}$.
4. $\sinh(2x) = 2\sinh(x)\cosh(x)$, denn $2\left(\frac{e^x - e^{-x}}{2}\right)\left(\frac{e^x + e^{-x}}{2}\right) = \frac{e^{2x}-e^{-2x}}{2}$.
5. $\cosh(x) \geq 1$, denn $\cosh(x) - 1 = 2\sinh(\frac{x}{2})^2 \geq 0$.
6. $\cosh : \mathbb{R} \to [1, +\infty)$ ist surjektiv und $\cosh : [0, +\infty) \to [1, +\infty)$ ist bijektiv, denn $\cosh(0) = 1$ und $\lim_{x\to\pm\infty} \cosh(x) = +\infty$ und $\cosh'(x) = \sinh(x) \geq 0$ für $x \geq 0$.
7. $\sinh : \mathbb{R} \to \mathbb{R}$ ist bijektiv, denn $\lim_{x\to\pm\infty} \sinh(x) = \pm\infty$ und $\sinh'(x) = \cosh(x) \geq 1 > 0$.
8. $\sinh^{-1}(y) = \ln(y + \sqrt{y^2+1})$, denn aus $y = \frac{e^x - e^{-x}}{2}$ folgt $(e^x)^2 - 2ye^x - 1 = 0$ und somit $0 < e^x = y + \sqrt{y^2+1}$.
9. $\cosh^{-1}(y) = \ln(y + \sqrt{y^2-1})$, denn aus $y = \frac{e^x + e^{-x}}{2}$ folgt $(e^x)^2 - 2ye^x + 1 = 0$ und somit $0 < e^x = y + \sqrt{y^2-1}$.
10. $\int \sqrt{t^2-1}\,dt = \frac{1}{2}\Big(t\sqrt{t^2-1} - \cosh^{-1}(t)\Big)$, denn Substitution $t = \cosh(x)$ liefert
$$\begin{aligned}\int \sqrt{t^2-1}\,dt &= \int \sinh(x)\,\sinh(x)\,dx = \int \frac{\cosh(2x)-1}{2}\,dx \\ &= \frac{\sinh(2x)}{4} - \frac{x}{2} = \frac{\sinh(x)\cosh(x) - x}{2} = \frac{1}{2}\left(t\,\sqrt{t^2-1} - \cosh^{-1}(t)\right).\end{aligned}$$
11. $t \mapsto (\cosh(t), \sinh(t))$ ist eine bijektive Parametrisierung des rechten Astes der gleichseitigen Hyperbel $\{(x, y) \in \mathbb{R}^2 : x^2 - y^2 = 1\}$, wegen (2), (6) und (7). Dies erklärt die Namensgebung, denn $t \mapsto (\cos(t), \sin(t))$ ist analog eine Parametrisierung des Einheitskreises, wobei der Parameter t den Winkel (d.h. die Länge des Bogens) oder auch die doppelte Fläche des zugehörigen Kreissektors mißt, siehe (13).



12. Die Fläche des Hyperbelsektors mit Spitze im Punkt (x,y) ist $\frac{\cosh^{-1}(x)}{2} = \frac{\sinh^{-1}(y)}{2}$. Folglich heißen die Umkehrfunktionen auch Areasinushyperbolicus $\mathrm{Arsinh} = \sinh^{-1}$, Areacosinushyperbolicus $\mathrm{Arcosh} = \cosh^{-1}$ und Areatangenshyperbolicus $\mathrm{Artanh} = \tanh^{-1}$. Offensichtlich ist diese Fläche gegeben durch

$$\frac{x\,y}{2} - \int_1^x y(t)\,dt = \frac{x\sqrt{x^2-1}}{2} - \int_1^x \sqrt{t^2-1}\,dt = \frac{1}{2}\cosh^{-1}(x).$$

13. Wir können auch die analoge Flächenberechnung für die Winkelfunktionen anstelle der Hyperbelfunktionen durchführen. Der Kreis $\{(x,y) : x^2+y^2=1\}$ wird ja durch $t \mapsto (\cos(t), \sin(t))$ parametrisiert. Die Fläche eines Kreisektors von $(1,0)$ bis $(x,y) = (\cos(t), \sin(t))$ ist somit (via Substitution $s = \cos(r)$) durch

$$\begin{aligned}
\frac{x\,y}{2} + \int_x^1 y(s)\,ds &= \frac{\cos(t)\,\sin(t)}{2} + \int_{\cos(t)}^1 \sqrt{1-s^2}\,ds \\
&= \frac{\sin(2t)}{4} + \int_t^0 \sqrt{1-\cos(r)^2}\,(-\sin(r))\,dr \\
&= \frac{\sin(2t)}{4} + \int_0^t \sin(r)^2\,dr \\
&= \frac{\sin(2t)}{4} + \int_0^t \frac{1}{1-\cos(2r)}2\,dr \\
&= \frac{\sin(2t)}{4} + \left(\frac{r}{2} - \frac{\sin(2r)}{4}\right)\Big|_{r=0}^t \\
&= \frac{\sin(2t)}{4} + \frac{t}{2} - \frac{\sin(2t)}{4} = \frac{t}{2}
\end{aligned}$$

gegeben, d.h. $\arcsin(y) = t$ mißt die doppelte Fläche des zugehörigen Kreissektors. Nach (2.1.2) ist die gleichzeitig die Länge des zugehörigen Bogens also der Name arcussinus voll gerechtfertigt, wohingegen das bei Mathematica gebräuchliche arcussinushyperbolicus ein Unsinn ist.

5.2.7 Bemerkung. Integration rationaler Funktionen.

Folgende rationale Ausdrücke können wir integrieren:

$$\int x^k\,dx = \frac{x^{k+1}}{k+1} + C \text{ für } k \neq -1$$
$$\int x^{-1}\,dx = \ln(x) + C$$
$$\int \frac{1}{1+x^2}\,dx = \arctan(x) + C$$
$$\int \frac{x}{1+x^2}\,dx = \frac{1}{2}\ln(1+x^2) + C = \ln(\sqrt{1+x^2}) + C$$

Weiters ist

$$\frac{1}{x^2-1} = \frac{1}{2}\Big(\frac{1}{x-1} - \frac{1}{x+1}\Big)$$

und somit

$$\begin{aligned}\int \frac{1}{x^2-1}\,dx &= \int \frac{1}{2}\Big(\frac{1}{x-1} - \frac{1}{x+1}\Big)\,dx = \frac{1}{2}\Big(\int \frac{1}{x-1}\,dx - \int \frac{1}{x+1}\,dx\Big)\\ &= \frac{1}{2}\Big(\ln(x-1) + C_1 - \ln(x+1) - C_2\Big) = \ln\left(\sqrt{\frac{x-1}{x+1}}\right) + C\end{aligned}$$

Diese Methode wollen wir nun auf allgemeine rationale Funktionen übertragen, d.h. auf Quotienten P/Q von Polynomen P und Q. Wir können natürlich zuerst Division mit Rest durchführen und damit ein Polynom abspalten, um eine rationale Funktion mit $\deg(P) < \deg(Q)$ zu erhalten.

Sei nun x_0 eine Nullstelle von Q. Dann können wir Q durch $(x - x_0)$ so oft wie möglich dividieren und erhalten $Q(x) = (x - x_0)^g Q_1(x)$ für eine natürliche Zahl $g \geq 1$ und ein Polynom Q_1 mit $Q_1(x_0) \neq 0$. Obiges Beispiel legt nun den Schluß nahe, daß wir von P/Q ein konstantes Vielfaches von $1/(x - x_0)^g$ abziehen können, um dadurch eine einfachere rationale Funktion (die x_0 nicht mehr als g-fache Nullstelle besitzt) zu erhalten. Wegen $\lim_{x\to x_0}(x-x_0)^g \frac{P(x)}{Q(x)} = \lim_{x\to x_0} \frac{P(x)}{Q_1(x)} = \frac{P(x_0)}{Q_1(x_0)}$, müßte dieser konstante Faktor durch $P(x_0)/Q_1(x_0)$ gegeben sein. Folglich betrachten wir die Differenz

$$\frac{P(x)}{Q_1(x)} - \frac{P(x_0)}{Q_1(x_0)} = \frac{P(x)Q_1(x_0) - P(x_0)Q_1(x)}{Q_1(x)Q_1(x_0)} = \frac{R(x)}{Q_1(x)}$$

mit $R(x) := \frac{P(x)Q_1(x_0) - P(x_0)Q_1(x)}{Q_1(x_0)}$. Offensichtlich ist $R(x_0) = 0$ und somit das Polynom R durch $(x - x_0)$ teilbar, d.h. $R(x) = (x - x_0) \cdot R_1(x)$ für ein Polynom R_1. Damit ist

$$\frac{P(x)}{Q(x)} = \frac{P(x)}{(x-x_0)^g Q_1(x)} = \frac{P(x_0)}{Q_1(x_0)} \cdot \frac{1}{(x-x_0)^g} + \frac{R_1(x)}{(x-x_0)^{g-1} \cdot Q_1(x)},$$

mit $\mathrm{grad}(R_1) < \mathrm{grad}(Q_1) + g - 1$ wegen $\lim_{x\to\infty} \frac{P(x)}{Q(x)} = 0 = \lim_{x\to\infty} \frac{1}{(x-x_0)^g}$. Induktion nach der Anzahl der Nullstellen (also dem Grad von Q) liefert somit:

5.2.8 Proposition. Partialbruchzerlegung im Komplexen.

Es seien P und Q Polynome mit komplexen Koeffizienten und mit $\mathrm{grad}(P) < \mathrm{grad}(Q)$. Weiters seien $x_1, \dots, x_n$ die verschiedenen Nullstellen von Q und $g_1, \dots, g_n$ ihre Vielfachheit. Dann ist

$$\frac{P(x)}{Q(x)} = \sum_{j=1}^{n} \sum_{g=1}^{g_j} \frac{p_{j,g}}{(x-x_j)^g}$$

mit gewissen eindeutig bestimmten Koeffizienten $p_{j,g} \in \mathbb{C}$. □

Wir können nun die Partialbruchzerlegung verwenden, um rationale Funktionen zu integrieren.

5.2.9 Beispiel des Integrals einer rationalen Funktion.
Es sei die rationale Funktion $f : x \mapsto \frac{1-2x+4x^2-x^4}{x-2x^2+x^3}$ gegeben. Division mit Rest liefert:

$$\frac{1-2x+4x^2-x^4}{x-2x^2+x^3} = -x-2+\frac{x^2+1}{x^3-2x^2+x}.$$

Die Nullstellen des Nenners sind 0 und 1 mit Vielfachheiten 1 und 2 also ist der Nenner $x^3-2x^2+x = x(x-1)^2$. Es muß somit nach (5.2.8) eine Darstellung der Form

$$\frac{x^2+1}{x^3-2x^2+x} = \frac{p_{1,1}}{(x-0)^1} + \frac{p_{2,1}}{(x-1)^1} + \frac{p_{2,2}}{(x-1)^2}$$

existieren. Ausmultiplizieren der Nenner liefert:

$$x^2+1 = p_{1,1}\,(x-1)^2 + p_{2,1}\,x(x-1) + p_{2,2}\,x \quad \forall x,$$

also

$$(p_{1,1}+p_{2,1}-1)\,x^2 + (-2p_{1,1}-p_{2,1}+p_{2,2})\,x + (p_{1,1}-1) = 0 \quad \forall x.$$

Es müssen somit alle Koeffizienten dieses Polynoms 0 sein, d.h.

$$\begin{aligned} p_{1,1}+p_{2,1} &= 1 \\ p_{2,2}-2p_{1,1}-p_{2,1} &= 0 \\ p_{1,1} &= 1 \end{aligned}$$

Die Lösung dieses linearen Gleichungssystems ist

$$p_{1,1}=1, \quad p_{2,1}=0 \text{ und } p_{2,2}=2,$$

d.h.

$$\frac{x^2+1}{x^3-2x^2+x} = \frac{1}{(x-0)^1} + \frac{0}{(x-1)^1} + \frac{2}{(x-1)^2} = \frac{1}{x} + \frac{2}{(x-1)^2}.$$

Eine andere Möglichkeit die Koeffizienten zu bestimmen, ist zuerst die $p_{i,g_i} = \lim_{x\to x_i} \frac{(x^2+1)(x-x_i)^{d_i}}{x^3-2x^2+x}$ zu bestimmen, also

$$\begin{aligned} p_{1,1} &= \lim_{x\to 0} \frac{x^2+1}{(x-1)^2} = 1 \\ p_{2,2} &= \lim_{x\to 1} \frac{x^2+1}{x} = 2 \end{aligned}$$

und dann mit dem Rest (der hier 0 ist) induktiv fortfahren

$$\frac{x^2+1}{x^3-2x^2+x} - \frac{1}{(x-0)^1} - \frac{2}{(x-1)^2} = \frac{x^2+1-(x-1)^2-2x}{x^3-2x^2+x} = \frac{0}{x-1}.$$

Somit ist

$$\begin{aligned} \int \frac{1-2x+4x^2-x^4}{x-2x^2+x^3}\,dx &= -\int x\,dx - \int 2\,dx + \int \frac{1}{x}\,dx + \int \frac{2}{(x-1)^2}\,dx \\ &= -\frac{x^2}{2} - 2\,x + \ln(x) - \frac{2}{x-1} + C. \end{aligned}$$

5.2.10 Proposition. Partialbruchzerlegung im Reellen.

Es seien P und Q Polynome mit reellen Koeffizienten und mit $\operatorname{grad}(P) < \operatorname{grad}(Q)$. *Weiters seien $x_1, \dots, x_k$ die verschiedenen reellen Nullstellen von Q und $g_1, \dots, g_k$ ihre Vielfachheit. Seien schließlich $a_1 \pm i\, b_1, \dots, a_j \pm i\, b_j$ die echt komplexen verschiedenen Nullstellen von Q und $h_1, \dots, h_j$ ihre Vielfachheit. Dann ist*

$$\frac{P(x)}{Q(x)} = \sum_{l=1}^{k} \sum_{g=1}^{g_l} \frac{p_{l,g}}{(x-x_l)^g} + \sum_{l=1}^{j} \sum_{g=1}^{h_l} \frac{r_{l,g} + s_{l,g} x}{(x^2 - 2a_l x + (a_l^2 + b_l^2))^g}$$

mit gewissen eindeutig bestimmten reellen Koeffizienten $p_{l,g}$, $r_{l,g}$ und $s_{l,g}$.

Beweis. Für jede komplexe Nullstelle $z = a + i\, b$ von Q ist auch die konjugierte Zahl $\overline{z} = a - i\, b$ Nullstelle und zwar mit der gleichen Vielfachheit g, denn aus $0 = Q(z) = \sum_j q_j z^j$ folgt durch Konjugieren $0 = \overline{Q(z)} = \sum_j q_j \overline{z}^j = Q(\overline{z})$.
Somit ist $Q(x) = ((x-z)(x-\overline{z}))^g\, Q_2(x)$ und wegen (5.2.8)

$$\frac{P(x)}{Q(x)} = \frac{\alpha}{(x-z)^g} + \frac{\beta}{(x-\overline{z})^g} + \frac{R(x)}{(x-z)^{g-1}(x-\overline{z})^{g-1} Q_2(x)}$$

wobei

$$\beta = \lim_{x \to \overline{z}} \frac{P(x)}{(x-z)^g Q_2(x)} = \lim_{y \to z} \frac{P(\overline{y})}{(\overline{y}-z)^g Q_2(\overline{y})} = \overline{\left(\lim_{y \to z} \frac{P(y)}{(y-\overline{z})^g Q_2(y)} \right)} = \overline{\alpha}.$$

Somit ist

$$\begin{aligned} \frac{\alpha}{(x-z)^g} + \frac{\beta}{(x-\overline{z})^g} &= \frac{\alpha(x-\overline{z})^g + \overline{\alpha}(x-z)^g}{((x-z)(x-\overline{z}))^g} \\ &= \frac{\sum_{k=0}^{g} \binom{g}{k} x^k 2\Re e(\alpha(-\overline{z})^{g-k})}{(x^2 - 2x\,\Re e(z) + |z|^2)^g} \end{aligned}$$

Sukzessive Division mit Rest durch $x^2 - 2x\,\Re e(z) + |z|^2$ liefert für den Zähler einen Ausdruck der Form

$$\sum_{j=1}^{g} (r_j + s_j\, x)\, (x^2 - 2x\,\Re e(z) + |z|^2)^{g-j}$$

und für den Bruch somit eine Summe der Form

$$\sum_{j=1}^{g} \frac{r_j + s_j\, x}{(x^2 - 2x\,\Re e(z) + |z|^2)^j}.$$

Induktion liefert nun das gewünschte Ergebnis. □

Beachte, daß eine Stammfunktion von $\frac{r+s\,x}{(x^2-2ax+(a^2+b^2))^d}$ wie folgt mittels Substitutionen $\frac{x-a}{b} = y$ und $y^2 + 1 = z$ bestimmt werden kann:

$$\begin{aligned} \int \frac{r+s\,x}{(x^2-2ax+(a^2+b^2))^d}\, dx &= \int \frac{r+s\,a+s\,b\,y}{(b^2(y^2+1))^d}\, b\, dy \\ &= \frac{r+sa}{b^{2d-1}} \int \frac{1}{(y^2+1)^d}\, dy + \frac{s}{2b^{2d-2}} \int \frac{2y}{(y^2+1)^d}\, dy \\ &= \frac{r+sa}{b^{2d-1}} \int \frac{1}{(y^2+1)^d}\, dy + \frac{s}{2b^{2d-2}} \int z^{-d}\, dz \end{aligned}$$

und weiters für $d > 0$ mittels partieller Integration

$$\begin{aligned} \int 1\, \frac{1}{(y^2+1)^d}\, dy &= y\, \frac{1}{(y^2+1)^d} - \int y\, (-d) \frac{2y}{(y^2+1)^{d+1}}\, dy = y\, \frac{1}{(y^2+1)^d} + 2d \int \frac{y^2+1-1}{(y^2+1)^{d+1}}\, dy \\ &= y\, \frac{1}{(y^2+1)^d} + 2d \int \frac{1}{(y^2+1)^d}\, dy - 2d \int \frac{1}{(y^2+1)^{d+1}}\, dy \end{aligned}$$

und somit erhalten wir nach Ersetzen von d durch $d-1$ die Rekursion

$$2(d-1)\int \frac{1}{(y^2+1)^d}\,dy = \frac{y}{(y^2+1)^{d-1}} + (2d-3)\int \frac{1}{(y^2+1)^{d-1}}\,dy.$$

5.3 Uneigentliche Integrale

5.3.1 Definition. Uneigentliche Integrale.
Wir wollen nun auch Integrale $\int_I f$ behandeln, bei denen entweder der Integrand f oder der Integrationsbereich I unbeschränkt ist. In diesem Fall spricht man von UNEIGENTLICHEN INTEGRALEN und definiert z.B. falls f bei b unbeschränkt aber auf jedem Teilintervall $[a,\beta]$ von $I := [a,b]$ mit $a<\beta<b$ D-integrierbar ist:

$$\int_a^b f := \lim_{\beta\to b-}\int_a^\beta f$$

Ist andererseits I unbeschränkt also z.B. $b=+\infty$ dann setzen wir

$$\int_a^{+\infty} f := \lim_{\beta\to+\infty}\int_a^\beta f$$

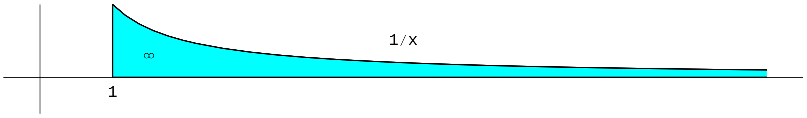
5.3.2 Beispiele uneigentlicher Integrale.

1. Es existiert das uneigentliche Integral $\int_0^1 \frac{1}{x}\,dx$ nicht (eigentlich), da

$$\lim_{t\to 0+}\int_t^1 \frac{1}{x}\,dx = \lim_{t\to 0+} \ln(1)-\ln(t) = +\infty.$$

2. Ebenso existiert $\int_1^{+\infty}\frac{1}{x}\,dx$ nicht (eigentlich), da

$$\lim_{t\to+\infty}\int_1^t \frac{1}{x}\,dx = \lim_{t\to+\infty} \ln(t)-\ln(1) = +\infty.$$

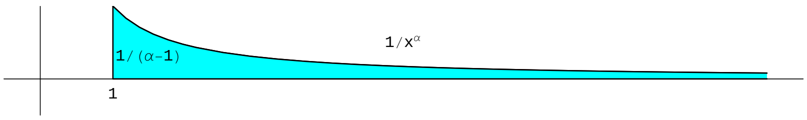


3. Hingegen existiert $\int_0^1 \frac{1}{x^\alpha}\,dx$ falls $0<\alpha<1$, denn

$$\lim_{t\to 0+}\int_t^1 \frac{1}{x^\alpha}\,dx = \lim_{t\to 0+}\frac{1}{1-\alpha} - \frac{1}{1-\alpha}t^{1-\alpha} = \frac{1}{1-\alpha}.$$

4. Ebenso existiert $\int_1^{+\infty}\frac{1}{x^\alpha}\,dx$ falls $1<\alpha$, denn

$$\lim_{t\to+\infty}\int_1^t \frac{1}{x^\alpha}\,dx = \lim_{t\to+\infty}\frac{1}{(1-\alpha)t^{\alpha-1}} - \frac{1}{1-\alpha} = \frac{1}{\alpha-1}.$$



5. Es ist

$$\int_0^1 \frac{1}{\sqrt{1-x^2}}\,dx = \lim_{t\to 1-}\int_0^t \frac{1}{\sqrt{1-x^2}}\,dx = \lim_{t\to 1-}\arcsin(t) = \arcsin(1) = \frac{\pi}{2}$$

$$\begin{aligned}\int_{-1}^1 \frac{1}{\sqrt{1-x^2}}\,dx &= \int_{-1}^0 \frac{1}{\sqrt{1-x^2}}\,dx + \int_0^1 \frac{1}{\sqrt{1-x^2}}\,dx \\ &= \int_1^0 \frac{1}{\sqrt{1-(-y)^2}}\,(-1)\,dy + \int_0^1 \frac{1}{\sqrt{1-x^2}}\,dx \\ &= 2\int_0^1 \frac{1}{\sqrt{1-x^2}}\,dx = 2\frac{\pi}{2} = \pi\end{aligned}$$

Allgemein ist (mittels Substitution $x = -y$)

$$\int_{-b}^0 f(x)\,dx = \int_b^0 f(-y)\,(-1)\,dy = \int_0^b f(-y)\,dy.$$

Für gerades f ist somit

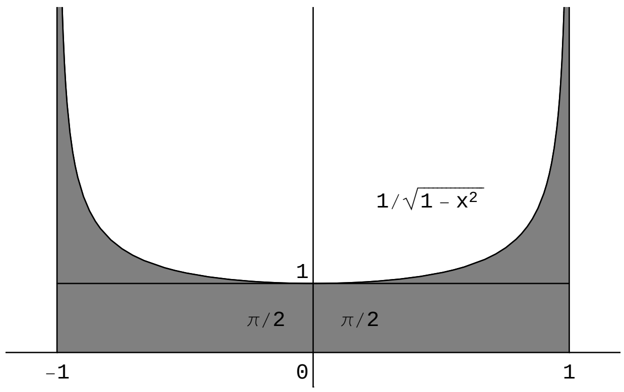
$$\int_{-b}^0 f(x)\,dx = \int_0^b f(x)\,dx,\ \text{also}\ \int_{-b}^b f(x)\,dx = \int_0^b f(x)\,dx + \int_0^b f(x)\,dx = 2\int_0^b f(x)\,dx$$

und für ungerades

$$\int_{-b}^0 f(x)\,dx = -\int_0^b f(x)\,dx,\ \text{also}\ \int_{-b}^b f(x)\,dx = -\int_0^b f(x)\,dx + \int_0^b f(x)\,dx = 0.$$

Obiges Integral kann durch die Substitutionen $x = 1/y$, $y = \cosh(t)$ und $s = e^t$ wie folgt umgeformt werden:

$$\begin{aligned}\int_0^1 \frac{1}{\sqrt{1-x^2}}\,dx &= \int_\infty^1 \frac{1}{\sqrt{1-(1/y)^2}}\,(-\frac{1}{y^2}\,dy) = \int_1^\infty \frac{1}{y\sqrt{y^2-1}}\,dy \\ &= \int_0^\infty \frac{1}{\cosh(t)\,\sinh(t)}\,\sinh(t)\,dt = \int_0^\infty \frac{1}{\cosh(t)}\,dt \\ &= \int_0^\infty \frac{2}{e^t+e^{-t}}\,dt = \int_1^\infty \frac{2}{s+1/s}\,\frac{1}{s}\,ds = 2\int_1^\infty \frac{1}{1+s^2}\,ds.\end{aligned}$$



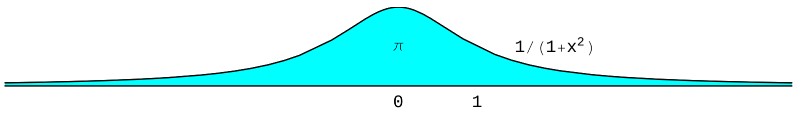
Beachte, daß Integranden mit $\sqrt{x^2\pm 1}$ wegen $\cosh(y)^2 - \sinh(y)^2 = 1$ eine Substitution der Form $x = \cosh(y)$ bzw. $x = \sinh(y)$ nahelegen. Wohingegen Integranden mit $\sqrt{1-x^2}$ wegen $\cos(y)^2 + \sin(y)^2 = 1$ eine Substitution der Form $x = \cos(y)$ (oder auch $x = \sin(y)$) nahelegen.

Wir können somit $\int_1^\infty \frac{1}{y\sqrt{y^2-1}}\,dy$ mittels partieller Integration weiter umformen, indem wir $\frac{1}{y} = \ln'(y)$ oder $\frac{1}{\sqrt{y^2-1}} = \operatorname{Arcosh}'(y)$ beachten:

$$\int \frac{1}{y\sqrt{y^2-1}}\,dy = \frac{\ln(y)}{\sqrt{y^2-1}} - \int \frac{-y\ln(y)}{(y^2-1)^{3/2}}\,dy$$
$$\int \frac{1}{y\sqrt{y^2-1}}\,dy = \frac{\operatorname{Arcosh}(y)}{y} - \int \frac{-\operatorname{Arcosh}(y)}{y^2}\,dy.$$

6. Es ist

$$\begin{aligned}\int_{-\infty}^{+\infty} \frac{dx}{1+x^2}\,dx &= \lim_{s\to+\infty}\lim_{t\to-\infty}\int_t^s \frac{1}{1+x^2}\,dx\\ &= \lim_{s\to+\infty}\lim_{t\to-\infty} \arctan(s)-\arctan(t)\\ &= \frac{\pi}{2} - \frac{-\pi}{2} = \pi.\end{aligned}$$



7. Was stimmt nicht mit folgender Rechnung?

$$\int_{-1}^1 \frac{1}{\sqrt[3]{x^4}}\,dx = \int_{-1}^1 x^{-\frac{4}{3}}\,dx = \frac{x^{-\frac{4}{3}+1}}{-\frac{4}{3}+1}\Bigg|_{x=-1}^1 = \frac{-3}{\sqrt[3]{x}}\Bigg|_{x=-1}^1 = -3-3 = -6.$$

Irgendetwas muß faul sein, denn der Integrand ist positiv und die untere Grenze ist die kleinere, also sollte das Integral positiv sein.

Der Fehler ist, daß wir beim 2-ten Gleichheitszeichen den Hauptsatz der Analysis angewandt haben, die Funktion $x \mapsto 1/\sqrt[3]{x}$ aber im Punkt 0 nicht definiert ist, und somit nicht auf ganz $[-1,1]$ eine Stammfunktion darstellt. Es besteht auch keine Möglichkeit diese Funktion zu einer stetigen oder differenzierbaren Funktion auf ganz $[-1,1]$ zu erweitern.

Dieses bestimmte Integral ist ein uneigentliches Integral, denn $x \mapsto \frac{1}{\sqrt[3]{x^4}}$ ist unbeschränkt bei 0, also ist

$$\begin{aligned}\int_{-1}^1 \frac{1}{\sqrt[3]{x^4}}\,dx &:= \int_{-1}^0 \frac{1}{\sqrt[3]{x^4}}\,dx + \int_0^1 \frac{1}{\sqrt[3]{x^4}}\,dx\\ &= \lim_{b\to 0-}\int_{-1}^b \frac{1}{\sqrt[3]{x^4}}\,dx + \lim_{a\to 0+}\int_a^1 \frac{1}{\sqrt[3]{x^4}}\,dx = +\infty+\infty = +\infty.\end{aligned}$$

5.3.3b Proposition. Vergleichstest für uneigentliche Integrale.
Es seien $f, g : [a,b] \to \mathbb{R}$ Funktionen mit $0 \le f \le g$ welche auf $[a,\beta]$ D-integrierbar sind für alle $a < \beta < b$. Falls das uneigentliche D-Integral $\int_a^b g$ existiert, so auch das uneigentliche D-Integral $\int_a^b f$ und es gilt $\int_a^b f \le \int_a^b g$.

Vergleiche dies mit (2.3.9) und (2.5.9) sowie Aufgabe (3.18).

Beweis. Es sei $F(\beta) := \int_a^\beta f$ und $G(\beta) := \int_a^\beta g$. Wegen $g \ge f \ge 0$ sind F und G monoton wachsend nach (5.1.10) und $F \le G$. Also ist $F(\beta) \le G(\beta) \le \lim_{\beta\to b-} G(\beta) = \int_a^b g < \infty$ und somit existiert $\int_a^b f = \lim_{\beta\to b-} F(\beta) \le \int_a^b g$. □

5.3.3a Allgemeine Konvergenz via Cauchy-Bedingung Die Cauchy-Bedingung für Folgen können wir auch auf allgemeinere Limiten und insbesonders auf uneigentliche Integrale übertragen:

Proposition.
Es seien X und Y metrische Räume, Y sei vollständig, $x_\infty \in X$ sei nicht isoliert in X und $f : X \setminus \{x_\infty\} \to Y$ eine Funktion.
Dann existiert $\lim_{x\to x_\infty} f(x)$ genau dann, wenn die Cauchy-Bedingung

$$d(f(x), f(x')) \to 0 \text{ für } x, x' \to x_\infty$$

erfüllt ist, d.h.

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall x, x' \in X \setminus x_\infty : d(x, x_\infty) < \delta, d(x', x_\infty) < \delta \Rightarrow d(f(x), f(x')) < \varepsilon.$$

Vergleiche dies mit (2.4.6) und (2.5.6).

Beweis. ($\Rightarrow$) 1. Möglichkeit dies zu sehen: Durch $f(x_\infty) := \lim_{x\to x_\infty} f(x)$ wird f zu einer in x_∞ stetigen Funktion erweitert. Somit ist $f \times f : X \times X \to Y \times Y$ definiert durch $(f \times f)(x_1, x_2) := (f(x_1), f(x_2))$ stetig und damit ist auch die Zusammensetzung $d \circ (f \times f) : (x, x') \mapsto d(f(x), f(x'))$ stetig, also ist

$$\begin{aligned}\lim_{x,x'\to x_\infty} d(f(x), f(x')) &= \lim_{(x,x')\to(x_\infty,x_\infty)} (d \circ (f \times f))(x, x') \\ &= (d \circ (f \times f))(x_\infty, x_\infty) = d(f(x_\infty), f(x_\infty)) = 0.\end{aligned}$$

2. Möglichkeit: Es sei $\eta := \lim_{x\to\times_\infty} f(x)$, dann gilt: $d(f(x), f(x')) \le d(f(x), \eta) + d(\eta, f(x')) \to 0$ für $x, x' \to x_\infty$.

($\Leftarrow$) Für jede Folge $(x_n)_n$ in $X \setminus \{x_\infty\}$ ist also $f(x_n)$ eine Cauchy-Folge in Y und konvergiert somit. Wir zeigen nun, daß dieser Grenzwert $\lim_{n\to\infty} f(x_n)$ nicht von der gewählten Folge abhängt. Sei dazu $(x'_n)_n$ eine weitere gegen x_∞ konvergente Folge. Dann konvergiert auch die Folge $(x''_n)_n = (x_0, x'_0, x_1, x'_1, \dots)$ gegen x_∞, also existiert $\eta := \lim_{n\to\infty} f(x''_n)$. Somit konvergieren aber auch die Teilfolgen $(f(x_n))_n$ und $(f(x'_n))_n$ beide gegen η, haben also den gleichen Gernzwert. Mit (3.1.3) folgt nun, daß $\lim_{x\to x_\infty} f(x)$ existiert und dieser Grenzwert ist. □

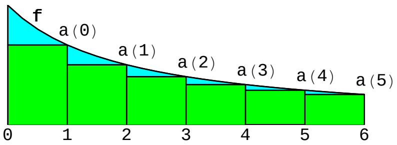
5.3.3 Folgerung. Cauchy-Kriterium für uneigentliche Integrale.
Es existiere $\int_a^\beta f$ für alle $a < \beta < b$. Dann existiert das uneigentliche Integral $\int_a^b f$ genau dann wenn $\int_{\beta_1}^{\beta_2} f \to 0$ für $\beta_1, \beta_2 \to b$.

Beweis. Die Cauchy-Bedingung für das uneigentliche Integral $\int_a^b f = \lim_{\beta\to b-} \int_a^\beta f$ ist:

$$\int_a^{\beta'} f - \int_a^\beta f = \int_\beta^{\beta'} f \to 0 \text{ für } \beta, \beta' \to b,$$

also folgt das Resultat aus (5.3.3a). □

5.3.4 Proposition. Integraltest für Reihenkonvergenz.
Es sei $f : [0, +\infty) \to [0, +\infty)$ uneigentlich D-integrierbar und $|a_n| \le f(x)$ für $n \le x \le n+1$. Dann konvergiert $\sum_k a_k$ absolut.



Beweis. Es ist $\sum_{k=0}^n |a_k| \le \int_0^{n+1} f \le \int_0^{+\infty} f < +\infty$. □

Beispiel.
Wir haben in (5.3.2) gesehen, daß das uneigentliche Integral $\int_1^\infty \frac{1}{x^\alpha}\,dx$ für $\alpha > 1$ existiert. Also ist

$\sum_k \frac{1}{k^\alpha}$ für diese α konvergent. Vergleich dies mit dem Beweis aus (2.5.14.1) unter Verwendung des Cauchy'sche Verdichtungssatzes (2.5.13).

5.3.5 Definition. Euler'sche Gammafunktion.
Unter der EULER'SCHE GAMMAFUNKTION versteht man die Funktion Γ welche für $x > 0$ definiert ist durch

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

Da die obere Grenze $+\infty$ ist, liegt hier ein uneigentliches Integral vor. Für $0 < x < 1$ ist $\lim_{t\to 0+} e^{-t}t^{x-1} = +\infty$, also ist der Integrand bei 0 unbeschränkt und das Integral auch aus diesem Grund ein uneigentliches.

Versuchen wir die Existenz von $\int_0^1 e^{-t}t^{x-1}\, dt = \lim_{a\to 0+} \int_a^1 e^{-t}t^{x-1}\, dt$ (für $0 < x < 1$) nachzuweisen. Es ist $0 \le e^{-t} \cdot t^{x-1} \le t^{x-1}$ für $0 < t \le 1$, also nach (5.3.3b)

$$\int_0^1 e^{-t}t^{x-1}\, dt \le \int_0^1 t^{x-1}\, dt = \lim_{a\to 0+} \frac{t^x}{x}\Big|_{t=a}^1 = \frac{1}{x} < \infty.$$

Nun zur Existenz von $\int_1^\infty e^{-t}t^{x-1}\, dt$. Um $\lim_{t\to\infty} \frac{t^{x-1}}{e^{t/2}}$ (für $x - 1 > 0$) zu bestimmen betrachten wir die $1/(x-1)$-Potenz und wenden darauf L'Hospital an:

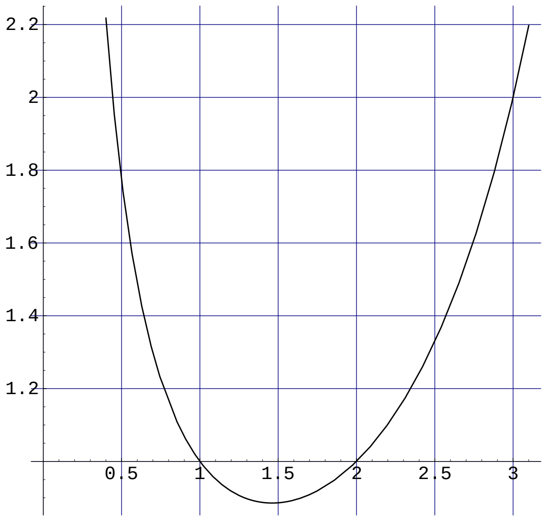
$$\lim_{t\to\infty} \frac{t}{e^{t/(2(x-1))}} = \lim_{t\to\infty} \frac{2(x-1)}{e^{t/(2(x-1))}} = 0.$$

Da $t \mapsto e^{-t/2}t^{x-1}$ stetig ist folgt nun die Existenz einer Schranke $C > 0$ mit $e^{-t/2}t^{x-1} \le C$ für alle $t \ge 1$ und somit nach (5.3.3c)

$$\begin{aligned}\int_1^{+\infty} e^{-t}t^{x-1}\, dt &= \int_1^{+\infty} e^{-t/2} \underbrace{e^{-t/2}t^{x-1}}_{\le C}\, dt \le C \int_1^{+\infty} e^{-t/2}\, dt \\ &= \lim_{b\to+\infty} C(-2\, e^{-t/2})|_{t=1}^b = 2\, C\, \frac{1}{\sqrt{e}} < \infty.\end{aligned}$$

Zusammengefaßt existiert also

$$\Gamma(x) = \int_0^1 e^{-t}\, t^{x-1}\, dt + \int_1^\infty e^{-t}\, t^{x-1}\, dt$$



5.3.6 Lemma. Gammafunktion als Fortsetzung von Faktorielle.
Es gilt folgende Rekursionsformel für die Euler'sche Gamma-Funktion:

$$\Gamma(x+1) = x\,\Gamma(x).$$

Insbesonders ist $\Gamma(n+1) := n!$ *für* $n \in \mathbb{N}$, *d.h.* Γ *definiert eine Erweiterung der Faktorielle-Funktion von* $\mathbb{N}$ *auf* $\{t \in \mathbb{R} : t > -1\}$.

Beweis. Es ist $\Gamma(1) = \int_0^\infty e^{-t}\,t^0\,dt = \lim_{\beta\to+\infty} -e^{-t}|_{t=0}^{\beta} = 1$. Die Rekursionsvorschrift folgt aus:

$$\begin{aligned}
\Gamma(x+1) &= \int_0^\infty e^{-t}\,t^x\,dt \\
&= \lim_{a\to 0+,b\to\infty} \int_a^b e^{-t}\,t^x\,dt \\
&= \lim_{a\to 0+,b\to\infty} \Big(-e^{-t}\,t^x\Big|_{t=a}^{b} - \int_a^b (-e^{-t})\,x\,t^{x-1}\,dt\Big) \\
&= 0 + \int_0^{+\infty} e^{-t}\,x\,t^{x-1}\,dt = x\,\Gamma(x). \quad \square
\end{aligned}$$

5.3.7 Bemerkung. $\Gamma(1/2)$**.**
Man kann zeigen, daß $\Gamma(x)\cdot\Gamma(1-x) = \frac{\pi}{\sin(\pi x)}$ für $0 < x \notin \mathbb{N}$ (siehe [13, S.198]). Insbesonders ist $\Gamma(1/2)^2 = \frac{\pi}{\sin(\pi/2)} = \pi$, also $\Gamma(1/2) = \sqrt{\pi}$.

Eine direktere Methode $\Gamma(1/2)$ zu berechnen besteht darin Substitution anzuwenden:

$$\Gamma(1/2) = \int_0^\infty e^{-t}\,t^{-1/2}\,dt = \int_0^\infty e^{-s^2}\frac{1}{s}\,2s\,ds = \int_{-\infty}^{+\infty} e^{-s^2}\,ds.$$

Wir werden in der Analysis 3 zeigen, daß das letzte Integral $\sqrt{\pi}$ ist.

5.3.8 Bemerkung. Integration von Kurven im Mehrdimensionalen.
Der Hauptsatz (5.2.2) läßt sich wie folgt auf's mehrdimensionale übertragen. Wenn $f = (f_1, \dots, f_m) : \mathbb{R} \supseteq I \to F := \mathbb{R}^m$ ist, dann können wir unter dem Integral $\int_I f$ in Analogie zu (2.3.4) und (4.1.18) den Vektor

$$\int_I f := \Big(\int_I f_1, \dots, \int_I f_m\Big) \in F$$

verstehen. Für stetig differenzierbares $f : I = [a,b] \to F$ gilt somit $f(b) - f(a) = \int_a^b f'(t)\,dt$ in F. Wir werden dies im Abschnitt (5.5) nochmals ausführlicher angehen.

5.4 Banach-Räume

Wir wollen uns nun von der 1-Dimensionalität unserer bisherigen Betrachtungen entgültig befreien und in höhere Dimensionen aufsteigen, d.h. Analysis für Kurven $f : \mathbb{R} \to \mathbb{R}^m$, für Funktionen $f : \mathbb{R}^n \to \mathbb{R}$ oder gleich für Abbildungen $f : \mathbb{R}^n \to \mathbb{R}^m$ treiben.

5.5.1 Vektorräume von Funktionen.
Wir haben bereits an einigen Stellen gesehen, daß für das Verständnis von niedrig dimensionalen Problemen – ja sogar von 1-dimensionalen Problemen – als Zwischenschritt (unendlich dimensionale) Räume von Funktionen hilfreich sein können:

1. In (2.1.2) haben wir einen Kreis(bogen) durch Polygonzüge approximiert. Um das exakt zu machen (vor allem was die Länge betrifft) brauchen wir einen Raum von Kurven.
2. Für die Stetigkeit von $f : \mathbb{R}^2 \to \mathbb{R}$ bei 0 benötigen wir die Konvergenz von $f(r\,x) \to f(0)$ für $r \to 0+$ gleichmäßig bzgl. $|x| = 1$, also von der Familie von Funktionen $f_r : y \mapsto f(r\,y)$ gegen die konstante Funktion $y \mapsto f(0)$, siehe (3.2.8).
3. Beim Vertauschen von Limiten benötigten wir in (3.2.16) die Gleichmäßigkeit eines der Limiten, d.h. die Konvergenz im Raum $B(Y, \mathbb{R})$ aus (4.2.8).
4. Die Tangente einer Kurve ist der (gleichmäßige) Grenzwert der Folge der gezoomten Kurven, siehe (4.1.2).
5. Konvergente Potenzreihen sind (gleichmäßige) Grenzwerte der Folge der durch die Partialsummen gegebenen Polynome, und als solche stetig, gliedweise differenzierbar und integrierbar, siehe (4.2.5), (4.2.14) und (4.2.16).
6. Die Stetigkeit des Integrals aus (5.1.12) ist eine Aussage über $\int : R([a, b], \mathbb{R}) \to \mathbb{R}$ bzgl. der gleichmäßigen Konvergenz am Raum $R(I, \mathbb{R})$ aus (5.1.7) aller reell-wertiger D-integrierbarer Funktionen am kompakten Intervall I.

Es sind uns für $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ bereits folgende Räume von Funktionen (kurz: Funktionenräume) untergekommen:

1. Unser Hauptbeispiel ist der Raum $\mathbb{K}^n$ aller endlichen Folgen $(x_0, \dots, x_{n-1})$ mit n Komponenten $x_0, \dots, x_{n-1} \in \mathbb{K}$, also aller Funktionen $x : \{0, \dots, n-1\} \to \mathbb{K}$.
2. Der Raum $B(X, \mathbb{K}) := \big\{f : X \to \mathbb{K} : f(X) \subseteq \mathbb{K} \text{ ist beschränkt}\big\}$ aller beschränkten Funktionen $X \to \mathbb{K}$ auf einer Menge X.
3. Oder allgemeiner der Raum $B(X, \mathbb{K}^n) := \big\{f : X \to \mathbb{K}^n : f(X) \subseteq \mathbb{K}^n \text{ ist beschränkt}\big\}$ aller beschränkten Abbildungen $X \to \mathbb{K}^n$ auf einer Menge X.
4. Speziell der Raum $\ell^\infty := B(\mathbb{N}, \mathbb{K})$ aller beschränkten Folgen in $\mathbb{K}$.
5. Der Raum $C(X, \mathbb{K})$ der stetigen Funktionen $X \to \mathbb{K}$ auf einem metrischen Raum X.
6. Oder allgemeiner der Raum $C(X, \mathbb{K}^n)$ der stetigen Abbildungen $X \to \mathbb{K}^n$ auf einem metrischen Raum X.
7. Der Raum $R([a, b], \mathbb{R})$ der D-integrierbaren Funktionen auf $[a, b]$.
8. Speziell der Raum $c := \{x \in \mathbb{K}^\mathbb{N} : \exists \lim_{n\to\infty} x_n\}$ aller konvergenten Folgen in $\mathbb{K}$, siehe (2.3.6).
9. Der Raum $\ell^1 := \{x \in \mathbb{K}^\mathbb{N} : \|x\|_1 := \sum_{n=0}^\infty |x_n| < \infty\}$ aller absolut summierbaren Folgen (bzw. absolut konvergenten Reihen).
10. Der Raum $\ell^2 := \{x \in \mathbb{K}^\mathbb{N} : (\|x\|_2)^2 := \sum_{n=0}^\infty |x_n|^2 < \infty\}$ aller quadratisch summierbaren Folgen(bzw. quadratisch konvergenten Reihen).

All diese Räume sind Vektorräume über $\mathbb{K}$.

Um Analysis zu treiben benötigen wir ungleich der linearen Algebra eine Distanzfunktion auf den zu untersuchenden (Funktionen-)Räumen E, damit wir von konvergenten Folgen (siehe (2.3.2)), deren Limiten und von stetigen Abbildungen (siehe (3.1.2)) sprechen können. Für Reihen in E benötigten wir zusätzlich eine Addition in E und eine Verträglichkeit der Metrik mit der Addition. Und für das Differenzieren (siehe (4.1.1)) und Integrieren (siehe (5.1.1)) von Abbildungen $f : \mathbb{R} \to E$ benötigen wir weiters eine Skalarmultiplikation um Ausdrücke der Form

$$f'(0) = \lim_{t\to 0} \frac{f(t) - f(0)}{t} \quad \text{und} \quad U(f, Z) = \sum_{I\in Z} \sup(f(I)) \cdot |I|$$

betrachten zu können.

Die Metrik sollte natürlich auch verträglich mit diesen Operationen sein. Was könnten wir dabei vernünftigerweise als Verträglichkeitsbedigungen an die Metrik $d : E \times E \to \mathbb{R}$ eines Vektorraums verlangen? Nun sie sollte wohl Translations-invariant sein, d.h.

$$d(x + z, y + z) = d(x, y) \text{ für alle } x, y, z \in E$$

erfüllen. Wenn wir speziell $z := -y$ setzen, so wäre $d(x, y) = d(x - y, 0) =: \|x - y\|$, also die Funktion $d : X\times X \to \mathbb{R}$ (mit zwei Variablen) durch die Funktion $\|\_\| : X \to \mathbb{R}$ (mit einer Variable) beschreibbar. Positivität der Metrik übersetzt sich (via $z := x - y$) in $\|z\| = 0 \Leftrightarrow z = 0$ und die Dreiecksungleichung wird (via $a := x - y$ und $b = y - z$) zu $\|a + b\| \leq \|a\| + \|b\|$. Die Symmetrie wird zu $\|z\| = \| - z\|$. Allgemeiner sollte die Verträglichkeit mit der Skalarmultiplikation wohl durch $\|\lambda\, x\| = |\lambda| \cdot \|x\|$ beschrieben werden. Wir geben also folgende

5.5.2 Definition. Norm und normierter Raum.
Eine Funktion $\|\_\| : E \to \mathbb{R}$ auf einem Vektorraum E über den Körper $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ heißt Norm, wenn die folgenden 3 Bedingungen erfüllt sind:

Definitheit: $\|f\| = 0 \Leftrightarrow f = 0$.

Positive Homogenität: $\|\lambda\, f\| = |\lambda| \cdot \|f\|$ für alle $\lambda \in \mathbb{K}$.

Dreiecksungleichung: $\|f + g\| \leq \|f\| + \|g\|$.

Vektorräume E die mit einer Norm versehen wurden heißen NORMIERTE RÄUME. Aus der Norm $\|\_\|$ erhalten wir offensichtlich eine (translationsinvariante) Metrik $d : E \times E \to \mathbb{R}$ welche durch $d(x, y) := \|x - y\|$ gegeben ist.

Es wäre natürlich naheliegend in Analogie zum Betrag $|x|$ für $x \in \mathbb{R}$, für $x \in \mathbb{C}$ und für $x \in \mathbb{R}^n$ aus (2.2.2) auch $|f|$ für Vektoren $f \in E$ also insbesonders für Funktionen f aus einem Funktionenraum E zu schreiben. Da üblicherweise aber $|f|$ die Funktion $x \mapsto |f(x)|$ (also die Zusammensetzung von Betragsfunktion und f) und nicht eine Zahl bezeichnet hat sich zur Unterscheidung die Bezeichnungsweise $\|f\|$ mit Doppelstrichen für Normen von f durchgesetzt. Diese sollte natürlich nicht als $\big|(|f|)\big|$ gelesen werden.

Für endliches X haben wir in Aufgabe (2.4) (siehe auch (2.2.2)) gezeigt, daß $x \mapsto \|x\|_1$ eine Norm ist und in (2.2.4), daß $x \mapsto \|x\|_2$ eine ist. Für $X = \mathbb{N}$ also den Raum ℓ^1 folgt dies nun, denn

$$\begin{aligned}\|x + y\|_1 &= \lim_{n\to\infty} \sum_{k=0}^{n} |x_k + y_k| = \sup\Big\{\sum_{k=0}^{n} |x_k + y_k| : n\Big\}\\ &\leq \sup\Big\{\sum_{k=0}^{n} |x_k| + \sum_{k=0}^{n} |y_k| : n\Big\} \leq \|x\|_1 + \|y\|_1,\end{aligned}$$

und ebenso für ℓ^2

$$\sum_{k=0}^{n} |x_k + y_k|^2 \le \left(\sqrt{\sum_{k=0}^{n} |x_k|^2} + \sqrt{\sum_{k=0}^{n} |y_k|^2} \right)^2 \le (\|x\|_2 + \|y\|_2)^2$$

$$\Rightarrow \quad \|x+y\|_2^2 = \sup\Big\{ \sum_{k=0}^{n} |x_k + y_k|^2 : n \in \mathbb{N} \Big\} \le (\|x\|_2 + \|y\|_2)^2.$$

In (2.2.2) haben wir gezeigt, daß $\|\_\|_\infty$ eine Norm ist, und zwar auf $B(X, \mathbb{R})$.

5.5.3 Definition. Banach-Raum.
Für metrische Räume war uns die Vollständigkeit sehr wichtig, d.h. die Gültigkeit des Cauchy-Kriteriums (2.4.6) für Konvergenz, siehe auch (1.6.5), (2.3.9), (3.4.1), (3.4.17) und (4.2.9).

Ein normierter Raum, für welchen die durch die Norm in (5.5.2) definierte Metrik $d : (f, g) \mapsto \|f - g\|$ vollständig ist, heißt BANACH-RAUM und wir wollen nun Analysis in Banach-Räumen treiben. Dies inkludiert insbesonders den $\mathbb{R}^n$, an welchen wir primär interessiert sind, ist aber auch wie oben angeführt für Funktionenräume von großen Interesse.

Jene LeserInnen/HöhrerInnen denen dieser Schritt vom 1-dimensionalen auf das unendlich-dimensionale (für den Moment) zu groß ist und Banach-Räume zu abstrakt scheinen, die mögen in all den folgenden Resultaten die allgemeinen Banach-Räume E, F, etc. durch endlich dimensionale, also (nach Wahl einer Basis) durch $\mathbb{R}^n$ (für $n \in \mathbb{N}$), ersetzen. In der Tat sieht man die meisten Konzepte bereits am $\mathbb{R}^2$ und für unsere Untersuchungen ist der wesentliche Schritt jener vom 1- auf das 2-dimensionale, wohingegen bei jenem vom 2- auf das ∞-dimensionale kaum mehr etwas Neues hinzu kommt.

5.5.4 Beispiele von Banach-Räumen.
In (2.4.6) haben wir gezeigt, daß $\mathbb{R}^n$ vollständig ist. In (4.2.9) haben wir gezeigt, daß $B(X, \mathbb{R}^n)$ vollständig ist: Sei nämlich $(f_n)_n$ eine Cauchy-Folge in $C(X, \mathbb{R}^n) \subseteq B(X, \mathbb{R}^n)$ (nach (3.3.5) und (3.3.4)). Also konvergiert f_n gleichmäßig gegen eine Funktion $f_\infty \in B(X, \mathbb{R}^n)$, da letzterer Raum vollständig ist nach (4.2.9). Nach (4.2.8) ist dann aber $f_\infty \in C(X, \mathbb{R}^n)$, also konvergiert $f_n \to f_\infty$ in $C(X, \mathbb{R}^n)$. Ganz analog zeigt man, daß $R([a, b], \mathbb{R}) \subseteq B([a, b], \mathbb{R})$ vollständig ist, wenn man $\Delta(f_\infty) \subseteq \bigcup_{n \in \mathbb{N}} \Delta(f_n)$ beachtet.

Und mittels (4.2.8) folgt, daß $C(X, \mathbb{R}^n)$ für kompaktes X vollständig ist. Somit gilt dies auch für die Spezialfälle $\ell^\infty = B(\mathbb{N}, \mathbb{R})$ und $c = C(\mathbb{N}_\infty, \mathbb{R})$.

Nun zu ℓ^1. Sei $x = (x^n)_n$ eine Cauchy-Folge in ℓ^1.

Achtung: Die Glieder x^n dieser Folge x sind Element $x^n \in \ell^1$, also seinerseits Folgen $x^n = (x^n_k)_{k \in \mathbb{N}}$ reeller Zahlen $x^n_k \in \mathbb{R}$. Um die beiden Indizes n und k nicht nebeneinander als $x_{n,k}$ schreiben zu müssen, haben wir n als oberen Index geschrieben, dieser bedeutet also keine Potenzierung. Eine andere Möglichkeit wäre auch gewesen x^n_k als $x_n(k)$ zu schreiben, und damit zum Ausdruck zu bringen, daß die Elemente von ℓ^1 auch als Funktionen $\mathbb{N} \to \mathbb{R}$ aufgefaßt werden können. Natürlich wäre ebenso $x(n, k)$ als Schreibweise möglich. Jeder möge hier seine eigenen Vorlieben entwickeln, für alle diese Schreibweisen gibt es gute Argumente und auch Gegen-Argumente.

Koordinate:	0.	1.	2.	3.	…	k.	…
$x^0 \in \ell^1$	x^0_0	x^0_1	x^0_2	x^0_3	…	x^0_k	…
$x^1 \in \ell^1$	x^1_0	x^1_1	x^1_2	x^1_3	…	x^1_k	…
$x^2 \in \ell^1$	x^2_0	x^2_1	x^2_2	x^2_3	…	x^2_k	…
$x^3 \in \ell^1$	x^3_0	x^3_1	x^3_2	x^3_3	…	x^3_k	…
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	
$x^n \in \ell^1$	x^n_0	x^n_1	x^n_2	x^n_3	…	x^n_k	…
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	
$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$		$\downarrow$	
$x^\infty \in \ell^1$	x^∞_0	x^∞_1	x^∞_2	x^∞_3	…	x^∞_k	…

Für jede Koordinate mit Index $k \in \mathbb{N}$ ist $(x_k^n)_n$ eine Cauchy-Folge in $\mathbb{R}$ (wegen $|x_k^n| \le \|x^n\|_1$) und somit konvergent. Sei $x_k^\infty := \lim_{n\to\infty} x_k^n$. Weiters ist $\|x^n\|_1$ eine Cauchy-Folge (wegen $\big|\,\|x^n\|_1 - \|x^m\|_1\,\big| \le \|x^n - x^m\|_1$) also beschränkt. Damit ist

$$\sum_{k=0}^n |x_k^\infty| \le \underbrace{\sum_{k=0}^n |x_k^\infty - x_k^m|}_{\le 1} + \underbrace{\sum_{k=0}^n |x_k^m|}_{\le \|x^m\|_1} \le 1 + \sup\{\|x^j\|_1 : j \in \mathbb{N}\}$$

für geeignet gewähltes m in Abhängigkeit von allen $k \in \{0,\dots,n\}$. Somit ist $\|x^\infty\|_1 < \infty$, also $x^\infty \in \ell^1$.

Weiters ist

$$\sum_{k=0}^n |x_k^\infty - x_k^j| \le \underbrace{\sum_{k=0}^n |x_k^\infty - x_k^m|}_{<\varepsilon} + \underbrace{\sum_{k=0}^n |x_k^m - x_k^j|}_{\le \|x^m - x^j\|_1} < \varepsilon + \varepsilon$$

für geeignet gewähltes m in Abhängigkeit von allen $k \in \{0,\dots,n\}$, also konvergiert $x^m \to x^\infty$ in ℓ^1.

Ganz analog kann man auch bei ℓ^2 vorgehen.

5.5.5 Lemma. Charakterisierung von Vollständigkeit.
Für normierte Räume E und jedes überall positive $\lambda \in \ell^1$ (d.h. $\lambda = (\lambda_n)_{n\in\mathbb{N}}$ mit $0 < \lambda_n \in \mathbb{R}$ und $\sum_{n=0}^\infty \lambda_n < \infty$) sind äquivalent

1. *Der Raum E ist vollständig;*

$\Leftrightarrow$ 2. *Jede absolut-konvergente Reihe in E konvergiert in E;*

$\Leftrightarrow$ 3. *Für jede beschränkte Folge b_n in E konvergiert die Reihe $\sum_n \lambda_n b_n$ in E;*

$\Leftrightarrow$ 4. *Jede Cauchy-Folge in E besitzt eine in E konvergente Teilfolge.*

Man/Frau rufe sich bei dieser Gelegenheit nochmals (2.5.8) und (4.2.10) in Erinnerung.

Eine Reihe $\sum_n x_n$ heißt ABSOLUT-KONVERGENT falls die Reihe $\sum_n \|x_n\|$ in $\mathbb{R}$ (absolut) konvergiert.

Beweis. ($1 \Rightarrow 2$) Es sei $\sum_n x_n$ absolut-konvergent, dann bilden die Partialsummen von $\sum_n x_n$ eine Cauchy-Folge, denn $\|\sum_{j=n}^{n+p} x_j\| \le \sum_{j=n}^{n+p} \|x_j\|$, also konvergiert $\sum_n x_n$ nach (1).

($2 \Rightarrow 3$) Es sei die Folge (b_n) beschränkt und (λ_n) absolut-summierbar. Dann ist $\sum_n \lambda_n b_n$ absolut-summierbar, denn $\sum_n \|\lambda_n b_n\| \le \|\lambda\|_1 \cdot \sup\{\|b_n\| : n \in \mathbb{N}\}$. Also konvergiert diese Reihe nach (2).

($3 \Rightarrow 4$) Es sei (x_i) Cauchy und λ absolut-summierbar. Dann gilt:

$$\begin{aligned}
&\forall k\ \exists n_k\ \forall i,j \ge n_k : \|x_i - x_j\| \le \lambda_k \quad (\text{o.B.d.A. } n_k \le n_{k+1})\\
\Rightarrow\quad &\Big\|\frac{1}{\lambda_k}\underbrace{(x_{n_{k+1}} - x_{n_k})}_{=:y_k}\Big\| \le 1\\
\Rightarrow\quad &\frac{1}{\lambda_k} y_k \text{ ist beschränkt}\\
\overset{(3)}{\Rightarrow}\quad &x_{n_j} = x_{n_0} + \sum_{k<j} \lambda_k \frac{1}{\lambda_k} y_k \text{ konvergiert.}
\end{aligned}$$

($4 \Rightarrow 1$) Es sei (x_n) eine Cauchy-Folge. Nach Voraussetzung existiert eine konvergente Teilfolge. Es sei $x_\infty := \lim_{k\to\infty} x_{n_k}$. Zu gegebenen $\varepsilon > 0$ wählen wir ein $N \in \mathbb{N}$ mit $\|x_n - x_m\| < \varepsilon$ für alle $n, m \ge N$ und ein $k \in \mathbb{N}$ mit $n_k \ge N$ und $\|x_\infty - x_{n_k}\| < \varepsilon$. Dann ist

$$\|x_n - x_\infty\| \le \|x_n - x_{n_k}\| + \|x_{n_k} - x_\infty\| < 2\varepsilon,$$

d.h. $x_n \to x_\infty$. $\square$

Wir können nun das Resultat (4.2.9) für $B(X,\mathbb{R}^n)$ verallgemeinern indem wir Räume von (beschränkten) Abbildungen mit Werten in Banach-Räumen betrachten:

5.5.6 Proposition. Banach-Raum der beschränkten Abbildungen.
Sei X eine Menge und E ein Banach-Raum. Dann ist der Raum

$$B(X,E) := \{f : X \to E \mid f(X) \text{ ist beschränkt in } E\}$$

ebenfalls ein Banach-Raum bzgl. der punktweisen Operationen und der Supremums-Norm

$$\|f\|_\infty := \sup\{\|f(x)\| : x \in X\}.$$

Die konvergenten Folgen bzgl. der zugehörigen Metrik sind gerade die gleichmäßig konvergenten.

Beweis. Daß $B(X,E)$ ein Vektorraum ist, folgt wie für $B(X,\mathbb{R}^n)$ in (4.2.9) wenn man ihm als Teilraum des Vektorraums aller Funktionen $f : X \to E$ auffaßt.

Daß $B(X,E)$ ein normierter Raum ist, zeigt man völlig analog zu (2.2.2).

Auch die Vollständigkeit zeigt man wie in (4.2.9). Sei dazu $(f_n)_n$ ein Cauchy-Folge in $B(X,E)$. Da für $x \in X$ die Punktevaluation $\operatorname{ev}_x : B(X,E) \to E$, $f \mapsto f(x)$ stetig (wegen $\|\operatorname{ev}_x(f)\| = \|f(x)\| \le \|f\|_\infty$) und linear ist (also gleichmäßig stetig ist), ist $f_n(x)$ ein Cauchy-Folge in E für jedes $x \in X$, und konvergiert somit. Es sei $f_\infty(x) := \lim_n f_n(x)$, dann gilt

$$\|f_n(x) - f_\infty(x)\| \le \underbrace{\|f_n(x) - f_k(x)\|}_{<\varepsilon \text{ für } n,k \ge n_0(\varepsilon)} + \underbrace{\|f_k(x) - f_\infty(x)\|}_{<\varepsilon \text{ für } k \ge k_0(\varepsilon,x)} \le 2\varepsilon$$

für $n \ge n_0(\varepsilon)$ (und geeignet gewählten k). Also konvergiert $f_n \to f$ in der Supremums-Norm.
Weiters ist

$$\|f_\infty(x)\| \le \|f_\infty(x) - f_n(x)\| + \|f_n(x)\| \le \|f_\infty - f_n\|_\infty + \|f_n\|_\infty < \infty$$

also gehört f_∞ zu $B(X,E)$. $\square$

Lemma. Vollständige Teilräume.
Es sei E ein Banach-Raum, F ein Teilvektorraum mit der Einschränkung der Norm von E als Norm. Falls F abgeschlossen ist, so ist F ebenfalls ein Banach-Raum.

Beweis. (Siehe (3.4.18)) Wenn $(y_n)_n$ eine Cauchy-Folge in F ist, dann auch in E und konvergiert somit in E wegen der Vollständigkeit von E. Dann liegt der Grenzwert wegen der Abgeschlossenheit von F in E sogar in F, also konvergiert sie in F. $\square$

Folgerung.
Für kompaktes X ist auch der Teilraum $C(X) := C(X,\mathbb{K})$ von $B(X,\mathbb{K})$ ein Banach-Raum, oder allgemeiner $C(X,E) \subseteq B(X,E)$ für Banach-Räume E.

Beweis. Nach (5.5.6) ist $B(X,E)$ vollständig, wenn E es ist. Nach (4.2.8) ist $C(X,E) := \{f : X \to E : f \text{ ist stetig}\}$ abgeschlossen in $B(X,E)$, also nach obigen Lemma selbst vollständig. $\square$

5.5.7 Produkte von Räumen.
Analog zum $\mathbb{R}^n := \mathbb{R} \times \ldots \times \mathbb{R}$ können wir auch Produkte $E_1 \times \ldots \times E_n$ von normierten Räumen $E_1,\ldots,E_n$ bilden. Mengenmäßig ist $E_1 \times \ldots \times E_n$ dabei das kartesische Produkt $\{(x_1,\ldots,x_n) : x_i \in E_i \text{ für alle } i\}$ (siehe (1.1.7)) und die Vektor-Operationen sind komponentenweise definiert, d.h.

$$\begin{aligned}(x_1,\ldots,x_n) + (y_1,\ldots,y_n) &:= (x_1+y_1,\ldots,x_n+y_n)\\ \lambda\cdot(x_1,\ldots,x_n) &:= (\lambda\, x_1,\ldots,\lambda\, x_n),\end{aligned}$$

also gerade so, daß die Projektionen $\mathrm{pr}_j : E_1 \times \ldots \times E_n \to E_j$, $(x_1,\ldots,x_n) \mapsto x_j$ linear sind. Als Norm können wir $\|(x_1,\ldots,x_n)\|_\infty := \max\{\|x_i\| : i \in \{1,\ldots,n\}\}$ verwenden. Wie in Aufgabe (2.8) für den $\mathbb{R}^n$ könnten wir aber genauso $\|(x_1,\dots,x_n)\|_1 := \sum_{i=1}^n \|x_i\|$ oder (das wegen Pythagoras geometrisch relevantere) $\|(x_1,\dots,x_n)\|_2 := \sqrt{\sum_{i=1}^n \|x_i\|^2}$ verwenden, welche die selben konvergenten- bzw. Cauchy-Folgen beschreiben, nämlich gerade jene Folgen deren Komponenten die entsprechende Eigenschaft haben. So sieht man auch sehr leicht ein, daß das Produkt $E_1 \times \ldots \times E_n$ von Banach-Räumen ebenfalls vollständig ist.

Für jeden normierten Raum E ist die Addition $a : E \times E \to E$, $(x,y) \mapsto x+y$ stetig, denn

$$\begin{aligned}\|a(x_1,y_1)-a(x_2,y_2)\| &= \|(x_1+y_1)-(x_2+y_2)\| \le \|x_1-x_2\| + \|y_1-y_2\| \\ &\le 2\max\{\|x_1-x_2\|, \|y_1-y_2\|\} = \|(x_1,y_1)-(x_2,y_2)\|,\end{aligned}$$

und ebenso ist die Skalarmultiplikation $m : \mathbb{K} \times E \to E$, $(\lambda, x) \mapsto \lambda \cdot x$ stetig, denn

$$\begin{aligned}\|m(\lambda_1,x_1)-m(\lambda_2,x_2)\| &= \|\lambda_1 x_1 - \lambda_2 x_2\| \le \|\lambda_1(x_1-x_2)\| + \|(\lambda_1-\lambda_2)x_2\| \\ &= |\lambda_1|\,\|x_1-x_2\| + |\lambda_1-\lambda_2|\,\|x_2\| \to 0\end{aligned}$$

für $\lambda_1 \to \lambda_2$ und $x_1 \to x_2$.

5.5.8 Definition. Stetige lineare Abbildungen.
Um verschiedene Banach-Räume zueinander in Beziehung zu setzen benötigen wir (strukturerhaltende) Abbildungen zwischen diesen. Wie in der linearen Algebra werden wir für diese die Linearität fordern. Die Verträglichkeit mit den Normen (und somit den Metriken) ist gerade die Stetigkeit. Wir wollen also stetig lineare Abbildungen zwischen normierten Räumen näher untersuchen.

Für lineare Abbildungen $f : \mathbb{R}^n \to F$ ist die Stetigkeit automatisch erfüllt, denn in (3.1.5) haben wir gezeigt, daß die Koordinaten-Projektionen $\mathrm{pr}_j : \mathbb{R}^n \to \mathbb{R}$, $x = (x^1,\dots,x^n) \mapsto x^j$ (siehe (3.1.5)) stetig sind und wegen

$$f(x) = f\Big(\sum_{j=1}^n x^j\, e_j\Big) = \sum_{j=1}^n x^j\, f(e_j) = \sum_{j=1}^n \mathrm{pr}_j(x)\, f(e_j),$$

ist f als Summe von Produkten stetiger Funktionen stetig.

Im unendlich-Dimensionalen ist dies nicht mehr richtig, wie die lineare Abbildung $\sum : E \to \mathbb{R}$, $x = (x^i)_i \mapsto \sum_{i=1}^\infty x_i$ auf dem bzgl. der Supremums-Norm normierten Raum $E := \{x = (x^i)_i \in \ell^\infty : x_i = 0$ für fast alle $i\}$ zeigt. Denn die Folge $n \mapsto \frac{1}{n}\chi_{\{1,\dots,n\}}$ konvergiert gegen 0 aber

$$\sum\Big(\frac{1}{n}\chi_{\{1,\dots,n\}}\Big) := \sum_{i=1}^\infty \frac{1}{n}\chi_{\{1,\dots,n\}}(i) = \sum_{i=1}^n \frac{1}{n} = 1 \not\to 0 = \sum(0).$$

5.5.9 Lemma. Charakterisierung stetig linearer Abbildungen.
Es sei $f : E \to F$ eine lineare Abbildung zwischen normierten Räumen. Dann sind folgende Aussagen äquivalent:

1. *f ist stetig;*
2. *f ist beschränkt auf allen beschränkten Mengen;*
3. *Das Bild $\{f(x) : \|x\| \le 1\}$ des Einheitsballs $\{x \in E : \|x\| \le 1\}$ ist beschränkt;*
4. *$\exists\, L \ge 0 \;\; \forall x \in E$: $\|f(x)\| \le L\|x\|$;*
5. *f ist* Lipschitz, *d.h. $\exists\, L \ge 0\; \forall x,y \in E : \|f(x)-f(y)\| \le L \cdot \|x-y\|$;*

Beweis. (1⇒2) Sei $B \subseteq E$ beschränkt, also existiert ein $r > 0$ mit $B \subseteq \{x : \|x\| \le r\} = r\cdot\{y : \|y\| \le 1\}$ (wobei wir $x = r\,y$ gesetzt haben). Angenommen $f(B)$ wäre nicht beschränkt. Dann gäbe es zu jedem $n \in \mathbb{N}$ ein $x_n \in E$ mit $\|x_n\| \le r$ und $\|f(x_n)\| \ge n$. Dann konvergiert $\frac{1}{n}x_n \to 0$ aber $\|f(\frac{1}{n}x_n)\| \ge 1$ und somit konvergiert $f(\frac{1}{n}x_n)$ nicht gegen $0 = f(0)$.

(2⇒3) Ist offensichtlich, da der Einheitsball nach Definition beschränkt ist.

(3⇒4) Nach Voraussetzung ist $f(B) \subseteq \{y : \|y\| \le L\}$ für ein $L > 0$, wobei $B := \{x : \|x\| \le 1\}$. Sei nun $x \in E$ beliebig. Dann ist $\frac{1}{\|x\|}x \in B$ und somit

$$\|f(x)\| = \left\| f\left(\|x\| \frac{1}{\|x\|} x\right)\right\| = \|x\| \cdot \left\| f\left(\frac{1}{\|x\|} x\right)\right\| \le \|x\|\, L.$$

(4⇒5) Es ist $\|f(x) - f(y)\| = \|f(x-y)\| \le L \cdot \|x - y\|$.

(5⇒1) Ist offensichtlich: zu $\varepsilon > 0$ wähle $\delta := \frac{\varepsilon}{L}$. □

Bemerkung. Kompaktheit.
Ein wesentlicher Unterschied zwischen endlich-dimensionalen und unendlich-dimensionalen Banach-Räumen E besteht darin, daß der Einheitsball $\{x \in E : \|x\| \le 1\}$ bei letzteren nicht mehr kompakt (wohl aber mit gleichen Beweis wie im endlich-dimensionalen beschränkt und abgeschlossen) ist, d.h. der Satz (3.3.4) von Bolzano und Weierstraß für diese nicht mehr gilt: Betrachte z.B. die Folge der (standard-Einheitsvektoren) $e^n \in \ell^\infty$, definiert durch

$$e^n_k := \begin{cases} 1 & \text{für } k = n \\ 0 & \text{andernfalls.} \end{cases}$$

Diese hat Norm $\|e^n\|_\infty = 1$ liegt also ganz im Einheitsball, kann aber in ℓ^∞ keine konvergente Teilfolge besitzten, denn $\|e^n - e^m\|_\infty = 1$ für alle $n \ne m$.

Dies ist natürlich ein großes Unglück, denn Kompaktheitsargumente haben wir oft sehr erfolgreich eingesetzt. Man kann sich hier dennoch ein wenig aus der Affäre ziehen, wenn man beachtet, daß die Folge e^n punktweise (= koordinatenweise) gegen 0 konvergiert. Dies läßt sich auf Banach-Räume übertragen, siehe z.B. [16, 7.28].

5.5.11 Folgerung. Banach-Raum der stetig linearen Abbildungen.
Es seien E und F normierte Räume. Dann ist $L(E,F)$ ein normierter Raum, wobei die Vektorraum-Operationen punktweise gegeben sind, d.h.

$$f + g : x \mapsto f(x) + g(x), \quad \lambda \cdot f : x \mapsto \lambda \cdot f(x),$$

und die Norm die sogenannte OPERATORNORM

$$\begin{aligned}\|f\| &:= \sup\Big\{\|f(x)\| : \|x\| \le 1\Big\} = \sup\Big\{\|f(x)\| : \|x\| = 1\Big\} = \sup\left\{\tfrac{\|f(x)\|}{\|x\|} : x \ne 0\right\} \\ &= \inf\Big\{K : \|f(x)\| \le K\|x\| \text{ für alle } x\Big\}\end{aligned}$$

ist.

Falls F vollständig ist, so gilt gleiches auch für $L(E,F)$.

Falls $F = \mathbb{K}$ ist so bezeichnen wir mit E^* den BANACH-RAUM $L(E,\mathbb{K})$ ALLER STETIGEN LINEAREN FUNKTIONALE auf E.

Falls $f : E \to F$ ein stetig linearer Operator ist, so bezeichnen wir mit $f^* : F^* \to E^*$ den ADJUNGIERTEN OPERATOR, der durch $f^*(\ell)(x) := \ell(f(x))$ gegeben ist.

Beweis. In der linearen Algebra lernt man, daß Summen und skalare Vielfache linearer Abbildungen wieder linear sind, und somit die Menge der linearen Abbildungen zwischen zwei Vektorräumen mit den punktweisen Operationen einen Vektorraum bilden, nämlich einen linearen Teilraum aller Abbildungen zwischen den beiden Vektorräumen.

Da die Addition $a : F \times F \to F$ und die skalar-Multiplikation $m_\lambda : F \to F$, $x \mapsto \lambda x$ nach (5.5.7) beide stetig sind, sind auch die Summe $f + g := a \circ (f,g)$ und skalare Vielfache $\lambda \cdot f = m_\lambda \circ f$ stetiger Abbildungen $f, g : E \to F$ wieder stetig, also bildet $L(E,F)$ einen linearen Teilraum des Vektorraums aller linearer Abbildungen.

Daß obige Beschreibungen der Operatornorm äquivalent sind, zeigt:

$$\begin{aligned}
\sup\{\|f(x)\| : \|x\| = 1\} &\le \sup\{\|f(x)\| : \|x\| \le 1\} \quad \text{(da mehr Elemente)}\\
&\le \sup\left\{\frac{\|f(x)\|}{\|x\|} : 0 \ne \|x\| \le 1\right\} \quad \left(\text{da } \|f(x)\| \le \frac{\|f(x)\|}{\|x\|} \text{ für } \|x\| \le 1\right)\\
&\le \sup\left\{\frac{\|f(x)\|}{\|x\|} : x \ne 0\right\} \quad \text{(da mehr Elemente)}\\
&\le \sup\{\|f(x)\| : \|x\| = 1\} \quad \left(\text{da} \quad \frac{\|f(x)\|}{\|x\|} = \left\|f(\frac{1}{\|x\|}x)\right\|\right),
\end{aligned}$$

$$\begin{aligned}
\inf\Big\{K : \|f(x)\| \le K \cdot \|x\| \text{ für alle } x\Big\} &= \inf\left\{K : \frac{\|f(x)\|}{\|x\|} \le K \text{für alle } x \ne 0\right\}\\
&= \inf\left\{K : \sup\left\{\frac{\|f(x)\|}{\|x\|} : x \ne 0\right\} \le K\right\}\\
&= \sup\left\{\frac{\|f(x)\|}{\|x\|} : x \ne 0\right\}.
\end{aligned}$$

Die Norm-Eigenschaften zeigt man nun wie in (2.2.2) für die Supremumsnorm, siehe auch (5.5.11):

$$\forall x : \|(f+g)(x)\| \le \|f(x)\| + \|g(x)\| \le (\|f\| + \|g\|)\,\|x\| \Rightarrow \|f+g\| \le \|f\| + \|g\|$$
$$\forall x : \|(\lambda\, f)(x)\| = |\lambda|\,\|f(x)\| \Rightarrow \|\lambda\, f\| = |\lambda|\,\|f\|$$

Eine andere Möglichkeit ist, dies auf (5.5.6) zurückzuspielen, denn wenn $A := \{x \in E : \|x\| \le 1\}$ den Einheitsball bezeichnet, so ist $\iota : L(E,F) \to B(A,F)$, $f \mapsto f|_A$ eine lineare und injektive(!) Abbildung und $\|f\| := \|f|_A\|_\infty = \|\iota(f)\|_\infty$. Da die Supremumsnorm nach (5.5.6) eine Norm ist gilt gleiches für die Operatornorm.

Nun sei F vollständig und $(f_i) \in L(E,F)$ ein Cauchy-Folge. Für jedes $x \in E$ konvergiert somit $f_i(x)$ gegen ein $f_\infty(x) \in F$, denn

$$\|f_i(x) - f_j(x)\| = \|(f_i - f_j)(x)\| \le \|f_i - f_j\|\,\|x\| \to 0 \text{ für } i,j \to \infty.$$

Da f_i punktweise gegen f_∞ konvergiert ist f_∞ linear, denn

$$f_\infty(x + \lambda\, y) = \lim_{i\to\infty} f_i(x + \lambda\, y) = \lim_{i\to\infty} f_i(x) + \lambda\, f_i(y) = \lim_{i\to\infty} f_i(x) + \lambda \lim_{i\to\infty} f_i(y) = f_\infty(x) + \lambda\, f_\infty(y).$$

Weiters ist für den Einheitsball $A \subseteq E$ die Folge $f_i|_A$ ein Cauchy-Folge in $B(A,F)$, und konvergiert somit gegen ein Funktion in $B(A,F)$. Da dies dann auch punktweise für $x \in A$ gelten muß, ist dieser Grenzwert $f_\infty|_A$. Somit ist f_∞ stetig nach (5.5.9) und $f_i \to f_\infty$ in $L(E,F)$, da es die Einschränkungen auf A in $B(A,F)$ tun. □

5.5.10 Bemerkung. Normen für stetig lineare Abbildungen.
Wir haben also für lineare Abbildungen $f : E \to F$ eine Zahl $\|f\| \in \mathbb{R} \cup \{+\infty\}$, die Operatornorm von f, die genau dann endlich ist, wenn f stetig ist. Wir wollen dies nun als Norm auf den Raum

$$L(E,F) := \{f : f \text{ ist eine stetig und lineare Abbildung von } E \text{ nach } F\}$$

aller stetig linearen Abbildungen $f : E \to F$ verwenden.

In der linearen Algebra lernt man, daß man nach Wahl einer (endlichen Basis) $B := (e_i)_{i=1}^n$ von E und $C := (f_j)_{j=1}^m$ von F vermöge der durch Koordinatendarstellung der Vektoren gegebenen Isomorphismen $\mathbb{K}^n \cong E$, $(x^1,\dots,x^n) \mapsto \sum_{i=1}^n x^i\, e_i$mit Umkehrabbildung $x \mapsto [x]_B : +(x^1,\dots,x^n)$ und $\mathbb{K}^m \cong F$, $(y^1,\dots,y^m) \mapsto \sum_{j=1}^m y^j\, f_j$ jede lineare Abbildung $a : E \to F$ als lineare Abbildung $\tilde{a} : \mathbb{K}^n \to \mathbb{K}^m$

auffassen kann, und somit mittels einer Matrix $[a]_{B,C} = [\tilde{a}] = (a_i^j)_{i=1,\dots n,j=1\dots m}$ beschrieben werden kann, welche durch

$$a(e_i) =: \sum_{j=1}^m a_i^j\, f_j$$

gegeben ist. Also ist

$$a(x) = a\Big(\sum_{i=1}^n x^i e_i\Big) = \sum_{i=1}^n x^i a(e_i) = \sum_{i=1}^n x^i \sum_{j=1}^m a_i^j f_j = \sum_{j=1}^m \Big(\sum_{i=1}^n a_i^j x^i\Big)\, f_j,$$

d.h. $\tilde{a}$ wirkt auf $(x^1,\dots,x^n) \in \mathbb{K}^n$ durch Matrix-Multiplikation mit $(a_i^j)_{i,j}$. Beachte, daß x^i, y^j, a_i^j hier keine Potenzen bedeuten, sondern wir bloß obere Indizes wie wir sie auch schon in (5.5.4) verwendet haben. Dies erlaubt einfacher zu erkennen über welche Indizes zu summieren ist, wenn man vereinbart, daß dies für Paare von gleichen oberen und unteren Index geschehen soll.

Versuchen wir also die Operatornorm einer (der einfachheithalber 2×2-)Matrix zu bestimmen, wobei wir vorerst am $\mathbb{R}^2 = \mathbb{R}\times\mathbb{R}$ die Maximums-Norm verwenden. Dann ist

$$\begin{aligned}\left\|\begin{pmatrix} a & b \\ c & d\end{pmatrix}\cdot\begin{pmatrix} x \\ y\end{pmatrix}\right\|_\infty &= \|(a\,x + b\,y, c\,x + d\,y)\|_\infty \\ &= \max\{|a\,x+b\,y|, |c\,x+d\,y|\} \\ &\le \max\{|x|,|y|\}\cdot\max\{|a|+|b|,|c|+|d|\} \\ &\le \max\{|a|+|b|,|c|+|d|\}\cdot\|(x,y)\|_\infty.\end{aligned}$$

also ist die Operatornorm $\le \max\{|a|+|b|,|c|+|d|\}$. Indem wir $(x,y) := (\mathrm{sign}(a),\mathrm{sign}(b))$ bzw. $(x,y) := (\mathrm{sign}(c),\mathrm{sign}(d))$ setzen, sehen wir, daß $\max\{|a|+|b|,|c|+|d|\}$ wirklich die Operatornorm ist.

Wenn wir anstelle der Maximums-Norm die geometrischere Euklid'ische Norm verwenden, dann ist die Operatornorm viel schwerer zu bestimmen: Der Einheitsball in $\mathbb{R}^2$ ist die Einheitskreisscheibe, und das Bild unter der Matrix ist eine elliptische Scheibe in $\mathbb{R}^2$. Die Operatornorm ist somit die größere der beiden Halbachse dieser Ellipse, siehe (6.4.11.3).

Wir können $n\times m$-Matrizen auch als Vektoren der Länge $n\cdot m$ auffassen, und die euklidischen Norm obiger Matrix ist dann $\sqrt{a^2+b^2+c^2+d^2}$.

Wir haben also verschiedene Normen am Raum $L(\mathbb{R}^2,\mathbb{R}^2)$ zur Verfügung, die aber alle die gleichen konvergenten Folgen und auch die gleichen Cauchy-Folgen beschreiben.

5.5.12 Bemerkung. Multilineare Abbildungen.
In (5.5.7) hatten wir die Stetigkeit der Addition $a : E\times E \to E$ gezeigt. Diese ist linear, denn $a((x_1,y_1)+\lambda(x_2,y_2)) = (x_1+\lambda x_2)+(y_1+\lambda y_2) = x_1+y_1+\lambda(x_2+y_2) = a(x_1,y_1)+\lambda\, a(x_2,y_2)$, also wäre das auch einfach aus $\|a(x_1,y_1)\| = \|x_1+y_1\| \le \|x_1\|+\|y_1\| \le 2\|(x_1,y_1)\|$ gefolgt.

Ebenso hatten wir die Stetigkeit der Skalarmultiplikation $m : \mathbb{K}\times E \to E$ gezeigt. Diese ist aber nicht linear, sondern nur bilinear, d.h. in jedem Faktor getrennt linear, d.h. für jedes $x\in E$ ist $\lambda\mapsto m(\lambda,x)$ linear und für jedes $\lambda\in\mathbb{K}$ ist $x\mapsto m(\lambda,x)$ linear.

Schon in der linearen Algebra benötigt man neben linearen auch bilineare und multilineare Abbildungen, wie z.B. die Determinante, das Kreuz-Produkt oder auch das skalare Produkt.

Wir benötigen also eine Verallgemeinerung von (5.5.9) auf multilineare Abbildungen:

Lemma. Stetigkeit multilinearer Abbildungen.
Es seien $E_1,\dots,E_m,F$ normierte Räume und $f : E_1\times\dots\times E_m \to F$ eine multilineare Abbildung.

Dann ist f genau dann stetig, wenn

$$\begin{aligned}\|f\| &:= \sup\{\|f(x_1,\dots,x_m)\| :\ \forall i : \|x_i\|\le 1\} \\ &= \sup\left\{\frac{\|f(x_1,\dots,x_m)\|}{\|x_1\|\dots\|x_m\|} : x_i\ne 0\right\} < \infty\end{aligned}$$

ist.

Beweis. (⇒) geht völlig analog zu (5.5.9) wenn man beachtet, daß $f(\lambda\,x) = \lambda^m\,f(x)$ gilt, d.h. f ist m-homogen.

(⇐) Mittels m-Homogenität sieht man sofort, daß f stetig bei 0 ist und die multi-Linearität hat dann die Stetigkeit überall zur Folge, was ich hier nur am typischen Fall $m = 2$ vorrechnen will: Es ist

$$\begin{aligned}\|f(x_1,y_1) - f(x_0,y_0)\| &\le \|f(x_1-x_0,y_1)\| + \|f(x_0,y_1-y_0)\| \\ &\le \|f\|\,\underbrace{\Big(\|x_1-x_0\|\,\|y_1\| + \|x_0\|\,\|y_1-y_0\|\Big)}_{\to 0}\end{aligned}$$

für $(x_1,y_1) \to (x_0,y_0)$. □

5.5.13 Bemerkung. Exponentialgesetz für multilineare Abbildungen.
Wir wollen nun bilineare Abbildungen $f : E_1 \times E_2 \to F$ als (lineare) Abbildungen auf E_1 mit Werten in einen Raum von (linearen) Abbildungen von E_2 nach F auffassen.

Wenn wir Abbildungen $f : E_1 \times E_2 \to F$ auf Bilinearität untersuchen, dann bedeutet dies ja einerseits für jedes (fixe) $x_1 \in E_1$ die Linearität von $\check f(x_1) : x_2 \mapsto f(x_1,x_2)$, $E_2 \to F$ zu zeigen, d.h. $\check f : x_1 \mapsto \check f(x_1)$ ist eine Abbildungen von E_1 mit Werten im Raum der linearen Abbildungen $E_2 \to F$, und andererseits für jedes (fixe) $x_2 \in E_2$ die Linearität von $x_1 \mapsto f(x_1,x_2)$, $E_1 \to F$ zu zeigen, d.h. $\check f(x_1 + \lambda y_1)(x_2) = f(x_1+\lambda y_1,x_2) = f(x_1,x_2)+\lambda\,f(y_1,x_2) = \check f(x_1)(x_2)+\lambda\,\check f(y_1)(x_2) = \Big(\check f(x_1)+\lambda\,\check f(y_1)\Big)(x_2)$ für alle $x_2 \in E$, d.h. $\check f(x_1 + \lambda\,x_2) = \check f(x_1) + \lambda\,\check f(x_2)$ als (lineare) Abbildungen $E_2 \to F$ zu zeigen (dies ist die Linearität von $\check f$).

Es ist also $f : E_1 \times E_2 \to F$ genau dann bilinear, wenn die durch $\check f(x_1)(x_2) := f(x_1,x_2)$ auf E_1 definierte Abbildung $\check f$ Werte im Vektorraum aller linearer Abbildungen $E_2 \to F$ hat und (auf E_1) linear ist.

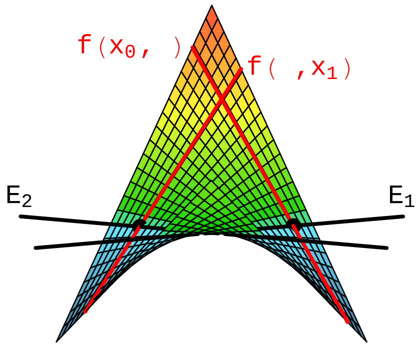
Beachte, daß wir $\check f$ auch für beliebige Abbildungen $f : X_1 \times X_2 \to Y$ betrachten können und zwar als Abbildung definiert auf X_1 mit Werten im Raum aller Abbildungen von X_2 nach Y durch $\check f(x_1) : x_2 \mapsto f(x_1,x_2)$, also $\check f(x_1)(x_2) := f(x_1,x_2)$.

Es unterscheidet sich $\check f$ von f nur dadurch, daß bei f beide Variablen $x_1 \in E_1$ und $x_2 \in E_2$ gleichzeitig als Paar (x_1,x_2) einzusetzen sind, bei $\check f$ hingegen zuerst x_1 und dann x_2 einzusetzen ist. Wir haben hier die Bindungsregel $\check f(x_1)(x_2) := (\check f(x_1))(x_2)$ verwendet, d.h. bei gleichen Vorrang ist eine Formel von links nach rechts auszuwerten, siehe z.B. $a - b - c := (a - b) - c$.

Als Memotechnik für die Notation $\check f$: Es wurde zwischen die, bei $f(x_1,x_2)$ als Paar (x_1,x_2) vereinten, Variablen x_1,x_2 ein Keil $\vee$ getrieben um $\check f(x_1)(x_2)$ zu erhalten.

Umgekehrt können wir aus einer Funktion g auf E_1 mit Werten im Raum der Funktionen von $E_2 \to F$ eine Abbildung $\hat g : E_1 \times E_2 \to F$ am Produkt machen. Wobei bei $\hat g(x_1,x_2) := g(x_1)(x_2)$ die beiden für g getrennten Variablen friedfertig unter einem Dach $\wedge$ als Paar vereint sind.

Wir zeigen nun, daß diese Zuordnungen einen (isometrischen) Isomorphismus beschreiben zwischen dem Banach-Raum $L(E_1,E_2;F)$der bilinearen stetigen Abbildungen $E_1 \times E_2 \to F$ einerseits und dem Banach-Raum $L(E_1,L(E_2,F))$ der linearen stetigen Abbildungen $E_1 \to L(E_2,F)$ andererseits.



Lemma. Bilineare versus lineare Abbildungen.
Es bezeichne $L(E_1, E_2; F)$ den Vektorraum der bilinearen stetigen Abbildungen $E_1 \times E_2 \to F$ versehen mit der Norm $\|f\| := \sup\{\|f(x,y)\| : \|x\| \leq 1, \|y\| \leq 1\}$. Dann haben wir einen Isomorphismus normierter Räume

$$L(E_1, E_2; F) \cong L(E_1, L(E_2, F))$$

gegeben durch $f \mapsto \check{f}$, $\hat{f} \leftarrow\!\shortmid f$.

Beweis. Daß $f \mapsto \check{f}$ wohldefiniert, linear und bijektiv ist mit Umkehrfunktion $\hat{g} \leftarrow\!\shortmid g$ sieht man leicht ein.

Daß dieser Isomorphismus die Länge bewahrt folgt sofort aus

$$\begin{aligned}\|f\| &= \sup\{\|f(x,y)\| : \|x\| \leq 1, \|y\| \leq 1\} = \sup\{\|\check{f}(x)(y)\| : \|x\| \leq 1, \|y\| \leq 1\} \\ &= \sup\{\|\check{f}(x)\| : \|x\| \leq 1\} = \|\check{f}\| \quad \square\end{aligned}$$

Wesentlicher Vorteil der Operatornorm im Gegensatz zu anderen Normen auf Räumen von Matrizen ist die

5.5.14 Folgerung. Die Algebra der stetig linearen Abbildungen.
Für die Operatornorm, wie sie in (5.5.11) definiert wurde. gilt: $\|id_E\| = 1$ und $\|f \circ g\| \leq \|f\| \cdot \|g\|$.

Beachte, daß $A := L(E, E)$ eine Algebra (mit 1) ist, d.h. ein Vektorraum zusammen mit einer bilinearen Abbildung $\circ : A \times A \to A$ die zusätzlich assoziativ ist (und eine Einheit $\mathrm{id} \in A$ besitzt).
Achtung, es gilt bei $\|f \circ g\| \leq \|f\| \cdot \|g\|$ nicht immer Gleichheit, wie das Beispiel $g(x) := (x, 0)$ und $f(x,y) := y$ zeigt.

Beweis.

$$\forall x : \|(f \circ g)(x)\| \leq \|f\| \, \|g\| \, \|x\| \Rightarrow \|f \circ g\| \leq \|f\| \, \|g\|. \quad \square$$

5.5.15 Definition. Banach-Algebra.
Eine BANACH-ALGEBRA ist eine normierte Algebra A die vollständig ist. Eines der Erz-Beispiele ist $L(E) := L(E, E)$ für jeden Banach-Raum E wobei die Einheit 1 der Multiplikation durch die Identität id_E auf E gegeben ist, insbesonders ist also der Raum der $n \times n$-Matrizen eine Banach-Algebra.

Das Erz-Beispiel einer kommutativen Banach-Algeba ist der Raum $B(X, \mathbb{K})$ der beschränkten Funktionen auf einer Menege X bzgl. der punkteweisen Multiplikation, wobei hier die Einheit 1 durch die konstante Funktion 1 gegeben ist.

Lemma. Stetigkeit von Potenzreihen auf Banach-Algebren.
Jede Potenzreihe mit Koeffizienten in $\mathbb{K}$ und Konvergenzradius $r > 0$ konvergiert auch auf dem Ball $\{x \in A : \|x\| < r\}$ in jeder Banach-Algebra A und stellt eine stetige Funktion dar.

Inbesonders können wir also e^A, $\cos(A)$, $\sin(A)$,... für jede quadratische Matrix bilden um eine neue Matrix zu erhalten.

Beweis. Es sei $\sum_n \lambda_n z^n$ eine Potenzreihe mit Konvergenzradius r. Es genügt wegen (5.5.5) zu zeigen, daß $\sum_n \lambda_n a^n$ absolut-konvergiert für alle $\|a\| < r$. Das ist offensichtlich der Fall, denn wegen $\| \sum_{k=n}^{n+p} \lambda_k a^k \| \leq \sum_{k=n}^{n+p} |\lambda_k| \, \|a\|^k$ ist die Cauchy-Bedingung erfüllt.
Die Stetigkeit ergibt sich nun genauso wie in (4.2.12). $\square$

5.5.16 Folgerung. Stetigkeit der Inversion.
Die Menge $\mathrm{Inv}(A)$ *der invertierbaren Elemente einer Banach-Algebra A ist offen und die* Inversion $\mathrm{inv} : a \mapsto a^{-1}$ *ist eine stetige Abbildung.*

Ein Element $a \in A$ heißt bekanntlich invertierbar (vgl. (1.2.4) und (1.2.5)), wenn eine Element $b \in A$ existiert, s.d. $a \cdot b$ und $b \cdot a$ die Einheit $1 \in A$ ist. Da diese b dann notwendigerweise eindeutig bestimmt ist, schreibt man a^{-1} für b.

Beweis. Die Potenzreihe $f(x) := \sum_{k=0}^{\infty} x^k$ konvergiert für $\|x\| < 1$ und klarerweise gilt $(1-x) \cdot f(x) = 1 = f(x) \cdot (1-x)$, denn

$$\begin{aligned}(1-x) \cdot f(x) &= (1-x) \cdot \lim_{n\to\infty} \sum_{k=0}^{n} x^k \\ &= \lim_{n\to\infty} (1-x) \cdot \sum_{k=0}^{n} x^k = \lim_{n\to\infty} 1 - x^{n+1} = 1.\end{aligned}$$

Folglich ist der Ball $\{y : \|y - 1\| < 1\}$ ganz in den invertierbaren Elementen enthalten und die Inversion $y \mapsto f(1-y)$ ist stetig auf dieser Umgebung. Für jedes andere invertierbare Element x_0 ist $\{y : \|y - x_0\| < \frac{1}{\|(x_0)^{-1}\|}\}$ eine Umgebung von invertierbaren Elementen y mit $y^{-1} = (x_0 \cdot (x_0)^{-1} \cdot y)^{-1} = ((x_0)^{-1} \cdot y)^{-1} \cdot (x_0)^{-1}$ da $\|(x_0)^{-1} \cdot y - 1\| \leq \|(x_0)^{-1}\| \cdot \|y - x_0\| < 1$. $\square$

Bemerkung. Invertierbarkeit für Matrizen.
Beache, daß (5.5.16) für Matrizen auch anders gezeigt werden kann:

Bekanntlich ist eine Matrix a genau dann invertierbar, wenn ihre Determinante $\det(a)$ nicht 0 ist. Da die Determinante eine Polynom in den Eintragungen der Matrix ist, ist $\det : a \mapsto \det(a)$ stetig und somit ist die Menge der Matrizen a mit $\det(a) \neq 0$ offen.

Die Inverse einer Matrix kann bekanntlich mittels det berechnet werden, indem man die Matrix der algebraischen Komplemente von A bildet, somit ist die Zuordnung $a \mapsto a^{-1}$ nach (3.1.4) und (3.1.5) stetig.

5.5.17 Bemerkung. Hilbert-Räume.
Wir sahen in (2.2.3), daß die geometrisch bedeutendste Metrik am $\mathbb{R}^n$ die euklidische Metrik ist, denn die zugehörige Norm erfüllt $\|x\|^2 = \langle x|x \rangle$, wobei $\langle \_|\_ \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $(x,y) \mapsto \sum_{i=1}^n x_i \, y_i$ eine symmetrische Bilinearform ist. Diese Bilinearform erlaubt uns nämlich Winkel $\angle xy$ zwischen Vektoren $x, y \neq 0$ durch

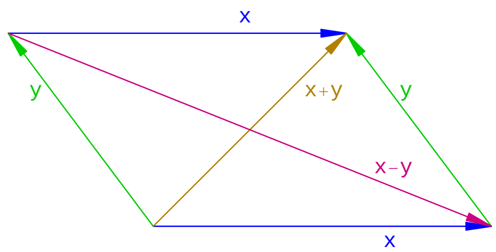
$$\cos(\angle xy) := \frac{\langle x|y \rangle}{|x| \, |y|}$$

zu definieren, und insbesonders Normalität vermöge $x \perp y :\Leftrightarrow \langle x|y \rangle = 0$.

Die selbe Konstruktion funkioniert auch beim Banach-Raum ℓ^2, wenn man $\langle x|y \rangle$ durch $\sum_{i=1}^{\infty} x_i \, y_i$ definiert. Etwas allgemeiner versteht man unter einen Hilbert-Raum E einen Banach-Raum der eine bilineare Funktion $\langle\_|\_\rangle : E\times E \to \mathbb{R}$ zuläßt, welche $\|x\|^2 = \langle x|x\rangle$ für alle $x \in E$ erfüllt. Man kann zeigen (siehe z.B. [16, 6.4]), daß dies genau dann der Fall ist, wenn die sogenannte Parallelogramm-Gleichung

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) \text{ für alle } x, y \in E$$

erfüllt ist.



In der linearen Algebra lernt man, daß jeder endlich dimensionale Vektorraum E isomorph zu seinem Dualraum E^* ist, wobei man einen Isomorphismus so erhalten kann, daß man die Elemente einer Basis von E auf die entsprechenden der dualen Basis abbildet. Für $E = \mathbb{R}^n$ wird dies gerade beschrieben durch das Transponieren der Spaltenvektoren $x \in E$ zu Zeilenvektoren x^t, also den Matrizendarstellungen von linearen Abbildungen $E \to \mathbb{R}$. Unter Verwendung des inneren Produkts $\langle\_|\_\rangle$ kann man diesen Isomorphismus auch Basis-unabhängig durch $E \ni x \mapsto \langle x|\_\rangle \in E^*$ beschreiben.

Man kann zeigen (siehe z.B. [16, 6.7]), daß auch für Hilbert-Räume durch $E \ni x \mapsto \langle x|\_\rangle \in E^*$ ein Isomorphismus $E \cong E^*$ gegeben ist, d.h. wir können lineare stetige Abbildungen $f : E \to \mathbb{R}$ durch jenen Vektor $x \perp \mathrm{Ker}(f)$ beschreiben, auf welchem $f(x) = 1$ ist und zwar vermöge $f(y) = \langle x|y\rangle$.

Unendlichdimensionale Hilbert-Räume werden in der Physik benötigt, um die Quantenmechanik zu formulieren, also um unsere Welt im kleinen in Griff zu bekommen.

5.5.18 Weitergehende Sätze aus der Funktionalanalysis für Banach-Räume. Im $\mathbb{R}^n$ haben wir viele Aussagen auf den 1-dimensionalen Fall zurückführen können, indem wie die Koordinaten $x^i \in \mathbb{R}$ der Vektoren $x = (x^1, \dots, x^n) \in \mathbb{R}^n$ betrachtet habe, siehe z.B. Aufgabe (2.11), (2.3.4), (2.4.5), (3.1.4), (4.1.1) und (5.3.8). Die Abbildungen $\mathrm{pr} : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto x^i$ die wir dabei verwendet haben sind linear (und somit stetig). Auch bei Funktionen-Räumen können wir Identitäten durch Auswerten mittels der stetigen linearen Funktionen $\mathrm{ev}_x : \mathbb{R}^X \to \mathbb{R}$, $f \mapsto f(x)$ verifizieren. Allerdings können wir Konvergenz so nicht beschreiben, denn $\mathrm{ev}_x(f_n) \to \mathrm{ev}_x(f_\infty)$ für alle $x \in X$ bedeutet nur punktweise Konvergenz, und z.B. auf $B(X,\mathbb{R})$ haben wir vor allem die gleichmäßige Konvergenz betrachtet.

Dennoch drängt sich die Frage auf, wie reichhaltig die Menge der stetig linearen Funktionen auf einen normierten Raum E ist. Für Hilbert-Räume habe wir das oben schon beantwortet. Die allgemeine Antwort liefert der Satz von Hahn-Banach (siehe z.B. [16, 7.2 und 7.7]), nach dem die stetig linearen Funktionen auf E die Punkte von E trennen, d.h. zu $x_1 \neq x_2$ in E existiert eine stetige lineare Abbildung $f : E \to \mathbb{R}$ mit $f(x_1) \neq f(x_2)$. Etwas allgemeiner können wir zu Punkten x die nicht in einer gegebenen abgeschlossenen konvexen Teilmenge $A \subseteq E$ (d.h. mit je zwei Punkten $a_0, a_1 \in A$ gehört auch die Verbindungsstrecke $\{a_0 + t(a_1 - a_0) : 0 \leq t \leq 1\}$ zu A) liegen eine stetige lineare Abbildung $f : E \to \mathbb{R}$ finden mit $f(x) \notin f(A)$.

Obwohl die punktweise Konvergenz zu schwach ist um die viel wichtigere gleichmäßige Konvergenz zur Folge zu haben, läßt sich für stetig lineare Abbildungen folgendes Prinzip der gleichmässigen Beschränktheit (siehe z.B. [16, 5.4]) zeigen: Es sei $\mathcal{F} \subseteq L(E,F)$ eine punktweise beschränkte Menge stetig linearer Abbildungen auf einen Banach-Raum E, dann ist $\mathcal{F}$ beschränkt im normierten Raum $L(E,F)$, d.h. $\{\|f\| : f \in \mathcal{F}\}$ ist beschränkt, insbesonders ist $\mathcal{F}$ gleichgradig stetig (bei 0), d.h.

$$\forall \varepsilon > 0 \ \exists \delta > 0;\ \forall f \in \mathcal{F};\ \forall x \in X : \|x\| < \delta \Rightarrow \|f(x)\| < \varepsilon.$$

Dies hat den Satz von Banach-Steinhaus (siehe z.B. [16, 5.5]) zur Folge: Konvergiert eine Folge stetig linearer Abbildungen auf einem Banach-Raum punktweise, so ist die Grenzfunktion ebenfalls

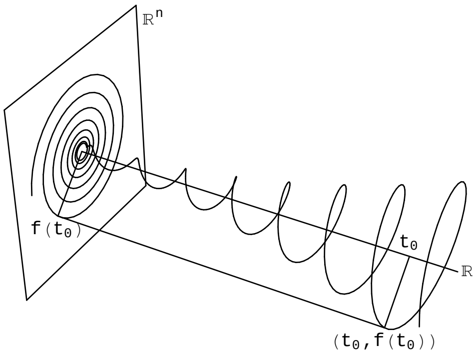
stetig und linear.
Die Linearität der Grenzfunktion f_∞ folgt dabei unmittelbar aus jener der Folgeglieder f_i. Die Stetigkeit ergibt sich aus (5.5.9), denn aus der punktweisen Konvergenz der f_i folgt klarerweise die punktweise Beschränktheit und somit die gleichmäßige Beschränktheit auf dem Einheitsball $B \subseteq E$, d.h. $L := \sup\{\|f_j\|_\infty : j \in \mathbb{N}\} < \infty$. Dann ist aber auch $\|f_\infty(x)\| \leq L$ für alle $x \in B$ und somit $\|f_\infty\|_\infty \leq L$.

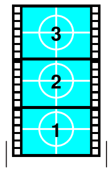
Im endlich-Dimensionalen lernt man, daß die Invertierbarkeit einer linearen Abbildung f durch die Aussage $\det(f) \neq 0$ beschrieben wird. Im unendlich-Dimensionalen hat man keine Determinante mehr zur Verfügung. Dennoch kann man zeigen, daß die inverse Abbildung einer bijektiven stetig linearen Abbildung zwischen Banach-Räumen selbst stetig ist. Dies ist der Satz über offene Abbildungen, siehe z.B. [16, 5.7].

Wen all dies genauer interessiert, der möge ein Buch über Funktionalanalysis (z.B. [16]) zur Hand nehmen.

5.5 Analysis von Kurven

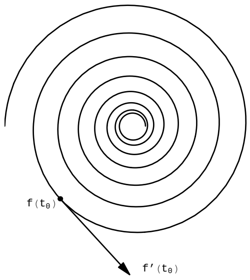
Wir wollen nun Analysis für Kurven treiben, d.h. für vektorwertigen Abbildungen $f : \mathbb{R} \supseteq J \to E$ die auf Intervallen $J \subseteq \mathbb{R}$ definiert sind und Werte im $\mathbb{R}^n$ oder allgemeiner in Banach-Räumen E haben. Um uns Bilder zu machen beachten wir zuerst, daß $f : \mathbb{R} \to \mathbb{R}^n$ per Definition (als Graph) eine Teilmenge von $\mathbb{R} \times \mathbb{R}^n = \mathbb{R}^{n+1}$ ist und somit bereits für $n = 2$ im 3-dimensionalen zu visualisieren ist. Für viele Zwecke reicht es uns allerdings, wenn wir die Parametrisierung ignorieren und nur das Bild $f(\mathbb{R}) \subset \mathbb{R}^n$ betrachten. Dies ist gerade durch das Bild der die Projektion $\mathrm{pr}_2 : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, $(t, y) \mapsto y$ (des Graphs) von f gegeben.

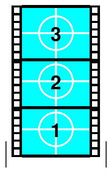




5.6.15 Definition. Differenzierbare Kurven.
Beginnen wir vorerst mit der einfach zu behandelnden Differenzierbarkeit von Kurven $f : [a, b] \to E$ mit Werten in Banach-Räumen E. Wie in (4.1.1) definieren wir die Ableitung (und damit die Differenzierbarkeit) von f im Punkt $x \in [a, b]$ durch

$$f'(x) := \lim_{h \to 0} \frac{f(x + h) - f(x)}{h} \in E.$$





In den Randpunkten $x \in \{a, b\}$ verlangen wir natürlich nur die einseitige Differenzierbarkeit, d.h. die Existenz von

$$f'(a) := \lim_{t \to a+} \frac{f(t) - f(a)}{t - a} \text{ und } f'(b) := \lim_{t \to b-} \frac{f(t) - f(b)}{t - b}.$$

Man kann nun auch für Kurven die meisten der grundlegenden Resultate (allem voran die Kettenregel (4.1.14) mit identischen Beweis, siehe auch (6.1.5)) wie in Abschnitt (4.1) zeigen und als Spezialfall:

6.1.6a Lemma. Kettenregel für lineare äußere Abbildung.
Es sei $f : \mathbb{R} \supseteq J \to E$ differenzierbar bei $x_0 \in J$ und $g : E \to F$ stetig und linear. Dann ist $g \circ f : \mathbb{R} \supseteq J \to E \to F$ differenzierbar bei x_0 mit Ableitung $(g \circ f)'(x_0) = g(f'(x_0))$.

Beweis.

$$\begin{aligned}(g \circ f)'(x_0) &= \lim_{h \to 0} \frac{(g \circ f)(x_0 + h) - (g \circ f)(x_0)}{h} \overset{g \text{ linear}}{=\!=} \lim_{h \to 0} g\left(\frac{f(x_0 + h) - f(x_0)}{h}\right) \\ &\overset{g \text{ stetig}}{=\!=} g\left(\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}\right) = g(f'(x_0)) \quad \square\end{aligned}$$

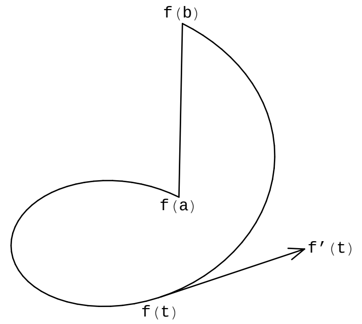
Folgerung. Die Ableitung kann komponentenweise berechnet werden.
Es sei $f = (f^1, f^2) : \mathbb{R} \supseteq J \to E_1 \times E_2 =: E$ *und* $x_0 \in J$. *Dann gilt:* $f : J \to E$ *ist differenzierbar bei* $x_0 \Leftrightarrow \ \forall i \in \{1,2\} : f^i : J \to E_i$ *ist differenzierbar bei* x_0.
Weiters gilt dann:

$$f'(x_0) = \Big((f^1)'(x_0), (f^2)'(x_0)\Big).$$

Beweis. Dieser zeigt man völlig analog zu dem entsprechenden Resultaten (2.3.4) für Konvergenz, (3.1.4) für Stetigkeit und (4.1.18) für Differenzierbarkeit von Funktionen $\mathbb{R} \to \mathbb{R}$, denn Differenzenquotienten können offensichtlich komponentenweise berechnet werden und die Konvergenz im Produkt ist ebenfalls komponentenweise. □

Bemerkung. Nicht-Übertragbarkeit des Mittelwertsatzes.
Beim Mittelwertsatz (4.1.5) bzw. Satz von Rolle (4.1.4) funktioniert die Verallgemeinerung allerdings bereits bei Kurven $f : [a,b] \to \mathbb{R}^2$ im 2-dimensionalen nicht mehr, wie das Beispiel $f(t) := (\cos(t), \sin(t))$ für $t \in [0, 2\pi]$ zeigt. Dabei ist $\frac{f(2\pi)-f(0)}{2\pi} = 0$ aber $\|f'(t)\| = 1$ für alle t.

Wir haben in (4.1.5) gesehen, daß es für Kurven im $\mathbb{R}^2$ allerdings Punkte gibt, wo die Ableitung noch parallel zur Sehne ist. Dies stimmt bereits im $\mathbb{R}^3$ nicht mehr, wie die Schraubenlinie $f : t \mapsto (\cos(t), \sin(t), t)$ zeigt. Dabei ist $\frac{f(2\pi)-f(0)}{2\pi} = (0,0,1)$ aber $f'(t) = (-\sin(t), \cos(t), 1)$ ist davon kein Vielfaches.



Beachte jedoch, daß

$$\frac{f'(\pi/2) + f'(3\pi/2)}{2} = \frac{(-1,0,1) + (1,0,1)}{2} = (0,0,1) = \frac{f(2\pi) - f(0)}{2\pi}$$

gilt, die Steigung der Sehne also sehr wohl ein Mittelwert von Ableitungen ist.

Wir können hier nur folgenden Spezialfall des Mittelwertsatzes zeigen:

5.6.16 Hilfssatz. Ableitung 0 impliziert konstant.
Es sei $f : [a,b] \to F$ *stetig und auf* (a,b) *differenzierbar und* $f' = 0$. *Dann ist* f *konstant.*

Dies verallgemeinert die Folgerung in (4.1.6).

1-ter Beweis. Es sei $a < t < s < b$ mit $f(s) \neq f(t)$. Dann existiert nach dem Satz von Hahn-Banach ein stetiges lineares Funktional $\ell \in L(F, \mathbb{R})$ mit $\ell(f(t)) \neq \ell(f(s))$. Nach dem klassischen Mittelwertsatz (4.1.5) und dem Spezialfall (6.1.6a) der Kettenregel ist

$$\frac{(\ell \circ f)(t) - (\ell \circ f)(s)}{t - s} = (\ell \circ f)'(\xi) = \ell(f'(\xi)) = \ell(0) = 0$$

für ein $\xi \in (t,s) \subseteq (a,b)$. □

Beachte, daß der folgende Beweis seine Gültigkeit behält, wenn nur die rechtsseitigen Ableitungen $\lim_{s->t+} \frac{f(s)-f(t)}{s-t}$ existieren und 0 sind.

2-ter Beweis. Sei vorerst $a < t_0 < s_0 < b$. Sukzessives Halbieren von $[t_0, s_0]$ und Ersetzen eines der beiden Randpunkte durch den Mittelpunkt liefert $|t_n - s_n| = |t_0 - s_0|/2^n$ und $\|f(t_n) - f(s_n)\| \geq \|f(t_0) - f(s_0)\|/2^n = \frac{\|f(t_0)-f(s_0)\|}{|t_0-s_0|}|t_n - s_n|$. Sei ξ das Zentrum dieser Intervallschachtelung. Da f differenzierbar bei ξ ist, gilt $\frac{f(r)-f(\xi)}{r-\xi} \to f'(\xi) = 0$ für $r \to \xi$ und somit

$$\frac{f(t_n) - f(s_n)}{t_n - s_n} = \underbrace{\frac{f(t_n) - f(\xi)}{t_n - \xi}}_{\to 0} \cdot \underbrace{\frac{t_n - \xi}{t_n - s_n}}_{0\leq \quad \leq 1} + \underbrace{\frac{f(\xi) - f(s_n)}{\xi - s_n}}_{\to 0} \cdot \underbrace{\frac{\xi - s_n}{t_n - s_n}}_{0\leq \quad \leq 1} \to 0$$

für $n \to \infty$. Also ist $f(t_0) = f(s_0)$, d.h. f ist konstant auf (a, b) und wegen der Stetigkeit von f auch auf $[a, b]$. □

Eine detaillierte Aussage über den Differenzenquotienten $\frac{f(b)-f(a)}{b-a}$ von Funktionen $f : [a, b] \to \mathbb{R}$ lieferte der Hauptsatz in der Form

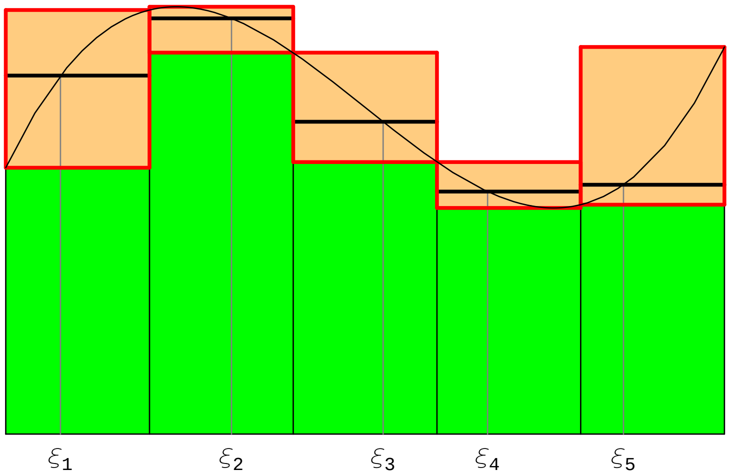
$$f(b) - f(a) = \int_a^b f'(x)\, dx = \int_0^1 f'(a + t\,(b - a))\,(b - a)\, dt.$$

Wir können also versuchen statt dem Mittelwertsatz den Hauptsatz zu übertragen, benötigen dafür aber natürlich das Integral für Kurven.

5.6.1 Definition. Riemann-Summen.
Bei der Definition (von Darboux) für Integrierbarkeit von Funktionen $f : J \to \mathbb{R}$ auf kompakten Intervallen $J \subseteq \mathbb{R}$ haben wir für die Ober- und Untersummen die Ordnung von $\mathbb{R}$ benötigt um von $\sup(f(I))$ und $\inf(f(I))$ und in der Folge von $\inf_Z O(f, Z)$ und $\sup_Z U(f, Z)$ sprechen zu können. Für Kurven $f : J \to E$ haben wir auf E keine Ordnung mehr zur Verfügung. In Analogie zu den Resultaten über Konvergenz (2.3.4), über Stetigkeit (3.1.4) und über Differenzierbarkeit (4.1.18) könnten wir Integrierbarkeit von $f : J \to \mathbb{R}^n$ ad hoc komponentenweise definieren, siehe (5.3.8). Für $f : J \to E$ mit Werten in Banach-Räumen E geht dies nicht mehr. Wir sollten jedenfalls versuchen sup und inf in der Definition von $\int f$ durch lim zu ersetzen. Anstelle für eine Zerlegung Z die Ober- und Untersummen $O(f, Z) := \sum_{I\in Z} \sup(f(I))\,|I|$ und $U(f, Z) := \sum_{I\in Z} \inf(f(I))\,|I|$ von f zu betrachten können wir die dazwischenliegenden Riemann-Summen

$$S(f, Z, \xi) := \sum_{I\in Z} |I|\, f(\xi_I)$$

für beliebige ‘Zwischenvektoren’ ξ von Z, d.h. $\xi = (\xi_I)_{I\in Z}$ mit $\xi_I \in I$ für alle $I \in Z$, definieren.



Diese Summen machen auch Sinn, wenn $f : J \to E$ Werte in einem beliebigen reellen Vektorraum E hat. Beachte dabei, daß wir die Faktoren $|I|$ vorgezogen haben, denn $|I| \in \mathbb{R}$, wohingegen $f(\xi_I) \in E$, und die Skalarmultiplikation wird üblicherweise von links geschrieben.

5.6.2 Lemma. D-Integrierbarkeit via Riemann-Summen.
Eine beschränkte Funktion $f : J \to \mathbb{R}$ ist genau dann D-integrierbar mit Integral φ, wenn für jedes $\varepsilon > 0$ eine Zerlegung Z_0 existiert, s.d. $|S(f, Z, \xi) - \varphi| < \varepsilon$ für alle Zerlegungen $Z \supseteq Z_0$ und Zwischenvektoren ξ von Z.

Beweis. ($\Rightarrow$) Nach Voraussetzung existiert zu $\varepsilon > 0$ eine Zerlegung Z_0 mit $|O(f, Z_0) - \varphi| < \varepsilon$ und $|U(f, Z_0) - \varphi| < \varepsilon$. Offensichtlich ist $U(f, Z_0) \le U(f, Z) \le S(f, Z, \xi) \le O(f, Z) \le O(f, Z_0)$, also auch $|S(f, Z, \xi) - \varphi| < \varepsilon$.

($\Leftarrow$) Aus $|S(f, Z_0, \xi) - \varphi| \le \varepsilon$ folgt $\varphi - \varepsilon < S(f, Z_0, \xi) < \varphi + \varepsilon$ für alle Zwischenvektoren ξ von Z_0 und somit $\varphi - \varepsilon \le U(f, Z_0) \le U(f) \le O(f) \le O(f, Z_0) \le \varphi + \varepsilon$. Folglich ist $|U(f) - \varphi| \le \varepsilon$ und $|O(f) - \varphi| \le \varepsilon$ für alle $\varepsilon > 0$, also $U(f) = \varphi = O(f)$. □

5.6.3 Definition. Riemann-Integral.
Die Beschreibung (5.6.2) des Integrals vermöge Riemann-Summen besagt anschaulich, daß $S(f, Z, \xi) \to \int f$ falls Z beliebig 'fein' wird. Um dies exakt zu machen (ohne den Begriff eines Netzes einzuführen) sollten wir die Menge aller PUNKTIERTEN ZERLEGUNGEN (d.h. Zerlegung zusammen mit zugehörigen Zwischenvektor) mit einer Metrik versehen. Dazu definieren wir das FEINHEITSMASS $|Z|$ einer punktierten Zerlegung als $|Z| := \max\{|I| : I \in Z\}$. Den Abstand zweier punktierter Zerlegungen Z und Z' definieren wir als

$$d(Z, Z') := \begin{cases} 0 & \text{falls } Z = Z' \text{ als punktierte Zerlegungen} \\ |Z| + |Z'| & \text{andernfalls.} \end{cases}$$

Dies ist offensichtlich eine Metrik, denn $d(Z, Z'') \le |Z| + |Z''| \le |Z| + |Z'| + |Z'| + |Z''| = d(Z, Z') + d(Z', Z'')$ für $Z \ne Z' \ne Z''$ und andernfalls ist die Dreiecksungleichung trivial.

Wir sind aber daran interessiert, was passiert, wenn $|Z|$ beliebig klein wird, also fügen wir der Menge aller punktierter Zerlegungen noch einen neuen Punkt ∞ (sozusagen die unendlich feine 'Zerlegung') hinzu, definieren seine Feinheit als $|\infty| = 0$ und verwenden für den Abstand die selbe Formel wie zuvor. Dann erhalten wir noch immer einen metrischen Raum. Beachte, daß der Abstand zweier verschiedener Punkte Z und Z' gerade $d(Z, \infty) + d(\infty, Z')$ ist, also wir immer den Umweg über ∞ gehen müssen. Man kann sich diesen Raum als Seeigel mit ∞ als Körper und verschieden langen Stacheln (zu jeder Zerlegung eine) vorstellen.

Es sei nun $f : J \to E$ eine Abbildung auf einem kompakten Intervall $J = [a, b] \subseteq \mathbb{R}$ mit Werten in einem Banach-Raum E. Für Zerlegungen Z von $[a, b]$ gegeben durch $a = t_0 < \dots < t_n = b$ (bzw. Teilintervallen $I \subseteq [a, b]$) und zugehörigen ZWISCHENVEKTOR ξ gegeben durch $\xi_i \in [t_{i-1}, t_i]$ für alle $1 \le i \le n$ (bzw. $\xi_I \in I$ für $I \in Z$) definieren wir die RIEMANN-SUMME von f bzgl. Z und ξ als

$$S(f, Z, \xi) := \sum_{i=1}^{n} (t_i - t_{i-1}) f(\xi_i) = \sum_{I \in Z} |I| \cdot f(\xi_I) \in E.$$

Wir nennen f RIEMANN-INTEGRIERBAR oder kurz R-INTEGRIERBAR auf J falls der Limes

$$\int_J f := \lim_{|Z| \to 0} S(f, Z, \xi) := \lim_{(Z,\xi) \to \infty} S(f, Z, \xi) \in E$$

existiert. Dieser Limes heißt dann das RIEMANN-INTEGRAL von f und man schreibt auch $\int_a^b f = \int_a^b f(x)\, dx$ dafür.

5.6.4 Cauchy-Kriterium für R-Integrierbarkeit.
Eine Abbildung $f : [a, b] \to E$ ist genau dann R-integrierbar, wenn

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall |Z_1|, |Z_2| < \delta \ \forall \xi_1, \xi_2 : \|S(f, Z_1, \xi_1) - S(f, Z_2, \xi_2)\| < \varepsilon.$$

Beweis. Dies folgt unmittelbar aus (5.3.3a). □

5.6.5 Lemma. Beschränktheit R-integrierbarer Kurven.
Riemann-integrierbare Funktionen $f : \mathbb{R} \supseteq [a,b] \to E$ sind beschränkt.

Beweis. Angenommen f ist nicht beschränkt. Somit gibt es für jede Zerlegung Z mindestens ein Intervall I_0, auf welchem f nicht beschränkt ist. Dann sind auch die Riemann-Summen $\xi \mapsto S(f,Z,\xi)$ unbeschränkt (und folglich nicht konvergent), denn dazu wähle zu $G > 0$ die ξ_I für $I \neq I_0$ beliebig und ξ_{I_0}, so daß

$$|I_0|\,\|f(\xi_{I_0})\| \geq G + \sum_{I \neq I_0} |I|\,\|f(\xi_I)\| \Rightarrow$$
$$\|S(f,Z,\xi)\| = \Big\|\sum_I |I|\, f(\xi_I)\Big\| \geq |I_0|\,\|f(\xi_{I_0})\| - \sum_{I\neq I_0} |I|\,\|f(\xi_I)\| \geq G. \quad \square$$

5.6.6. Differenz zweier Riemann-Summen Um die Cauchy-Bedingung aus (5.6.5) anwenden zu können benötigen wir eine Abschätzung für die Differenz zweier Riemann-Summen wobei wir o.B.d.A. annehmen können, daß dabei eine der Zerlegung eine Verfeinerung der anderen ist.

Lemma.
Es seien $Z' \supseteq Z$ Zerlegungen von $[a,b]$ und $\Omega_f(Z) := \max\{\Omega(f|_I) : I \in Z\}$ die Oszillation *von f bzgl. der Zerlegung Z. Dann ist*

$$\|S(f,Z,\xi) - S(f,Z',\xi')\| \leq (b-a)\cdot \Omega_f(Z).$$

Beweis.

$$\begin{aligned} S(f,Z,\xi) - S(f,Z',\xi') &= \sum_{I\in Z}\Big(|I|\,f(\xi_I) - \sum_{I\supseteq I'\in Z'} |I'|\,f(\xi'_{I'})\Big)\\ &= \sum_{I\in Z}\sum_{I\supseteq I'\in Z'} |I'|\,\Big(f(\xi_I) - f(\xi'_{I'})\Big)\\ \Rightarrow\quad \|S(f,Z,\xi) - S(f,Z',\xi')\| &\leq \sum_{I\in Z}\sum_{I\supseteq I'\in Z'} |I'|\cdot \underbrace{\|f(\xi_I) - f(\xi'_{I'})\|}_{\leq \Omega(f|_I)}\\ &\leq \sum_{I\in Z}\sum_{I\supseteq I'\in Z'} |I'|\cdot \Omega_f(Z) = (b-a)\,\Omega_f(Z). \quad \square \end{aligned}$$

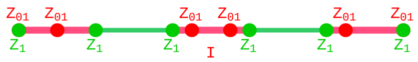
5.6.7 Proposition. R-integrierbar versus D-integrierbar.
Es sei $f : J \to E$ eine Abbildung auf einem kompakten Intervall $J = [a,b] \subseteq \mathbb{R}$ mit Werten in einem Banach-Raum E.

Dann ist f genau dann Riemann-integrierbar mit Riemann-Integral $\int f$, wenn f beschränkt ist und für jedes $\varepsilon > 0$ eine Zerlegung Z_0 von J existiert, s.d. $\|S(f,Z,\xi) - \int f\| < \varepsilon$ für alle Zerlegungen $Z \supseteq Z_0$.

Falls $E = \mathbb{R}$ ist, so ist dies nach (5.6.2) zur Darboux-Integrierbarkeit äquivalent.

Beweis. ($\Rightarrow$) Die Beschränktheit von f haben wir in (5.6.5) gezeigt. Nach Voraussetzung existiert zu $\varepsilon > 0$ ein $\delta > 0$, s.d. $\|S(f,Z,\xi) - \int f\| < \varepsilon$ aus $d((Z,\xi),\infty) < \delta$ folgt. Da Zerlegungen Z mit $|Z| < \delta$ existieren (man Teile J in hinreichend viele gleich große Intervalle) ist diese Richtung gezeigt.

($\Leftarrow$) Sei Z_0 eine Zerlegung wie gefordert, N die Anzahl ihrer Teilintervalle. Sei nun Z_1 eine beliebige Zerlegung mit $d((Z_1,\xi),\infty) = |Z_1| \leq \delta \leq \frac{\varepsilon}{N\,\Omega(f)}$ und $Z_{01} = Z_0 \cup Z_1$ die gemeinsame Verfeinerung. Sei ξ_1 eine Zwischenvektor von Z_1 und ξ_{01} ein Zwischenvektor von Z_{01} der auf allen $I \in Z_1$ die nicht unterteilt wurden mit ξ_1 übereinstimmt.



Dann ist nach Voraussetzung $\|S(f, Z_{01}, \xi_{01}) - \int f\| < \varepsilon$ und weiters $\|S(f, Z_1, \xi_1) - S(f, Z_{01}, \xi_{01})\| < \varepsilon$, denn höchstens N viele Intervalle I von Z_1 zerfallen in Z_{01} in Teilintervalle der Länge $\leq \delta$ und für diese gilt

$$\|S(f|_I, Z_1, \xi_1) - S(f|_I, Z_{01}, \xi_{01})\| \leq |I| \cdot \Omega(f) \leq \delta \cdot \Omega(f) \leq \frac{\varepsilon}{N}$$

nach obigen Lemma (5.6.6). □

5.6.8 Lemma. Integrieren vertauscht mit stetig linearen Abbildungen.
Es sei $f : \mathbb{R} \supseteq J \to E_1$ R-integrierbar auf dem kompakten Intervall $J = [a, b]$ und $\ell : E_1 \to E_2$ stetig und linear. Dann ist $\ell \circ f : J \to E_2$ R-integrierbar und

$$\int_J (\ell \circ f) = \ell\Big(\int_J f\Big).$$

Beweis. Der Beweis ist analog zu (6.1.6a). Da ℓ linear ist vertauscht es mit Riemann-Summen, und weil es stetig ist mit Limiten:

$$\int_J (\ell \circ f) = \lim_{|Z|\to 0} S(\ell \circ f, Z, \xi) \xlongequal{\ell \text{ linear}} \lim_{|Z|\to 0} \ell\Big(S(f, Z, \xi)\Big) \xlongequal{\ell \text{ stetig}} \ell\Big(\lim_{|Z|\to 0} S(f, Z, \xi)\Big) = \ell\Big(\int_J f\Big).$$ □

5.6.9 Lemma. Komponentenweise Integration.
Es sei $f = (f^1, f^2) : [a, b] \to E_1 \times E_2 =: E$. Dann gilt: $f : [a, b] \to E$ ist R-integrierbar $\Leftrightarrow$ $\forall i \in \{1, 2\}$: $f^i : [a, b] \to E_i$ ist R-integrierbar.
Weiters gilt dann:

$$\int_a^b f = \Big(\int_a^b f^1, \int_a^b f^2\Big).$$

Beweis. Dieser geht nun völlig analog zu dem entsprechenden Resultat für Konvergenz, Stetigkeit und Differenzierbarkeit, denn Riemann-Summen können offensichtlich koordinatenweise berechnet werden und die Konvergenz im Produkt ist ebenfalls koordinatenweise. □

Folgerung.
Es sei $f = (f^1, \dots, f^q) : [a, b] \to \mathbb{R}^q$. Dann gilt: $f : [a, b] \to \mathbb{R}^q$ ist R-integrierbar $\Leftrightarrow$ $\forall i \in \{1, \dots, q\}$: $f^i : [a, b] \to \mathbb{R}$ ist R-integrierbar.
Weiters ist dann:

$$\int_a^b f = \Big(\int_a^b f^1, \dots, \int_a^b f^q\Big).$$

Vergleiche dies mit der ad hoc Definition (5.3.8).

Beweis. Induktion und obiges Lemma (5.6.9). □

Wir können diese Folgerung nun verwenden um die meisten Resultate aus (5.1) und (5.2) auf Kurven zu übertragen – allen voran natürlich:

5.6.10 Folgerung. Lebesgue-Integrabilitätskriterium.
Es ist $f : [a, b] \to \mathbb{R}^q$ R-integrierbar genau dann, wenn f beschränkt und fast überall stetig ist (Dies verallgemeinert (5.1.4)).

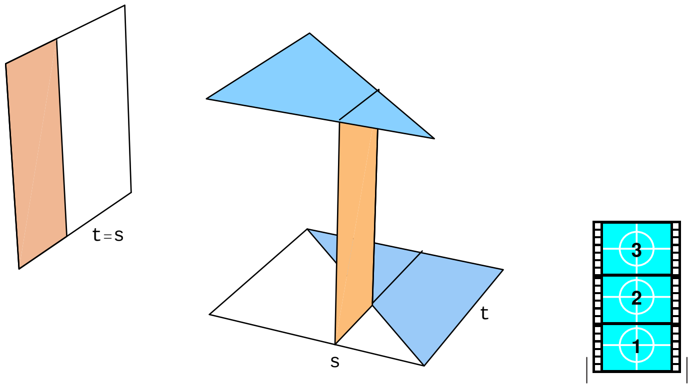
Beweis. Dies folgt sofort aus dem 1-dimensionalen Satz, da Stetigkeit, Beschränktheit und Integrierbarkeit komponentenweise festgestellt werden können. Beachte dazu, daß $\{x : f \text{ ist unstetig bei } x\} = \bigcup_i \{x : f^i \text{ ist unstetig bei } x\}$ gilt. □

Wir haben dieses Kriterium für viele Aussagen über D-Integrierbarkeit wesentlich verwendet. Um diese Resultate möglichst auch im Banach-Raum-wertigen Fall zu erhalten müssen wir nun anders vorgehen, denn folgendes Beispiel zeigt die nicht Übertragbarkeit des Lebesgue'schen Integrabilitätskriteriums.

5.6.11 Beispiel. Nicht Erweiterbarkeit des Integrabilitätskriteriums.
Es sei $f : [0,1] \to B([0,1],\mathbb{R})$ gegeben durch $f(t) := \chi_{[t,1]}$, also

$$f(t)(s) := \begin{cases} 1 & \text{für } t \le s \\ 0 & \text{andernfalls.} \end{cases}$$

Dann ist f R-integrierbar mit Integral $\int_0^1 f = \mathrm{id} \in B([0,1],\mathbb{R})$ aber f ist nirgends stetig.



Wäre f stetig bei $t_0 \in [0,1]$ so auch $\chi_{[0,t_0]} = \mathrm{ev}_{t_0} \circ f : [0,1] \to B([0,1],\mathbb{R}) \to \mathbb{R}$, denn $\mathrm{ev}_{t_0} : B([0,1],\mathbb{R}) \to \mathbb{R}$, $h \mapsto h(t_0)$ ist stetig. Oder direkter: $\|f(t_1) - f(t_0)\|_\infty = \|\chi_{[t_1,1]} - \chi_{[t_0,1]}\|_\infty = 1$ für $t_1 \neq t_0$.

Das Integral läßt sich auch leicht erraten. Denn wenn f integrierbar ist, dann auch $\chi_{[0,t]} = \mathrm{ev}_t \circ f$ ($\neq f(t) = \chi_{[t,1]}$) und $\mathrm{ev}_t\left(\int_0^1 f\right) = \int_0^1 \mathrm{ev}_t \circ f = \int_0^1 \chi_{[0,t]} = t$ nach (5.6.8).

Diese Funktion $\mathrm{id} : t \mapsto t$ ist in der Tat das Riemann-Integral $\int_0^1 f \in B([0,1],\mathbb{R})$, denn sei Z eine Zerlegung und ξ ein zugehöriger Zwischenvektor. Für $s \in [0,1]$ sei $k = k_s$ maximal gewählt mit $\xi_k \le s$. Dann ist $t_{k-1} \le \xi_k \le s < \xi_{k+1} \le t_{k+1}$ und somit ist

$$\left|\Big(\mathrm{id} - S(f,Z,\xi)\Big)(s)\right| = \left|s - \sum_{i \le k} 1 \cdot (t_i - t_{i-1})\right| = |s - t_k| \le 2|Z| \quad \Rightarrow \quad \|\mathrm{id} - S(f,Z,\xi)\|_\infty \le |Z|\delta \to 0 \qquad \square$$

5.6.12 Beispiel. Zusammensetzen integrierbarer Abbildungen.
In (5.1.5) haben wir gezeigt, daß die Zusammensetzung $g \circ f$ von integrierbaren Funktionen f mit stetig beschränkten Funktionen g selbst integrierbar ist. Daß dies selbst für $g : x \mapsto \|x\|$ im unendlich-Dimensionalen nicht mehr gilt zeigt folgendes Beispiel:

Es sei $f : [0,1] \to B(\mathbb{Q} \cap [0,1],\mathbb{R})$ gegeben durch $f(t) := \chi_t$, also

$$f(t)(s) := \begin{cases} 1 & \text{für } t = s \\ 0 & \text{andernfalls.} \end{cases}$$

Dann ist f R-integrierbar mit Integral 0, weiters ist $\|\text{\_}\|_\infty$ stetig und auf $f(I) \subseteq \{h : \|h\|_\infty \le 1\}$ beschränkt, aber $t \mapsto \|f(t)\|_\infty$ ist die Dirichletfunktion, und somit nicht R-integrierbar.

Analog zum vorigen Beispiel (5.6.11) erhalten wir als Kandidaten für $\int_0^1 f$ die Funktion $s \mapsto \int_0^1 (\mathrm{ev}_s \circ f) = 0$, da

$$\mathrm{ev}_s \circ f = \begin{cases} \chi_{\{s\}} & \text{für } s \in \mathbb{Q} \\ 0 & \text{andernfalls.} \end{cases}$$

Man zeigt wie zuvor in (5.6.11), daß die Riemann-Summen in der Supremums-Norm gegen diese Funktion $s \mapsto 0$ konvergieren.

Bei fixen t ist $f(t) = 0$ falls $t \notin \mathbb{Q}$ und $f(t) = \chi_{\{t\}}$ andernfalls, d.h.

$$\|f(t)\|_\infty = \begin{cases} 1 & \text{für } t \in \mathbb{Q} \\ 0 & \text{andernfalls.} \end{cases}$$

5.6.13 Proposition. Grundlegendes über das R-Integral.

(1) $f : [a, b] \to E$ stetig $\Rightarrow f$ R-integrierbar (Dies verallgemeinert (5.1.2)).

(2) R-Integrieren ist linear (Dies verallgemeinert (5.1.7)).

(3) $f : [a, b] \to E$ R-integrierbar, $[c, d] \subseteq [a, b] \Rightarrow f|_{[c,d]}$ R-integrierbar (Dies verallgemeinert (5.2.1)).

(4) $f : [a, c] \to E$ R-integrierbar $\Leftarrow f|_{[a,b]}$ und $f|_{[b,c]}$ R-integrierbar (Dies verallgemeinert (5.1.9)).

(5) Für $a \le b \le c$ gilt

$$\int_a^b f(t)\,dt + \int_b^c f(t)\,dt = \int_a^c f(t)\,dt,$$

wobei wir nur voraussetzen müssen, daß eine der beiden Seiten existiert.Für $a < b$ definieren wir folglich $\int_b^a f(t)\,dt := -\int_a^b f(t)\,dt$ und erhalten obige Additivität für alle $a, b, c \in \mathbb{R}$ (Dies verallgemeinert (5.1.9))

Beweis. (1) Sei $\varepsilon > 0$ gegeben. Da $[a, b]$ kompakt ist, ist f auf $[a, b]$ gleichmäßig stetig, also existiert ein δ mit $\|f(t) - f(s)\| < \varepsilon$ für $|t - s| < \delta$. Es sei Z eine Partition von $[a, b]$ mit $|Z| < \delta$ – dann ist $\Omega_f(Z) \le \varepsilon$ – und sei Z' eine Verfeinerung mit zugehörigen Zwischenvektoren ξ und ξ'. Dann ist

$$\|S(f, Z, \xi) - S(f, Z', \xi')\| \le |b - a|\,\Omega_f(Z) \le |b - a|\,\varepsilon$$

nach (5.6.6). Wenn also für zwei punktierte Zerlegungen Z^1 und Z^2 die Feinheitsmaße kleiner als δ sind, so gilt bezüglich einer gemeinsamen Verfeinerung Z:

$$\|S(f, Z^1) - S(f, Z^2)\| \le \|S(f, Z^1) - S(f, Z)\| + \|S(f, Z^2) - S(f, Z)\| \le 2\varepsilon|b - a|,$$

somit erfüllen die Riemann-Summen die Cauchy-Bedingung und konvergieren nach (5.6.4).

(2) Folgt sofort aus $S(f+\lambda\, g, Z, \xi) = S(f, Z, \xi)+\lambda\, S(g, Z, \xi)$, oder abstrakter: f, g R-integrierbar $\xRightarrow{(5.6.9)}$ (f, g) R-integrierbar $\xRightarrow{(5.6.8)}$ $f + \lambda\, g = h \circ (f, g)$ R-integrierbar, wobei $h : E \times E \to E$, $(x, y) \mapsto x + \lambda\, y$ stetig und linear ist. Weiters ist

$$\int_a^b (f + \lambda\, g) = \int_a^b h \circ (f, g) \overset{(5.6.8)}{=\!=\!=} h\Big(\int_a^b (f, g)\Big) \overset{(5.6.9)}{=\!=\!=} h\Big(\int_a^b f, \int_a^b g\Big) = \int_a^b f + \lambda \int_a^b g.$$

(3) Je zwei punktierte Zerlegungen Z_i von $[c, d]$ können durch fixe punktierte Zerlegungen von $[a, c]$ und $[d, b]$ zu Zerlegungen $\tilde{Z}_i$ von $[a, b]$ gemacht werden und dann gilt

$$S(f, \tilde{Z}_1) - S(f, \tilde{Z}_2) = S(f, Z_1) - S(f, Z_2),$$

also ist die Cauchy-Bedingung erfüllt und $\lim_{|Z|\to 0} S(f|_{[c,d]}, Z)$ existiert nach (5.6.4).

(4) Unter Verwendung der Definition können wir das wie folgt zeigen: Sei zu $\varepsilon > 0$ ein $\delta > 0$ derart gewählt, daß für Zerlegungen von $[a, b]$ und $[b, c]$ mit Feinheitsmaß $< \delta$ die entsprechenden Riemann-Summen die Integrale bis auf ε approximieren. Sei Z eine beliebige Zerlegung von $[a, c]$ mit Feinheit δ. Sei Z' die Zerlegung $Z \cup \{b\}$. Dann ist $\|\int_a^b f - S(f, Z' \cap [a, b], \xi^-)\| < \varepsilon$ und $\|\int_b^c f - S(f, Z' \cap$

$[b,c],\xi^+)\| < \varepsilon$ für beliebig gewählte Zwischenvektoren ξ^+ und ξ^-. Falls $Z = Z'$ (d.h. $b \in Z$) so ist $S(f,Z,\xi) = S(f,Z'\cap[a,b],\xi^-) + S(f,Z'\cap[b,c],\xi^+)$ mit zusammenpassenden Zwischenvektoren ξ^-, ξ^+ und ξ. Andernfalls sei $b \in I \in Z$ und $b \in I_+, I_- \in Z'$. Dann ist $S(f,Z,\xi) - |I|f(\xi_I) = S(f,Z'\cap[a,b],\xi^-) + S(f,Z'\cap[b,c],\xi^+) - |I_-|f(\xi^-_{I_-}) - |I_+|f(\xi^+_{I_+})$, wobei die Zwischenvektoren auf allen Intervallen ungleich I_-, I_+ und I zusammenpassend gewählt seien. Alle drei Ausdrücke der Form $|I'|f(\xi_{I'})$ haben Norm $\le |I'|\cdot\|f|_{I'}\|_\infty \le \delta\,\|f\|_\infty$, also existiert $\int_a^c f$ und ist gleich $\int_a^b f + \int_b^c f$.

Mittels der Beschreibung (5.6.7) kann dies kürzer bewiesen werden:

$$\begin{aligned}
&\xRightarrow{(5.6.7)} \exists\, Z_0^-, Z_0^+ \;\; \forall Z^- \supseteq Z_0^-, Z^+ \supseteq Z_0^+ \;\; \forall \xi^-,\xi^+ : \\
&\qquad \Big\|\int_a^b f - S(f,Z^-,\xi^-)\Big\| \le \varepsilon, \quad \Big\|\int_b^c f - S(f,Z^+,\xi_+)\Big\| < \varepsilon \\
&\Rightarrow \forall Z \supseteq Z_0 := Z_0^- \cup Z_0^+ \;\; \forall \xi : \Big\|\int_a^b f + \int_b^c f - S(f,Z,\xi)\Big\| \le \\
&\qquad \le \Big\|\int_a^b f - S(f,Z^-,\xi^-)\Big\| + \Big\|\int_b^c f - S(f,Z^+,\xi_+)\Big\| \le 2\varepsilon \\
&\qquad \text{mit } Z^- := Z\cap[a,b], \quad Z^+ := Z\cap[b,c].
\end{aligned}$$

(5) Wie bereits in (()3) und (()4) oben gezeigt zieht die Existenz einer Seite jene der anderen nach sich. Nun können wir den Beweis wie im reellen Fall (5.19) führen indem wir nur solche Zerlegungen Z von $[a,c]$ zu betrachten, die b als Teilungspunkt enthalten. Oder aber man verwendet den Satz von Hahn-Banach um die Aussage direkt auf den reellen Fall zurückzuführen. □

5.6.14 Proposition. Mittelwertsatz für integrierbare Kurven.

(1) $f : [a,b] \to E$ R-integrierbar, $f([a,b]) \subseteq A \subseteq E$ mit A abgeschlossen und konvex $\Rightarrow \int_a^b f \in (b-a)\cdot A$.

(2) $f : [a,b] \to E$ R-integrierbar $\Rightarrow \left\|\int_a^b f\right\| \le |b-a| \sup\Big\{\|f(t)\| : t \in [a,b]\Big\}$.

(3) $f : [a,b] \to \mathbb{R}^q$ R-integrierbar $\Rightarrow \left\|\int_a^b f\right\| \le \int_a^b \|f(x)\|\,dx$.

(4) $f : [a,b] \to E$ stetig $\Rightarrow \left\|\int_a^b f\right\| \le \int_a^b \|f(x)\|\,dx \le |b-a|\cdot\sup\Big\{\|f(t)\| : t\in[a,b]\Big\}$ (Verallgemeinert (5.1.11)).

(5) Der Raum $R([a,b],E)$ aller R-integrierbaren Kurven $[a,b] \to E$ ist abgeschlossen in $B([a,b],E)$ und somit vollständig.

(6) Integrieren $\int_a^b : R([a,b],E) \to E$, $f \mapsto \int_a^b f$ ist stetig .

Dabei heißt eine Teilmenge $A \subseteq E$ eines Vektorraums E KONVEX, wenn $a,b \in A$, $0 < \lambda < 1 \Rightarrow a + \lambda(b-a) \in A$.

Operationen, die wir für Elemente einer Menge bereits erklärt haben, können wir auch auf Mengen erweitern, also z.B. $\lambda\cdot A := \{\lambda\cdot a : a \in A\}$.

Den entsprechenden Satz (5.1.11) für das 1-dimensionale haben wir direkt aus der Monotonie des Integrals gefolgert. Diese macht im vektorwertigen Fall keinen Sinn mehr, und wie schon beim Mittelwertsatz (4.1.5) der Differentialrechnung dürfen wir nicht mehr erwarten, daß $\frac{1}{b-a}\int_a^b f$ in $f([a,b])$ liegt.

Beweis. (1) Beachte, daß $\frac{1}{b-a}S(f,Z,\xi) = \frac{1}{b-a}\sum_{I\in Z}|I|\,f(\xi_I) = \sum_{I\in Z}\frac{|I|}{b-a}f(\xi_I)$ in A liegt, da die Koeffizienten $|I|/(b-a) \ge 0$ sind und $\sum_{i\in Z}\frac{|I|}{b-a} = 1$ ist, also eine konvex-Kombination von Vektoren

$f(\xi_I) \in A$ vorliegt. Da A abgeschlossen vorausgesetzt wurde, ist folglich auch $\int_a^b f = \lim_{|Z|\to 0} S(f,Z) \in (b-a)\,A$.

(2) Da f nach (5.6.7) beschränkt ist, liegt $f([a,b])$ im abgeschlossenen Ball $B_{\|f\|_\infty}(0)$, also einer konvexen Menge, denn mit $\|\sum_i \lambda_i\, x_i\| \le \sum_i \|\lambda_i\, x_i\| = \sum_i \lambda_i \cdot \|x_i\| \le \sum_i \lambda_i\, r = r$ falls $\|x_i\| \le r$ und $\lambda_i \ge 0$ für alle i und $\sum_i \lambda_i = 1$ ist. Also ist $\int_a^b f \in (b-a)\, B_{\|f\|_\infty}(0)$ nach 1, also $\|\int_a^b f\| \le (b-a)\,\|f\|_\infty$.

(3) & (4) Falls f stetig ist, so ist auch $\|f\| : t \mapsto \|f(t)\|_\infty$ stetig (beachte, daß die Schreibweise $\|f\|$ analog zu $|f|$ ist, aber leicht als $\|f\|_\infty := \sup\{\|f(x)\| : x \in X\}$ verstanden werden kann) und somit R-integrierbar. Für $E = \mathbb{R}^n$ ist dies wegen dem Lebesgue'schen Integrabilitätskriterium (5.6.10) automatisch erfüllt. Für die Riemann-Summen gilt

$$\|S(f,Z,\xi)\| \le \sum_{I\in Z} |I| \cdot \|f(\xi_I)\| = S(\|f\|, \mathcal{Z}) \le |b-a|\, \|f\|_\infty.$$

Somit gilt die entsprechende Ungleichung auch für das Riemann-Integral.

(5) & (6) Es seien $f_n \in R([a,b],E)$ und $f_n \to f_\infty$ in $B([a,b],E)$. Wegen (2) ist $(\int_a^b f_n)_n$ eine Cauchy-Folge in E also konvergent. Sei $\varphi := \lim_{n\to\infty} \int_a^b f_n$. Zu $\varepsilon > 0$ existiert somit ein $n \in \mathbb{N}$ mit $\|f_n - f_\infty\|_\infty < \varepsilon$ und $\|\int_a^b f_n - \varphi\| < \varepsilon$ und weiters ein $\delta > 0$ mit $\|\int_a^b f_n - S(f_n, Z)\| < \varepsilon$ für alle punktierten Zerlegungen Z mit Feinheitsmaß $|Z| < \delta$. Somit ist

$$\begin{aligned}\|\varphi - S(f_\infty, Z)\| &\le \left\|\varphi - \int_a^b f_n\right\| + \left\|\int_a^b f_n - S(f_n,Z)\right\| + \underbrace{\|S(f_n,Z) - S(f_\infty,Z)\|}_{\le (b-a)\,\|f_n - f_\infty\|}\\ &\le \varepsilon + \varepsilon + (b-a)\,\varepsilon = (2+b-a)\,\varepsilon,\end{aligned}$$

also f_∞ integrierbar und $\int_a^b f_\infty = \lim_{n\to\infty} \int_a^b f_n$. □

5.6.17 Hauptsatz der Analysis für Kurven.
Es sei $f : [a,b] \to E$ stetig und $\varphi(x) := \int_a^x f$. Dann ist φ ist differenzierbar und $\varphi' = f$.

Es sei umgekehrt $f : \mathbb{R} \supseteq [a,b] \to E$ stetig differenzierbar (oder auch nur mit R-integrierbarer Ableitung), dann ist

$$\int_a^b f' = f(b) - f(a).$$

Dies verallgemeinert (5.2.2). Die Bedeutung der Integration vektorwertigen Kurven liegt vor allem in diesen Satz der uns in die Lage versetzt aus Eigenschaften der Ableitung f' auf solche der Funktion f zu schließen und nicht wie bei skalarwertigen Funktionen in der Flächenberechnung.

Beweis.

$$\varphi(x+h) - \varphi(x) - f(x)h = \int_x^{x+h} (f(t) - f(x))\, dt$$

$$\left\|\int_x^{x+h} (f(t) - f(x))\, dt\right\| \overset{(5.6.14.4)}{\le} |h|\, \max\Big\{\big\|f(t) - f(x)\big\| : t \in \overline{x, x+h}\Big\} \to 0$$

EXAM

Mittels Hahn-Banach und (4.6.9) kann die zweite Aussage sofort auf den reellen Satz zurückgespielt werden, denn danach genügt es

$$\ell(f(b) - f(a)) = \ell\Big(\int_a^b f'\Big) = \int_a^b \ell \circ f' = \int_a^b (\ell \circ f)'$$

zu zeigen, was nach dem klassischen Hauptsatz (5.2.2) klarerweise gilt.

Ein anderer Beweis für stetig differenzierbares f besteht darin, $f_1(t) := f(a) + \int_a^t f'(s)\, ds$ zu betrachten. Dann ist $f_1(a) = f(a)$ und $f_1' = f'$ nach dem 1. Teil. Nach (5.6.16) ist also $f_1 = f$. □

5.6.17a Bemerkung. Stetig einseitig differenzierbar impliziert differenzierbar.
Da wir im zweiten Teil des Beweises (für $g' = 0 \Rightarrow g$ konstant) nur die rechtsseitige Ableitung benötigen zeigt dies, daß ein stetiges $f : (a,b) \to E$, welches eine stetige rechtseitige Ableitung $f' : t \mapsto \lim_{s\to 0+} \frac{f(t+s)-f(t)}{s}$ besitzt, bereits (stetig) differenzierbar ist.

5.6.18 Folgerung. Mittelwertsatz der Differentialrechnung.
Es sei $f : [a,b] \to E$ (stetig) differenzierbar und sei $f'([a,b]) \subseteq A$ mit einer abgeschlossenen konvexen Teilmenge $A \subseteq E$. Dann liegt der Differenzenquotient $\frac{f(b)-f(a)}{b-a}$ in A.
Insbesonders ist $\|f(b) - f(a)\| \leq (b-a)\,\|f'\|_\infty$.

Beweis. Für f mit R-integrierbarer Ableitung ist $f(b) - f(a) = \int_a^b f' \in (b-a)\,A$ nach (5.6.17) und (5.6.14.1).

EXAM Für allgemeines differenzierbares f folgt das Resultat aus dem 1-dimensionalen, denn falls $\frac{f(b)-f(a)}{b-a} \notin A$ so existiert nach dem Satz von Hahn-Banach ein stetig lineares $\ell : E \to \mathbb{R}$ mit $\ell\left(\frac{f(b)-f(a)}{b-a}\right) \notin \ell(A)$. Andererseits ist aber

$$\ell\left(\frac{f(b)-f(a)}{b-a}\right) = \frac{(\ell\circ f)(b) - (\ell\circ f)(a)}{b-a} = (\ell\circ f)'(\xi) = \ell(f'(\xi)),$$

für ein $\xi \in [a,b]$ nach dem einfach zu beweisenden Pendant zu (5.6.8) für differenzierbare Kurven (siehe auch (6.1.5)).

Der Spezialfall folgt wenn man für A den abgeschlossenen Ball $B_{\|f'\|_\infty}(0)$ verwendet. $\square$

5.6.19 Substitutionsformel.
Es sei $f : \mathbb{R} \supseteq [\bar a, \bar b] \to F$ stetig, $g : [a,b] \to [\bar a, \bar b]$ stetig differenzierbar. Dann ist

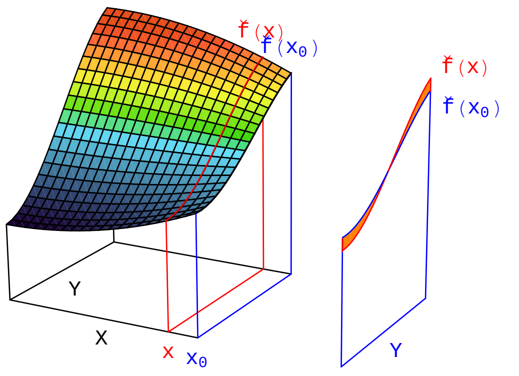
$$\int_a^b (f\circ g)\cdot g' = \int_{g(a)}^{g(b)} f.$$

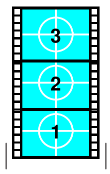
Dies verallgemeinert (5.2.3.5).

Beweis. Nach dem Hauptsatz (5.6.17) ist $\varphi(x) := \int_{g(a)}^x f(s)\,ds$ eine Stammfunktion von f auf $[\bar a, \bar b]$, d.h. $\varphi' = f$. Die Kettenregel (siehe auch (6.1.5)) für Funktionen $[a,b] \to [\bar a, \bar b] \to F$ liefert $(\varphi\circ g)' = (\varphi'\circ g)\cdot g' = (f\circ g)\cdot g'$, und damit

$$\int_a^b (f\circ g)\cdot g' = \int_a^b (\varphi\circ g)' \overset{(5.6.17)}{=\!=} (\varphi\circ g)|_a^b = \varphi(g(b)) - \varphi(g(a)) \overset{(5.6.17)}{=\!=} \int_{g(a)}^{g(b)} \varphi' = \int_{g(a)}^{g(b)} f. \quad \square$$

6.1.16 Proposition. Exponentialgesetz für stetige Funktionen.
Es seien X und I metrische Räume und I kompakt sowie E ein Banach-Raum. Dann ist $f : X \times I \to E$ genau dann stetig, wenn $\check f : X \to C(I,E)$ wohldefiniert und stetig ist.





Für eine Verallgeminerung dieses Resultats auf (gewisse) nicht-kompakte I siehe z.B. [15, 2.4.8].

Beweis. ($\Rightarrow$) Es sei $f : X \times I \to E$ stetig. Dann ist insbesonders für jedes $x \in X$ die Einschränkung $\check{f}(x) : I \to E$, $t \mapsto f(x,t)$ stetig, also $\check{f} : X \to C(I,E)$ wohldefiniert. Bleibt zu zeigen, daß $\check{f}$ stetig ist. Sei dazu $x_0 \in X$ und $\varepsilon > 0$. Da f stetig ist existiert für jedes $t_0 \in I$ ein $\delta_{t_0} > 0$ s.d. aus $d((x,t),(x_0,t_0)) = \max\{d(x,x_0), d(t,t_0)\} < \delta_{t_0}$ folgt, daß $\|f(x,t) - f(x_0,t_0)\| < \varepsilon$ ist. Die Familie $\{U_{\delta_t}(t) : t \in I\}$ ist eine Überdeckung mit offenen Bällen des kompakten Raums I, also existiert nach dem Satz (5.1.4b) von Heine-Borel eine endliche Teilüberdeckung $\{U_{\delta_{t_i}}(t_i) : 1 \le i \le N\}$. Sei $\delta := \min\{\delta_{t_i} : 1 \le i \le N\}$. Dann folgt aus $d(x,x_0) < \delta$ und $t \in T$ die Existenz eines $i \le N$ mit $d(t,t_i) < \delta_i$ und somit ist $\|f(x,t) - f(x_0,t_i)\| < \varepsilon$ und $\|f(x_0,t) - f(x_0,t_i)\| < \varepsilon$, also

$$\|f(x,t) - f(x_0,t)\| < \|f(x,t) - f(x_0,t_i)\| + \|f(x_0,t_i) - f(x_0,t)\| < 2\varepsilon$$

und folglich ist für alle $x \in X$ mit $d(x,x_0) < \delta$:

$$\|\check{f}(x) - \check{f}(x_0)\|_\infty = \sup\Big\{\|f(x,t) - f(x_0,t)\| : t \in I\Big\} \le 2\varepsilon.$$

($\Leftarrow$) Sei nun $\check{f} : X \to C(I,E)$ wohldefiniert und stetig, d.h. für jedes fixe $x \in X$ ist $\check{f}(x) : I \to E$, $t \mapsto f(x,t)$ stetig und $\check{f}(x)$ konvergiert gegen $\check{f}(x_0)$ für $x \to x_0$ gleichmäßig auf I. Mit anderen Worten: f ist einerseits partiell stetig in der zweiten Variable und andererseits stetig in der erstem Variable und zwar gleichmäßig bzgl. der zweiten. Nach (3.2.8) hat dies die Stetigkeit von f zur Folge.

□

6.1.17 Folgerung. Stetigkeit des Evaluierens.
Es sei I ein kompakter metrischer Raum und E ein Banach-Raum. Dann ist Evalualieren $\mathrm{ev} : C(I,E) \times I \to E$ *gegeben durch* $(f,x) \mapsto f(x)$ *stetig.*

Beweis. Nach (5.5.6) ist $C(I,E)$ ein Banach-Raum. Die zu ev assoziierte Abbildung $\check{\mathrm{ev}} : C(I,E) \to C(I,E)$ ist durch $\check{\mathrm{ev}}(f)(x) := \mathrm{ev}(f,x) = f(x)$ gegeben, also ist $\check{\mathrm{ev}}(f) = f$, d.h. $\check{\mathrm{ev}} = \mathrm{id}$ und dies ist stetig. Nach (6.1.16) ist somit ev stetig. □

6.1.18 Die Abbildung Untenstern.
Für metrische Räume X, Y und Z und stetige Abbildungen $f : Y \to Z$ sei $f_* : C(X,Y) \to C(X,Z)$ definiert durch $f_*(g) := f \circ g$, also durch Zusammensetzen mit f von links.

Folgerung. Stetigkeit von f_*.
Es sei I kompakt, E und F Banach-Räume, $U \subseteq E$ offen und $f : E \supseteq U \to F$ stetig. Dann ist $C(I,U)$ in $C(I,E)$ offen und $f_* : C(I,E) \supseteq C(I,U) \to C(I,F)$ stetig.

Wie benötigen dieses Lemma in (6.2.7) nur im Fall, wo $I \subseteq \mathbb{R}$ ein kompaktes Intervall ist.

Beweis. Beachte zuerst, daß $C(I,U)$ in $C(I,E)$ offen ist, denn wenn $c_0 \in C(I,U)$, dann ist $K := c_0(I)$ kompakt, da I es ist, und somit ist der Abstand $\delta := \inf\{\|x-y\| : x \notin U, y \in K\} > 0$, d.h. jedes c mit $\|c - c_0\| < \delta$ erfüllt $c(I) \subseteq U$: Die Funktion $d : y \mapsto \inf\{\|x-y\| : x \notin U\}$ ist stetig und positiv, denn für $|d(y_1) - d(y_2)| \le \|y_1 - y_2\|$, also nimmt sie auf der kompakten Menge K ein Minimum $\delta > 0$ an.

Nach (6.1.16) genügt zu zeigen, daß $(f_*)^\wedge : C(I,U) \times I \to U \to F$ stetig ist, was wegen

$$(f_*)^\wedge = f \circ \mathrm{ev} : (g,x) \mapsto f(g(x)) = (f \circ g)(x) = f_*(g)(x) = (f_*)^\wedge(g,x)$$

nach (6.1.17) offensichtlich ist. □

6.1.19 Folgerung. Stetigkeit der Variablenvertauschung.
Es sei I kompakt und E und F Banach-Räume. Die Abbildung $\tilde{\ } : C(I, L(E,F)) \to L(E, C(I,F))$, *definiert durch* $f \mapsto \tilde{f}(: x \mapsto (t \mapsto f(t)(x)))$ *– also Vertauschen von Variablen – ist linear und stetig.*

1. Beweis. Es sei $f \in C(I, L(E,F))$, dann ist $\tilde f(x) = \mathrm{ev}_x \circ f \in C(I,F)$ und $\tilde f$ ist linear, da $\mathrm{ev}_t \circ \tilde f = f(t) \in L(E,F)$. Weiters ist $\tilde f$ stetig, denn

$$\begin{aligned}\|\tilde f\| &= \sup\{\|\tilde f(x)\|_\infty : \|x\| = 1\} = \sup\{\|(\mathrm{ev}_x \circ f)(t)\| : \|x\| = 1,\, t \in I\}\\ &= \sup\{\|f(t)(x)\| : \|x\| = 1,\, t \in I\} = \sup\{\|f(t)\| : t \in I\} = \|f\|_\infty.\end{aligned}$$

Dies zeigt zugleich auch, daß die lineare Abbildung $\tilde{}$ Längen bewahrt, also insbesonders stetig ist mit Operatornorm 1. □

2. Beweis. Es ist $\tilde{} : C(I, L(E,F)) \to L(E, C(I,F))$ nach (5.5.13) genau dann linear und stetig, wenn die assoziierte Abbildung $C(I, L(E,F)) \times E \to C(I,F)$, $(f,x) \mapsto (t \mapsto f(t)(x))$ bilinear und stetig ist. Bilinearität ist offensichtlich und Stetigkeit ist nach (6.1.16) äquivalent mit jener der assoziierten Abbildung $C(I,L(E,F)) \times E \times I \to F$, $(f,x,t) \mapsto f(t)(x)$. Diese ist die Zusammensetzung $\mathrm{ev} \circ (\mathrm{ev} \times \mathrm{id}) : C(I, L(E,F)) \times E \times I \to L(E,F) \times E \to F$ und somit stetig nach (6.1.17) und (5.5.12), da $\|\mathrm{ev}(f,x)\| = \|f(x)\| \le \|f\| \cdot \|x\|$. □

Diese Abbildung ist injektiv, aber im Allgemeinen nicht surjektiv, denn sei $E = \ell^1$, $F = \mathbb{R}$ und $I = [0, 2\pi]$ und $g \in L(E, C(I,F))$ gegeben durch $\lambda = (\lambda_n)_{n=1}^\infty \mapsto \Big(t \mapsto \sum_n \lambda_n \sin(nt)\Big)$. Die Abbildung g ist linear und wohldefiniert, da die Reihe in $C(I,F)$ absolut konvergiert, und

$$\Big\| t \mapsto \sum_n \lambda_n \sin(nt) \Big\|_\infty \le \sum_n |\lambda_n|\, \|t \mapsto \sin(nt)\|_\infty = \|\lambda\|_\infty.$$

Sie liegt aber nicht im Bild von $\tilde{}$, denn dann wäre ihr Urbild $\tilde g : t \mapsto (\lambda \mapsto \sum_n \lambda_n \sin(nt))$ und würde $\frac{\pi}{2n}$ auf die Abbildung die e_n auf $\sin(\pi/2) = 1$ abbildet, also auf ein Abbildung $\tilde g(\frac{\pi}{2n})$ mit Norm ≥ 1. Hingegen ist $\tilde g(0) = 0$, also wäre $\tilde g$ nicht stetig.

6.1.21a Differenzierbarkeit von Kurven in $C(I, \mathbb{R})$.
Nach (6.1.16) ist für stetige $f : U \times I \to \mathbb{R}$ mit $U \subseteq \mathbb{R}$ offen die assoziierte Kurve $\check f : U \to C(I,\mathbb{R})$ stetig. Wir wollen nun ihre Differenzierbarkeit untersuchen. Dazu verwenden wir die stetig linearen Abbildungen $\mathrm{ev}_t : C(I,\mathbb{R}) \to \mathbb{R}$, $h \mapsto h(t)$ für $t \in I$. Falls die Ableitung $\check f'(x)$ existiert, so wäre nach dem Spezialfall (6.1.6a) der Kettenregel die Zusammensetzung $\mathrm{ev}_t \circ \check f : U \to C(I,\mathbb{R}) \to \mathbb{R}$ differenzierbar bei x mit Ableitung

$$(\mathrm{ev}_t \circ \check f)'(x) = \mathrm{ev}_t(\check f'(x)) = \check f'(x)(t).$$

Es ist $(\mathrm{ev}_t \circ \check f)(x) = \mathrm{ev}_t(\check f(x)) = \check f(x)(t) = f(x,t)$, also existiert die Ableitung von $\mathrm{ev}_t \circ \check f$ an der Stelle x genau dann, wenn für alle $t \in I$ die erste partielle Ableitung $\partial_1 f$ an der Stelle (x,t) existiert, und es ist $\check f'(x)(t) = (\mathrm{ev}_t \circ \check f)'(x) = \partial_1 f(x,t) = (\partial_1 f)^\vee(x)(t)$. Wenn wir zusätzlich wollen, daß $\check f$ stetig differenzierbar ist, so müssen wir also zusätzlich zur Existenz der ersten partiellen Ableitung $\partial_1 f$ nach (6.1.16) auch deren Stetigkeit fordern.

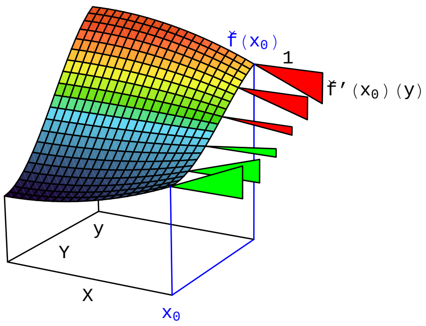
Um zu sehen, daß diese Anforderungen an f ausreichen für die stetige Differenzierbarkeit von $\check f$, betrachten wir nun:

$$\frac{\check f(x+h) - \check f(x)}{h}(t) = \frac{f(x+h,t) - f(x,t)}{h} \overset{(5.2.2)}{=\!=\!=} \frac{1}{h}\int_0^1 \frac{d}{ds} f(x + s\,h, t)\, ds = \int_0^1 \partial_1 f(x + s\,h, t)\, ds.$$

Da $\partial_1 f$ als stetig vorausgesetzt wurde ist $(h,s,t) \mapsto \partial_1 f(x + s\,h, t)$ stetig und somit nach (6.1.16) $h \mapsto \Big((s,t) \mapsto \partial_1 f(x + s\,h, t)\Big)$ stetig nach $C([0,1] \times I, \mathbb{R})$, d.h. $\partial_1 f(x + s\,h, t)$ konvergiert gegen $\partial_1 f(x + s\,0, t) = \partial_1 f(x,t)$ für $h \to 0$ und zwar gleichmäßig für $s \in [0,1]$ und $t \in I$. Wegen der Stetigkeit des Integrierens $\int_0^1 : C([0,1] \times I, \mathbb{R}) \to C(I,\mathbb{R})$ nach (5.1.12) konvergiert $\int_0^1 \partial_1 f(x + s\,h, t)\, ds$ gegen $\int_0^1 \partial_1 f(x,t)\, ds = \partial_1 f(x,t)$ und zwar gleichmäßig für $t \in I$. Dies zeigt die Differenzierbarkeit von $\check f$ im Punkte x mit Ableitung $\check f'(x) = (\partial_1 f)^\vee(x)$. Zusammengefaßt gilt also die

Proposition.
Es sei $U \subseteq \mathbb{R}$ offen und $f : U \times I \to \mathbb{R}$ eine stetige Abbildung. So ist $\check f : U \to C(I, \mathbb{R})$ genau dann

stetig differenzierbar, wenn $\partial_1 f : U \times I \to \mathbb{R}$ existiert und stetig ist. Unter diesen Voraussetzungen gilt $\check{f}'(x)(t) = \partial_1 f(x,t)$. □



5.6.20 Bemerkung. 1-dimensionale Mannigfaltigkeiten.
(Die Bilder) differenzierbare(r) Kurven $c : I \to E$ werden in der DIFFERENTIALGEOMETRIE als sogenannte 1-dimensionale Mannigfaltigkeiten, einem Spezialfall von mehrdimensionalen MANNIGFALTIGKEITEN (wie sie für die allgemeine Relativitätstheorie benötigt werden, also um unsere Welt im großen in Griff zu bekommen) behandelt und insbesonders Begriffe wie Krümmung, Torsion, Umlaufzahl diskutiert, siehe z.B. [14, Kapitel I].

In der ALGEBRAISCHEN TOPOLOGIE (siehe z.B. [22]) wird versucht unter anderen mittels Kurven Löcher in Räumen einzufangen und via geschlossener Kurven Räumen Gruppen zuzuordnen. Diese Gruppen sagen dann relativ viel über das Aussehen des Raumes aus. Bekanntestes Beispiel ist der JORDAN'SCHE KURVENSATZ oder schärfer das SATZ VON SCHÖNFLIES, welches besagt, daß jede EINFACH GESCHLOSSENE KURVE in $\mathbb{C}$, d.h. stetige injektive Abbildung vom Einheitskreis $S^1 := \{z \in \mathbb{C} : |z| = 1\} \to \mathbb{C}$ sich zu einer bijektiven stetigen Abbildung $f : \mathbb{C} \to \mathbb{C}$ mit stetiger Inversen erweitern läßt.

In der KNOTENTHEORIE werden einfach geschlossenen Kurven im $\mathbb{R}^3$ untersucht.

6 Differentiation im mehrdimensionalen

In diesen Kapitel wollen wir die Differentialrechnung von allgemeinen Abbildungen $f : E \to F$ zwischen Banach-Räumen E und F und damit insbesonders zwischen $\mathbb{R}^p$ und $\mathbb{R}^q$ entwickeln.

6.1 Differenzierbarkeit

6.1.1a Bemerkung. Stetigkeit für Abbildungen zwischen Banach-Räumen.
Eine Abbildung $f : E \supseteq U \to F$ zwischen Banach-Räumen E und F mit U offen in E ist nach (3.2.3) an der Stelle $x_0 \in U$ genau dann stetig, wenn

$$f(x_0) = \lim_{x \to x_0} f(x) = \lim_{h \to 0} f(x_0 + h)$$

gilt, wobei wir $h := x - x_0$ gesetzt haben. Analog zu unserer Untersuchung in (3.2.7) für Funktionen $f : \mathbb{R}^2 \to \mathbb{R}$ können wir h "polar"-zerlegen in $h = t\,v$ mit $t := \|h\| \geq 0$ und $v := \frac{1}{\|h\|}\,h$. Dann ist $\lim_{h\to 0} f(x_0 + h) = f(x_0)$ genau dann, wenn $f(x_0 + t\,v) \to f(x_0)$ für $t \to 0+$ gilt und zwar gleichmäßig bzgl. $\|v\| = 1$. Egal aus welcher Richtung v mit $\|v\| = 1$ wir uns also x_0 annähern, es muß also der Funktionswert gegen $f(x_0)$ konvergieren und zwar mit einer allen v gemeinsamen "Minimalgeschwindigkeit".

6.1.1b Definition. Richtungsableitung von Abbildungen.
Für $E = \mathbb{R}$ haben wir in (4.1.1) im Falle $F = \mathbb{R}$ und in (5.6.15) für allgemeines F die Ableitung als

$$f'(x_0) := \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{h \to 0} \frac{1}{h}\Big(f(x_0 + h) - f(x_0)\Big) \in F$$

definiert.

Für mehrdimensionales E können wir $\frac{f(x_0+h)-f(x_0)}{h}$ nicht mehr bilden, da wir durch Vektoren nicht dividieren können. Allerdings haben wir in (4.1.13) für $v \in E$ bereits die Richtungsableitung

$$d_v f(x_0) := \lim_{t \to 0} \frac{f(x_0 + t\,v) - f(x_0)}{t} \in F$$

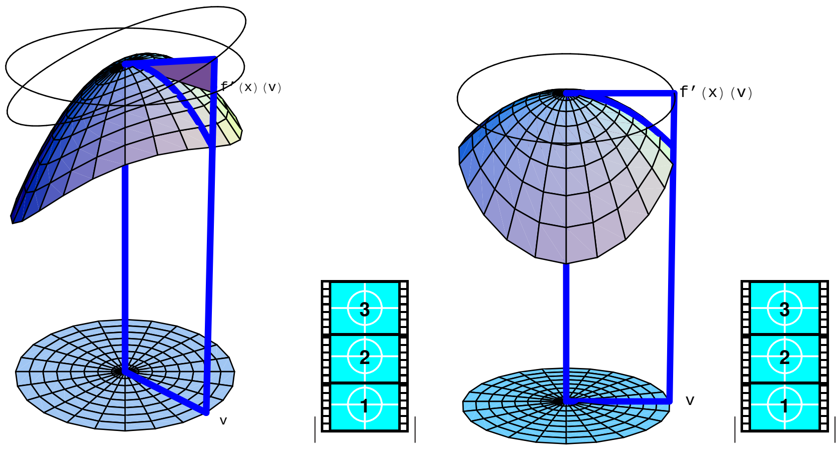
behandelt.

Wir wollen nun etwas allgemeiner für Banach-Räume E und F und $U \subseteq E$ offen die Richtungsableitung $d_v f(x_0)$ von $f : E \supseteq U \to F$ an der Stelle $x_0 \in U$ in Richtung $v \in E$ durch

$$d_v f(x_0) := \lim_{t \to 0+} \frac{f(x_0 + t\,v) - f(x_0)}{t} \in F$$

definieren. Beachte, daß die Frage nach der Existenz dieses Limes für alle $v \in E$ Sinn macht, denn für t nahe 0 ist $x_0 + t\,v \in U$ da $x_0 \in U$ und $U \subseteq E$ offen ist. Für diese Richtungsableitung gilt allgemein $d_{s\,v} f(x) = s\,d_v f(x)$ für $s \geq 0$ (d.h. $v \mapsto d_v f(x)$ ist $\mathbb{R}^+$-homogen), denn für $s = 0$ sind beide Seiten 0, und für $s > 0$ haben wir

$$d_{s\,v} f(x) = \lim_{t \to 0+} \frac{f(x + tsv) - f(x)}{t} = s \lim_{ts \to 0+} \frac{f(x + tsv) - f(x)}{ts} = s\,d_v f(x).$$



An einigen Beispielen haben wir in (4.1.13) gesehen, daß $v \mapsto d_v f(x_0)$ sogar linear ist und das bringt uns zu folgender

6.1.1c Vorbemerkung zur Ableitung.
Lineare Approximation: In (4.1.2) sahen wir, daß die durch die Ableitung beschriebene Tangente jene Gerade $g : h \mapsto f'(x_0) \cdot h$ ist, welche die nach 0 verschobene Funktion $\tilde f : h \mapsto f(x_0 + h) - f(x_0)$ gut approximiert, d.h.

$$\lim_{h\to 0} \frac{\tilde f(h) - g(h)}{h} = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h} - f'(x_0) = 0.$$

Diese Idee läßt sich nun auf Funktionen $f : E \to F$ zwischen Banach-Räumen übertragen. Dazu verschieben wir die Funktion ebenfalls so, daß der Punkt $(x_0, f(x_0)) \in E \times F$ im Nullpunkt zu liegen kommt, d.h. wir ersetzen x durch $x - x_0 =: h$ und y durch $y - f(x_0)$ und somit $f : x \mapsto f(x)$ durch die nach $(0,0)$ verschobene Funktion $\tilde f : h \mapsto x := x_0 + h \mapsto f(x) = f(x_0 + h) \mapsto f(x_0 + h) - f(x_0)$:

$$\begin{tikzcd} x \arrow[d, mapsto] & h + x_0 & E \arrow[r, "f"] \arrow[d, phantom, ""] & F \arrow[d, "\cong"] & y \arrow[d, mapsto] \\ x - x_0 & h \arrow[u, mapsto] & E \arrow[u, "\cong"] \arrow[r, "\tilde f"] & F & y - f(x_0) \end{tikzcd}$$

Die Ableitung $g : E \to F$ sollte nun jene lineare Funktion $g : E \to F$ sein, die $\tilde f$ auf die gleiche Weise wie in (4.1.2) approximiert, d.h. für welche

$$\lim_{h\to 0} \frac{\|(\tilde f - g)(h)\|}{\|h\|} = 0, \quad \text{oder äquivalent} \quad \lim_{h\to 0} \frac{(\tilde f - g)(h)}{\|h\|} = 0$$

ist.

Eindeutigkeit von g: Es kann nur eine lineare Abbildung g dieser Art geben, denn wenn wir bei der letztgenannten Grenzwertbildung die Variable h nur längs einer (Halb-)Geraden durch 0 variieren, d.h. nur h der Gestalt tv mit $\mathbb{R}^+ \ni t \to 0$ und fixen v betrachten. dann ist auch dieser eingeschränkte Grenzwert 0, d.h.

$$0 = \lim_{t\to 0+} \frac{f(x_0+tv) - f(x_0) - g(tv)}{\|tv\|} = \lim_{t\to 0+} \frac{1}{\|v\|} \Big(\frac{f(x_0+tv) - f(x_0)}{t} - g(v) \Big).$$

Multiplizieren wir dies mit der Konstanten $\|v\|$ so erhalten wir

$$g(v) = \lim_{t\to 0+} \frac{f(x_0+tv) - f(x_0)}{t} = (f_{x_0,v})'(0) = d_v f(x_0),$$

wobei $f_{x_0,v} : \mathbb{R} \to F$ die Abbildung f längs der (Halb-)Geraden $t \mapsto x_0 + tv$ ist, also $f_{x_0,v}(t) := f(x_0 + tv)$. Somit stimmt $g(v)$ mit der Richtungsableitung $d_v f(x_0)$ überein.

Stetigkeit von f versus jener von g: Beachte, daß f genau dann bei x_0 stetig ist, wenn g es ist, denn insbesonders muß der Zähler des obigen Grenzwertes auch gegen 0 gehen (da es der Nenner tut), d.h. $f(x+v) - f(x) - g(v) \to 0$ für $v \to 0$. Somit gilt $f(x+v) \to f(x)$ genau dann, wenn $g(v) \to 0$ für $v \to 0$.

Lineare Abbildungen $g : E \to F$ zwischen endlich dimensionalen Vektorräumen sind nach (5.5.8) automatisch stetig, für unendlich dimensionale Banach-Räume ist es nun naheliegend die Stetigkeit von g zu fordern.

Zusammengefaßt geben wir also folgende

6.1.1 Definition. Ableitung von Abbildungen.
Es sei $f : E \supseteq U \to F$ mit Banach-Räumen E und F und $U \subseteq E$ offen. Dann heißt f DIFFERENZIERBAR BEI $x_0 \in E$ falls eine stetige lineare Abbildung $g : E \to F$ existiert, welche die nach 0 verschobene Abbildung $\tilde{f} : h \mapsto f(x_0 + h) - f(x_0)$ im folgenden Sinn approximiert:

$$\lim_{h\to 0} \frac{\|\tilde{f}(h) - g(h)\|}{\|h\|} = 0, \text{ oder äquivalent auch } \lim_{h\to 0} \frac{1}{\|h\|}\Big(\tilde{f}(h) - g(h)\Big) = 0.$$

Die lineare Abbildung g ist, sofern sie existiert, eindeutig festgelegt (denn $g(v) = d_v f(x_0)$ für alle $v \in E$), heißt ABLEITUNG VON f BEI x_0 und wird mit $f'(x_0)$ bezeichnet.

Eine Abbildung $f : E \supseteq U \to F$ heißt DIFFERENZIERBAR, falls f bei allen Punkten $x_0 \in U$ differenzierbar ist. Ihre ABLEITUNG ist die Abbildung

$$f' : E \supseteq U \to L(E, F), \quad x_0 \mapsto f'(x_0).$$

Beachte, daß dies ein wesentlicher Unterschied zum Fall von Kurven (d.h. der Situation $E = \mathbb{R}$) ist, wo die Ableitung f' Werte im selben Raum wie f hat.

Bemerkung. Vergleich mit der 1-dimensionalen Situation.
Ist insbesonders $E = \mathbb{R}$, dann können wir die linearen (stetigen) Abbildungen $g : E \to F$ mit ihren Werten $g(1) \in F$ vermöge $g(v) = g(v \cdot 1) = v \cdot g(1)$ identifizieren:

$$\begin{aligned} L(\mathbb{R}, F) &\to F, \quad g \mapsto g(1) \\ L(\mathbb{R}, F) &\leftarrow F, \quad (t \mapsto tv) \leftarrow\!\shortmid v. \end{aligned}$$

In diesem Sinn ist die Ableitung $f'(x) \in L(E, F)$ einer Abbildung $f : \mathbb{R} \supseteq U \to F$ bei x als Element in $F \cong L(\mathbb{R}, F)$ auffaßbar. Dieses ist durch den (rechtsseitigen) Differentialquotienten

$$f'(x)(1) = d_1 f(x) = \lim_{t\to 0+} \frac{f(x+t) - f(x)}{t}$$

gegeben. Somit ist die hier gegebene Definition der Ableitung im Falle $E = \mathbb{R}$ gleichbedeutend mit (der rechtsseitigen Variante) jener aus (4.1.1) bzw. aus (5.6.15) für Kurven.

In Analogie zum Fall $E = \mathbb{R}$ schreibt man oft auch allgemein $g \cdot v$ an Stelle von $g(v)$ für lineare Abbildungen $g : E \to F$ und $v \in E$, insbesonders also $f'(x) \cdot v$ an Stelle von $f'(x)(v)$ für die Ableitung von f an der Stelle x in Richtung v.

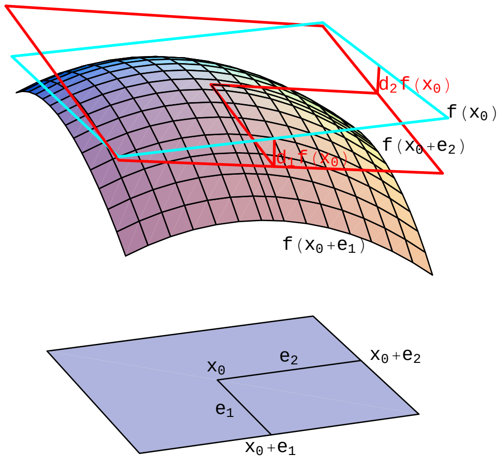
6.1.1d Bemerkung. Partielle Ableitungen.
Falls $E = \mathbb{R}^p$ ist, dann haben wir ausgezeichnete Richtungen, nämlich jene die durch die Standard-Basis $e_1, \dots, e_p \in \mathbb{R}^p$ gegeben sind. Folglich haben wir auch die ausgezeichneten Richtungsableitungen

$$\partial_i f(x) := d_{e_i} f(x) = \lim_{t\to 0+} \frac{f(x + t\, e_i) - f(x)}{t},$$

die wir als (rechtsseitige) PARTIELLE ABLEITUNGEN

$$\frac{\partial f}{\partial x_i}(x_1,\dots,x_p) = \frac{\partial}{\partial x_i} f(x_1,\dots,x_p)$$

bezeichnet haben.



Es ist

$$\frac{\partial f}{\partial x_i}(x_1,\dots,x_p) := \frac{d}{dt}\Big|_{t=0} f(x_1,\dots,x_{i-1},x_i+t,x_{i+1},\dots,x_p),$$

also die Ableitung von f an der Stelle x_i wobei alle anderen Koordinaten $x_1,\dots,x_{i-1},x_{i+1},\dots,x_p$ festgehalten sind.

Umgekehrt können wir eine allgemeine Richtungsableitung $d_v f(x)$ auch als 1-te partielle Ableitung der Funktion $g:(t,x,v)\mapsto f(x+t\,v)$ an der Stelle $(0,x,v)$ interpretieren, denn

$$d_v f(x) = \lim_{t\to 0+}\frac{f(x+t\,v)-f(x)}{t} = \lim_{t\to 0+}\frac{g(t,x,v)-g(0,x,v)}{t} = \partial_1 g(0,x,v) =: \frac{\partial}{\partial t}\Big|_{t=0} g(t,x,v).$$

Man schreibt oft etwas mißbräuchlich $\frac{d}{dt}|_{t=0}$ anstelle von $\frac{\partial}{\partial t}|_{t=0}$.

Daß die Existenz der Richtungsableitungen für die Differenzierbarkeit nicht ausreicht zeigt folgende Proposition.

6.1.2 Proposition. Differenzierbarkeit via Richtungsableitungen.
Eine Abbildung $f: E \supseteq U \to F$ ist genau dann differenzierbar bei $x \in U$, wenn folgende Punkte allesamt erfüllt sind:

(1) Die Richtungsableitung $d_v f(x)$ existiert für alle $v \in E$;

(2) $v \mapsto d_v f(x)$ ist linear und stetig;

(3) $\frac{f(x+tv)-f(x)}{t} \to d_v f(x)$ für $t \to 0+$ gleichmäßig bzgl. $\|v\| = 1$.

Verbal formuliert muß also der die Richtungsableitung definierende Grenzwert gleichmäßig bzgl. der Richtung existieren und linear und stetig von dieser abhängen.

Beweis. ($\Rightarrow$) Nach obigen Bemerkungen existieren für eine differenzierbare Abbildung f alle Richtungsableitungen und es ist $f'(x): v \mapsto d_v f(x)$ linear und stetig.

($\Leftarrow$) Wir setzen $g(v) := d_v f(x)$. Wenn wir die Polarzerlegung $h = t\,v$ mit $t := \|h\| \geq 0$ und $v := \frac{1}{\|h\|} h$ verwenden so übersetzt sich der die Ableitbarkeit definierende Limes wegen $g(h)/\|h\| = g(tv)/t = t\,g(v)/t = g(v)$ wie folgt:

$$0 = \lim_{h\to 0} \frac{f(x+h) - f(x) - g(h)}{\|h\|} = \lim_{t\to 0+, \|v\|=1} \Big(\frac{f(x+t\,v) - f(x)}{t} - g(v) \Big) \quad \Leftrightarrow$$

$$\Leftrightarrow \quad g(v) = \lim_{t\to 0+} \frac{f(x+t\,v) - f(x)}{t} \text{ glm. bzgl. } \|v\| = 1 \quad \square$$

6.1.3 Elementare Beispiele differenzierbarer Abbildungen.

(1) **Stetig lineare Abbildungen sind differenzierbar.** Nicht unerwartet ist jede stetig lineare Abbildung f differenzierbar mit $f'(x) = f$, denn es ist $f(x+v) - f(x) - f(v) = 0$. Die Ableitung f' einer stetig linearen Abbildung $f : E \to F$ ist somit die konstante Abbildung $f' : E \to L(E,F)$ die überall den Wert

$$f'(x) = f \in L(E,F)$$

besitzt (eine nicht konstante lineare Abbildung).

Gleichmäßigkeit der Richtungsableitungsbildung genügt nicht. Falls f nur linear und nicht notwendig stetig ist, dann existiert $d_v f(x) = f(v)$ und der Limes ist gleichmäßig, aber f ist nicht differenzierbar. Also können wir in der Definition der Ableitung die Stetigkeit von $f'(x)$ nach (6.1.1) nur durch die Stetigkeit von f bei x ersetzen. Sie aber nicht vollständig streichen.

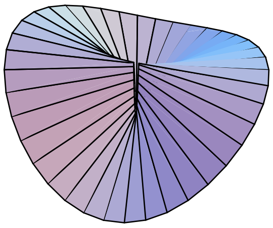
(2) Es sei $f(x,y,z) := x \cdot y^2 + \sin(z)$. Dann sind die partiellen Ableitungen

$$\begin{aligned}
\frac{\partial f(x,y,z)}{\partial x} &= \partial_1 f(x,y,z) = y^2 \\
\frac{\partial f(x,y,z)}{\partial y} &= \partial_2 f(x,y,z) = 2xy \\
\frac{\partial f(x,y,z)}{\partial z} &= \partial_3 f(x,y,z) = \cos(z)
\end{aligned}$$

Wir werden in (6.1.24) Argumente liefern, die zeigen, daß auch die Ableitung $f'(x,y,z) : \mathbb{R}^3 \to \mathbb{R}$ existiert und ihre Matrixdarstellung $[f'(x,y,z)] = (y^2, 2xy, \cos(z))$ ist, also gerade diese partiellen Ableitungen als Eintragung besitzt. Wir wollen dies denoch hier direkt nachrechnen. Daß die Ableitung $g := f'(x,y,z) : \mathbb{R}^3 \to \mathbb{R}$ tatsächlich durch $g(h,k,l) := y^2\,h + 2\,x\,y\,k + \cos(z)\,l$ gegeben ist, folgt aus

$$\begin{aligned}
&\lim_{(h,k,l)\to 0} \frac{f(x+h, y+k, z+l) - f(x,y,z) - g(h,k,l)}{\|(h,k,l)\|} = \\
&= \lim_{(h,k,l)\to 0} \frac{(x+h)\,(y+k)^2 + \sin(z+l) - x\,y^2 - \sin(z) - (y^2\,h + 2\,x\,y\,k + \cos(z)\,l)}{\sqrt{h^2+k^2+l^2}} \\
&= \lim_{(h,k,l)\to 0} \frac{x\,k^2 + 2\,h\,y\,k + h\,k^2 + \sin(z+l) - \sin(z) - \cos(z)\,l}{\sqrt{h^2+k^2+l^2}} \\
&= \lim_{(h,k,l)\to 0} \underbrace{(x\,k + 2\,h\,y + h\,k)}_{\to 0} \underbrace{\frac{k}{\sqrt{h^2+k^2+l^2}}}_{\text{beschränkt}} + \underbrace{\left(\frac{\sin(z+l) - \sin(z)}{l} - \cos(z) \right)}_{\to 0} \underbrace{\frac{l}{\sqrt{h^2+k^2+l^2}}}_{\text{beschränkt}} \\
&= 0
\end{aligned}$$

(3) **Partielle Ableitungen genügen nicht.** Es sei $f(x,y) := \sin(2\varphi)$ für $r \neq 0$, wobei (r,φ) die Polarkoordinaten von (x,y) sind, und $f(0) := 0$, d.h. $f(x,y) = 2\sin(\varphi)\cos(\varphi) = \frac{2xy}{x^2+y^2}$ für $(x,y) \neq 0$, siehe (3.2.7). Dann ist $\partial_1 f(0) = 0 = \partial_2 f(0)$, aber $\partial_{(1,1)} f(0)$ existiert nicht, da $t \mapsto f(t,t)$ unstetig ist bei 0. Insbesonders existiert also $f'(0)$ nicht.



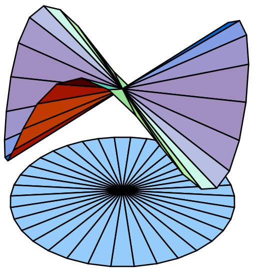
(4) **Richtungsableitungen genügen nicht.** Es sei

$$\begin{aligned} f(x,y) &:= r\,\sin(3\varphi) = r\Big(\sin(\varphi)\,\cos(2\varphi) + \cos(\varphi)\,\sin(2\varphi)\Big) \\ &= r^3\Big(3\sin(\varphi)\cos(\varphi)^2 - \sin(\varphi)^3\Big)/r^2 = \frac{3x^2y - y^3}{x^2+y^2} = \frac{3x^2 - y^2}{x^2+y^2}\,y. \end{aligned}$$

Beachte, daß f homogen ist und die Differenzenquotienten somit

$$\frac{f(tv) - f(0)}{t} = f(v) \text{ also } d_v f(0) = \lim_{t\to 0+} \frac{f(tv) - f(0)}{t} = f(v)$$

erfüllen. Es ist aber f nicht linear, denn $f(1,0) = 0$, $f(0,1) = -1$ aber $f(1,1) = 1$, und somit existiert $f'(0)$ nicht. Wir können also in der Definition der Differenzierbarkeit die Linearität von $f'(x)$ nicht weglassen.



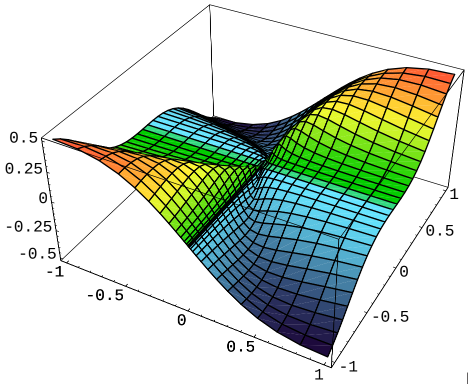
(5) **Linearität der Richtungsableitung genügt nicht.** Es sei $f(x,y) := \frac{xy^3}{x^2+y^4}$. Vergleiche dies mit $(x,y) \mapsto \frac{x\,y^2}{x^2+y^4}$ aus (3.2.8). Dann ist f stetig bei 0, denn

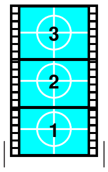
$$|f(x,y)| = \left|\frac{xy^2}{x^2+y^4}\,y\right| \leq \frac{|y|}{2},$$

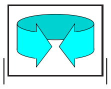
und

$$d_{(x,y)} f(0,0) = \frac{d}{dt}\Big|_{t=0} f(tx,ty) = \lim_{t\to 0} \frac{t^2xy^3}{x^2+t^2y^4} = 0,$$

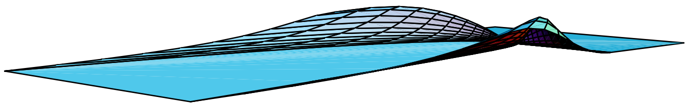
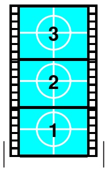
folglich ist $\partial_v f(0)$ linear (und stetig) in v aber dennoch existiert $f'(0)$ nicht, denn der Limes ist nicht gleichmäßig. Setze $x := t$, $y := \sqrt{1-t^2}$, dann ist der Limes $1/2$. Beachte, daß wir uns hier längs einer Ellipse an 0 annähern, sei nämlich $x_1 := tx - 1/2 = t^2 - 1/2$, $y_1 := ty = t\sqrt{1-t^2}$ und somit ist $x_1^2 + y_1^2 = 1/4$. Wir können in (6.1.2) also auch nicht die Gleichmäßigkeit des Limes weglassen.

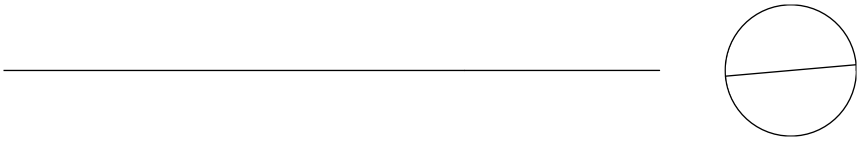
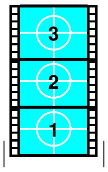






Ein weiteres Beispiel dafür, daß die Bedingung (3) in der Proposition (6.1.2) wirklich notwendig ist zeigt folgendes Beispiel einer Funktion $\mathbb{R}^2 \to \mathbb{R}$, wobei die zweite Animation den Schnitt mit einer sich um 0 drehenden vertikalen Ebene zeigt.

6.1.4 Lemma. Bilineare Abbildungen sind differenzierbar.
Es sei $f : E_1 \times E_2 \to F$ stetig und bilinear. Dann ist f differenzierbar mit Ableitung

$$f'(x_1, x_2)(v_1, v_2) := f(v_1, x_2) + f(x_1, v_2).$$

Beispiele bilinearer Abbildungen sind:

1. Die Multiplikation $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(t, s) \mapsto ts$;
2. Die Skalarmultiplikation $\mathbb{R} \times E \to E$, $(t, v) \mapsto tv$;
3. Die Multiplikation $A \times A \to A$, $(x, y) \mapsto x \bullet y$ einer Banach-Algebra wie z.B. die Komposition auf $L(E, E)$ oder die Faltung $\star$ auf ℓ^1 gegeben durch $(x \star y)(k) := \sum_{i+j=k} x(i)\, y(j)$;
4. Die Auswertung (oder Evaluation) $\mathrm{ev} : L(E, F) \times E \to F$, $(A, x) \mapsto A(x)$;
5. Die Komposition $\circ : L(F, G) \times L(E, F) \to L(E, G)$, $(T, S) \mapsto T \circ S$;
6. Die Determinante $\det : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$;
7. Das Kreuzprodukt $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$;
8. Das innere Produkt $\langle \_|\_ \rangle$ am $\mathbb{R}^n$ oder am ℓ^2, gegeben durch $\langle x|y \rangle = \sum_i x^i\, y^i$;
9. Allgemeiner die Form $4\langle x|y \rangle := \|x+y\|^2 - \|x-y\|^2$, falls die Norm die Parallelogramm-Gleichung $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$ erfüllt, d.h. E ein Hilbert-Raum ist. Es gilt dann die Cauchy-Schwarz Ungleichung $|\langle x|y \rangle| \leq \|x\|\, \|y\|$, siehe (2.2.3).

Beweis. Es ist

$$f\Big(\underbrace{(x_1,x_2)}_{=x}+\underbrace{(v_1,v_2)}_{=v}\Big)-f\underbrace{(x_1,x_2)}_{=x}-\Big(\underbrace{f(x_1,v_2)+f(v_1,x_2)}_{=:g(v)}\Big)=f\underbrace{(v_1,v_2)}_{=v}$$

und somit ist

$$\lim_{v\to 0}\frac{\|f(x+v)-f(x)-g(v)\|}{\|v\|}=\lim_{v\to 0}\frac{\|f(v)\|}{\|v\|}=0,$$

da $|f(v_1,v_2)|\le\|f\|\,\|v_1\|\,\|v_2\|\le\|f\|\,\|v\|^2$ nach (5.5.12).

Die hinter diesen Beweis stehende Idee ist, daß jede bilineare Abbildung f nach Definition partiell linear ist, also die partiellen Ableitungen

$$\partial_1 f(x_1,x_2)(v_1)=f(v_1,x_2)\quad\text{und}\quad\partial_2 f(x_1,x_2)(v_2)=f(x_1,v_2),$$

existieren. Aus (6.1.24) wird daraus ebenfalls das Resultat folgen, wenn wir die nachfolgende Kettenregel verwenden. □

6.1.5 Kettenregel.
Es sei $f:E\supseteq U\to F$ differenzierbar bei $x\in U$ und $g:F\supseteq V\to G$ differenzierbar bei $f(x)\in V$. Dann ist $g\circ f:E\supseteq U\supseteq f^{-1}(V)\to V\to G$ differenzierbar bei x und es gilt

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

Diese Resultat ist nicht überraschend, denn natürlich sollte die lineare Approximation einer Zusammensetzung $g\circ f$ die Zusammensetzung der entsprechenden linearen Approximationen an g und an f sein. Es ist dies eine Verallgemeinerung von (4.1.14).

Der nachfolgende Beweis ist eine leicht modifizierte Version von jenem in (4.1.14), wobei man an Stelle $\lim_{v\to 0}\frac{r(v)}{v}$ den Ausdruck $\lim_{v\to 0}\frac{1}{\|v\|}r(v)$ verwendet.

Beweis. Es bezeichne

$$\begin{aligned}r_f(v)&:=f(x+v)-f(x)-f'(x)(v)\\ r_g(w)&:=g(f(x)+w)-g(f(x))-g'(f(x))(w)\end{aligned}$$

dann ist $\lim_{v\to 0}r_f(v)/\|v\|=0$ und $\lim_{w\to 0}r_g(w)/\|w\|=0$ und wenn wir $w:=f(x+v)-f(x)=f'(x)(v)+r_f(v)$ setzten hat somit

$$\begin{aligned}r_{g\circ f}(v)&:=(g\circ f)(x+v)-(g\circ f)(x)-\Big(g'(f(x))\circ f'(x)\Big)(v)\\
&=g(f(x+v))-g(f(x))-\Big(g'(f(x))\circ f'(x)\Big)(v)\\
&=g'(f(x))\Big(f(x+v)-f(x)\Big)+r_g\Big(f(x+v)-f(x)\Big)-\Big(g'(f(x))\circ f'(x)\Big)(v)\\
&=g'(f(x))\Big(\underline{f'(x)(v)}+r_f(v)\Big)+r_g\Big(f'(x)(v)+r_f(v)\Big)-\underline{\Big(g'(f(x))\circ f'(x)\Big)(v)}\\
&=g'(f(x))(r_f(v))+r_g\Big(f'(x)(v)+r_f(v)\Big)\end{aligned}$$

die geforderte Eigenschaft:

$$\begin{aligned}\frac{\|r_{g\circ f}(v)\|}{\|v\|}&\le\frac{\|g'(f(x))(r_f(v))\|}{\|v\|}+\frac{\|r_g\big(f'(x)(v)+r_f(v)\big)\|}{\|f'(x)(v)+r_f(v)\|}\Big(\frac{\|f'(x)(v)\|}{\|v\|}+\frac{\|r_f(v)\|}{\|v\|}\Big)\\
&\le\|g'(f(x))\|\cdot\frac{\|r_f(v)\|}{\|v\|}+\frac{\|r_g\big(f'(x)(v)+r_f(v)\big)\|}{\|f'(x)(v)+r_f(v)\|}\Big(\|f'(x)\|+\frac{\|r_f(v)\|}{\|v\|}\Big)\\
&\to 0,\end{aligned}$$

denn nach (5.5.11) ist $\|a(x)\| \le \|a\| \cdot \|x\|$ für alle linearen a wobei $\|a\| := \sup\{\|a(x)\| : \|x\| \le 1\}$ die Operatornorm von a bezeichnet. □

Die Kettenregel stimmt auch für Richtungsableitungen, wenn die äußere Abbildung g als differenzierbar vorausgesetzt wird.

6.1.6 Folgerung. Spezialfall der Kettenregel.
Falls $f : E \supseteq U \to F$ differenzierbar bei $x \in U$ und $g : F \to G$ linear und stetig ist, so ist

$$(g \circ f)'(x) = g \,\circ\, f'(x).$$

Achtung: $g \,\circ\, f'(x)$ bedeutet nicht $(g \circ f')(x)$ sondern $g \circ (f'(x))$ nach der Metaregel, daß die näher beieinander stehenden f' und (x) vor den entfernteren g und f' berechnet werden sollen.

Beweis.

$$(g \circ f)'(x) \overset{(6.1.5)}{=\!=\!=} g'(f(x)) \circ f'(x) \overset{(6.1.3.1)}{=\!=\!=} g \circ f'(x). \quad \square$$

Bemerkung. Differenzierbarkeit bzgl. äquivalenter Normen.
Die Differenzierbarkeit einer Abbildung $f : E \supseteq U \to F$ ist für äquivalente Normen gleichbedeutend. Im Wertebereich ist das ohnehin klar und im Definitionsbereich auch leicht einzusehen. Es folgt aber auch aus der Kettenregel (6.1.5), wenn wir als eine der beiden Abbildungen jeweils die Identität bzg. der verschiedenen Normen verwenden.

6.1.7 Komponenten bzgl. der Werte der Ableitung.
Wir wollen nun die Resultate aus (2.3.4), (3.1.4) und (5.3.8) auch für das Differenzieren via Komponenten herleiten. Wir benötigen dies insbesonders für $F = F_1 \times F_2$ bzw. mittels Induktion für $F = \mathbb{R}^m \cong \mathbb{R}^{m-1} \times \mathbb{R} \cong \mathbb{R} \times \ldots \times \mathbb{R} \times \mathbb{R}$. Eine Abbildung wie z.B. die Ableitung $g := f'(x_0) : E \to F_1 \times \ldots \times F_m$ wird natürlich durch ihre Komponenten $g^1 : E \to F_1, \ldots,\ g^m : E \to F_m$ eindeutig festgelegt. Eine allgemeine Abbildung $g : E \to F_1 \times \ldots \times F_m$ ist genau dann linear resp. stetig, wenn die Komponenten $g^i := \mathrm{pr}_i \circ g : E \to F_i$ für $i \in \{1, \ldots, m\}$ für alle $i \in \{1, \ldots, m\}$ es sind (Ersteres zeigt man in der linearen Algebra, letzteres gilt, da die Konvergenz im Produkt die komponentenweise ist), d.h. $g \mapsto (g^1, \ldots, g^m)$ ist ein Isomorphismus

$$\begin{aligned} L(E, F_1 \times \ldots \times F_m) &\cong L(E, F_1) \times \ldots \times L(E, F_m), \\ g &\mapsto (\mathrm{pr}^1 \circ g, \ldots, \mathrm{pr}^m \circ g), \\ (v \mapsto (g^1(v), \ldots, g^m(v))) &\leftarrow\!\shortmid (g^1, \ldots, g^m) \end{aligned}$$

von Banach-Räumen, denn

$$\begin{aligned} \|g\| &= \sup\bigl\{\|g(v)\| : \|v\| \le 1\bigr\} = \sup\Bigl\{\max\bigl\{\|g^j(v)\| : 1 \le j \le m\bigr\} : \|v\| \le 1\Bigr\} \\ &= \sup\Bigl\{\|g^j(v)\| : 1 \le j \le m, \|v\| \le 1\Bigr\} \\ &= \max\Bigl\{\sup\bigl\{\|g^j(v)\| : \|v\| \le 1\bigr\} : 1 \le j \le m\Bigr\} = \max\{\|g^j\|\}. \end{aligned}$$

Hierbei haben wir verwendet, daß

$$\sup\Bigl\{h(x, y) : x \in X, y \in Y\Bigr\} = \sup\Bigl\{\sup\bigl\{h(x, y) : x \in X\bigr\} : y \in Y\Bigr\}$$

für Funktionen $g : X \times Y \to \mathbb{R}$ ist. Beachte, daß wenn wir mit $\mathrm{ins}_j : F_j \to F_1 \times \ldots \times F_m$ die j-te Insertionsabbildung $v \mapsto (0, \ldots, 0, v, 0, \ldots, 0)$ bezeichnen, so ist $g = \sum_{j=1}^m \mathrm{ins}_j \circ g^j$. Und $g \leftrightarrow (g^1, \ldots, g^m)$ ist bijektiv, denn $\sum_j \mathrm{ins}_j \circ \mathrm{pr}_j = \mathrm{id}_{F_1 \times \ldots \times F_m}$ und $\mathrm{pr}_j \circ \mathrm{ins}_j = \mathrm{id}$ sowie $\mathrm{pr}_i \circ \mathrm{ins}_j = 0$ für $j \ne i$.

Lemma. Komponentenweise Ableitung.
Es sei $f = (f^1, \ldots, f^m) : E \supseteq U \to F_1 \times \ldots \times F_m$ und $x \in U$. Dann ist f genau dann differenzierbar

bei x, wenn die Komponenten $f^j : E \supseteq U \to F_j$ differenzierbar bei x sind für alle $j \in \{1, \dots, m\}$. Für alle $v \in E$ gilt dann

$$f'(x)(v) = \Big((f^1)'(x)(v), \dots, (f^m)'(x)(v)\Big),$$

also

$$f'(x) = \Big((f^1)'(x), \dots, (f^m)'(x)\Big).$$

Damit die Memotechnik der Matrizenmultiplikation funktioniert, sollten wir Vektoren in $F_1 \times \ldots \times F_m$ besser als Spaltenvektoren schreiben. Wir sagen dann die "Matrix"- oder Komponenten-Darstellung $[f'(x)]$ von $f'(x)$ ist

$$[f'(x)] = \begin{pmatrix} (f^1)'(x) \\ \vdots \\ (f^m)'(x) \end{pmatrix}$$

Dies verallgemeinert (4.1.18).

Beweis. ($\Rightarrow$) Da $f^j := \mathrm{pr}_j \circ f$, wobei $\mathrm{pr}_j : F_1 \times \ldots \times F_m \to F_j$ die kanonische Projektion bezeichnet, folgt dies aus (6.1.6).

($\Leftarrow$) folgt, da das Restglied $r(h) := f(x+h) - f(x) - f'(x)(h)$ komponentenweise berechnet werden kann und Limiten in einem Produkt komponentenweise berechnet werden können. □

6.1.8 Folgerung. Linearität des Differenzierens.
Es seien $f, g : E \supseteq U \to F$ differenzierbar bei $x \in U$ und $\lambda \in \mathbb{R}$, dann ist auch $f + \lambda g$ differenzierbar bei x und es gilt $(f + \lambda\, g)'(x) = f'(x) + \lambda\, g'(x)$.

Beweis. Es ist $a : (f, g) \mapsto f + \lambda\, g$, $F \times F \to F$ stetig und linear und $(f, g) : E \to F \times F$ differenzierbar mit Ableitung $(f, g)'(x) = (f'(x), g'(x))$, also $(f + \lambda\, g)'(x) = (a \circ (f, g))(x) = a \circ (f, g)'(x) = a \circ (f'(x), g'(x)) = f'(x) + \lambda\, g'(x)$. □

Dieser Beweis entspricht dem 2-ten Beweis von (3.1.6) für Stetigkeit und verallgemeinert (4.1.19), siehe auch die Bemerkung zum Beweis in (4.1.19).

6.1.9 Folgerung. Leibniz'sche Produktregel.
Es sei $b : F_1 \times F_2 \to F$ stetig und bilinear, sowie $f_j : E \supseteq U \to F_j$ differenzierbar, dann ist $b \circ (f_1, f_2) : E \supseteq U \to F$ differenzierbar und

$$\Big(b \circ (f_1, f_2)\Big)'(x)(v) = b\Big(f_1'(x)(v), f_2(x)\Big) + b\Big(f_1(x), f_2'(x)(v)\Big).$$

Insbesonders gilt das für die Multiplikation b einer Banach-Algebra wie z.B. $\mathbb{R}$, $C(I, \mathbb{R})$ oder $L(E, E)$.

Ist insbesonders b die Multiplikation von $\mathbb{R}$, so lautet diese Regel für $f_1, f_2 : E \to \mathbb{R}$:

$$(f_1 \cdot f_2)'(x) = f_2(x) \cdot f_1'(x) + f_1(x) \cdot f_2'(x).$$

Ist $f_2(x) \neq 0$ so gilt auch

$$\left(\frac{f_1}{f_2}\right)'(x) = \frac{f_2(x)\, f_1'(x) - f_1(x)\, f_2'(x)}{f_2(x)^2}.$$

Dies verallgemeinert (4.1.15) und (4.1.17) mit einem wie in (4.1.15) angedeuteten Beweis.

Beachte, daß die Reihenfolge $f_2(x) \cdot f_1'(x)$ hier eher angebracht ist als $f_1'(x) \cdot f_2(x)$, denn $g := f_1'(x)$ ist eine (lineare) Abbildung und $\lambda := f_2(x)$ ein Skalar, also $\lambda \cdot g$ die Abbildung $v \mapsto \lambda \cdot g(v)$.

Beweis. Nach (6.1.4) ist b differenzierbar mit Ableitung $b'(y_1, y_2)(v_1, v_2) = b(v_1, y_2) + b(y_1, v_2)$. Somit ist

$$\begin{aligned}
(b \circ (f_1, f_2))'(x)(v) &= b'\Big((f_1, f_2)(x)\Big)\Big((f_1, f_2)'(x)(v)\Big) \\
&= b'\Big(f_1(x), f_2(x)\Big)\Big(f_1'(x)(v), f_2'(x)(v)\Big) \\
&= b\Big(f_1'(x)(v), f_2(x)\Big) + b\Big(f_1(x), f_2'(x)(v)\Big)
\end{aligned}$$

Insbesonders ist für $b(y_1, y_2) := y_1 \cdot y_2$:

$$\begin{aligned}(f_1 \cdot f_2)'(x)(v) &= f_1'(x)(v) \cdot f_2(x) + f_1(x) \cdot f_2'(x)(v)\\ &= (f_2(x) \cdot f_1'(x))(v) + (f_1(x) \cdot f_2'(x))(v) = \Big(f_2(x) \cdot f_1'(x) + f_1(x) \cdot f_2'(x)\Big)(v).\end{aligned}$$

Die Quotientenregel folgt nun, denn nach (4.1.3) ist $\frac{d}{dx}\frac{1}{x} = -\frac{1}{x^2}$ und somit nach Kettenregel (6.1.5) $\frac{d}{dx}\frac{1}{f_2(x)} = -\frac{1}{f_2(x)^2}\, f_2'(x)$ und schlußendlich

$$\begin{aligned}\left(\frac{f_1}{f_2}\right)'(x) &= \frac{d}{dx} f_1(x) \cdot \frac{1}{f_2(x)} + f_1(x) \cdot \frac{d}{dx}\frac{1}{f_2(x)} = \frac{f_1'(x)}{f_2(x)} - \frac{f_1(x)\, f_2'(x)}{f_2(x)^2}\\ &= \frac{f_2(x)\, f_1'(x) - f_1(x)\, f_2'(x)}{f_2(x)^2}. \quad \square\end{aligned}$$

6.1.10 Komponenten bzgl. der Argumente der Ableitung.
Nun wollen wir das zu (6.1.7) duale aber schwierigere (siehe z.B. (3.2.7)) Problem in Angriff nehmen. Wir benötigen dies insbesonders für $E = E_1 \times E_2$ und mittels Induktion für $E = \mathbb{R}^m \cong \mathbb{R}^{m-1} \times \mathbb{R} \cong \mathbb{R} \times \ldots \times \mathbb{R} \times \mathbb{R}$. Für differenzierbare Abbildungen $f : E_1 \times \ldots \times E_m \to F$ ist die Ableitung bei x eine lineare Abbildung $g := f'(x) : E_1 \times \ldots \times E_m \to F$, also ist

$$\begin{aligned}g(v_1, \ldots, v_m) &= g\Big((v_1, 0, \ldots, 0) + (0, v_2, 0, \ldots, 0) + \cdots + (0, \ldots, 0, v_m)\Big)\\ &= \underbrace{g(v_1, 0, \ldots, 0)}_{=:g_1(v_1)} + \underbrace{g(0, v_2, 0, \ldots, 0)}_{=:g_2(v_2)} + \cdots + \underbrace{g(0, \ldots, 0, v_m)}_{=:g_m(v_m)},\end{aligned}$$

d.h. $g = \sum_j g_j \circ \mathrm{pr}_j$, wobei die Komponenten $g_j : E_j \to F$ als "Einschränkungen" $g_j = g \circ \mathrm{ins}_j : E_j \hookrightarrow E_1 \times \ldots \times E_m \to F$ von g linear und stetig sind. Die natürlichen INSERTIONen oder auch INJEKTIONen sind dabei durch

$$\mathrm{inj}_j = \mathrm{ins}_j : E_j \to E_1 \times \ldots \times E_m, \quad v \mapsto (0, \ldots, 0, v, 0, \ldots, 0)$$

gegeben. Allgemein ist eine Abbildung $g : E_1 \times \ldots \times E_m \to F$ gegeben durch $g(v_1, \ldots, v_m) := \sum_{j=1}^m g_j(v_j)$ genau dann linear bzw. stetig, wenn die Komponenten $g_j := g \circ \mathrm{ins}_j : E_j \to F$ für alle $j \in \{1, \ldots, m\}$ es sind (Ersteres zeigt man in der linearen Algebra und letzteres folgt aus der Stetigkeit der Addition $F \times F \to F$). Wir haben insbesonders einen Isomorphismus (von Banach-Räumen wenn wir auf $E_1 \times \ldots \times E_m$ die Norm $\|(v_1, \ldots, v_m)\| = \sum_{j=1}^m \|v_j\|$ verwenden):

$$\begin{aligned}L(E_1 \times \ldots \times E_m, F) &\cong L(E_1, F) \times \ldots \times L(E_m, F),\\ &\qquad g \mapsto (g \circ \mathrm{ins}_1, \ldots, g \circ \mathrm{ins}_m) = (g_1, \ldots, g_m)\\ ((v^1, \ldots, v^m) \mapsto \sum_j g_j(v^j)) &\mapsfrom (g_1, \ldots, g_m)\end{aligned}$$

denn

$$\begin{aligned}\|g\| &= \sup\Big\{\|g(v_1, \ldots, v_m)\| : \sum_{j=1}^m \|v_j\| \le 1\Big\}\\ &= \sup\Big\{\Big\|\sum_j g_j(v_j)\Big\| : \sum_{j=1}^m \|v_j\| \le 1\Big\}\\ &\le \sup\Big\{\sum_j \|g_j\|\,\|v_j\| : \sum_{j=1}^m \|v_j\| \le 1\Big\}\\ &\le \sup\Big\{\max\{\|g_j\| : j\} \cdot \sum_j \|v_j\| : \sum_{j=1}^m \|v_j\| \le 1\Big\}\\ &= \max\big\{\|g_j\| : j\big\} \le \|g\|, \text{ wegen } \|g_j\| = \|g \circ \mathrm{ins}_j\| \le \|g\| \cdot \|\mathrm{ins}_j\| = \|g\|.\end{aligned}$$

Dies führt zu folgender

Definition. Partielle Ableitungen.
Es sei $f : E_1 \times \ldots \times E_m \supseteq U \to F$ eine Abbildung und $x = (x^1, \ldots, x^m) \in U$. Unter der j-ten PARTIELLEN ABLEITUNG $\partial_j f(x)$ von f bei x versteht man die Ableitung der Funktion $f \circ \mathrm{ins}_j^x : E_j \supseteq (\mathrm{ins}_j^x)^{-1}(U) \to F$ bei x^j, wobei die AFFINE INSERTIONSABBILDUNG durch

$$\mathrm{ins}_j^x : E_j \to E_1 \times \ldots \times E_m, \quad z \mapsto (x^1, \ldots, x^{j-1}, z, x^{j+1}, \ldots, x^m)$$

gegeben ist, d.h. ins_j^x ist affin mit linearen Teil $\mathrm{ins}_j := \mathrm{ins}_j^0$ und konstanten Teil $(\ldots, x^{j-1}, 0, x^{j+1}, \ldots)$. Vermöge einer Translation um $-x^j$ nach 0 können wir $\partial_j f(x)$ auch als Ableitung von $(f \circ (\mathrm{ins}_j + x))$ an der Stelle 0 definieren.

Im Spezialfall $E_1 = \cdots = E_m = \mathbb{R}$ entspricht dies der in (6.1.1) gegebenen Definition der partiellen Ableitungen wenn man den Isomorphismus $L(E_i, F) = L(\mathbb{R}, F) \cong F$ verwendet.

6.1.11 Folgerung. Differenzierbar impliziert partiell-differenzierbar.
Es sei $f : E_1 \times \ldots \times E_m \supseteq U \to F$ differenzierbar bei $x = (x^1, \ldots, x^m) \in U$. Dann existieren alle partielle Ableitungen $\partial_j f(x)$ für $j = 1, \ldots, m$ und es ist

$$f'(x)(h^1, \ldots, h^m) = \sum_{j=1}^{m} \partial_j f(x)(h^j).$$

Wir sagen auch die Matrixdarstellung $[f'(x)]$ von $f'(x)$ ist

$$[f'(x)] = \Big(\partial_1 f(x), \ldots, \partial_m f(x)\Big)$$

Beweis. Wegen der Kettenregel (6.1.5) ist

$$\begin{aligned}\partial_j f(x)(h^j) &= (f \circ \mathrm{ins}_j^x)'(x^j)(h^j) \\ &= \Big(f'(\mathrm{ins}_j^x(x^j)) \circ (\mathrm{ins}_j^x)'(x^j)\Big)(h^j) = f'(x)(\mathrm{ins}_j^0(h^j)),\end{aligned}$$

denn die Ableitung der durch $h^j \mapsto (x^1, \ldots, x^{j-1}, h^j, x^{j+1}, \ldots, x^m)$ gegebenen affinen Funktion $\mathrm{ins}_j^x : E_j \to E_1 \times \ldots \times E_m$ ist der lineare Anteil

$$\mathrm{ins}_j : h^j \mapsto (0, \ldots, 0, h^j, 0, \ldots, 0);$$

oder auch

$$\partial_j f(x^j)(h^j) = (f \circ (\mathrm{ins}_j + x))'(0)(h^j) = \Big(f'(0 + x) \circ (\mathrm{ins}_j'(0) + 0)\Big)(\mathrm{ins}_j(h^j)) = f'(x)(h^j). \quad \square$$

Bemerkung.
Wir haben in (6.1.3.3) und (6.1.3.4) gesehen, daß die Umkehrung hier nicht gilt, d.h. wir aus der Existenz der partiellen Ableitungen nicht auch die Differenzierbarkeit von f schließen können. Siehe jedoch (6.6.5) für die Äquivalenz unter stärkeren Voraussetzungen.

6.1.12 Komponenten bzgl. Argumente und Werte der Ableitung.
Aus der linearen Algebra wissen wir, daß sich lineare Abbildungen $g : \mathbb{R}^m \to \mathbb{R}^n$ durch Matrizen $[g] = (g_i^j)_{i=1,\ldots,m;j=1,\ldots,n}$ beschreiben lassen, dabei sind die Spalten $(g_i^j)_{j=1}^n$ gerade die Bilder $g(e_i)$ der standard-Basisvektoren e_i und $g(x) = \big(\sum_{i=1}^m g_i^j x^i\big)_{j=1,\ldots,n}$ für $x = (x^i)_{i=1}^m$.

Wir können nun genauso vorgehen um (stetig) lineare Abbildungen $g : E_1 \times \ldots \times E_m \to F_1 \times \ldots \times F_n$ in einfachere Komponenten $g_i^j : E_i \to F_j$ zu zerlegen, d.h. wir kombinieren (6.1.7) und (6.1.10) zu einen zusammengesetzten Isomorphismus

$$\begin{aligned} L(E_1 \times \ldots \times E_m, F_1 \times \ldots \times F_n) &= L(E_1 \times \ldots \times E_m, \prod_{j=1}^n F_j) \\ &\cong \prod_{j=1}^n L(E_1 \times \ldots \times E_m, F_j) \cong \prod_{j=1}^n \prod_{i=1}^m L(E_i, F_j). \end{aligned}$$

Für lineare Abbildungen $g : E_1 \times \ldots \times E_m \to F_1 \times \ldots \times F_n$ sind die Komponenten g_i^j durch

$$\begin{tikzcd} (\ldots,0,x^i,0,\ldots) & E_1 \times \ldots \times E_m \arrow[r, "g"] \arrow[dr, "g^j"] & F_1 \times \ldots \times F_n \arrow[d, "\mathrm{pr}_j"] & g(\ldots,0,x^i,0,\ldots) \arrow[d, mapsto] \\ x^i \arrow[u, mapsto] & E_i \arrow[u, "\mathrm{inj}_i"] \arrow[r, "g_i^j"'] & F_j & g^j(\ldots,0,x^i,0,\ldots) \end{tikzcd}$$

definiert. Dann ist $g(x) = \Big(\mathrm{pr}_1(g(x)), \ldots, \mathrm{pr}_n(g(x))\Big)$ und

$$\mathrm{pr}_j(g(x)) = \mathrm{pr}_j\Big(g\Big(\sum_i \mathrm{inj}_i(x^i)\Big)\Big) = \sum_i \big(\mathrm{pr}_j \circ g \circ \mathrm{inj}_i\big)(x^i) =: \sum_i g_i^j(x^i),$$

d.h.

$$g(x) = \begin{pmatrix} g_1^1 & \ldots & g_m^1 \\ \vdots & \ddots & \vdots \\ g_1^n & \ldots & g_m^n \end{pmatrix} \cdot \begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix}.$$

In Analogie zum skalaren Fall $E_1 = \cdots = E_m = \mathbb{R} = F_1 = \cdots = F_n$ nennen wir $(g_i^j)_{i=1,\ldots,m;j=1,\ldots,n}$ ebenfalls Matrixdarstellung oder Komponentendarstellung $[g]$ von g. Somit erhalten wir:

Folgerung. Jacobi-Matrix.
Es sei $f = (f^1, \ldots, f^m) : E_1 \times \ldots \times E_n \supseteq U \to F_1 \times \ldots \times F_m$ differenzierbar bei $x \in U$, dann existieren für alle i und alle j die partiellen Ableitungen $\partial_i f^j$ der j-ten Komponente f^j und die Matrixdarstellung $[f'(x)]$ von $f'(x)$ ist gegeben durch

$$\begin{pmatrix} \partial_1 f^1(x) & \ldots & \partial_n f^1(x) \\ \vdots & \ddots & \vdots \\ \partial_1 f^m(x) & \ldots & \partial_n f^m(x) \end{pmatrix},$$

d.h. für $v = (v^1, \ldots, v^n) \in E_1 \times \ldots \times E_n$ ist

$$f'(x)(v) = \begin{pmatrix} \partial_1 f^1(x) & \ldots & \partial_n f^1(x) \\ \vdots & \ddots & \vdots \\ \partial_1 f^m(x) & \ldots & \partial_n f^m(x) \end{pmatrix} \cdot \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

Insbesonders heißt diese Matrix im Falle $E_i = \mathbb{R} = F_j$ für alle i und j Jacobi-Matrix von f an der Stelle x.

Beweis. Das Resultat folgt durch Anwenden von (6.1.7) und dann (6.1.12) auf jede der Komponenten

f^j, denn

$$
\begin{aligned}
f'(x)(v) &\xlongequal{(6.1.7)} \Big((f^1)'(x)(v), \dots, (f^m)'(x)(v)\Big)\\
&\xlongequal{(6.1.12)} \Big(\sum_i \partial_i f^1(x)(v^i), \dots, \sum_i \partial_i f^m(x)(v^i)\Big)\\
&= \begin{pmatrix} \partial_1 f^1(x) & \dots & \partial_n f^1(x)\\ \vdots & \ddots & \vdots \\ \partial_1 f^m(x) & \dots & \partial_n f^m(x)\end{pmatrix} \cdot \begin{pmatrix} v^1 \\ \vdots \\ v^n\end{pmatrix} \qquad \square
\end{aligned}
$$

Wir hätten natürlich gerne die Umkehrung zur Verfügung, aber (6.1.3) zeigt, daß diese nicht allgemein gilt, siehe jedoch (6.1.24)

6.1.13 Stetigkeit der Ableitung.

Wie wir in (4.1.4) gesehen haben, ist das Verschwinden der Ableitung $f'(x_0)$ eine notwendige Bedingung dafür, daß $x_0 \in U$ ein Stelle ist, wo $f : E \supseteq U \to \mathbb{R}$ ein Minimum oder Maximum hat. Um allerdings entscheiden zu können, ob ein solches nun wirklich vorliegt wurde in der Schule die zweite Ableitung benutzt. Wir wollen nun den Weg ebnen höhere Ableitungen zu definieren. Gleichzeitig werden wir Bedingungen finden, die uns erlauben aus der Existenz der Richtungs- oder der partiellen Ableitungen auf die Differenzierbarkeit schließen zu können.

Falls $f : E \supseteq U \to F$ differenzierbar ist, so existiert für jeden Punkte $x \in U$ die Ableitung $f'(x) \in L(E, F)$ von f bei x. Wir erhalten somit eine Abbildung $f' : E \supseteq U \to L(E, F)$.

Unter der zweiten Ableitung werden wir in (6.3.1) (im wesentlichen) die Ableitung dieser Abbildung verstehen.

Wir müssen also von Differenzierbarkeit oder zumindest Stetigkeit von Abbildungen $g : E \supseteq U \to L(E, F)$ sprechen können. Da E und F als Banach-Räume vorausgesetzt sind, ist auch die Menge $L(E, F)$ aller stetig linearen Abbildungen von E nach F nach (5.5.11) ein Banach-Raum.

6.1.14 Definition. Stetig differenzierbar.

Es heißt $f : E \supseteq U \to F$ STETIG DIFFERENZIERBAR (kurz C^1), falls f differenzierbar ist (d.h. $f'(x)$ für alle $x \in U$ existiert) und die Ableitung $f' : E \supseteq U \to L(E, F)$ stetig ist.

Beachte, daß die Kettenregel (6.1.5) und somit auch die daraus hergeleitete Produktregel auch für stetig differenzierbare Abbildungen gilt, denn die Ableitung der Zusammensetzung $f \circ g$ ist nach (6.1.5) durch $(f \circ g)'(x) = f'(g(x)) \circ g'(x)$ gegeben. Also ist $(f \circ g)' = \mathrm{komp} \circ (f' \circ g, g')$ und sowohl $f' \circ g$ als auch g' sind nach Voraussetzung stetig und $\mathrm{komp} : L(F, G) \times L(E, F) \to L(E, G)$, $(a, b) \mapsto a \circ b$ nach (6.1.4).

Zur Abschreckung den klassischen Beweis dieses Sachverhalts im Spezialfall $E = F = G = \mathbb{R}^2$, d.h. für $f = (f_1, f_2) : \mathbb{R}^2 \to \mathbb{R}^2$ und $g = (g_1, g_2) : \mathbb{R}^2 \to \mathbb{R}^2$ Nach (6.1.11) ist

$$
[g'(x)] = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x_1, x_2) & \frac{\partial g_1}{\partial x_2}(x_1, x_2)\\ \frac{\partial g_2}{\partial x_1}(x_1, x_2) & \frac{\partial g_2}{\partial x_2}(x_1, x_2)\end{pmatrix}
$$

$$
[f'(y)] = \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(y_1, y_2) & \frac{\partial f_1}{\partial y_2}(y_1, y_2)\\ \frac{\partial f_2}{\partial y_1}(y_1, y_2) & \frac{\partial f_2}{\partial y_2}(y_1, y_2)\end{pmatrix}
$$

und nach der Kettenregel (6.1.5) die Ableitung der Zusammensetzung somit

$$
\begin{aligned}
[f'(g(x))] \cdot [g'(x)] &= \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(g_1(x_1, x_2), g_2(x_1, x_2)) & \frac{\partial f_1}{\partial y_2}(g_1(x_1, x_2), g_2(x_1, x_2))\\ \frac{\partial f_2}{\partial y_1}(g_1(x_1, x_2), g_2(x_1, x_2)) & \frac{\partial f_2}{\partial y_2}(g_1(x_1, x_2), g_2(x_1, x_2))\end{pmatrix} \cdot \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x_1, x_2) & \frac{\partial g_1}{\partial x_2}(x_1, x_2)\\ \frac{\partial g_2}{\partial x_1}(x_1, x_2) & \frac{\partial g_2}{\partial x_2}(x_1, x_2)\end{pmatrix}\\
&= \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(g_1(x_1,x_2),g_2(x_1,x_2)) \cdot \frac{\partial g_1}{\partial x_1}(x_1,x_2) + \frac{\partial f_1}{\partial y_2}(g_1(x_1,x_2),g_2(x_1,x_2)) \cdot \frac{\partial g_2}{\partial x_1}(x_1,x_2) \\ \qquad \frac{\partial f_1}{\partial y_1}(g_1(x_1,x_2),g_2(x_1,x_2)) \cdot \frac{\partial g_1}{\partial x_2}(x_1,x_2) + \frac{\partial f_1}{\partial y_2}(g_1(x_1,x_2),g_2(x_1,x_2)) \cdot \frac{\partial g_2}{\partial x_2}(x_1,x_2) \\ \frac{\partial f_2}{\partial y_1}(g_1(x_1,x_2),g_2(x_1,x_2)) \cdot \frac{\partial g_1}{\partial x_1}(x_1,x_2) + \frac{\partial f_2}{\partial y_2}(g_1(x_1,x_2),g_2(x_1,x_2)) \cdot \frac{\partial g_2}{\partial x_1}(x_1,x_2) \\ \qquad \frac{\partial f_2}{\partial y_1}(g_1(x_1,x_2),g_2(x_1,x_2)) \cdot \frac{\partial g_1}{\partial x_2}(x_1,x_2) + \frac{\partial f_2}{\partial y_2}(g_1(x_1,x_2),g_2(x_1,x_2)) \cdot \frac{\partial g_2}{\partial x_2}(x_1,x_2) \end{pmatrix}
\end{aligned}
$$

Beachte hier auch die klassische stark interpretationsbedürftige Schreibweise $\frac{\partial f_1}{\partial y_1}\Big(g_1(x_1,x_2), g_2(x_1,x_2)\Big)$ für $\partial_1 f_1(g(x))$.

Da $\operatorname{ev} : L(E,F)\times E \to F$ bilinear und stetig ist ($\|\operatorname{ev}(a,x)\| = \|a(x)\| \le \|a\|\cdot\|x\|$) folgt aus der Stetigkeit von $f' : U \to L(E,F)$ jene von $\widehat{f'} : U\times E \to L(E,F)\times E \to F$.

Für endlich dimensionales $E \cong \mathbb{R}^m$ gilt auch die Umkehrung, denn dann ist $L(E,F)\cong L(\mathbb{R}^m,F)\cong L(\mathbb{R},F)^m \cong F^m$ und aus der Stetigkeit von $\widehat{f'} : U\times E \to F$ folgt jene von $\operatorname{ev}_{e_i}\circ f' : U \to L(E,F) \to F$ und somit auch von $f' : U \to F^m$ (punktweise Konvergenz).

Im allgemeinen würden wir für die umgekehrte Richtung folgende Zusammensetzung betrachten:

$$f' : U \to C(E, U\times E) \to L(E,F),\ x \mapsto \Big(v \mapsto (x,v)\Big) \mapsto \Big(v \mapsto \widehat{(f')}(x)(v)\Big)$$

Damit können wir aber die Stetigkeit von f' nicht zeigen, denn die von $C(E,U\times E)\to L(E,F)$ gilt nicht, da wir in $L(E,F)$ gleichmässig auf beschränkt Mengen testen müssen, aber in $C(E,U\times E)$ es nur für kompakt Mengen wissen. Hingegen ist für die Topologie auf $L(E,F)$ der gleichmäßigen Konvergenz auf kompakt Mengen die Abbildung $U \to L(E,F)$ genau dann stetig, wenn $U\times E\to F$ es ist.

Siehe auch das folgende Beispiel.

6.1.15 Smolyanov's Beispiel.

Definition. Smolyanov's Funktion.
Es sei $f : \ell^2 \to \mathbb{R}$ definiert durch $f := \sum_{k\ge 1}\frac{1}{k^2} f_k$, wobei $f_k(x) := \varphi(k(kx_k-1))\cdot\prod_{j<k}\varphi(jx_j)$ und $\varphi : \mathbb{R}\to[0,1]$ glatt ist mit $\varphi(0)=1$ und $\varphi(t)=0$ für $|t|\ge\frac14$. Es sei $A := \{x\in\ell^2 : |kx_k| \le \frac14\ \forall k\}$. Dies ist eine abgeschlossene Teilmenge von ℓ^2.

Behauptung.
Die Funktion $f : \ell^2 \to \mathbb{R}$ ist differenzierbar.

Beweis. Beachte, daß für $x\in\ell^2$ höchstens ein $f_k(x)$ ungleich 0 sein kann. In der Tat impliziert $f_k(x)\ne 0$, daß $|kx_k-1|\le\frac{1}{4k}\le\frac14$, und daher $kx_k\ge\frac34$ und also $f_j(x)=0$ für $j>k$.

Für $x\notin A$ existiert ein $k>0$ mit $|kx_k|>\frac14$ und die Menge dies Punkte erfüllend dieser Bedingung ist offen. Es folgt, daß $\varphi(kx_k)=0$ und daher $f=\sum_{j<k}\frac{1}{j^2}f_j$ glatt ist auf dieser offenen Menge.

Andererseits sei $x\in A$. Dann ist $|kx_k-1|\ge\frac34>\frac14$ und daher $\varphi(k(kx_k-1))=0$ für alle k und also $f(x)=0$. Es sei $v\in\ell^2$ so, daß $f(x+v)\ne 0$. Dann existiert ein eindeutiges k so, daß $f_k(x+v)\ne0$ und deshalb ist $|j(x_j+v_j)|<\frac14$ für $j<k$ und $|k(x_k+v_k)-1|<\frac{1}{4k}\le\frac14$. Wegen $|kx_k|\le\frac14$ ist $|kv_k|\ge 1-|k(x_k+v_k)-1|-|kx_k|\ge 1-\frac14-\frac14=\frac12$. Daher ist $|f(x+v)|=\frac{1}{k^2}|f_k(x+v)|\le\frac{1}{k^2}\le(2|v_k|)^2\le 4\|v\|^2$. Also ist $\frac{\|f(x+v)-0-0\|}{\|v\|}\le 4\|v\|\to 0$ für $\|v\|\to 0$, d.h. f ist differenzierbar bei x mit Ableitung 0. □

Behauptung.
Die Ableitung $f' : \ell^2 \to (\ell^2)' := L(\ell^2,\mathbb{R})$ ist nicht stetig.

Beweis. Sei $a\in\mathbb{R}$ mit $\varphi'(a)\ne 0$. Dann ist $f'(t\,e^k)(e^k)=\frac{d}{dt}\frac{1}{k^2}f_k(t\,e^k)=\frac{d}{dt}\frac{1}{k^2}\varphi(k^2\,t-k)=\varphi'(k\,(k\,t-1))=\varphi'(a)$ wenn $t=t_k:=\frac1k\left(\frac{a}{k}+1\right)$, und geht somit nach 0 für $k\to\infty$. Jedoch ist $f'(0)(e^k)=0$, da $0\in A$. □

Behauptung.
Die Ableitung $(f')^\wedge : \ell^2\times\ell^2 \to \mathbb{R}$ ist stetig.

Beweis. Wir haben zu zeigen, daß $f'(x^n)(v^n)\to f'(x)(v)$ für $(x^n,v^n)\to(x,v)$. Für $x\notin A$ ist dies offensichtlich erfüllt, da dann ein k existiert mit $|kx_k|>\frac14$ und daher $f=\sum_{j\le k}\frac{1}{j^2}f_j$ lokal um x.

Wenn $x\in A$ dann ist $f'(x)=0$ und es verbleibt also den Fall zu betrachten, wo $x^n\notin A$. Es sei $\varepsilon>0$ gegeben. Lokal um x^n gibt es höchstens einen nicht-verschwindenden Summanden f_k ($x^n\notin A \Rightarrow\ \exists\, k :$

$|kx^k| > 1/4$, wähle k minimal. $\Rightarrow$ $\forall j < k : |jx^j| \leq 1/4 \Rightarrow |j(jx^j - 1)| \geq 3j/4$ (1-te Faktor) $\Rightarrow f_j = 0$ lokal für $j < k$ und $f_j = 0$ für $j > k$ (2-te Faktor)). Somit können wir Ableitung wie folgt bestimmen:

$$|f'(x^n)(v^n)| = \left|\frac{1}{k^2} f_k'(x^n)(v^n)\right| \overset{(!)}{\leq} \frac{\|\varphi'\|_\infty}{k^2} \left(k^2|v_k^n| + \sum_{j<k} j|v_j^n|\right).$$

Da $v \in \ell^2$, finden wir ein K_1 so, daß $(\sum_{j\geq K_1} |v_j|^2)^{1/2} \leq \frac{\varepsilon}{2\|\varphi'\|_\infty}$. Also schließen wir aus $\|v^n - v\|_2 \to 0$, daß $|v_j^n| \leq \frac{\varepsilon}{\|\varphi'\|_\infty}$ für $j \geq K_1$ und große n. Für die endlich vielen kleinen n können wir K_1 so vergrößern, daß für diese n und $j \geq K_1$ auch $|v_j^n| \leq \frac{\varepsilon}{\|\varphi'\|_\infty}$ ist. Weiters gibt es ein konstantes $K_2 \geq 1$ so, daß $\|v^n\|_\infty \leq \|v^n\|_2 \leq K_2$ für alle n. Nun wähle $N \geq K_1$ so groß, daß $N^2 \geq \frac{1}{\varepsilon}\|\varphi'\|_\infty K_2 K_1^2$. Offensichtlich ist $\sum_{n<N} \frac{1}{n^2} f_n$ glatt. Es verbleibt jene n zu betrachten, für welche der nicht-verschwindende Term Index $k \geq N$ hat. Für jene Terme haben wir

$$\begin{aligned}
|f'(x^n)(v^n)| &= |\frac{1}{k^2} f_k'(x^n)(v^n)| \\
&\leq \|\varphi'\|_\infty \left(|v_k^n| + \frac{1}{k^2}\sum_{j<k} j|v_j^n|\right) \\
&\leq |v_k^n|\|\varphi'\|_\infty + \|\varphi'\|_\infty \frac{1}{k^2} \sum_{j<K_1} j|v_j^n| + \frac{1}{k^2} \sum_{K_1\leq j<k} j|v_j^n|\,\|\varphi'\|_\infty \\
&\leq \varepsilon + \|\varphi'\|_\infty \frac{K_1^2}{N^2}\|v^n\|_\infty + \frac{1}{k^2} \sum_{K_1\leq j<k} j\,\varepsilon \\
&\leq \varepsilon + \varepsilon + \varepsilon = 3\varepsilon
\end{aligned}$$

Dieses zeigt die Stetigkeit. □

Folgende Konsequenz aus dem Hauptsatz (5.6.17) wird uns des öfteren nützlich sein.

6.1.23a Lemma. Richtungsdifferenzenquotient als Integral der Ableitung.
Es sei $f : E \supseteq U \to F$ eine Abbildung mit $U \subseteq E$ offen, $v \in E$ und $d_v f$ existiere und sei stetig und $x \in U$. Dann ist $x + t\,v \in U$ und

$$\frac{f(x + t\,v) - f(x)}{t} = \int_0^1 d_v f(x + s\,t\,v)\,ds$$

für alle hinreichend kleinen $|t|$.

Beweis. Da $t \mapsto x + t\,v$ stetig ist und 0 auf $x \in U$ abbildet existiert ein $\delta > 0$ mit $x + t\,v \in U$ für alle $|t| \leq \delta$.

Nun betrachten wir die Kurve $c : [-\delta, \delta] \to F$, $s \mapsto f(x + sv)$. Da $d_v f(x + s\,v)$ existiert, ist c differenzierbar mit $c'(s) = d_v f(x + s\,v)$. Da $d_v f$ stetig vorausgesetzt ist, ist c' stetig und nach dem Hauptsatz (5.6.17) und der Substitutionsformel (5.6.19) angewandt auf c ist

$$\frac{f(x + tv) - f(x)}{t} = \frac{c(t) - c(0)}{t} = \int_0^t \frac{c'(r)}{t}\,dr = \int_0^t \frac{d_v f(x + rv)}{t}\,dr = \int_0^1 d_v f(x + stv)\,ds.$$

Dies erhält man auch durch Anwendung des Hauptsatzes (5.6.17) auf die Kurve $c : s \mapsto f(x + s\,t\,v)$ mit Ableitung (wegen der Kettenregel (6.1.5))

$$c'(s) = \frac{d}{ds} f(x + s\,t\,v) = f'(x + s\,t\,v)\left(\frac{d}{ds} stv\right) = f'(x + s\,t\,v)(t\,v) = t\,f'(x + s\,t\,v)(v),$$

denn damit ist

$$\frac{f(x + tv) - f(x)}{t} = \frac{c(1) - c(0)}{t} = \int_0^1 \frac{c'(s)}{t}\,ds = \int_0^1 f'(x + tsv)(v)\,ds.$$

□

6.1.23 Proposition. Differenzierbarkeit via Richtungsableitungen.
Eine Abbildung $f : E \supseteq U \to F$ ist genau dann C^1, wenn die Richtungsableitungen eine wohldefinierte und stetige Abbildung $E \supseteq U \to L(E, F)$, $x \mapsto (v \mapsto d_v f(x))$ liefern.

Beweis. ($\Rightarrow$) Nach (6.1.2) existieren die Richtungsableitungen und $d_v f(x) = f'(x)(v)$, also definieren sie gerade die Ableitung $f' : E \supseteq U \to L(E, F)$.

($\Leftarrow$) Wegen Proposition (6.1.2) setzen wir $f'(x)(v) := d_v f(x)$ und müssen nur zeigen, daß

$$\frac{f(x + tv) - f(x)}{t} - f'(x)(v) \to 0 \text{ für } t \to 0+ \text{ glm. für } \|v\| = 1.$$

Für $F = \mathbb{R}$ ist nach dem Mittelwertsatz (4.1.5) angewendet auf $c : s \mapsto f(x + sv)$ mit Ableitung $c'(s) = d_v f(x + sv)$ für ein $0 < \vartheta < 1$:

$$\frac{f(x + tv) - f(x)}{t} = \frac{1}{t}\Big(c(t) - c(0)\Big) = c'(\vartheta t) = d_v f(x + \vartheta tv).$$

Wegen der vorausgesetzten Stetigkeit existiert für jedes $\varepsilon > 0$ ein $\delta > 0$, so daß für alle $\|w\| < \delta$ folgendes gilt: $\|d_v f(x + w) - d_v f(x)\| \le \varepsilon\|v\|$. Für $\|v\| = 1$ und $0 < t < \delta$ erhalten wir also

$$\left\|\frac{f(x + tv) - f(x)}{t} - f'(x)(v)\right\| \le \sup\Big\{\|d_v f(x + \vartheta tv) - d_v f(x)\| : 0 < \vartheta < 1\Big\} \le \varepsilon.$$

Für allgemeines F müssen wir anders vorgehen und verwenden (6.1.23a):

$$\frac{f(x + tv) - f(x)}{t} = \int_0^1 d_v f(x + stv)\, ds,$$

Da $f' : U \to L(E, F)$ stetig vorausgesetzt ist, existiert zu $\varepsilon > 0$ ein $\delta > 0$ mit $\|f'(x + h) - f'(x)\| < \varepsilon$ für alle $\|h\| < \delta$, also ist

$$\begin{aligned}\left\|\frac{f(x + tv) - f(x)}{t} - f'(x)(v)\right\| &= \left\|\int_0^1 f'(x + tsv)(v)\, ds - \int_0^1 f'(x)(v)\, ds\right\| \\ &\le \int_0^1 \Big\|f'(x + tsv) - f'(x)\Big\| \cdot \|v\|\, ds \\ &\le \int_0^1 \varepsilon \cdot \|v\|\, ds = \varepsilon\end{aligned}$$

für alle $\|v\| \le 1$ und $|t| < \delta$. □

6.1.24 Proposition. Differenzierbarkeit via partieller Ableitungen.
Eine Abbildung $f : E_1 \times \ldots \times E_m \supseteq U \to F$ ist genau dann stetig differenzierbar auf U, wenn sämtliche partielle Ableitungen $\partial_j f : U \to L(E_j, F)$ für $j \in \{1, \ldots, m\}$ existieren und stetig sind.

Insbesonders ist für $E_1 = \cdots = E_m = \mathbb{R}$ eine Abbildung $f : \mathbb{R}^m \supseteq U \to F$ genau dann C^1, wenn die partiellen Ableitungen $\partial_i f : U \to F$ für $i \in \{1, \ldots, m\}$ alle existieren und stetig sind.

Beweis. ($\Rightarrow$) Dies folgt aus (6.1.12), denn $\partial_j f(x) = (f \circ \mathrm{ins}_j^x)'(x^j) = f'(x) \circ \mathrm{ins}_j = \mathrm{ins}_j^*(f'(x)) = (\mathrm{ins}_j^* \circ f')(x) \in L(E_j, F)$, und $(\mathrm{ins}_j^*) : L(E_1 \times \ldots \times E_m, F) \to L(E_j, F)$ ist eine stetige Kontraktion.

($\Leftarrow$) Die wesentliche Beweisidee sieht man bereits im Fall $m = 2$ und $E_1 = E_2 = F = \mathbb{R}$. Wir führen den Beweis also (vorerst) nur in dieser Situation. Beachte, daß dann $L(E_i, F) = L(\mathbb{R}, \mathbb{R}) \cong \mathbb{R}$ ist. Es genügt wegen (6.1.2) zu zeigen, daß $d_v f(x) := \lim_{t \to 0+} \frac{f(x+tv)-f(x)}{t}$ glm. für $\|v\| \le 1$ existiert und $v \mapsto d_v f(x)$ für jedes $x \in U$ linear und stetig ist. O.B.d.A. sei $x = (0, 0)$ und $v = (v_1, v_2)$. Dann ist

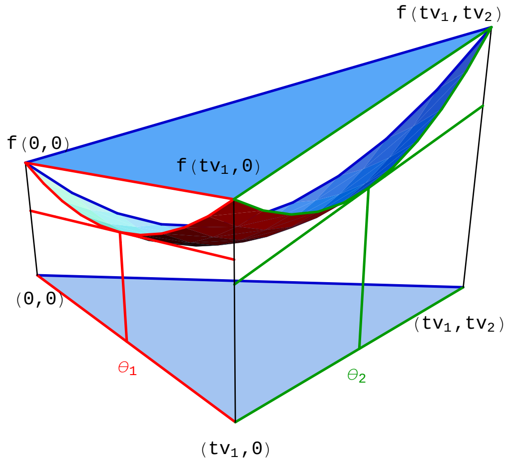
$$\frac{f(tv_1, tv_2) - f(0, 0)}{t} = \frac{f(tv_1, tv_2) - f(tv_1, 0)}{tv_2} v_2 + \frac{f(tv_1, 0) - f(0, 0)}{tv_1} v_1.$$

Wegen dem Mittelwertsatz (4.1.5) angewendet auf $s \mapsto f(s,0)$ existiert zu t und v_1 ein ϑ_1 mit $0 < \vartheta_1 < 1$ und

$$\frac{f(tv_1,0)-f(0,0)}{tv_1} = \partial_1 f(\vartheta_1 tv_1, 0).$$

Ebenso existiert – wegen dem Mittelwertsatz angewendet auf $s \mapsto f(tv_1,s)$ zu t und $v=(v_1,v_2)$ – ein ϑ_2 mit $0<\vartheta_2<1$ und

$$\frac{f(tv_1,tv_2)-f(tv_1,0)}{tv_2} = \partial_2 f(tv_1, \vartheta_2 tv_2).$$



Insgesamt erhalten wir also

$$\frac{f(tv_1,tv_2)-f(0,0)}{t} = \partial_2 f(tv_1,\vartheta_2 tv_2)\cdot v_2 + \partial_1 f(\vartheta_1 tv_1,0)\cdot v_1$$

und wegen der Stetigkeit von $\partial_1 f$ und $\partial_2 f$ konvergiert dieser Ausdruck für $\|v\| = \sqrt{\|v_1\|^2+\|v_2\|^2} \le 1$ gleichmäßig gegen

$$\partial_1 f(0,0)\cdot v_1 + \partial_2 f(0,0)\cdot v_2.$$

Da dieser Ausdruck offensichtlich linear in $v=(v_1,v_2)$ ist, ist f differenzierbar nach (6.1.2) und $f'(x)(v) = \sum_i \partial_i f(x)\, v_i$. Da $\partial_i f$ stetig ist, ist es auch f', also ist f stetig differenzierbar.

Nun für allgemeines F und $m=2$ (Für $m>2$ folgt das Resultat dann einfach mittels Induktion). Wegen (6.1.23a) angewandt auf die partiellen Funktionen $x \mapsto f(tv_1,x)$ und $x \mapsto f(x,0)$ von f ist

$$\begin{aligned}\frac{f(t\,v_1,t\,v_2)-f(0,0)}{t} &= \frac{f(t\,v_1,t\,v_2)-f(t\,v_1,0)}{t} + \frac{f(t\,v_1,0)-f(0,0)}{t}\\ &= \int_0^1 \partial_2 f(t\,v_1,s\,t\,v_2)(v_2)\,ds + \int_0^1 \partial_1 f(s\,t\,v_1,0)(v_1)\,ds\end{aligned}$$

also

$$\begin{aligned}&\Big\|\frac{f(t\,v_1,t\,v_2)-f(0,0)}{t} - \partial_1 f(0,0)(v_1) - \partial_2 f(0,0)(v_2)\Big\| =\\ &= \Big\|\int_0^1 \Big(\partial_2 f(t\,v_1,s\,t\,v_2)-\partial_2 f(0,0)\Big)(v_2)\,ds + \int_0^1 \Big(\partial_1 f(s\,t\,v_1,0)-\partial_1 f(0,0)\Big)(v_1)\,ds\Big\|\\ &\le \int_0^1 \big\|\partial_2 f(t\,v_1,s\,t\,v_2)-\partial_2 f(0,0)\big\|\cdot\|v_2\|\,ds + \int_0^1 \big\|\partial_1 f(s\,t\,v_1,0)-\partial_1 f(0,0)\big\|\cdot\|v_1\|\,ds < 2\varepsilon,\end{aligned}$$

falls $\|v_1\|\le 1$, $\|v_2\|\le 1$ und $|t|<\delta$, wobei $\delta>0$ wegen der Stetigkeit von $\partial_1 f$ und $\partial_2 f$ so gewählt werden kann, daß $\|\partial_1 f(x,y)-\partial_1 f(0,0)\| < \varepsilon$ und $\|\partial_2 f(x,y)-\partial_2 f(0,0)\| < \varepsilon$ für alle (x,y) mit $\|(x,y)\| = \max\{\|x\|,\|y\|\} < \delta$. □

Beispiel.
In (6.1.3.2) haben wir die partiellen Ableitungen von $f : (x, y, z) \mapsto x \cdot y^2 + \sin(z)$ als

$$\frac{\partial f(x,y,z)}{\partial x} = \partial_1 f(x,y,z) = y^2$$
$$\frac{\partial f(x,y,z)}{\partial y} = \partial_2 f(x,y,z) = 2xy$$
$$\frac{\partial f(x,y,z)}{\partial z} = \partial_3 f(x,y,z) = \cos(z)$$

bestimmt und verwendet um zu zeigen, daß die Ableitung $f'(x, y, z)$ diese als Matrixkoeffizienten besitzt. Mittels (6.1.24) können wir uns die aufwendigen Abschätzungen sparen, denn die partiellen Ableitungen sind offensichtlich stetig.

6.1.21 Proposition. Differenzierbare Abbildungen nach $C(I, F)$.
Es seien E und F Banach-Räume, I ein kompakter metrischer Raum, $U \subseteq E$ offen und $f : U \times I \to F$ eine stetige Abbildung. So ist $\check{f} : U \to C(I, F)$ genau dann stetig differenzierbar, wenn $\partial_1 f : U \times I \to L(E, F)$ existiert und als Abbildung $U \to L(E, C(I, F))$ wohldefiniert und stetig ist. Unter diesen Voraussetzungen gilt $\check{f}'(x)(v)(t) = \partial_1 f(x, t)(v)$ für alle $x \in U$, $v \in E$ und $t \in I$.

Beachte, daß Wohldefiniertheit und Stetigkeit von $\partial_1 f : U \times I \to L(E, F)$ nach (6.1.16) äquivalent zu jener von $(\partial_1 f)^\vee : U \to C(I, L(E, F))$, dies aber nach (6.1.19) nicht äquivalent zu der Wohldefiniertheit und Stetigkeit von $U \to L(E, C(I, F))$ ist: Wähle dazu z.B. als $\check{f}$ die lineare Abbildung g aus Beispiel in (6.1.19). Dann ist $\check{f}$ stetig differenzierbar, aber $\partial_1 f(x, t)(v) = f(v, t) = g(v)(t)$ nicht stetig von $E \times I \to L(E, F)$.

Beweis. Im Spezialfall eines kompaktes Intervalls I und $E = F = \mathbb{R}$ haben wir dies in (6.1.21) gezeigt. In jenen Beweis können wir ohne Probleme das kompakte Intervall I durch einen kompakten metrischen Raum ersetzen und den Zielraum $\mathbb{R}$ von f durch einen Banach-Raum F.

Mehr Beachtung müssen wir der Verallgemeinerung von $U \subseteq \mathbb{R}$ auf $U \subseteq E$ schenken. Aus dem obigen Spezialfall folgt vorerst nur die Äquivalenz der Existenz der Richtungsableitung $d_v\check{f}(x)$ von $\check{f}$ an der Stelle x mit jener der (partiellen) Richtungsableitung $d_{(v,0)}f(x, t)$ für alle $t \in I$ sowie die Identität

$$d_v\check{f}(x)(t) = d_{(v,0)}f(x, t) = \partial_1 f(x, t)(v).$$

Nach (6.1.23) genügt für $\check{f} \in C^1$ die Wohldefiniertheit und Stetigkeit von $x \mapsto (v \mapsto d_v\check{f}(x))$, $U \to L(E, C(I, F))$ zu überprüfen. Wegen obiger Formel ist dies offensichtlich äquivalent zur Wohldefiniertheit und Stetigkeit von $\partial_1 f$ aufgefaßt als Abbildung $U \to L(E, C(I, F))$. □

6.1.20 Lemma. Differenzierbarkeit von f_*.
Es sei I kompakt, E und F Banach-Räume, $U \subseteq E$ offen und $f : E \supseteq U \to F$ stetig differenzierbar. Dann ist $f_* : C(I, E) \supseteq C(I, U) \to C(I, F)$ ebenfalls stetig differenzierbar mit Ableitung

$$(f_*)'(g)(h)(t) = f'\big(g(t)\big)\big(h(t)\big) \text{ für } g \in C(I, U),\ h \in C(I, E) \text{ und } t \in I.$$

1. Beweis. Nach (6.1.23) genügt es die Existenz, Wohldefiniertheit und Stetigkeit der Richtungsableitungen $C(I, E) \supseteq C(I, U) \to L(C(I, E), C(I, F))$ zu zeigen. Die Existenz ist eine Aussage über die Kurve $\varphi : t \mapsto f_*(g + t\,h)$ für $g \in C(I, U)$ und $h \in C(I, E)$. Nach (6.1.21) folgt dies aus der Existenz und Stetigkeit von $\partial_1\hat{\varphi} : \mathbb{R} \times I \to C(I, F)$, also von $(t, s) \mapsto f(g(s) + t\,h(s))$. Da f differenzierbar ist, ist $\partial_1\hat{\varphi}(t, s) = f'(g(s) + t\,h(s)) \cdot h(s)$ und somit $d_h f_*(g) = (f' \circ g) \cdot h$. Dies ist wohldefiniert und stetig von $C(I, U) \to L(C(I, E), C(I, F))$, da $(f')_* : C(I, U) \to C(I, L(E, F))$ stetig ist und die Evaluation $\mathrm{ev} : L(E, F) \times E \to F$ bilinear und stetig ist, also auch $\mathrm{ev}_* : C(I, L(E, F)) \times C(I, E) \cong C(I, L(E, F) \times E) \to C(I, F)$ und somit $C(I, U) \to C(I, L(E, F)) \to L(C(I, E), C(I, F))$, $g \mapsto f' \circ g \mapsto h \mapsto (f' \circ g) \cdot h$ stetig ist. □

2. Beweis. Wegen (6.1.21) genügt zu zeigen, daß für $\widehat{f_*} = f \circ \mathrm{ev} : C(I,U) \times I \to U \to F$ $\partial_1 \widehat{f_*}$ existiert und als Abbildung $C(I,U) \to L(C(I,E), C(I,F))$ wohldefiniert und stetig ist. Ersteres ist offensichtlich, denn $d_h f_*(g)(s) = \frac{d}{dt}|_{t=0} f(g(s) + t\,h(s)) = f'(g(s)) \cdot h(s)$ und letzteres, da $(f')_* : C(I,U) \to C(I, L(E,F))$ und $\mathrm{ev}_* : C(I, L(E,F)) \times C(I,E) \cong C(I, L(E,F) \times E) \to C(I,F)$ nach (6.1.18) stetig (und letztere bilinear) sind und somit

$$C(I,U) \xrightarrow{(f')_*} C(I, L(E,F)) \xrightarrow{(\mathrm{ev}_*)^\vee} L(C(I,E), C(I,F))$$

stetig ist. □

6.1.22 Vertauschbarkeit vom Differenzieren und Integrieren.
Gegeben sei eine Funktion $f : \mathbb{R} \times [a,b] \to \mathbb{R}$. Für jedes fixes $t \in [a,b]$ sei die partielle Funktion $x \mapsto f(x,t)$ differenzierbar also $\partial_1 f(x,t) = \frac{\partial}{\partial x} f(x,t)$ existent und für jedes fixe $x \in \mathbb{R}$ sei die partielle Funktion $t \mapsto f(x,t)$ über $[a,b]$ integrierbar also $g(x) := \int_a^b f(x,t)\,dt$ existent. Die Frage die sich stellt ist inwieweit diese beiden Operationen miteinander kommutieren, also

$$\int_a^b \partial_1 f(x,t)\,dt = \frac{d}{dx} \int_a^b f(x,t)\,dt \text{ für alle } x \in \mathbb{R} \text{ gilt.}$$

Offensichtlich vertauscht die Bildung von Differenzenquotient und von Riemann-Summen miteinander:

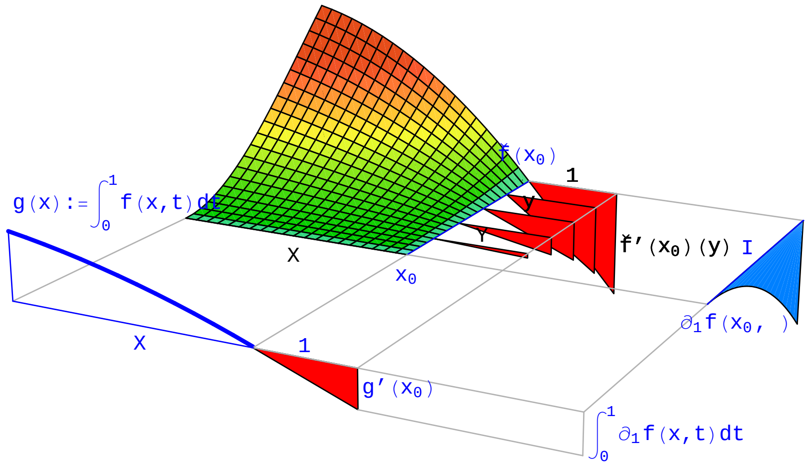
$$\begin{aligned} S\left(t \mapsto \frac{f(x+h,t) - f(x,t)}{h}, Z\right) &= \sum_{I \in Z} \frac{f(x+h,t_I) - f(x,t_I)}{h}\,|I| \\ &= \frac{\sum_{I\in Z} f(x+h,t_I)\,|I| - \sum_{I \in Z} f(x,t_I)\,|I|}{h} \\ &= \frac{S(t \mapsto f(x+h,t), Z) - S(t \mapsto f(x,t), Z)}{h} \end{aligned}$$

und unter geeigneten Voraussetzungen sollten nach (3.2.8) auch die beiden Limesbildungen des Integrals und der Ableitung miteinander vertauschen.

Versuchen wir nun die dafür notwendigen Bedingungen zu finden. Wir haben $g(x)$ als Integral $\int_a^b (\check{f}(x))$ von $\check{f}(x) : t \mapsto f(x,t)$ definiert, d.h. g als Zusammensetzung von $x \mapsto \int_a^b \circ \check{f}$. Damit wir keine Probleme beim Integrieren haben, sollte $\check{f}$ Werte in $C([a,b], \mathbb{R})$ haben. Auf diesen Raum ist Integrieren $\int_a^b$ nach (5.1.12) linear und stetig, also nach dem Spezialfall (6.1.6) (siehe auch (5.6.15)) der Kettenregel (6.1.5) die Zusammensetzung $\int_a^b \circ \check{f}$ differenzierbar mit Ableitung $(\int_a^b \circ \check{f})'(x) = \int_a^b (\check{f}'(x))$ sofern $\check{f} : \mathbb{R} \to C([a,b], \mathbb{R})$ differenzierbar ist. Dies haben wir in (6.1.21) äquivalent durch die Existenz und Stetigkeit von $\partial_1 f : \mathbb{R} \times [a,b] \to \mathbb{R}$ beschrieben.

Wenn wir $\int_0^1 g$ als Mittelwert der Funktion g interpretieren, so besagt diese Formel, daß die infinitesimale Änderung $\frac{d}{dx} \int_0^1 f(x,t)\,dt$ (bzgl. x) der Mittelwerte $\int_0^1 f(x,t)\,dt$ von $t \mapsto f(x,t)$ ident mit dem Mittelwert $\int_0^1 \partial_1 f(x,t)\,dt$ der infinitesimalen Änderung $\partial_1 f(x,t)$ von f bzgl. der 1-ten Variable ist.

Proposition.
Es sei $f : E \times \mathbb{R} \supseteq U \times [a,b] \to F$ stetig mit offenen $U \subseteq E$ und $\partial_1 f : U \times [a,b] \to L(E,F)$ existiere und sei stetig. Dann ist $g : U \to F$ gegeben durch $x \mapsto \int_a^b f(x,t)\,dt$ differenzierbar mit Ableitung $g'(x) = \int_a^b \partial_1 f(x,t)\,dt$.



1-ter Beweis. Die Funktion $g : x \mapsto \int_a^b f(x,t)\,dt$ ist durch Zusammensetzung $U \to C([a,b],F) \to F$ von $f^\vee : x \mapsto (t \mapsto f(x,t))$ und $h \mapsto \int_a^b h$ gegeben. Die zweite Abbildung $\int_a^b$ ist nach (5.6.13.2) linear und nach (5.6.14.4) stetig. Die erste Funktion $f^\vee$ ist nach (6.1.21) stetig differenzierbar mit Ableitung $(f^\vee)'(x)(v)(t) = \partial_1 f(x,t)(v)$, denn aus der Wohldefiniertheit und Stetigkeit von $\partial_1 f : U \times [a,b] \to L(E,F)$ folgt nach (6.1.16) jene von $(\partial_1 f)^\vee : U \to C([a,b], L(E,F))$ und nach (6.1.19) jene von $(\partial_1 f)^\vee : U \to C([a,b], L(E,F)) \to L(E, C([a,b],F))$. Nach der Kettenregel (6.1.6) ist somit die Zusammensetzung g differenzierbar mit Ableitung

$$\begin{aligned}
g'(x)(v) &= \Big(\int_a^b \circ f^\vee\Big)'(x)(v) = \int_a^b \Big((f^\vee)'(x)(v)\Big)\\
&= \int_a^b \widetilde{(\partial_1 f)^\vee(x)}(v)(t)\,dt = \int_a^b \partial_1 f(x,t)(v)\,dt\\
&= \int_a^b \mathrm{ev}_v\Big(\partial_1 f(x,t)\Big)\,dt = \mathrm{ev}_v\Big(\int_a^b \partial_1 f(x,t)\,dt\Big) = \Big(\int_a^b \partial_1 f(x,t)\,dt\Big)(v) \qquad \square
\end{aligned}$$

2-ter Beweis. Direkter, d.h. unter Vermeidung der Differentiation der Kurve g in $C(I,F)$, aber etwas weniger offensichtlich ist folgender Beweis. Es ist

$$\begin{aligned}
g(x+v) - g(x) - \int_a^b \partial_1 f(x,t)\,dt \cdot v &=\\
&= \int_a^b f(x+v,t)\,dt - \int_a^b f(x,t)\,dt - \mathrm{ev}_v\Big(\int_a^b \partial_1 f(x,t)\,dt\Big)\\
&\overset{(5.6.8)}{=\!=} \int_a^b f(x+v,t)\,dt - \int_a^b f(x,t)\,dt - \int_a^b \mathrm{ev}_v\Big(\partial_1 f(x,t)\Big)\,dt\\
&\overset{(5.6.13.2)}{=\!=} \int_a^b f(x+v,t) - f(x,t) - \partial_1 f(x,t)\cdot v\,dt\\
&\overset{(6.1.23a)}{=\!=} \int_a^b \int_0^1 \Big(\partial_1 f(x+sv,t) - \partial_1 f(x,t)\Big)\cdot v\,ds\,dt
\end{aligned}$$

Nun ist aber $\partial_1 f : U \times [a,b] \to L(E,F)$ stetig, also $\partial_1 f(\_,t) : U \to L(E,F)$ stetig und zwar gleichmäßig bzgl. $t \in [a,b]$ nach (3.2.8), d.h.

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall t \in [a,b]\ \forall h : \|h\| < \delta \Rightarrow \|\partial_1 f(x+h,t) - \partial_1 f(x,t)\| \le \varepsilon.$$

Also ist nach ist nach (5.6.14.2) die Norm der rechten Seite kleiner oder gleich $(b-a)\varepsilon\|v\|$ und dies zeigt die Proposition. $\square$

6.2 Lösung von Gleichungen

Wir wollen nun Gleichungen $f(x) = y$ in differenzierbarer Weise nach x auflösen. Eindeutige Lösbarkeit für alle y ist dabei zur Invertierbarkeit von f äquivalent. Für lineare Gleichungen $f(x) = y$ (d.h. $f : E \to F$ ist linear) wird das in der linearen Algebra behandelt, und im endlich-dimensionalen die Invertierbarkeit von f durch das Nicht-Verschwinden der Determinante $\det(f)$ beschrieben.

6.2.2. Inverse Funktionen Wir wollen nun solche Gleichungen $f(x) = y$ für nicht-lineare f behandeln. Sei also $g(y) := x$ eine Lösung bei gegebener rechter Seite y. Falls nicht nur f sondern auch g differenzierbar ist, so gilt wegen $f(g(y)) = y$ nach der Kettenregel (6.1.5) $f'(g(y)) \circ g'(y) = \mathrm{id}$. Wenn die Lösungen eindeutig sind, d.h. f injektiv ist, dann ist wegen $f(g(f(x))) = f(x)$ auch $g(f(x)) = x$ also $g'(f(x)) \circ f'(x) = \mathrm{id}$ und speziell für $x = g(y)$ ist $g'(y) \circ f'(g(y)) = \mathrm{id}$, d.h. $g'(y)$ ist die inverse (lineare Abbildung) zu $f'(g(y))$.

Die Idee des folgenden Satzes ist die Umkehrung dieser Beobachtung, nämlich wenn die Gleichung $f'(x_0)(x) = y$ die durch die lineare Approximation $f'(x_0)$ von f gegeben ist, eindeutig für alle y lösbar ist, dann sollte auch die nicht-lineare Gleichung $f(x) = y$ zumindest für y nahe $f(x_0)$ eindeutig mit x nahe x_0 lösbar sein.

Satz über inverse Funktionen.
Es sei $f : E \supseteq U \to F$ C^1, $U \subseteq E$ offen, $x_0 \in U$ und $f'(x_0) : E \to F$ ein Isomorphismus. Dann ist f ein LOKALER C^1-DIFFEOMORPHISMUS *, d.h. es gibt offene Umgebungen $V \subseteq U$ von x_0 und W von $f(x_0)$ so daß $f : V \to W$ bijektiv ist und die Inverse $f^{-1} : W \to V$ C^1 ist mit Ableitung $(f^{-1})'(y) = f'(f^{-1}(y))^{-1}$. Vergleiche dies mit (4.1.20),*

Eine Menge $U \subseteq E$ heißt OFFENE UMGEBUNG von $x_0 \in E$, wenn U offen ist und $x_0 \in U$ liegt.

Beweis. Zuerst vereinfachen wir das Problem. Durch eine Translation (wir ersetzen f durch die Abbildung $x \mapsto f(x_0 + x) - f(x_0)$) erreichen wir $x_0 = 0$ und $f(x_0) = 0$. Durch Zusammensetzen mit $f'(x_0)^{-1}$ erreichen wir $f'(x_0) = \mathrm{id}$, denn nach (6.1.6) ist $(f'(x_0)^{-1} \circ f)'(x_0) = f'(x_0)^{-1} \circ f'(x_0) = \mathrm{id}$. Da die lineare Approximation id von f invertierbar ist, sollten wir das Restglied

$$r(x) := f(x_0 + x) - f(x_0) - f'(x_0)(x) = f(x) - x$$

betrachten. Es ist $r(0) = 0$ und $r'(0) = f'(0) - \mathrm{id} = 0$, also existiert ein $\delta > 0$ mit $\|x\| \le \delta \Rightarrow x \in U$ da U offen ist und $\|f'(x) - \mathrm{id}\,\| = \|r'(x)\| \le 1/2$ wegen der Stetigkeit von r'. Somit ist nach dem Mittelwertsatz (5.6.18)

$$\|r(x_2) - r(x_1)\| \le \max\Big\{\|r'(x_1 + t(x_2 - x_1))(x_2 - x_1)\|\Big\} \le \frac{1}{2}\|x_2 - x_1\|$$

für $\|x_j\| \le \delta$. Speziell für $x_1 := 0$ erhalten wir $\|r(x_2)\| \le \|x_2\|/2$ für $\|x_2\| \le \delta$.

Wir behaupten nun, daß für alle $\|y\| < \delta/2$ ein eindeutiges $\|x\| < \delta$ mit $f(x) = y$ existiert. Dazu wandeln wir diese Gleichung in die Fixpunktgleichung $x = g_y(x) := y + x - f(x) = y - r(x)$ um. Für $\|y\| < \delta/2$ und $\|x\| \le \delta$ ist $\|g_y(x)\| \le \|y\| + \|r(x)\| < \delta/2 + \|x\|/2 \le \delta$ und weiters ist $\|g_y(x_2) - g_y(x_1)\| = \|r(x_2) - r(x_1)\| \le \|x_2 - x_1\|/2$. Also folgt aus dem Banach'schen Fixpunktsatz (3.4.17) die Existenz eines eindeutig bestimmten Fixpunktes $\|x\| \le \delta$ und wegen $\|x\| = \|g_y(x)\| < \delta$ liegt dieser im Inneren.

Es sei nun $W := \{y : \|y\| < \delta/2\}$ und $V := f^{-1}(W) \cap \{x : \|x\| < \delta\}$. Dann existiert für jedes $y \in W$ ein eindeutiges $\|x\| < \delta$ mit $f(x) = y$, also liegt $x \in V$ und $f : V \to W$ ist bijektiv.

Es ist f^{-1} (Lipschitz-)stetig, denn für $y_j \in W$ und $x_j \in V$ mit $f(x_j) = y_j$ gilt: $\|x_2 - x_1\| \le \|r(x_1) - r(x_2)\| + \|f(x_2) - f(x_1)\| \le \|y_2 - y_1\| + \frac{1}{2}\|x_2 - x_1\|$, also $\|f(y_2) - f(y_1)\| = \|x_2 - x_1\| \le 2\|y_2 - y_1\|$. Und schließlich ist f^{-1} differenzierbar, denn

$$\begin{aligned}\big\|f^{-1}(y_2) - f^{-1}(y_1) - f'(x_1)^{-1}(y_2 - y_1)\big\| &= \big\|x_2 - x_1 - f'(x_1)^{-1}(f(x_2) - f(x_1))\big\| \\ &\le \|f'(x_1)^{-1}\| \cdot \underbrace{\frac{\|f'(x_1)(x_2 - x_1) - f(x_2) + f(x_1)\|}{\|x_2 - x_1\|}}_{\to 0} \cdot \underbrace{\|x_2 - x_1\|}_{\le 2\|y_2 - y_1\|}.\end{aligned}$$

Beachte dabei, daß $f'(x_1)^{-1}$ existiert und mittels geometrischer Reihe berechnet werden kann nach (5.5.16) oder Aufgabe (6.23), da $\|f'(x_1) - \mathrm{id}\| < 1$.

Daß f^{-1} C^1 ist folgt, da $(f^{-1})' = \mathrm{inv} \circ f' \circ f^{-1}$ und alle Komponenten u.a. nach (5.5.16) stetig sind. $\square$

Bemerkung. Berechenbarkeit der Umkehrfunktion.
Beachte, daß der Beweis des Satzes nicht nur eine Existenzaussage zeigt, sondern gleichzeitig eine Methode zur näherungsweisen Bestimmung der lokalen Lösung angibt, nämlich für y nahe $f(x_0)$ iterativ die Funktion g_y auf den Startwert x_0 anzuwenden. Da wir im Beweis die Funktion f durch $\tilde f : x \mapsto f'(x_0)^{-1}(f(x+x_0) - f(x_0))$ ersetzt haben (um $x_0 = 0$, $f(x_0) = 0$ und $f'(x_0) = \mathrm{id}$ zu erhalten) bedeutet dies, daß $\tilde y = \tilde f(\tilde x) \Leftrightarrow \tilde x = g_{\tilde y}(\tilde x) := \tilde y + \tilde x - \tilde f(\tilde x)$, also

$$f(\tilde x + x_0) = f'(x_0)(\tilde y) + f(x_0) \qquad \Leftrightarrow \qquad \tilde x = g_{\tilde y}(\tilde x) = \tilde y + \tilde x - f'(x_0)^{-1}(f(\tilde x + x_0) - f(x_0))$$

bzw. wenn wir $x := \tilde x + x_0$ und $y := f'(x_0)(\tilde y) + f(x_0)$ setzen:

$$\begin{aligned} f(x) = y \qquad \Leftrightarrow \qquad x &= x_0 + g_{\tilde y}(x - x_0) \\ &= x + f'(x_0)^{-1}(y - f(x_0)) - f'(x_0)^{-1}(f(x) - f(x_0)) \\ &= x + f'(x_0)^{-1}(y - f(x)) \end{aligned}$$

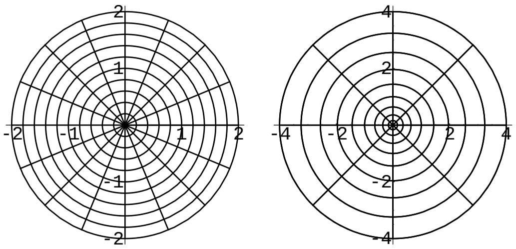
also ist die Lösung x von $f(x) = y$ durch die Iterationsfunktion $x \mapsto x - f'(x_0)^{-1}(f(x) - y)$ erhaltbar. Dies ist das vereinfachte Newtonverfahren aus (6.2.11) angewandt auf die Nullstellengleichung $f(x) - y = 0$.

6.2.3 Beispiele inverser Abbildungen.

1. Quadrieren: Es sei $f : \mathbb{C} \to \mathbb{C}$ geben durch $f(z) := z^2$. Also ist f als Abbildung von $\mathbb{R}^2 \cong \mathbb{C}$ gegeben durch $f(x, y) = (x^2 - y^2, 2xy)$. Die Jacobi-Matrix von f ist somit

$$[f'(x, y)] = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}$$

mit Determinante $\det(f'(x, y)) = 4(x^2 + y^2) \neq 0$ für $(x, y) \neq (0, 0)$. Nach dem Satz über inverse Funktionen existiert somit lokal um Punkte $\neq 0$ die (eindeutig bestimmte) Umkehrfunktion $\sqrt{}$. Global ist jedoch zu jeder Lösung z von $z^2 = w$ auch $-z$ eine weitere Lösung. Wenn wir beachten, daß f in Polarkoordinaten durch $(r, \varphi) \mapsto (r^2, 2\varphi)$ gegeben ist, so sehen wir, daß wir für $W := \{u + i\,v \in \mathbb{C} : u > 0\}$ die Halbebene und für $V := \{x + i\,y \in \mathbb{C} : y \neq 0 \text{ oder } x > 0\}$ die Ebene vermindert um die negative x-Achse verwenden können um einen Diffeomorphismus $f : W \to V$ zu erhalten.



In diesem Fall können wir die Gleichung $z^2 = w$, also das (nicht-lineare) Gleichungssystem

$$\begin{aligned} x^2 - y^2 &= u \\ 2\,x\,y &= v \end{aligned}$$

sogar explizit auflösen, denn die erste Gleichung liefert $x^2 = y^2 + u > 0$ und wenn wir dies in die quadrierte zweite einsetzen $(y^2+u)y^2 = \left(\frac{v}{2}\right)^2$, also die BIQUADRATISCHE GLEICHUNG

$$
\begin{aligned}
(y^2)^2 + u\,y^2 - \left(\frac{v}{2}\right)^2 = 0 \quad &\Leftrightarrow \quad y^2 = \frac{-u \pm \sqrt{u^2+v^2}}{2} \\
&\Rightarrow \quad y = \pm\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}} \\
\text{und} \quad x = \frac{v}{2\,y} &= \pm\frac{v}{\sqrt{2}\,\sqrt{\sqrt{u^2+v^2}-u}} \\
&= \pm\frac{v\,\sqrt{\sqrt{u^2+v^2}+u}}{\sqrt{2}\,\sqrt{(u^2+v^2)-u^2}} \\
&= \pm\,\mathrm{sign}(v)\,\sqrt{\frac{\sqrt{u^2+v^2}+u}{2}}.
\end{aligned}
$$

Beachte aber, daß wir diesem Ausdruck für die Umkehrfunktion f^{-1} wegen $\mathrm{sign}(v)$ und den Wurzeln nicht sofort ansehen, daß er C^1 in (u,v) ist.

2. POLARKOORDINATEN:
Es sei $f : \mathbb{R}^2 \supseteq \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}^2$, $(r,\varphi) \mapsto (r\,\cos(\varphi), r\,\sin(\varphi))$. Dann ist f C^1 und

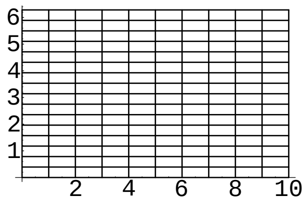
$$
f'(r,\varphi) = \begin{pmatrix} \cos(\varphi) & -r\,\sin(\varphi) \\ \sin(\varphi) & r\,\cos(\varphi) \end{pmatrix}
$$

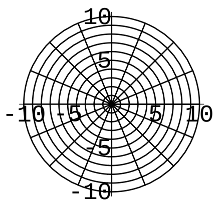
invertierbar, da $\det(f'(r,\varphi)) = r > 0$, aber

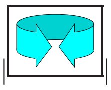
$$
f(r,\varphi + 2\,k\,\pi) = f(r,\varphi) \text{ für alle } k \in \mathbb{Z}.
$$

Also ist f nicht global injektiv. Dies steht im Kontrast zum 1-dimensionalen Satz (4.1.20) über inverse Funktionenen. Wenn wir nun $W := \{(x,y) \in \mathbb{R}^2 : y \neq 0 \text{ oder } x > 0\}$ und $V := \{(r,\varphi) \in \mathbb{R}^2 : r > 0, -\pi < \varphi < \pi\}$ setzen, dann ist $f : V \to W$ ein Diffeomorphismus und wir können die Umkehrfunktion (dies aus kartesischen Koordinaten die zugehörigen Polarkoordinaten ausrechnet) auch explizit angeben: $r = \sqrt{x^2+y^2}$ und

$$
\varphi = \begin{cases} \arctan(\frac{y}{x}) & \text{für } x > 0 \\ \mathrm{arccot}(\frac{x}{y}) & \text{für } y > 0 \\ \mathrm{arccot}(\frac{x}{y}) - \pi & \text{für } y < 0 \end{cases}
$$







Betrachten wir die Zusammensetzung

$$
\tilde{f} : \mathbb{R}^2 \to \mathbb{R}^2, \qquad (s,\varphi) \mapsto (e^s, \varphi) \xmapsto{f} (e^s\,\cos(\varphi), e^s\,\sin(\varphi)),
$$

so ist nach der Kettenregel (6.1.5)

$$
\tilde{f}'(s,\varphi) = \begin{pmatrix} \cos(\varphi) & -e^s\,\sin(\varphi) \\ \sin(\varphi) & e^s\,\cos(\varphi) \end{pmatrix} \cdot \begin{pmatrix} e^s & 0 \\ 0 & 1 \end{pmatrix}
$$

und somit $\det(\tilde{f}'(s,\varphi)) = e^s \cdot e^s = e^{2s} > 0$ für alle s. Dennoch ist $\tilde{f}$ nicht injektiv.

3. Ähnliches gilt auch für Kugelkoordinaten

$$f : (r, \varphi, \theta) \mapsto r\Big(\cos(\theta)\,\cos(\varphi), \cos(\theta)\,\sin(\varphi), \sin(\theta)\Big)$$

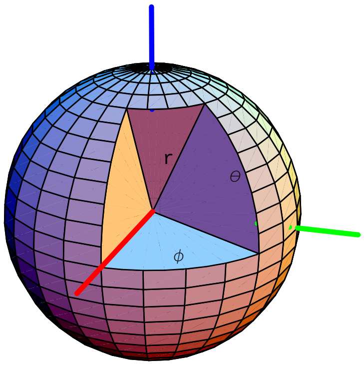
(wobei r der Radius, φ die Längengrade und ϑ die Breitengrade sind) mit Jacobi-Matrix

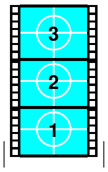
$$f'(r, \varphi, \theta) = \begin{pmatrix} \cos(\theta)\,\cos(\varphi) & -r\,\cos(\theta)\,\sin(\varphi) & -r\,\sin(\theta)\,\cos(\varphi) \\ \cos(\theta)\,\sin(\varphi) & r\,\cos(\theta)\,\cos(\varphi) & -r\,\sin(\theta)\,\sin(\varphi) \\ \sin(\theta) & 0 & r\,\cos(\theta) \end{pmatrix}$$

und Determinante (durch Entwicklung nach der letzten Zeile)

$$\begin{aligned} \det(f'(r, \varphi, \theta)) &= \sin(\theta)\,r^2\,\sin(\theta)\,\cos(\theta)\,(\sin(\varphi)^2 + \cos(\varphi)^2) + 0 + \\ &\quad + r\,\cos(\theta)\,r\,\cos(\theta)^2\,(\sin(\varphi)^2 + \cos(\varphi)^2) \\ &= r^2\,\cos(\theta) \neq 0, \text{ für } -\pi/2 < \theta < \pi/2 \text{ und } r \neq 0. \end{aligned}$$

Folglich existiert eine stetig differenzierbare (lokale) Umkehrfunktion die aus den kartesischen Koordinaten (x, y, z) eines Punktes (der nicht auf der Halbebene $y = 0 \wedge x \leq 0$ liegt) die Kugelkoordinaten (r, φ, ϑ) berechnet.





Explizit ist $r = \sqrt{x^2 + y^2 + z^2}$, $\vartheta = \arcsin(z)$ und $(\cos(\vartheta), \varphi)$ die Polarkoordinaten von (x, y).

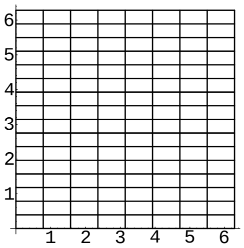
4. Es sei $f : \mathbb{R}^2 \to \mathbb{R}^2$ gegeben durch

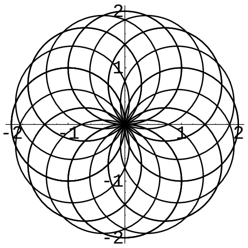
$$f(x, y) = (\sin(x) - \cos(y), \sin(y) - \cos(x))$$

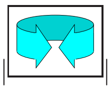
mit Jacobi-Matrix

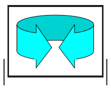
$$\begin{pmatrix} \cos(x) & \sin(y) \\ \sin(x) & \cos(y) \end{pmatrix}$$

und Determinante $\cos(x)\,\cos(y) - \sin(x)\,\sin(y) = \cos(x + y)$. Dies ist $\neq 0$ falls x nahe $-y$ liegt. Folglich gibt es lokale C^1-Lösungen.









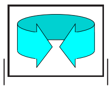
Hier werden die Geraden $x = x_0$ auf verkehrt durchlaufene Kreise $y \mapsto (\sin(x_0), -\cos(x_0)) - (\cos(-y), \sin(-y))$ mit Mittelpunkt $(\sin(x_0), -\cos(x_0)) = \left(\cos\left(\frac{3\pi}{2} + x_0\right), \sin\left(\frac{3\pi}{2} + x_0\right)\right)$ und Radius 1 abgebildet. Die Umkehrfunktion kann man mit einigen Aufwand auch explizit bestimmen (Hiweis: Mathematika).

5. Schlußendlich noch ein Beispiel, wo sich die Umkehrfunktion nicht mehr explizit bestimmen läßt: $f : \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) \mapsto (x + \sin(y)/2, y - \cos(x))$. Die Jacobi-Matrix ist

$$\begin{pmatrix} 1 & \cos(y)/2 \\ \sin(x) & 1 \end{pmatrix}$$

mit Determinante $1 - \frac{\sin(x)\cos(y)}{2} \geq \frac{1}{2}$.

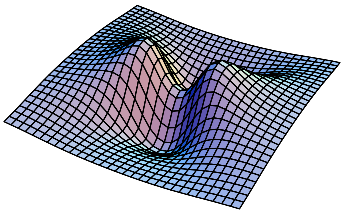
Wir zeigen zusätzlich, daß die Umkehrfunktion global existiert Das Gleichungssystem $f(x, y) = (u, v)$ ist – wenn wir $x = u + r$ und $y = v + s$ setzen – äquivalent zu $r = -\sin(v + s)/2$, $s = \cos(u + r)$. Und somit ist $r = -\sin(v + \cos(u + r))/2$, eine Fixpunktgleichung die nach Banach'schen Fixpunktsatz für jedes $(u, v) \in \mathbb{R}^2$ eine eindeutig bestimmte Lösung $r = r(u, v) \in \mathbb{R}$ besitzt. Somit ist $(u, v) \mapsto (x, y) = (u + r(u, v), v + \cos(u + r(u, v)))$ die global definierte und nach inversen Funktionensatz stetig differenzierbare Umkehrfunktion zu f.

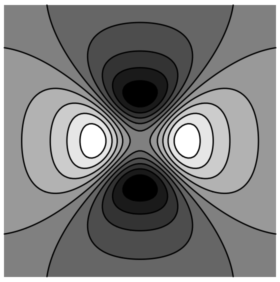


6.2.4 Implizite Funktionen. Nun geht es uns darum so etwas wie die NIVEAULINIEN

$$\{(x, y) : f(x, y) = z\}$$

eines Gebirges $f : \mathbb{R}^2 \to \mathbb{R}$ zu diskutieren, oder wenn wir f als Luftdruck interpretieren die ISOBAREN oder bei Interpretation von f als Temperatur die ISOTHERMEN. Wir versuchen dazu bei vorgegebenen z die Lösungen (x, y) von $f(x, y) = z$ mittels Funktionen $x \mapsto y$ (oder Funktionen $y \mapsto x$)zu beschreiben.





Wie beim Satz (6.2.2) über inverse Funktionen ist die Idee, daß die nicht-lineare Gleichung $f(x, y) = z$ lokal nach y lösbar sein sollte wenn es die approximierende lineare Gleichung $z = f'(x_0, y_0) \cdot (x, y) = \partial_1 f(x_0, y_0) \cdot x + \partial_2 f(x_0, y_0) \cdot y$ ist, also die lineare Abbildung $\partial_2 f(x_0, y_0)$ invertierbar ist.

Satz über implizite Funktionen.
Es sei $f : E \times F \to F$ lokal um (x_0, y_0) C^1 und $\partial_2 f(x_0, y_0) : F \to F$ ein Isomorphismus. Dann existieren offene Umgebungen U von x_0, V von y_0 und W von $f(x_0, y_0)$, so daß für jedes $z \in W$ und $x \in U$ ein eindeutiges $y \in V$ existiert mit $f(x, y) = z$. Weiters ist die Abbildung $(x, z) \mapsto y$, $U \times W \to V$ C^1.

Beachte, daß wir die Ableitungen von $h : (x, z) \mapsto y$ durch Differenzieren der Gleichung $f(x, y) = z$ unter Anwendung der Kettenregel (6.1.5) errechnet werden können, denn $\frac{\partial}{\partial x}$ liefert

$$\partial_1 f(x, h(x, z)) + \partial_2 f(x, h(x, z))\, \partial_1 h(x, z) = 0$$

und $\frac{\partial}{\partial z}$ liefert

$$\partial_1 f(x, h(x, z)) \cdot 0 + \partial_2 f(x, h(x, z))\, \partial_2 h(x, z) = \mathrm{id}\,.$$

Beweis. Die Idee des folgenden Beweises besteht darin den Satz (6.2.2) über inverse Funktionen anzuwenden. Dazu benötigen wir eine Abbildung g mit invertierbarer Ableitung und ergänzen deshalb $f : E \times F \to F$ zu $g = (\mathrm{pr}_1, f) : E \times F \to E \times F$ gegeben durch $(x, y) \mapsto (x, f(x, y))$. Dann ist

$$g'(x,y) = \begin{pmatrix} 1 & 0 \\ \partial_1 f(x,y) & \partial_2 f(x,y) \end{pmatrix}$$

invertierbar mit Inverser

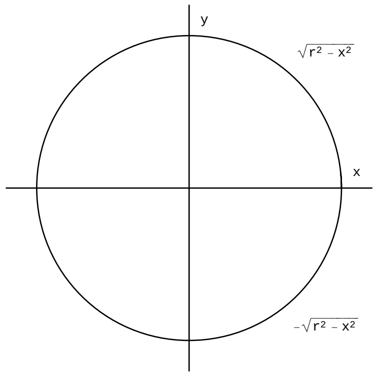
$$\begin{pmatrix} 1 & 0 \\ -\partial_2 f(x,y)^{-1} \cdot \partial_1 f(x,y) & \partial_2 f(x,y)^{-1} \end{pmatrix},$$

(Man erhält dies durch einen allgemeinen Ansatz), also existieren nach dem Satz (6.2.2) über inverse Funktionen Umgebungen U, V und O von x_0, y_0 und $g(x_0, y_0)$, so daß $g : U \times V \to O$ ein Diffeomorphismus ist.

Nun wähle Umgebungen U_1 von x_0 und W von $f(x_0, y_0)$ mit $U_1 \times W \subseteq O$. Dann ist für jedes $x \in U_1$ und $z \in W$ der Punkt $(x, z) \in O$, also existiert ein eindeutiger Punkt $(u, y) \in U \times V$ mit $(x, z) = g(u, y) = (u, f(u, y))$, d.h. $u = x \in U_1$ und $f(x, y) = z$, also ein eindeutiger Punkt $y \in V$ mit $f(x, y) = z$. Weiters ist $(x, z) \mapsto y$ durch $(u, y) = g^{-1}(x, z) =: (h_1(x, z), h_2(x, z))$ gegeben, wobei h_j die Komponenten von g^{-1} bezeichnet. Folglich ist diese Zuordnung h_2 und somit C^1. □

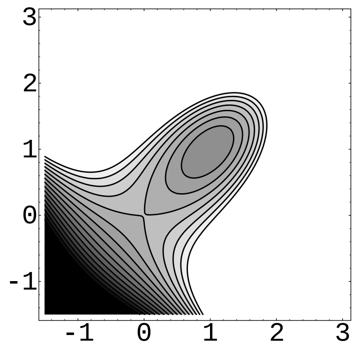
6.2.5 Beispiele implizit gegebener Funktionen.

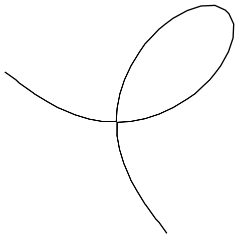
1. Der Einheitskreis ist durch die Gleichung $x^2 + y^2 = 1$ gegeben, d.h. als Lösungsmenge der impliziten Gleichung $f(x, y) := x^2 + y^2 = 1$. Die Ableitung von f ist $f'(x, y) = (2x, 2y)$ d.h. für $y_0 \neq 0$ ist y nach (6.2.4) lokal um y_0 nach x auflösbar und für $x_0 \neq 0$ ist x lokal um x_0 nach y auflösbar. Wir können dies sogar explizit machen: $y = \pm\sqrt{1 - x^2}$ bzw. $x = \pm\sqrt{1 - y^2}$.



   Diese Funktionen sind allerdings bei $x = \pm 1$ bzw. $y = \pm 1$ nicht differenzierbar (die Ableitung wäre $\mp\infty$). Global können wir den Kreis jedenfalls nicht durch eine Funktion $y = g(x)$ beschreiben, denn zu jedem $|x| < 1$ existieren zwei y mit $f(x, y) = 1$. Eine globale Parametrisierung durch $t \mapsto (\cos(t), \sin(t))$ haben wir allerdings schon.

2. Das Kartesische Blatt ist durch $x^3 + y^3 = 3xy$ gegeben, d.h. als Lösung der impliziten Gleichung $f(x, y) := x^3 + y^3 - 3xy = 0$. Die Ableitung von f ist $f'(x, y) = (3(x^2 - y), 3(y^2 - x))$, also ist nach dem Satz (6.2.4) über impliziten Funktionen y lokal um y_0 nach x nahe x_0 auflösbar falls $y_0^2 \neq x_0$. Andernfalls ist $y_0^2 = x_0$ und wegen $f(x_0, y_0) = 0$ ist $y_0^6 + y_0^3 - 3y_0^3 = 0$, d.h. $(x_0, y_0) = (0, 0)$ oder $(x_0, y_0) = (\sqrt[3]{2}, \sqrt[3]{4})$.





Ohne die Lösungskurve $x \mapsto y(x)$ explizit zu kennen, können wir durch Differenzieren der impliziten Gleichung $0 = f(x, y(x))$ dennoch deren Ableitung berechnen:

$$0 = \partial_1 f(x, y(x)) + \partial_2 f(x, y(x))\, y'(x) = 3x^2 - 3y(x) + (3y(x)^2 - 3x)\, y'(x) \Rightarrow$$
$$y'(x) = \frac{y(x) - x^2}{y(x)^2 - x}.$$

Mittels der Cardano'sche Formeln kann man diese kubische Gleichung auch explizit nach x (oder y) auflösen und erhält für die obere Hälfte für $-\infty < x < \sqrt[3]{4}$:

$$y = \sqrt[3]{\frac{2}{\sqrt{1 - \frac{4}{x^3}} - 1}} + x \sqrt[3]{\frac{\sqrt{1 - \frac{4}{x^3}} - 1}{2}}.$$

Besser ist allerdings Polarkoordinaten zu verwenden, d.h. $x = r\cos(\varphi)$ und $y = r\sin(\varphi)$ zu setzen und somit $r^3(\sin(\varphi)^3 + \cos(\varphi)^3) = 3r^2 \sin(\varphi)\cos(\varphi)$, also

$$r = \frac{3\sin(\varphi)\cos(\varphi)}{\sin(\varphi)^3 + \cos(\varphi)^3}$$

zu erhalten. Oder, damit dieser Ausdruck rational wird, können wir einen neuen Parameter $s = \tan(\varphi)$ einführen und erhalten

$$x = r\cos(\varphi) = \frac{3\sin(\varphi)\cos(\varphi)^2}{\sin(\varphi)^3 + \cos(\varphi)^3} = \frac{3\tan(\varphi)}{\tan(\varphi)^3 + 1} = \frac{3s}{1 + s^3}$$
$$y = r\sin(\varphi) = \frac{3\sin(\varphi)^2\cos(\varphi)}{\sin(\varphi)^3 + \cos(\varphi)^3} = \frac{3\tan(\varphi)^2}{\tan(\varphi)^3 + 1} = \frac{3s^2}{1 + s^3}.$$

Beachte jedoch, daß diese Parametrisierung bei $s = -1$ vom Pol im 2-ten Quadraten auf jenen im 4-ten Quadranten springt und der Nullpunkt längs des senkrechten Kurventeils nur für $s \to \pm\infty$ erreicht wird. Eine andere Parametrisierung die dieses Problem vermeidet erhalten wir, wenn wir als Parameter $t = \tan(\varphi/2)$ einführen. Dann ist

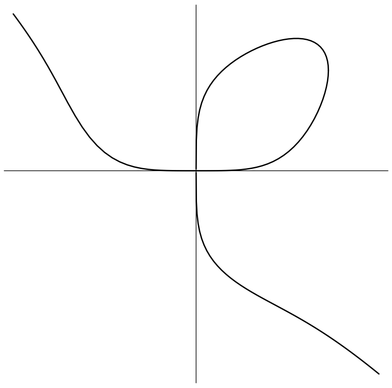
$$\sin(\varphi) = \frac{2t}{1 + t^2}$$
$$\cos(\varphi) = \frac{1 - t^2}{1 + t^2}$$
$$r = \frac{6t(1 - t^2)(1 + t^2)}{8t^3 + (1 - t^2)^3}$$
$$x = \frac{-6(t^2 - 1)^2 t}{(t^2 - 2t - 1)(t^4 + 2t^3 + 2t^2 - 2t + 1)}$$
$$y = \frac{12(t^2 - 1)t^2}{(t^2 - 2t - 1)(t^4 + 2t^3 + 2t^2 - 2t + 1)}$$

für $1 - \sqrt{2} < t < 1 + \sqrt{2}$.

3. Betrachten wir noch folgende implizite Gleichung die sich nun nicht mittels Wurzeln lösen läßt:

$$f(x,y) := x^5 + y^5 - 5\,x\,y = 0.$$

Die Ableitung von f ist $f'(x,y) = (5(x^4 - y), 5(y^4 - x))$, also ist nach dem Satz (6.2.4) über impliziten Funktionen y lokal um y_0 nach x nahe x_0 auflösbar falls $y_0^4 \neq x_0$. Andernfalls ist $y_0^4 = x_0$ und wegen $f(x_0, y_0) = 0$ ist $0 = y_0^{20} + y_0^5 - 5y_0^5 = y_0^{15}(y_0^4 - 4)$, d.h. $y_0 = 0 = x_0$ oder ($y_0 = \sqrt[2]{2}$ und $x_0 = 2$).

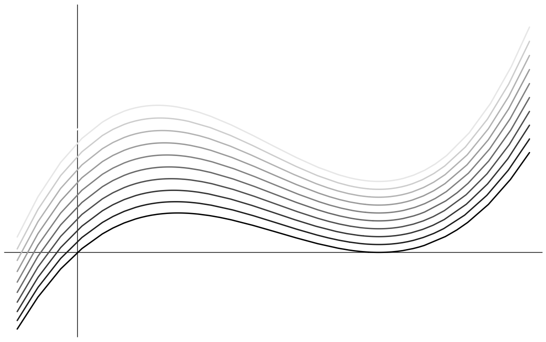


4. Auf Grund des Fundamentalsatzes der Algebra wissen wir, daß jedes Polynom $p_a(x) := p(x) := a_n\,x^n + a_{n-1}\,x^{n-1} + \dots + a_1\,x + a_0$ von Grad $n > 0$ (komplexe) Nullstellen besitzt. Es ist natürlich interessant wie sich die Nullstellen ändern, wenn man die Koeffizienten (leicht) ändert, denn in den angewandten Wissenschaften lassen sich die Koeffizienten ja üblicherweise nur durch Messungen näherungsweise bestimmen. Im Fall $n = 2$ kennen wir die Formel $x = \frac{a_1 \pm \sqrt{a_1^2 - 4a_0\,a_2}}{2a_0}$, welche die beiden Nullstellen als zwei Funktionen $\mathbb{C}^3 \to \mathbb{C}$ in den 3 Koeffizienten ausdrückt. Für $n > 4$ gibt es bekanntlich keine solche Formel. Wir wollen also in der impliziten Gleichung

$$p(a,x) := \sum_{k=0}^{n} a_k\,x^k = 0$$

x nach a auflösen. Offensichtlich ist $p : \mathbb{R}^{n+1} \times \mathbb{R} \to \mathbb{R}$ C^1 und $\partial_2 p(a,x) = \sum_{k=1}^n a_k\,k\,x^{k-1} = \sum_{k=0}^{n-1}(k+1)\,a_{k+1}\,x^k = p_a'(x)$. Für eine Nullstelle x_0 eines Polynoms p ist $p'(x_0) = 0$ genau dann, wenn x_0 eine Mehrfachnullstelle ist. Denn genau dann läßt sich $p(x)$ durch $x - x_0$ dividieren, d.h. $p(x) = (x - x_0)\,p_1(x)$ für ein Polynom p_1 vom Grad $n-1$ und somit ist x_0 genau dann eine Mehrfachnullstelle von p, wenn $p'(x_0) = p_1(x_0) = 0$ ist, denn $p'(x) = 1\,p_1(x) + (x - x_0)\,p_1'(x)$.

Lokal hängen also reelle Einfachnullstellen auf stetig differenzierbare Weise von den Koeffizienten des Polynoms ab.



Obwohl es keine Formel für diese Lösungsfunktion $x = x(a)$ gibt so können wir dennoch deren Ableitung bestimmen indem wir $p(a, x(a)) = 0$ differenzieren und erhalten $x'(a) = -\partial_2 p(a, x(a))^{-1} \circ \partial_1 p(a, x(a)) = -\frac{\partial_1 p(a, x(a))}{p_a'(x(a))}$. Insbesonders erhalten wir als partielle Ableitung nach dem i-ten Koeffizenten

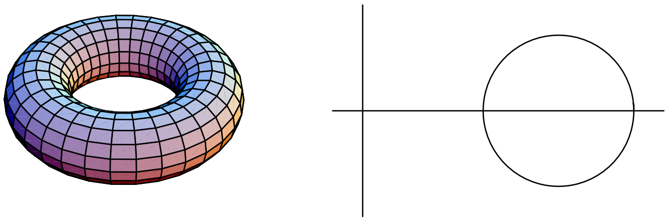
$$\partial_i x(a) = -\frac{x(a)^i}{p_a'(x(a))}.$$

Für den Fall von reellen Mehrfachnullstellen hat es vor kurzem Fortschritte gegeben, siehe [18].

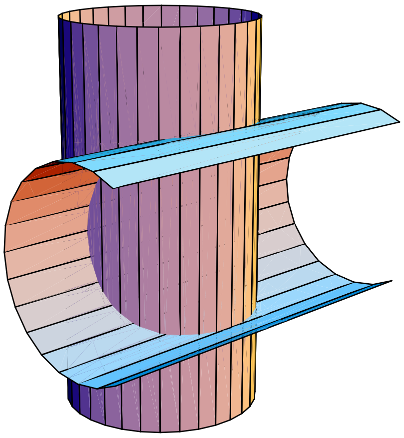
5. Es ist die Menge der Lösungen der impliziten Gleichung $f(x,y,z) := z^2 + (\sqrt{x^2+y^2} - 5)^2 = 4^2$ gesucht, ein sogenannter TORUS. Die Ableitung ist

$$f'(x,y,z) = \Big(2\big(\sqrt{x^2+y^2}-5\big)\frac{x}{\sqrt{x^2+y^2}}, 2\big(\sqrt{x^2+y^2}-5\big)\frac{y}{\sqrt{x^2+y^2}}, 2z\Big)$$

Für $z_0 \neq 0$ ist also z nahe z_0 eindeutig aus (x,y) berechenbar (in der Tat läßt sich z explizit als $z = \pm\sqrt{4^2 - (\sqrt{x^2+y^2}-5)^2}$ berechnen), beschreibt also eine Fläche im $\mathbb{R}^3$. Ist andererseits $z_0 = 0$ und $f(x_0,y_0,z_0) = 0$, so ist $\sqrt{x_0^2+y_0^2} = 5 \pm 2 \neq 0$ und somit $f'(x,y,z) \neq (0,0,0)$ also sicher eine der 3 Variablen aus den anderen beiden auf C^1-Weise berechenbar (Z.B. aus $x^2+y^2 = (5+\sqrt{4^2-z^2})^2$).



6. Es ist die Schnittmenge 2 normal aufeinander stehende Zylinder, deren Achsen sich nicht treffen, zu bestimmen.



Der 1. Zylinder ist z.B. durch $(x+1)^2+y^2 = 4$ gegeben und der 2. Zylinder durch $(x-1)^2+z^2 = 4$. Gesucht sind also die Lösungen der impliziten Gleichung $f(x,y,z) := ((x+1)^2+y^2, (x-1)^2+z^2) = (4,4)$. Die Ableitung von f ist

$$f'(x,y,z) = \begin{pmatrix} 2x+2 & 2y & 0 \\ 2x-2 & 0 & 2z \end{pmatrix}.$$

Wir fassen f als Abbildung $f : \mathbb{R}\times\mathbb{R}^2 \to \mathbb{R}^2$ auf. Dann ist $\partial_2 f(x,y,z) = \begin{pmatrix} 2y & 0 \\ 0 & 2z \end{pmatrix}$ genau dann invertierbar, wenn $yz \neq 0$ ist. Andererseits folgt aus $y = 0$, daß $(x+1)^2 = 4$, d.h. $x = -1 \pm 2$ und somit ist wegen $(x-1)^2 + z^2 = 4$ im $+$-Fall $z = \pm 2$ und sonst existiert kein passendes z. Aus $z = 0$ folgt hingegen, daß $(x-1)^2 = 4$, d.h. $x = 1 \pm 2$ ist und somit ist wegen $(x+1)^2+y^2 = 4$ im $-$-Fall $y = \pm 2$ und sonst existiert kein passendes y. D.h. abgesehen von den 4 Punkten $(1,0,\pm 2)$ und $(-1,\pm 2,0)$ ist also y und z lokal eindeutig aus x ausrechenbar.

6.2.6 Differenzierbarkeit des Invertierens.
Sehr relevant für die angewandten Wissenschaften ist die Frage wie sich die Lösungen eines (linearen) Gleichungssystems ändern, wenn wir dessen Koeffizienten leicht ändern. Da das Gauß'sche Eliminiationsverfahren nicht stetig ist, ist die folgende Antwort keineswegs offensichtlich.

Betrachten wir vorerst (der Kompliziertheit ;-) halber) den Fall von (invertierbaren) 2×2-Matrizen A. Wir wollen insbesonders die Differenzierbarkeit von $\mathrm{inv} : A \mapsto A^{-1}$ zeigen. Dazu berechnen wir die Richtungsableitung

$$\mathrm{inv}'(A)(B) := d_B\,\mathrm{inv}(A) = \lim_{t\to 0+} \frac{(A+tB)^{-1}-A^{-1}}{t} = \frac{d}{dt}\Big|_{t=0} (A+t\,B)^{-1}.$$

Sei dazu

$$A := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \text{und} \qquad B := \begin{pmatrix} h & k \\ l & m \end{pmatrix}.$$

Dann ist

$$\begin{aligned}
(A+t\,B)^{-1} &= \begin{pmatrix} a+t\,h & b+t\,k \\ c+t\,l & d+t\,m \end{pmatrix}^{-1} \\
&= \frac{1}{(a+t\,h)(d+t\,m)-(b+t\,k)(c+t\,l)} \begin{pmatrix} d+t\,m & -(b+t\,k) \\ -(c+t\,l) & (a+t\,h) \end{pmatrix} \\
&= \begin{pmatrix} \frac{d+t\,m}{(a+t\,h)(d+t\,m)-(b+t\,k)(c+t\,l)} & -\frac{b+t\,k}{(a+t\,h)(d+t\,m)-(b+t\,k)(c+t\,l)} \\ -\frac{c+t\,l}{(a+t\,h)(d+t\,m)-(b+t\,k)(c+t\,l)} & \frac{a+t\,h}{(a+t\,h)(d+t\,m)-(b+t\,k)(c+t\,l)} \end{pmatrix} \\
\Rightarrow \quad \frac{d}{dt}(A+t\,B)^{-1} &= \frac{1}{((a+t\,h)(d+t\,m)-(b+t\,k)(c+t\,l))^2} \cdot \\
&\quad \cdot \begin{pmatrix} -d^2\,h + c\,d\,k + b\,d\,l - b\,c\,m + 2\,d\,k\,l\,t - 2\,d\,h\,m\,t + k\,l\,m\,t^2 - h\,m^2\,t^2 \\ b\,d\,h - a\,d\,k - b^2\,l + a\,b\,m - 2\,b\,k\,l\,t + 2\,b\,h\,m\,t - k^2\,l\,t^2 + h\,k\,m\,t^2 \\ c\,d\,h - c^2\,k - a\,d\,l + a\,c\,m - 2\,c\,k\,l\,t + 2\,c\,h\,m\,t - k\,l^2\,t^2 + h\,l\,m\,t^2 \\ -b\,c\,h + a\,c\,k + a\,b\,l - a^2\,m + 2\,a\,k\,l\,t - 2\,a\,h\,m\,t + h\,k\,l\,t^2 - h^2\,m\,t^2 \end{pmatrix} \\
\Rightarrow \quad d_B\,\mathrm{inv}(A) &= \frac{1}{(a\,d-b\,c)^2} \begin{pmatrix} -d^2\,h + c\,d\,k + b\,d\,l - b\,c\,m & b\,d\,h - a\,d\,k - b^2\,l + a\,b\,m \\ c\,d\,h - c^2\,k - a\,d\,l + a\,c\,m & -b\,c\,h + a\,c\,k + a\,b\,l - a^2\,m \end{pmatrix}
\end{aligned}$$

Wer noch immer nicht genug Koordinaten gesehen hat, der möge die Rechnung nun für 3×3-Matrizen durchführen. Um zu verstehen, was dieser Ausdruck wirklich bedeutet, behandeln wir das Problem nun Koordinaten-frei, also warum nicht gleich für Banach-Algebren.

Folgerung.
Es sei A eine Banach-Algebra. Dann ist die Inversion $\mathrm{inv} : \mathrm{Inv}(A) \to A$, $a \mapsto a^{-1}$ C^1. *Weiters gilt*

$$\mathrm{inv}'(x)(h) = -x^{-1}\cdot h\cdot x^{-1}.$$

Beachte die entsprechende Formel in $\mathbb{R}$ und für kommutative Banach-Algebren:

$$\frac{d}{dx}x^{-1}\cdot h = \frac{d}{dx}\frac{1}{x}\cdot h = -\frac{1}{x^2}\cdot h = -x^{-1}\cdot h\cdot x^{-1}.$$

1-ter Beweis. Es ist x^{-1} durch die Gleichung $x\cdot x^{-1} = 1$ gegeben, also $\iota : x \mapsto x^{-1}$ durch die implizite Gleichung $\mu(x,\iota(x)) = 1$ wobei $\mu : A\times A \to A$ die Multiplikation ist. Da μ bilinear ist, ist $\partial_2\mu(x,y)(h) = \mu(x,h)$ und somit ist $\partial_2\mu(x,y)$ invertierbar mit inverser $h \mapsto \mu(x^{-1},h)$. Nach dem Satz (6.2.4) über impliziten Funktionen ist also $\iota : \mathrm{inv}(A) \to A$ differenzierbar mit Ableitung

$$\iota'(x)(h) = -\Big(\big(\partial_2\mu(x,\iota(x))\big)^{-1}\cdot\big(\partial_1\mu(x,\iota(x))\big)\Big)(h) = -x^{-1}\cdot h\cdot x^{-1}.$$

□

2-ter Beweis. Die Ableitung von inv können wir im Prinzip auch durch gliedweises Differenzieren der von Neumann Potenzreihe $f : x \mapsto (1-x)^{-1} = \sum_{k=0}^{\infty} x^k$ erhalten, aber falls x und h nicht kommutiert, so erhalten wir als Ableitung des k-ten Summanden $p : x \mapsto x^k$:

$$p'(x)(h) = hx^{k-1} + xhx^{k-2} + \cdots + x^{k-2}hx + x^{k-1}h.$$

An der Stelle 0 ist dies einfacher, denn 0 kommutiert mit allen $h \in A$ und somit erhalten wir $f'(0)(h) = 0 + h + 0 + \dots$ für die Ableitung der von Neumann Reihe f bei 0. Eine Anwendung der Kettenregel (6.1.5) auf $\mathrm{inv}(x) = f(1-x)$ liefert also $\mathrm{inv}'(1)(h) = f'(0) \cdot (-h) = -h$. Wegen $x^{-1} = x_0^{-1} \cdot (x \cdot x_0^{-1})^{-1}$ folgt nach Anwendung der Kettenregel auf

$$\mathrm{inv} : x \mapsto x \cdot x_0^{-1} \mapsto \mathrm{inv}\big(x \cdot x_0^{-1}\big) \mapsto x_0^{-1} \cdot \mathrm{inv}(x \cdot x_0^{-1}),$$

daß

$$\mathrm{inv}'(x_0)(h) = x_0^{-1} \cdot \Big(\mathrm{inv}'(1) \cdot x_0^{-1}\Big) = -x_0^{-1} \cdot h \cdot x_0^{-1}. \quad \square$$

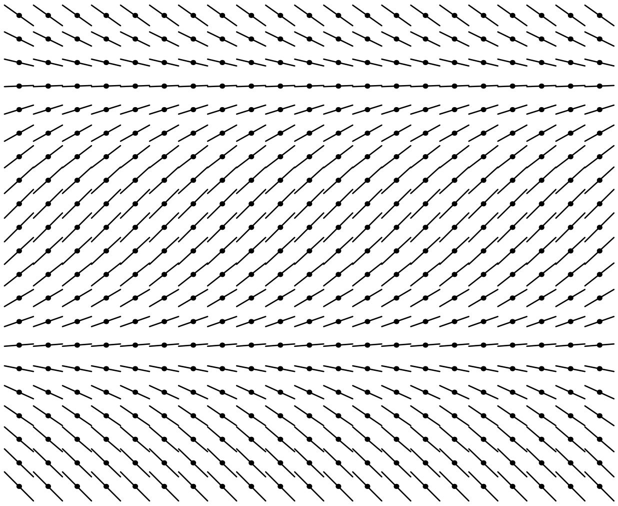
Gewöhnliche Differentialgleichungen

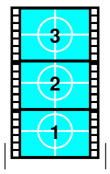
6.2.8a Bemerkung. Gewöhnliche (zeitunabhängige) Differentialgleichung (1-ter Ordnung). Um den nächsten Typ von Gleichung zu motivieren betrachten wir folgendes Problem. Ein Objekt bewege sich im Raum. Wir kennen zwar nicht den Ort an dem sich das Objekt zu jedem gegebenen Zeitpunkt befindet, wissen aber für jeden Ort, die Geschwindigkeit mit der das Objekt dort vorbeikommt (falls es dort überhaupt hin kommt). Können wir daraus den Aufenthaltsort zu jeden Zeitpunkt bestimmen?

Der einfachheithalber sei das Objekt vorerst ein Punkt der sich auf der Zahlengeraden bewegt. Diese Bewegung wird also durch eine Funktion $t \mapsto x(t)$ beschreiben, die jedem Zeitpunkt $t \in \mathbb{R}$ die Position $x(t) \in \mathbb{R}$ auf der Zahlengeraden zuordnet. Was wir kennen ist zu jedem Ort $x \in \mathbb{R}$ auf der Zahlengeraden die Geschwindigkeit (sagen wir $f(x)$) mit der das Objekt vorbekommt, d.h. Für jeden Zeitpunkt $t \in \mathbb{R}$ sollte die Geschwindigkeit $x'(t)$ gleich f an der Stelle $x(t)$ sein, also

$$x'(t) = f(x(t)) \text{ für alle } t \in \mathbb{R} \text{ oder kurz } x' = f \circ x$$

gelten. Dies ist eine sogenannte GEWÖHNLICHE DIFFERENTIALGLEICHUNG für die Funktion x welche durch die Funktion f beschrieben wird. Vom Graph der Funktion $x : \mathbb{R} \to \mathbb{R}$ wissen wir also nur, daß wenn er eine Höhe x_0 trifft, so muß seine Tangente den Richtungsvektor $(1, f(x_0))$ haben. Wir kennen also alle möglichen Tangenten und suchen die dazupassende Funktion $t \mapsto x(t)$.



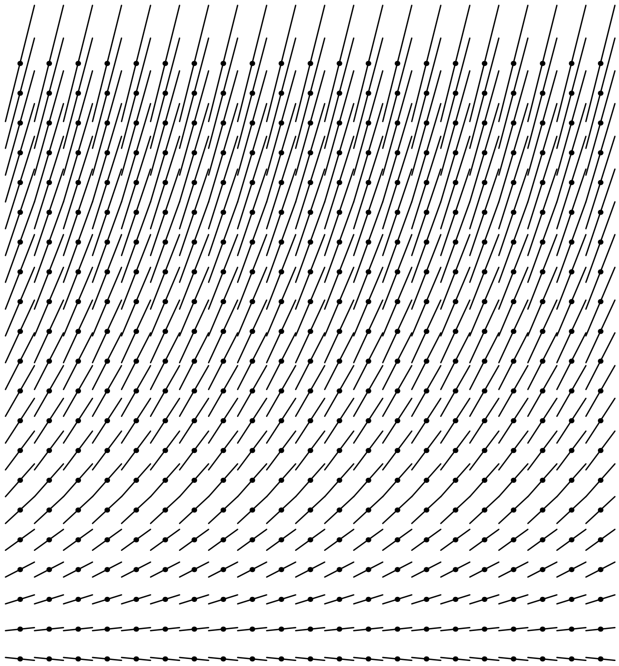


6.2.15 Beispiel. Exponentielles Wachstum.
Sei z.B. $f(x) := x$, die entsprechende Differentialgleichung lautet dann

$$x' = x,$$

d.h. der Zuwachs oder Gewinn $x'(t)$ ist proportional zu $x(t)$, also Kapitalismus in Reinkultur. Wir kennen die Lösung $x : t \mapsto e^t$. Da $x \mapsto x' - x$ linear ist, muß auch jedes Vielfache $t \mapsto c\, e^t$ mit $c \in \mathbb{R}$ eine Lösung sein. Wenn wir noch den Anfangs- oder Startwert $x(0) := x_0$ kennen, so ist $c = c\, e^0 = x(0) = x_0$, also $t \mapsto x_0\, e^t$ eine Lösung.

Es ist von vornherein nicht klar, ob es nicht weitere Lösungen gibt. Sei x irgendeine Lösung mit $x(0) = x_0$. Dann betrachten wir $c : t \mapsto x(t)e^{-t}$. Es ist $c'(t) = x'(t)\, e^{-t} - x(t)\, e^{-t} = 0$, da $x' = x$ vorausgesetzt ist. Also ist c konstant und somit $c(t) = c(0) = x(0)\, e^{-0} = x_0$, d.h. $x(t) = c(t)\, e^t = x_0\, e^t$. Also ist dies die einzige Lösung.



5.4.3 Beispiel. Nicht globale Lösungen.

Es sei nun $f(x) := 1 + x^2$. Eine Lösung x der Differentialgleichung $x' = f \circ x$ mit Anfangsbedingung $x(0) = 0$ kennen wir, nämlich $t \mapsto \tan(t)$. Dies ist auch die einzige Lösung: Sei nämlich x eine beliebige Lösung und $c(t) := \arctan(x(t))$. Dann ist

$$c'(t) = \arctan'(x(t))\, x'(t) = \frac{1}{1 + x(t)^2}\,(1 + x(t)^2) = 1,$$

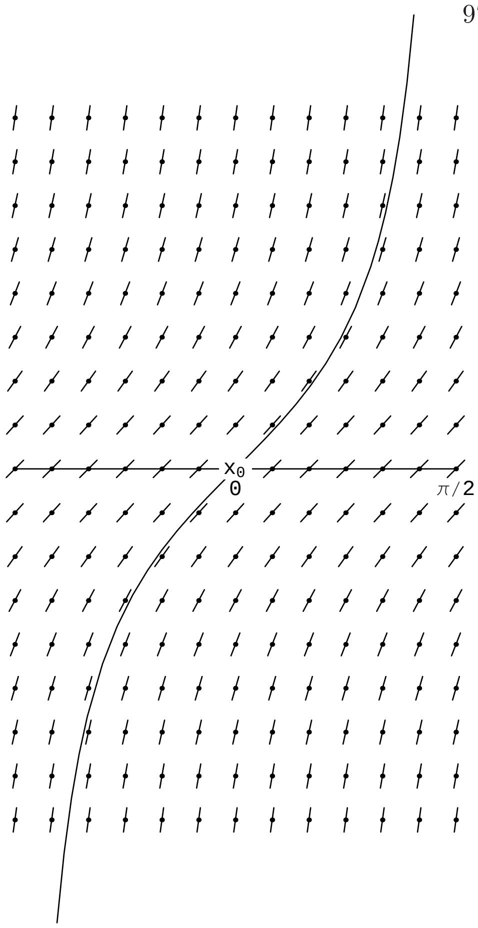
also

$$c(t) = c(0) + \int_0^t c' = t,$$

und somit

$$x(t) = \tan(t)$$

die einzige Lösung. Die Lösung existiert somit nur auf dem Intervall $(-\pi/2, +\pi/2)$ (und erreicht $\pm\infty$ zum Zeitpunkt $\pm\pi/2$) ganz im Unterschied zur global definierten Funktion f. Sie entwischt also in endlicher Zeit $\pi/2$ nach unendlich.



6.5.3b Bemerkung. Methode der Separation der Variablen.
Wie können wir nun symstematischer beim Lösen solcher Differentialgleichungen vorgehen? Wir rechnen dazu einmal wild drauf los wobei wir – wie schon öfters nützlich – mit $\frac{dx}{dt}$ wie mit einem Quotienten rechnen, und erhalten sukzessive:

$$\begin{aligned}\frac{dx}{dt} = x' = 1 + x^2 \quad &\Rightarrow \quad \frac{dx}{1+x^2} = dt \quad \Rightarrow \quad \arctan(x) = \int \frac{1}{1+x^2}\,dx = \int dt = t + C\\ &\Rightarrow \quad x(t) = \tan(t + C) \quad \Rightarrow \quad 0 = x(0) = \tan(0 + C) \quad \Rightarrow \quad C = 0\\ &\Rightarrow \quad x(t) = \tan(t).\end{aligned}$$

Diese Methode der Separation der Variablen hat somit zum Ziel geführt.

Sei nun f beliebig. So können wir mit der gleichen Methode vorgehen:

$$\begin{aligned}&\frac{dx}{dt} = x'(t) = f(x(t)) \Rightarrow \quad \frac{1}{f(x)}\,dx = dt \Rightarrow \quad G(x) := \int \frac{1}{f(x)}\,dx = \int dt = t + C\\ \Rightarrow \quad &x = G^{-1}(t + C) \text{ mit } C = G(x_0).\end{aligned}$$

Dies setzt natürlich voraus, daß wir sowohl $G := \int 1/f$ bestimmen können als auch dessen Umkehrfunktion G^{-1}. Indem wir nun die Probe machen, können wir uns überzeugen, daß diese etwas suspekte Rechnung wirklich eine Lösung $x(t) := G^{-1}(t + C) = G^{-1}(t + G(x_0))$ geliefert hat:

$$x'(t) = (G^{-1})'(t + C) = \frac{1}{G'(G^{-1}(t + C))} = f(G^{-1}(t + C)) = f(x(t)).$$

6.2.8a Bemerkung. Lösbarkeit zeitunabhängiger gewöhnlicher Differentialgleichungen.
Nun interessieren uns aber in Wirklichkeit keine Bewegungen auf der Zahlengerade sondern im (mindestens 3-dimensionalen) Raum E, d.h. f ist nun eine Funktion $f : E \supseteq U \to E$ und gesucht ist eine

Kurve $x : \mathbb{R} \supseteq I \to U$ mit $x' = f \circ x$, d.h. $x'(t) = f(x(t))$ für alle $t \in I$ und gegebenen Anfangswert $x(0) = x_0$, o.B.d.A. sei $x_0 = 0$. Die Methode der Separation der Variablen können wir nun nicht mehr anwenden, denn x ist vektorwertig, und somit macht $\int \dots dx$ keinen Sinn. Wir müssen also einen abstrakten lokalen Existenz und Eindeutigkeitssatz zeigen.

Wir können die Gleichung $x' = f \circ x$ auch als $(d - f_*)(x) = 0$ schreiben, wenn d das Differenzieren $x \mapsto x'$ bezeichnet. Es ist also eine Gleichung $\varphi(x) = 0$ zu lösen, wobei $\varphi := d - f_*$ eine Abbildung ist, die aus Kurven x neue Kurven $x' - f \circ x$ macht und die Lösung x eine Kurve ist. Dies ist nicht unähnlich zum Satz (6.2.4) über implizite Funktionen, wo wir zu gegebener Funktion $f : (t, x) \mapsto f(t, x)$ eine (lokal definierte) Funktion $x : t \mapsto x(t)$ gesucht haben mit $f(t, x(t)) = 0$, also $f \circ (\mathrm{id}, x) = 0$.

Der wesentlich Unterschied liegt nun darin, daß wir die impliziten Gleichungen zumindest lokal punktweise lösen konnten, d.h. zu jedem t getrennt die Lösung x_t von $f(t, x_t) = 0$ bestimmen konnten. Bei Differentialgleichungen geht dies nicht mehr, denn in der Gleichung für $x(t)$ kommt auch $x'(t)$ vor und dieses involviert Werte von x in der Nähe von t. Man kann also nur mehr die Funktion als ganzes und nicht mehr einzelne Funktionswerte getrennt berechnen.

Ein weiteres Problem ist, daß die Lösung x nur auf einem im vorhinein nicht bekannten Intervall J zu existieren braucht, wir also nicht sagen können in welchem Raum $C^1(J, U)$ wir die Lösung suchen. Sei $[-a, a] \subseteq J$. Dann können wir die Lösung umparametrisieren und erhalten $x_a : [-1, 1] \to U$, $t \mapsto x(a\,t)$. Für diese Kurve x_a gilt: $x_a'(t) = a\,x'(a\,t) = a\,f(x(a\,t)) = a\,f(x_a(t))$. Wenn wir für Kurven x mit $\bar f(a, x)$ die Kurve $t \mapsto x'(t) - a\,f(x(t))$ bezeichnet, so ist $\bar f$ eine Abbildung von $\mathbb{R} \times C^1([-1, 1], U) \to C([-1, 1], E)$ und x_a eine Lösung der impliziten Gleichung $\bar f(a, x_a) = 0$.

Um den Satz (6.2.4) über impliziten Funktionen anwenden zu können, müssen wir also nur die stetige Differenzierbarkeit von $\bar f$ und die Invertierbakeit von $\partial_2 \bar f(a, x)$ zeigen. Da $d : C^1([-1, 1], E) \to C([-1, 1], E)$ linear und stetig ist und $f_* : C([-1, 1], U) \to C([-1, 1], E)$ nach (6.1.20) C^1 ist, ist $\bar f : (a, x) \mapsto x' - a\,f \circ x$ ebenfalls C^1 und $\partial_2 \bar f(0, 0)(v) = v' - 0 \cdot (f_*)'(0) \cdot v = v'$. Nach dem Hauptsatz (5.6.17) ist somit $\partial_2 \bar f(0, 0)$ invertierbar, wenn wir für v nur solche Kurven zulassen, die $v(0) = 0$ erfüllen.

Nach dem Satz (6.2.4) über impliziten Funktion erhalten wir $\varepsilon > 0$ und $\varepsilon' > 0$ so daß für alle $|a| \le \varepsilon$ ein eindeutig bestimmtes $x_a \subseteq C^1(I, U)$ mit $\|x_a\|_\infty < \varepsilon'$ existiert, s.d. $x_a(0) = 0$ und $\bar f(a, x_a) = 0$.

Nun ist $x : t \mapsto x_\varepsilon(\frac{t}{\varepsilon})$ für $|t| \le \varepsilon$ die eindeutige Lösung der ursprünglichen Gleichung, denn

$$x'(t) - f(x(t)) = \tfrac{1}{\varepsilon} x_\varepsilon'(\tfrac{t}{\varepsilon}) - f(x_\varepsilon(\tfrac{t}{\varepsilon})) = \tfrac{1}{\varepsilon}\left(x_\varepsilon'(\tfrac{t}{\varepsilon}) - \varepsilon\, f(x_\varepsilon(\tfrac{t}{\varepsilon}))\right) = \tfrac{1}{\varepsilon}\left(\bar f(\varepsilon, x_\varepsilon)(\tfrac{t}{\varepsilon})\right) = 0.$$

Wir habe also folgendes Theorem gezeigt:

Theorem.
Es sei $f : E \supseteq U \to E$ C^1 und $x_0 \in U$. Dann existiert ein offenes Intervall $I \subseteq \mathbb{R}$ um 0 und eine eindeutig bestimmte Kurve $x : I \to U$ mit $x(0) = x_0$ und $x'(t) = f(x(t))$ für alle $t \in I$. □

Ganz analog zeigen wir nun einen Existenz- und Eindeutigkeitssatz für parameterabhängige gewöhnliche Differentialgleichungen 1-ter Ordnung mit spezieller Anfangsbedingung.

6.2.8 Theorem. Parameterabhängigkeit von Differentialgleichungen.
Es sei $f : E \times P \supseteq W \to E$ C^1 und $(0, p_0) \in W$. Dann existieren offene Umgebungen I von 0 und V von p_0, so daß für jedes $p \in V$

$$x'(t) = f(x(t), p) \text{ und } x(0) = 0$$

eine eindeutige Lösung $x_p : I \to U$ hat. Weiters ist $(t, p) \mapsto x_p(t)$ eine C^1-Abbildung $I \times V \to U \subseteq E$.

Beweis. Es sei x_p eine Lösung auf $[-a, a]$ für irgendein kleines $a > 0$ der Differential-Gleichung

$$x'(t) = f(x(t), p) \text{ und } x(0) = 0.$$

Wir setzen $\bar x_{p,a}(s) := x_p(a\,s)$. Dann existiert $\bar x$ auf $I := [-1, 1]$ und erfüllt

$$\bar x'(s) = a\,f(\bar x(s), p) \text{ und } \bar x(0) = 0.$$

Es sei

$$\bar{f}(p,a;\bar{x})(s) := \bar{x}'(s) - a\, f(\bar{x}(s),p),$$

dann ist

$$\bar{f} : (P\times\mathbb{R})\times\{\bar{x}\in C^1(I,E) : x(0)=0\}\to C(I,E).$$

Wegen (6.1.20) ist $\bar{f}$ C^1 und für die 2-te partielle Ableitung erhalten wir

$$\partial_2\bar{f}(p,a;x)(h)(s) = h'(s) - a\,\partial_1 f(\bar{x}(s),p)\cdot h(s).$$

Für $(p,a)=(p_0,0)$ ist $\partial_2\bar{f}(p_0,0,\bar{x})\cdot h = h'$ und daher ist $\partial_2\bar{f}(p_0,0,\bar{x}) : \{h\in C^1(I,E) : h(0) = 0\}\to C(I,E)$ wegen des Hauptsatzes (5.6.17) invertierbar (!). Nach dem Satz (6.2.4) über impliziten Funktion erhalten wir $\varepsilon>0$ und eine Umgebung V von p_0, so daß für alle (a,p) mit $|a|\le\varepsilon$ und $p\in V$ ein eindeutig bestimmtes $\bar{x}\in C^1(I,E)$ existiert, mit $\bar{x}(0)=0$ und $\bar{f}(p,a,\bar{x})=0$.

Man beachte für $|r|\le 1$ die Homogenitäts-Relation

$$\bar{f}(p,r\,a,\bar{x}\circ r)(s) = (\bar{x}\circ r)'(s) - ra\,f(\bar{x}(rs),p) = r\bar{x}'(rs) - ra\,f(\bar{x}(rs),p) = r\,\bar{f}(p,a,\bar{x})(r\,s)$$

Somit ist $\bar{x}_{p,ra}(s)=\bar{x}_{p,a}(rs)$. Für $|t|\le\varepsilon$ sei nun $x_p(t) := x(t,p) := \bar{x}_{p,\varepsilon}(s) = \bar{x}_{p,\varepsilon}(t/\varepsilon) = \bar{x}_{p,\varepsilon t/t_0}(t_0/\varepsilon)$, wobei $s\,\varepsilon := t$, $s_0\,\varepsilon := t_0$ und $t = t_0\,t/t_0$ für $|t|\le|t_0|\le\varepsilon$. Dieser Ausdruck ist C^1 in (t,p), da $\mathrm{ev}_{s_0}\circ\bar{x} : (a,p)\mapsto\bar{x}_{p,a}\mapsto\bar{x}_{p,a}(s_0)$ C^1 ist. Und weiters ist x_p eine Lösung der Differentialgleichung, denn

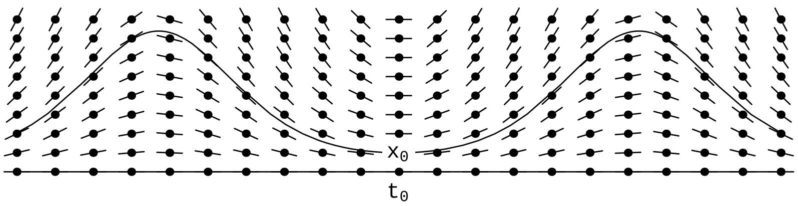
$$\frac{\partial}{\partial t}x_p(t) = \frac{\partial}{\partial t}\bar{x}_{p,\varepsilon}\left(\frac{t}{\varepsilon}\right) = \frac{1}{\varepsilon}\bar{x}'_{p,\varepsilon}\left(\frac{t}{\varepsilon}\right) = \frac{1}{\varepsilon}\varepsilon f(\bar{x}_{p,\varepsilon}(s),p) = f(x_p(t),p). \quad \square$$

6.2.7 Definition. Gewöhnliche zeitabhängige Differentialgleichung.
Für die Anwendungen werden die Differentialgleichungen durch Funktionen f beschrieben sein, die zusätzlich auch von der Zeit t abhängen. Man denke z.B. an das Problem die Fahrt eines Heißluftballons zu bestimmen, wenn man an jedem Punkt der Erde (oder besser des Luftraums) die Windgeschwindigkeit kennt, sich diese im Laufe des Tages und damit der Fahrt aber klarerweise ändert.

Unter einer GEWÖHNLICHEN DIFFERENTIALGLEICHUNG 1-TER ORDNUNG versteht man deshalb Gleichungen der Form

$$x'(t) = f(t,x(t)) \quad \text{mit} \quad x(t_0)=x_0,$$

wobei $f:\mathbb{R}\times E\supseteq I\times U\to E$ für ein offenes Intervall I in $\mathbb{R}$, welches t_0 enthält und eine offene Umgebung U von x_0 in einem Banach Raum E ist (so ein f heißt (ZEITABHÄNGIGES VEKTORFELD auf U) und wobei die Lösung $x : I\supseteq J\to U\subseteq E$ eine auf einem offenen Intervall $J\subseteq I$ differenzierbare Kurve sein soll, welches t_0 enthält.



Im Falle, daß $f(t,x)$ nicht von der 2-ten Variable abhängt (also ein ORTSUNABHÄNGIGES VEKTORFELD ist), d.h. $x'(t)=f(t)$ zu lösen ist, wissen wir nach (5.2.2), daß $x : t\mapsto x_0+\int_{t_0}^t f(s)\,ds$ die eindeutig bestimmte Lösung ist, falls f stetig ist.

Den Fall, daß $f(t,x)$ nicht von der 1-ten Variable abhängt (also ein ZEITUNABHÄNGIGES VEKTORFELD ist) haben wir bereits behandelt.

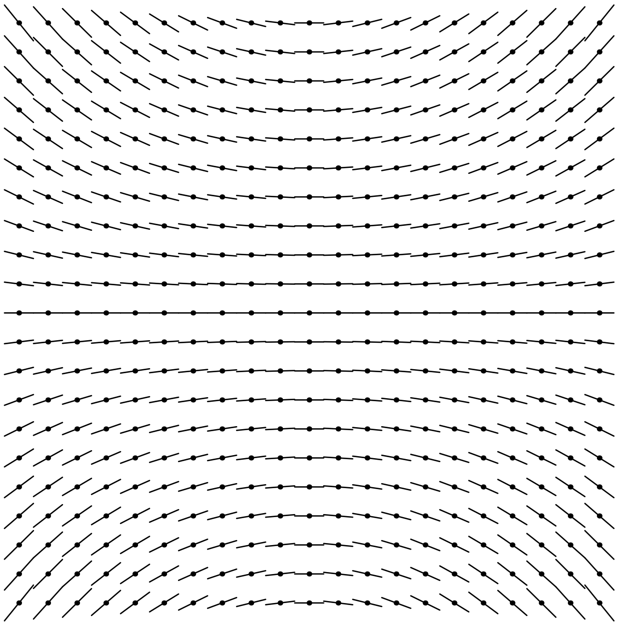
6.2.19 Beispiel einer zeitabhängigen Differentialgleichung.
Einfachstes nicht-triviales Beispiel einer zeitabhängigen Differentialgleichung ist wohl $f(t,x)=t\cdot x$, also die (Differentialgleichung)

$$x'(t) = \frac{dx}{dt} = t\cdot x(t).$$

Wir können wieder die Methode der Separation der Variablen anwenden und erhalten

$$\frac{dx}{dt} = t \cdot x \quad \Rightarrow \quad \frac{dx}{x} = t\,dt \quad \Rightarrow \quad \underbrace{\int \frac{1}{x}\,dx}_{=\log(x)} = \underbrace{\int t\,dt}_{\frac{t^2}{2}+C} \quad \Rightarrow \quad x = e^{\log(x)} = e^{\frac{t^2}{2}+C} = c\,e^{\frac{t^2}{2}}$$

mit einer Konstanten $c := e^C$.
Einsetzen von $x(t) := c\,e^{\frac{t^2}{2}}$ in die Differentialgleichung zeigt, daß diese eher heuristische Methode wieder zu einer Lösung geführt hat.



Klassisch ist folgendes Resultat, welches man durch direktes Zurückführen auf den Banach'schen Fixpunktssatz zeigt, was den Vorteil hat, daß man damit die Lösung direkt iterativ näherungsweise ausrechnen kann.

EXAM

5.4.2 Satz von Picard-Lindelöf.
Es sei $f : [-a, a] \times B \to E$ stetig und bzgl. der zweiten Variable eine Lipschitz *(d.h. $\exists L > 0\, \forall t, x_1, x_2 : \|f(t, x_1) - f(t, x_2)\| \le L\,\|x_1 - x_2\|$), wobei B der abgeschlossene Ball im Banach-Raum E mit Mittelpunkt x_0 und Radius b sei. Weiters, sei $a_1 := \min\{a, \frac{b}{\|f\|_\infty}\}$, wobei $\|f\|_\infty := \sup\{\|f(t, x)\| : |t| \le a, x \in B\}$.*
Dann existiert eine eindeutige (iterative erhaltbare) Lösung $x : [-a_1, a_1] \to B$ der Differentialgleichung

$$x'(t) = f(t, x(t)) \text{ mit } x(0) = x_0.$$

Beweis. Aus $\|f(t,x_1)-f(t,x_2)\| \le L\|x_1-x_2\|$ folgert man $\|f(t,x)-f(t,x_0)\| \le L\,b$ und da $t \mapsto f(t,x_0)$ stetig ist erhalten wir $\|f(t,x_0)\| \le \|f(.,x_0)\|_\infty$ und somit

$$\|f(t,x)\| \le \|f(t,x) - f(t,x_0)\| + \|f(t,x_0)\| \le L \cdot b + \|f(.,x_0)\|_\infty < \infty$$

Um den Banach'schen Fixpunkt-Satz (3.4.17) anzuwenden, übersetzen wir die Differential-Gleichung in eine äquivalente(!) Integral Gleichung auf dem abgeschlossenen Ball $X := \{x : \|x(t) - x_0\| \le b\}$ des vollständig metrischen Raumes $C([-a_1, a_1], E)$.
($\Rightarrow$) Aus $x'(t) = f(t, x(t))$ und $x(0) = x_0$ erhalten wir mittels dem Hauptsatz (5.6.17) $x(t) - x(0) = \int_0^t x'(s)\,ds = \int_0^t f(s, x(s))\,ds$. Also ist $x(t) = x_0 + \int_0^t f(s, x(s))\,ds$.
($\Leftarrow$) Umgekehrt erhalten wir aus $x(t) = x_0 + \int_0^t f(s, x(s))\,ds$

$$\frac{x(t+r) - x(t)}{r} = \frac{1}{r}\Big(\int_0^{t+r} f(s, x(s))\,ds - \int_0^t f(s, x(s))\,ds\Big) = \frac{1}{r}\int_t^{t+r} f(s, x(s))\,ds,$$

welches wegen dem Hauptsatz gegen $f(t, x(t))$ konvergiert.

Der entsprechende Integral-Operator $I : x \mapsto I(x) : t \mapsto x_0 + \int_0^t f(s, x(s))\, ds$ läßt X invariant, denn für $x \in X$ und $|t| \le a_1$ gilt

$$|I(x)(t) - x_0| = \left| \int_0^t f(s, x(s))\, ds \right| \le a_1 \, \|f\|_\infty \le b.$$

Und wenn man die äquivalente Norm $\|x\| := \max\{\|x(t)\| e^{-2L|t|} : t \in [-a_1, a_1]\}$ verwendet, ist I eine $\frac{1}{2}(1 - e^{-La_1})$-Kontraktion, denn

$$\begin{aligned}
\left| e^{-2\,L\,|t|} \left(I(x_1) - I(x_2)\right)(t) \right| &= \left| e^{-2\,L\,|t|} \int_0^t f(s, x_1(s)) - f(s, x_2(s))\, ds \right| \\
&\le e^{-2\,L\,|t|}\, L \cdot \left| \int_0^t |x_1(s) - x_2(s)|\, e^{-2\,L\,|s|}\, e^{2\,L\,|s|}\, ds \right| \\
&\le e^{-2\,L\,|t|}\, L \cdot \left| \int_0^t \|x_1 - x_2\|\, e^{2\,L\,|s|}\, ds \right| \\
&\le e^{-2\,L\,|t|}\, \frac{1}{2}\, \|x_1 - x_2\| \left(e^{2\,L\,|t|} - 1 \right) \\
&= \frac{1}{2}\, \|x_1 - x_2\| \left(1 - e^{-2\,L\,|t|} \right) \\
&\le \frac{1}{2}\, \|x_1 - x_2\|.
\end{aligned}$$

Also ist der Banach'sche Fixpunktsatz anwendbar und wenn man mit einer beliebigen Funktion in X, z.B. der konstanten Funktion x_0, startet und rekursiv $x_{n+1} := I(x_n)$ definiert, dann konvergiert x_n in X gegen den eindeutigen Fixpunkt x_∞, d.h. der Lösung der Differentialgleichung. □

Bemerkung. Äquivalenz der Lösbarkeitssätze verschiedener Gleichungstypen.
Die Beweise der lokalen Existenz und Eindeutigkeitsätze beruhen allesamt auf dem Banach'schen Fixpunktsatz (3.4.17), und zwar haben wir zuerst den Satz (6.2.2) über inverse Funktionen gezeigt, daraus dann (6.2.4) über implizite Funktionen gefolgert und daraus seinerseits (6.2.8), (6.2.8b) und (6.2.8c) über gewöhnliche Differentialgleichungen.

Man könnte aber auch umgekehrt vorgehen. Zuerst den Beweis von (5.4.2) führen, damit dann den Satz (6.2.4) über impliziten Funktionen im Fall $f : \mathbb{R} \times F \to F$ zeigen, wobei die lokale Lösung y durch die Differentialgleichung $\partial_1 f(t, y(t)) + \partial_2 f(t, y(t)) \cdot y'(t) = 0$, also $y'(t) = -\partial_2 f(t, y(t))^{-1} \cdot \partial_1 f(t, y(t))$, gegeben ist. Im allgemeinen Fall $f : E \times F \to F$ können wir für jedes $x \in E$ (mit $\|x\| = 1$) die lokale Lösung y der impliziten Gleichung als Lösung der Differentialgleichung $y'(tx) = -\partial_2 f(tx, y(tx))^{-1} \circ \partial_1 f(tx, y(tx))$ erhalten, brauchen dafür aber die differenzierbare Abhängigkeit der Lösung der Differentialgleichung vom Parameter $x \in E$. Dies ist mit obiger Methode aber eher aufwendig. Jedenfalls können wir diese Methode anwenden um die Lösung der impliziten Gleichung mittels der guten (iterativen) Methoden für Differentialgleichungen zu lösen.

Der Satz (6.2.2) über inverse Funktionen folgt nun seinerseits aus jenen über implizite Funktionen, denn f^{-1} ist durch die Gleichung $f \circ f^{-1} = \mathrm{id}$, also die implizite Gleichung $g(x, f^{-1}(x)) := f(f^{-1}(x)) - x = 0$ gegeben, wobei $\partial_2 g(x, y) := f'(y)$ ist. Da die lokale Lösung eindeutig ist, ist f lokal injektiv, also folgt daraus auch $f^{-1} \circ f = \mathrm{id}$, denn $f \circ (f^{-1} \circ f) = (f \circ f^{-1}) \circ f = \mathrm{id} \circ f = f$.

6.2.8b Folgerung. Zeitabhängige Differentialgleichung mit Parameter.
Das Theorem (6.2.8) gilt auch für zeitabhängige Differential-Gleichungen

$$x'(t) = f(t, x(t), p) \text{ mit } x(0) = 0.$$

Beweis. Es sei x_p eine Lösung der Differentialgleichung

$$x_p'(t) = f(t, x_p(t), p) \text{ und } x_p(0) = 0.$$

Für $\bar{x}(t,p) := (t, x_p(t))$ ist dann

$$\begin{aligned}\frac{\partial}{\partial t}\bar{x}(t,p) &= \Big(1, \frac{\partial}{\partial t}x(t,p)\Big) = (1, f(t, x(t,p), p)) = (1, f(\bar{x}(t,p), p)) =: \bar{f}(\bar{x}(t,p), p)\\ \text{mit}\quad \bar{x}(0,p) &= (0, x(0,p)) = (0,0)\end{aligned}$$

Also ist die zeitabhängige Differentialgleichung äquivalent zu einer Zeit-unabhängigen Differentialgleichung mit Parametern $(t,p) \in \mathbb{R} \times P$. □

6.2.8c Folgerung. Abhängigkeit von Anfangsbedingung der Differentialgleichung.
Das Theorem (6.2.8) gilt auch für zeitabhängige Differential-Gleichungen mit allgemeiner Anfangsbedingung

$$x'(t) = f(t, x(t), p) \text{ und } x(s) = y.$$

Beweis. Es sei $x_{p,y,s}$ eine Lösung von

$$x'(t) = f(t, x(t), p) \text{ und } x(s) = y.$$

Wir betrachten für

$$\bar{x}(t,p,y,s) := x(s+t, p, y, s) - y = x_{p,y,s}(t+s) - y$$

die Differentialgleichung

$$\begin{aligned}\frac{\partial}{\partial t}\bar{x}(t,p,y,s) &= \frac{\partial}{\partial t}\Big(x(s+t,p,y,s) - y\Big) = f(s+t, x(s+t,p,y,s), p) = f(s+t, \bar{x}(t,p,y,s)+y, p)\\ &= \bar{f}(t, \bar{x}(t,p,y,s), p, y, s) \quad \text{wobei} \quad \bar{f}(t, \bar{x}, p, y, s) := f(s+t, \bar{x}+y, p)\end{aligned}$$

und Anfangswert $\bar{x}(0,p,y,s) = x(s,p,y,s) - y = 0$ ist. Also ist das Problem auf eine zeitabhängige Gleichung mit spezieller Anfangsbedingung $0 \mapsto 0$ und Parameter $(p,y,s) \in P \times E \times \mathbb{R}$ reduziert. □

5.4.4 Folgerung. Lineare Differentialgleichung.
Es sei $f(t,x) := a(t)\cdot x + b(t)$, dann existieren die Lösungen dort wo $a : J \to L(E,E)$ und $b : J \to E$ C^1 sind.

Beweis. Es sei $x : J_0 \to E$ eine Lösung mit maximalen Definitionsinterval $J_0 \subseteq J$. Angenommen $t_1 := \sup J_0 \in J$. Dann existiert ein $\delta > 0$, s.d. eine Lösung $y : (t_1-\delta, t_1+\delta) \to E$ mit $y(t_0) = 0$ sowie eine Lösung $g : (t_1-\delta, t_1+\delta) \to GL(E) \subseteq L(E,E)$ von $g'(t) = a(t)\circ g(t)$ mit $g(t_1) = \mathrm{id}$ existiert. Sei nun $t_1 - \delta < t_2 < t_1$ und

$$\tilde{x}(t) := y(t) + g(t)\cdot g(t_2)^{-1}\cdot\Big(x(t_2) - y(t_2)\Big) \text{ für } t_1-\delta < t < t_1+\delta.$$

Dann ist $\tilde{x}(t_2) = y(t_2) + \mathrm{id}\cdot\big(x(t_2) - y(t_2)\big) = x(t_2)$ und $\tilde{x}'(t) = y'(t) + g'(t)g(t_2)^{-1}(x(t_2) - y(t_2)) = a(t)\,y(t) + b(t) + a(t)\,g(t)\,g(t_2)^{-1}\,(x(t_2) - y(t_2)) = a(t)\,\tilde{x}(t) + b(t)$, ein Widerspruch zur Maximalität von J_0. □

Bemerkung.

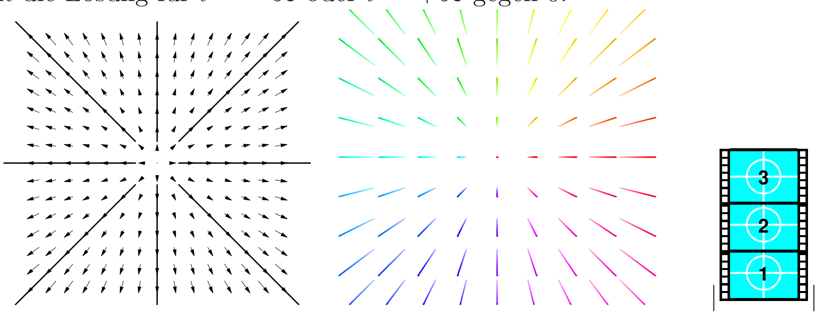
1. Falls $a \in \mathbb{R}$ konstant und $b = 0$ ist (also eine HOMOGENE LINEARE 1-DIMENSIONALE DIFFERENTIALGLEICHUNG MIT KONSTANTEN KOEFFIZIENTEN gegeben ist), dann ist die Lösung $x(t) = x(0)\,e^{at}$, siehe (6.2.15).
2. Dies gilt auch für $a \in L(E,E)$ (d.h. HOMOGENE LINEARE MEHRDIMENSIONALE DIFFERENTIALGLEICHUNGEN MIT KONSTANTEN KOEFFIZIENTEN), wobei $e^a := \exp(a) := \sum_{k=0}^{\infty}\frac{a^k}{k!}$ für alle $a \in L(E,E)$ konvergiert und $x(t) = e^{at}x_0$ die Anwendung der linearen Abbildung e^{at} auf den Anfangsvektor $x_0 \in E$ ist, siehe Aufgabe (5.53).

3. Falls $v \neq 0$ ein Eigenvektor zum Eigenwert λ von $a \in L(E,E)$ ist, d.h. $a(v) = \lambda\, v$ gilt, so liefert die Methode der VARIATION DER KONSTANTEN den Ansatz $x(t) := c(t)\, v$. Dies ist genau dann eine Lösung der Differentialgleichung $x'(t) = a(x(t))$, wenn $c'(t)\, v = x'(t) = a(x(t)) = a(c(t)\, v) = c(t)\, \lambda\, v$, also $c'(t) = \lambda\, c(t)$ und somit $c(t) = c(0)\, e^{\lambda\, t}$ ist.

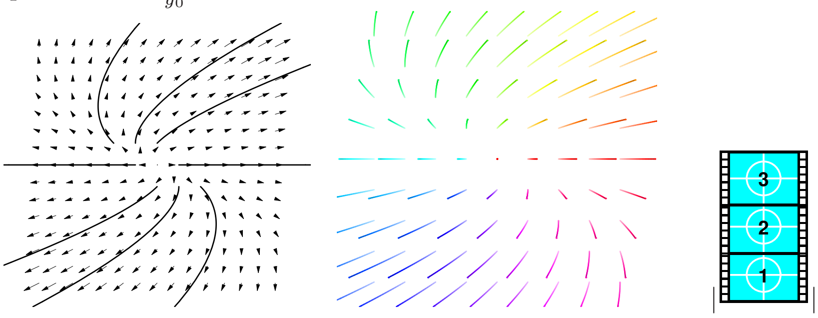
   Beachte, daß jede Linearkombination von Lösungen einer homogenen linearen Differentialgleichung selbst wieder eine Lösung ist. Wenn also $v_1, \dots, v_n$ Eigenvektoren von a mit Eigenwerten $\lambda_1, \dots, \lambda_n$ sind, so ist auch $x(t) = \sum_{k=1}^n c_i\, e^{\lambda_i\, t}\, v_i$ eine Lösung zum Anfangswert $\sum_i c_i\, v_i$. Wenn die v_i eine Basis bilden, so kann daraus jede Lösung berechnet werden.

4. Betrachten wir nun insbesonders den 2-dimensionalen Fall einer homogenen lineare Differentialgleichung
$$\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$
   Es seien λ und μ die beiden Eigenwerte dieser 2×2-Matrix A und v und w zugehörige Eigenvektoren. Dann können die folgenden Fälle auftreten:

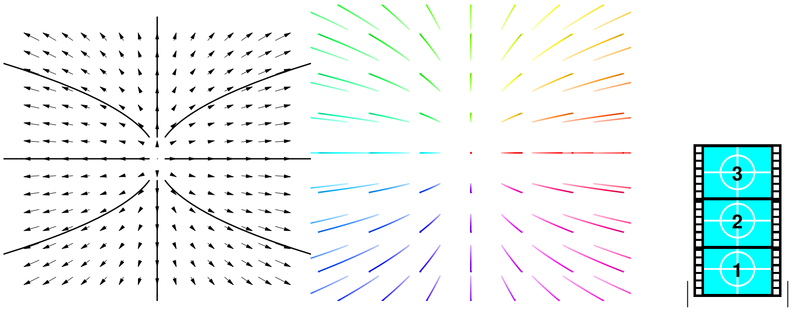
   - $\lambda = \mu = 0$ und $v \neq w$. Dann ist $A = 0$ und alle Lösungen konstant.
   - $\lambda = \mu = 0$ und $v = w$. Dies ist z.B. für $a = c = d = 0$, $b = 1$ der Fall, also für die Differentialgleichung $x'(t) = y(t)$, $y'(t) = 0$, d.h. $y(t) = y_0$ und $x(t) = x_0 + t\, y_0$. Die Lösungskurven sind also affine Geraden parallel zum Eigenvektor.
   - $\lambda \neq 0 = \mu$. Dies ist z.B. für $a = \lambda$, $b = c = d = 0$ der Fall, also für die Differentialgleichung $x'(t) = \lambda\, x(t)$, $y'(t) = 0$, d.h. $x(t) = x_0\, e^{\lambda t}$ und $y(t) = y_0$. Die Lösungskurven sind also durch die Exponentialfunktion parametrisierte Halbgeraden parallel zum Eigenvektor v. Je nach dem ob $\lambda > 0$ oder $\lambda < 0$ geht die Lösung für $t \to -\infty$ oder $t \to +\infty$ gegen die w-Achse.
   - $\lambda = \mu \neq 0$ und $v \neq w$. Dann ist $A = \lambda$ id und die Lösungen sind $x(t) = x_0\, e^{\lambda t}$, $y(t) = y_0\, e^{\lambda t}$ parametrisieren also gerade die Halbstrahlen durch 0. Je nach dem ob $\lambda > 0$ oder $\lambda < 0$ geht die Lösung für $t \to -\infty$ oder $t \to +\infty$ gegen 0.



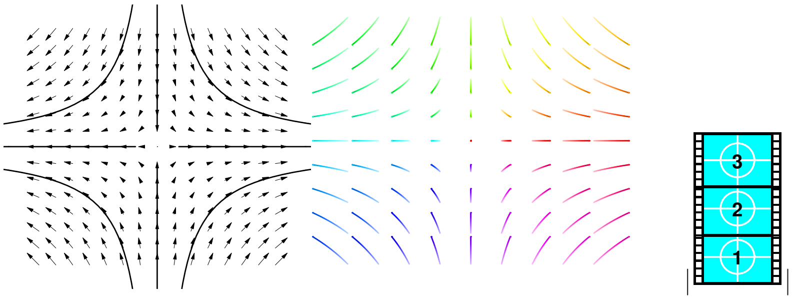
   - $\lambda = \mu \neq 0$ und $v = w$. Dies ist z.B. für $a = d = \lambda$, $b = 1$, $c = 0$ der Fall, also für $x' = \lambda\, x + y$, $y' = \lambda\, y$, d.h. $y(t) = y_0\, e^{\lambda\, t}$ und $x'(t) = c(t)\, e^{\lambda\, t}$ wobei $c'(t) = y_0$ ist, also $c(t) = x_0 + t\, y_0$. Die Lösungskurven gehen für $t \to \pm\infty$ gegen 0 und ∞ und schneiden dabei die y-Achse zum Zeitpunkt $t := -\frac{x_0}{y_0}$.



   - $\lambda > \mu > 0$ oder $\lambda < \mu < 0$. Dies ist z.B. für $a = \lambda$, $d = \mu$, $b = c = 0$ der Fall. Die Lösungen $x(t) = x_0\, e^{\lambda t}$, $y(t) = y_0\, e^{\mu t}$ beschreiben halbe verallgemeinerte Parabeln $(x/x_0)^\mu = (y/y_0)^\lambda$ die tangential an die y-Achse liegen.



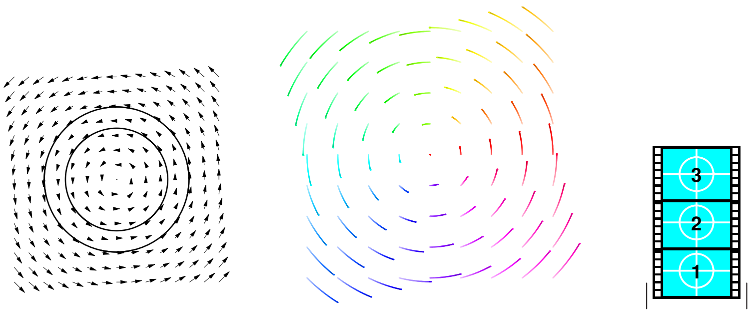
- $\lambda > 0 > \mu$. Wie zuvor sind die Lösungen $x(t) = x_0\, e^{\lambda t}$ und $y(t) = y_0\, e^{\mu t}$ und wegen $\lambda\, \mu < 0$ beschreiben diese verallgemeinerte Hyperbeln $1 = (x/x_0)^{-\mu} \cdot (y/y_0)^{\lambda}$.



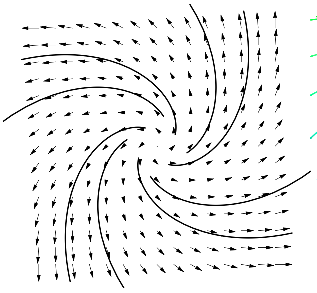
- $\mu = \overline{\lambda} \in \mathbb{C}$ und $\Re e(\lambda) = 0$. Dies ist z.B. für $a = 0 = d$, $b = -1$, $c = 1$ der Fall, also für die Differentialgleichung $x'(t) = -y(t)$, $y'(t) = x(t)$. Damit ist $x''(t) = -y'(t) = -x(t)$ und somit $x(t) = u\, \cos(t) + v\, \sin(t)$ und $y(t) = -x'(t) = u\, \sin(t) - v\, \cos(t)$. Dann ist $x_0 = x(0) = u$ und $y_0 = y(0) = -v$. Sei nun $(x_0, y_0) = r(\cos(\varphi), \sin(\varphi))$, dann ist

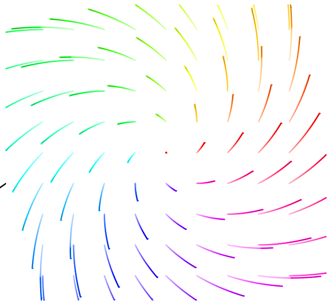
$$x(t) = r\, \cos(\varphi)\, \cos(t) - r\, \sin(\varphi)\, \sin(t) = r\, \cos(\varphi + t) \text{ und}$$
$$y(t) = r\, \cos(\varphi)\, \sin(t) + r\, \sin(\varphi)\, \cos(t) = r\, \sin(\varphi + t).$$

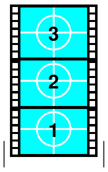
Die Lösungen parametrisieren also konzentrische Kreis um 0. Mittels der komplexen Lösungen $t \mapsto e^{\pm it} \binom{\pm i}{1}$ können wir das ebenfalls erhalten indem wir deren Summe $t \mapsto \binom{i(e^{it}-e^{-it})}{e^{it}+e^{-it}} = 2\binom{-\sin(t)}{\cos(t)}$ bzw. das $(-i)$-fache der Differenz $t \mapsto \binom{i(e^{it}+e^{-it})}{e^{it}-e^{-it}} = 2i\binom{\cos(t)}{\sin(t)}$ betrachten, siehe (4.2.6).



- $\mu = \overline{\lambda} \in \mathbb{C}$ und $\Re e(\lambda) \neq 0$. Dies ist z.B. für $a = c = d = 1$ und $b = -1$ der Fall mit Eigenwerten $1 \pm i$. Komplexe Eigenvektoren sind $(\pm i, 1)$ also sind $(x(t), y(t)) = e^{(1+i)t}(i, 1)$ und $(x(t), y(t)) = e^{(1-i)t}(-i, 1)$ zwei komplexe Lösungen. Die Summe dieser Lösungen liefert $e^t(-\sin(t), \cos(t))$ und die Differenz $i\, e^t(\cos(t), \sin(t))$, siehe (4.2.6). Also ist die reelle Lösung mit Anfangswert $r(\cos(\varphi), \sin(\varphi))$ durch $t \mapsto r\, e^t\big(\cos(\varphi + t), \sin(\varphi + t)\big)$ gegeben, parametrisiert somit eine Spirale die für $t \to \pm\infty$ gegen 0 bzw. ∞ konvergiert und dabei unendlich oft um 0 herumläuft.





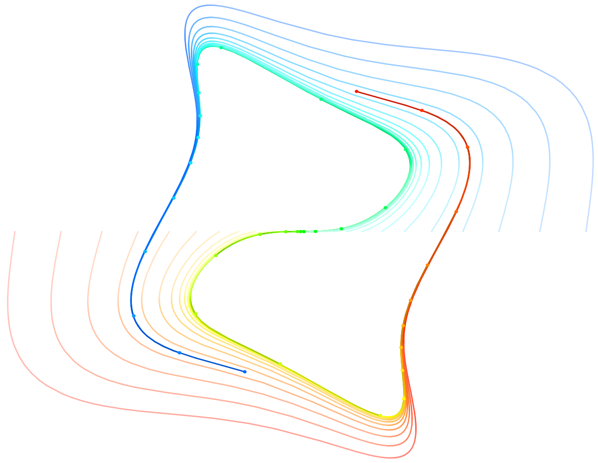


5. Ist $a : J \to \mathbb{R}$ beliebig und $x'(t) = a(t) \cdot x(t)$ eine sogenannte 1-DIMENSIONALE HOMOGENE LINEARE DIFFERENTIALGLEICHUNG MIT NICHT-KONSTANTEN KOEFFIZIENTEN, dann folgt aus $\frac{dx}{dt} = x'(t) = a(t)\, x(t)$ formal $\ln(x) = \int \frac{1}{x}\, dx = \int a(t)\, dt$ und somit
$$x(t) = e^{\int a(t)\, dt}.$$
Einsetzen zeigt, daß dies in der Tat die Lösung der Differentialgleichung ist.

6. Ist wie zuvor a Skalar-wertig, b beliebig und $y'(t) = a(t) \cdot y(t) + b(t)$ eine sogenannte 1-DIMENSIONALE INHOMOGENE LINEARE DIFFERENTIALGLEICHUNG MIT NICHT-KONSTANTEN KOEFFIZIENTEN, so erhält man die Lösung y der inhomogenen Gleichung $y'(t) = a(t)\, y(t) + b(t)$ durch VARIATION DER KONSTANTEN der Lösung x der homogenen Gleichung $x'(t) = a(t)\, x(t)$, d.h. durch den Ansatz $y(t) := c(t)\, x(t)$. Dann ist
$$a(t)\, y(t) + b(t) = y'(t) = c(t)\, x'(t) + c'(t)\, x(t) = c(t)\, a(t)\, x(t) + c'(t)\, x(t) = a(t)\, y(t) + c'(t)\, x(t),$$
also $c'(t) = \frac{b(t)}{x(t)}$, d.h. $c(t) = \int \frac{b(t)}{x(t)}\, dt$.

7. Für lineare Differentialgleichungen mit konstanten Koeffizienten können wir auch erfolgreich die JORDAN'SCHE NORMALFORM einsetzen wonach die lineare Abbildung a durch gegeignete Wahl einer Basis auf Block-diagonal-Form bringen läßt. Die Differentialgleichung sieht dann blockweise wie folgt aus:
$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \\ x_3'(t) \\ \vdots \\ x_n'(t) \end{pmatrix} = \begin{pmatrix} \lambda & 0 & \dots & \dots & 0 \\ 1 & \lambda & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & \lambda \end{pmatrix} \cdot \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$
Mittels Variation der Konstanten erhalten wir rekursiv:
$$\begin{aligned} x_1'(t) = \lambda x_1(t) &\Rightarrow x_1(t) = x_1(0)\, e^{\lambda t} \\ x_2'(t) = x_1(t) + \lambda x_2(t) &\Rightarrow x_2(t) = \Big(x_2(0) + x_1(0)\, t\Big)\, e^{\lambda t} \\ x_3'(t) = x_2(t) + \lambda x_3(t) &\Rightarrow x_3(t) = \left(x_3(0) + x_2(0)\, t + x_1(0)\, \frac{t^2}{2}\right) e^{\lambda t} \\ &\vdots \\ x_n'(t) = x_{n-1}(t) + \lambda x_n(t) &\Rightarrow x_n(t) = \left(\sum_{k=1}^{n} x_k(0) \frac{t^k}{k!}\right) e^{\lambda t}. \end{aligned}$$

8. Ist $x'(t) = f(x(t))$ eine zeitunabhängige nicht-lineare Differentialgleichung und $f(x_0) = 0$, so ist $x : t \mapsto x_0$ eine STATIONÄRE LÖSUNG ein sogenannter FIXPUNKT DER DIFFERENTIALGLEICHUNG. Indem man die Ableitung $A := f'(x_0)$ von f betrachtet, kann man im GENERISCHEN FALL (d.h. wo kleine Änderungen der Matrix die Eigenschaften der Eigenwerte wie z.B. deren Positivität nicht ändert) zeigen, daß sich die Lösungen der Differentialgleichung $x'(t) = f(x(t))$ nahe x_0 im wesentlichen so wie jene der linearen Differentialgleichung $x'(t) = A \cdot x(t)$ verhalten.



Wie untersuchen z.B. die nicht-lineare Differentialgleichung

$$\begin{aligned} x' &= x(4-y^2)-3y \\ y' &= y(4-x^2)+3x \end{aligned}$$

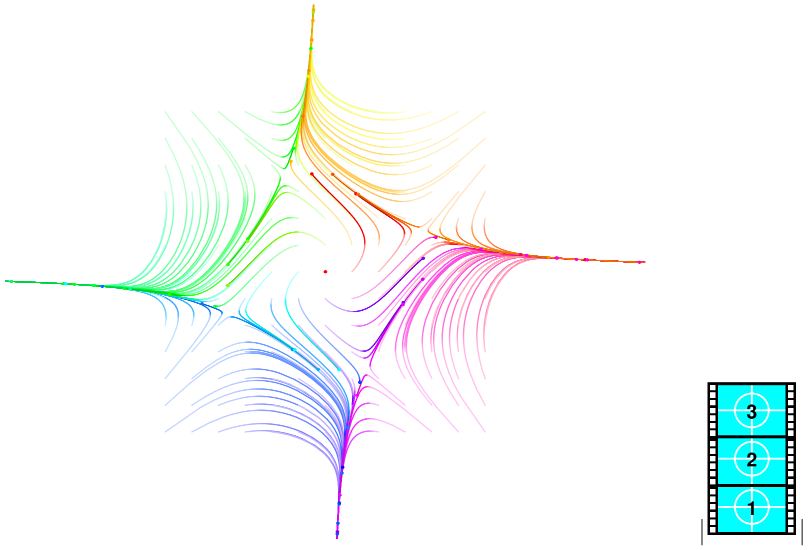
Die stationären Punkte sind die Lösungen der Gleichung $f(x,y)=0$ also des Gleichungssystems

$$\begin{aligned} 0 &= x(4-y^2)-3y \\ 0 &= y(4-x^2)+3x \end{aligned}$$

und somit $(\sqrt{10}, \frac{1}{2}\sqrt{10})$, $(-\frac{1}{2}\sqrt{10}, \sqrt{10})$, $(-\sqrt{10}, -\frac{1}{2}\sqrt{10})$, $(\frac{1}{2}\sqrt{10}, -\sqrt{10})$ und $(0,0)$. Die Ableitung f', deren Eigenwerte und deren Eigenvektoren sind in diesen Punkten gegeben durch:

Punkt	Jacobi-Matrix	Eigenwerte	Eigenvektoren
$(\sqrt{10}, \frac{1}{2}\sqrt{10})$	$\begin{pmatrix} 3/2 & -13 \\ -7 & -6 \end{pmatrix}$	$-25/2,\ 8$	$\binom{13}{14},\ \binom{-2}{1}$
$(-\frac{1}{2}\sqrt{10}, \sqrt{10})$	$\begin{pmatrix} -6 & 7 \\ 13 & 3/2 \end{pmatrix}$	$-25/2,\ 8$	$\binom{-14}{13},\ \binom{1}{2}$
$(-\sqrt{10}, -\frac{1}{2}\sqrt{10})$	$\begin{pmatrix} 3/2 & -13 \\ -7 & -6 \end{pmatrix}$	$-25/2,\ 8$	$\binom{13}{14},\ \binom{-2}{1}$
$(\frac{1}{2}\sqrt{10}, -\sqrt{10})$	$\begin{pmatrix} -6 & 7 \\ 13 & 3/2 \end{pmatrix}$	$-25/2,\ 8$	$\binom{-14}{13},\ \binom{1}{2}$
$(0,0)$	$\begin{pmatrix} 4 & -3 \\ 3 & 4 \end{pmatrix}$	$4-3i,\ 4+3i$	$\binom{-i}{1},\ \binom{i}{1}$

Die ersten 4 Punkte sind also "Sattelpunkte" d.h. es gibt zwei gegenüberliegende Richtungen (Eigenvektoren mit positiven Eigenwerten) wo die Lösungen herausfliesen und zwei gegenüberliegende Richtungen (Eigenvektoren mit negativen Eigenwerten) wo die Lösungen hineinfliesen. Der letzte Punkt ist eine Quelle aus der alle Lösungen herausspiralen.



5.4.5 Bemerkung. Differentialgleichungen höherer Ordnung.
Unter einer GEWÖHNLICHEN DIFFERENTIALGLEICHUNG n-TER ORDNUNG versteht man eine Gleichung der Form

$$x^{(n)}(t) = f(t, x(t), x'(t), \dots, x^{(n-1)}(t)),$$

wobei f eine lokal definierte Abbildung $\mathbb{R} \times E^n \to E$ ist und die Lösung x lokal von $\mathbb{R}$ nach E n-mal differenzierbar sein soll.

Betrachten wir als ganz einfaches Beispiel die Gleichung 2-ter Ordnung

$$x''(t) = g \text{ mit einer Konstanten } g.$$

Wenn wir $y := x'$ setzen, dann lautet die Gleichung $y'(t) = g$ und hat als Lösung $y(t) = g\,t + C_1$, und die Konstante C_1 kann aus der Anfangsbedingung $y_0 = y(t_0) = x'(t_0)$ als $C_1 := y(0) - g\,t_0$ berechnet werden. Schließlich hat die Differentialgleichung 1-ter Ordnung

$$x'(t) = y(t) = g\,(t - t_0) + y_0$$

als Lösung

$$x(t) = g\,\frac{(t - t_0)^2}{2} + y_0\,t + C_0$$

und die Konstante C_0 kann aus der Anfangsbedingung $x_0 = x(t_0)$ als $C_0 := x_0 - y_0\,t_0$ berechnet werden. Die allgemeine Lösung obiger Differentialgleichung 2-ter Ordnung ist somit durch

$$x(t) = \frac{g}{2}\,(t - t_0)^2 + x'(t_0)\,(t - t_0) + x(0)$$

gegeben. Diese Differentialgleichung tritt in der Physik an prominenter Stelle auch wirklich auf, wie wir im nachfolgenden Beispiel (5.4.6) sehen werden.

Ganz analog kann man auch bei allgemeinen Differentialgleichungen n-ter Ordnung vorgehen: Wenn man $y_0(t) := x(t)$, $y_1(t) := x'(t)$, …, $y_{n-1}(t) := x^{(n-1)}(t)$ für eine Lösung x setzt, dann ist $y : t \mapsto (y_0(t), \dots, y_{n-1}(t))$ eine Lösung der gewöhnlichen Differentialgleichung 1. Ordnung:

$$y'(t) = g(t, y(t)),$$

wobei $g : \mathbb{R} \times E^n \to E^n$ gegeben ist durch

$$g(t, y_0, \dots, y_{n-1}) := \Big(y_1, \dots, y_{n-1}, f(t, y_0, \dots, y_{n-1})\Big).$$

Umgekehrt liefert auch jede Lösung $y = (y_0, \dots, y_{n-1})$ von $y'(t) = g(t, y(t))$ eine Lösung $x := y_0$ von $x^{(n)}(t) = f(t, x(t), \dots, x^{(n-1)}(t))$.

Ist insbesonders $f(t, x_0, \ldots, x_{n-1}) := \sum_{k=0}^n a_k(t)\, x_k(t)$, so ist $g(t, y)$ linear in y und wird durch die Matrix

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & 1 \\ a_0 & a_1 & \ldots & a_{n-2} & a_{n-1} \end{pmatrix}$$

gegeben. Deren Eigenwerte λ sind gerade die Lösungen der Gleichung

$$\det \begin{pmatrix} -\lambda & 1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & -\lambda & 1 \\ a_0 & a_1 & \ldots & a_{n-2} & a_{n-1} - \lambda \end{pmatrix} = 0$$

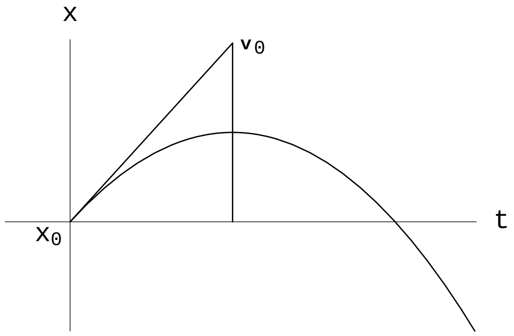
und somit (durch Entwicklung nach der letzten Zeile) von

$$\begin{aligned} 0 &= a_0 \cdot 1 - a_1 \cdot (-\lambda) + - \cdots + (-1)^{n-1}(a_{n-1} - \lambda)\, (-\lambda)^{n-1} \\ &= a_0 + a_1\, \lambda + \cdots + a_{n-1}\, \lambda^{n-1} - \lambda^n. \end{aligned}$$

5.4.6 Beispiel. Gravitationsfelder.
Bekanntlich bewegen sich Objekte auf die keine Kräfte einwirken gleichförmig, d.h. ihre Geschwindigkeit $x'(t)$ ist konstant, also $x'(t) = v_0$ und somit $x(t) = t\, v_0 + x_0$, wobei $x_0 = x(0)$ und $v_0 = x'(0)$ die Anfangsbedingung ist. Wirkt hingegen eine Kraft F auf ein Objekt der Masse m, so gilt das

Newton'sche Kraftgesetz: Kraft ist Masse mal Beschleunigung,

d.h. das Objekt erfährt eine zur Kraft proportionale Beschleunigung. Ist insbesonders die Kraft konstant (z.B. ist die Erdanziehungskraft $g \cdot m$ nahe der Erdoberfläche, wobei $g \approx 9.81 m/s^2$ die Erdbeschleunigung bezeichnet) so lautet das Newton'sche Kraftgesetz $g \cdot m = F = m \cdot x''(t)$, wobei $x''(t)$ die Beschleunigung des Objekts beschreibt, welches sich zum Zeitpunkt t and der Stelle $x(t)$ befindet. Die Flugbahn wird also durch die Differentialgleichung 2-ter Ordnung $x''(t) = g$ beschrieben und die Lösungen sind nach (5.4.5) Parabeln $x(t) = \frac{g}{2}\,(t - t_0)^2 + x'(t_0)\,(t - t_0) + x(0)$.



Entfernt sich das Objekt hingegen wesentlich von der Erdoberfläche, so nimmt die Anziehungskraft der Erde (mit dem Quadrat der Entfernung $x(t)$ zum Erdmittelpunkt) ab, ist also durch $F = -\frac{\gamma M m}{x(t)^2}$ gegeben, wobei γ die Gravitationskonstante und M die Masse der Erde bezeichnet. Die zugehörige Differentialgleichung ist somit

$$x''(t) = -\frac{\gamma M}{x(t)^2}.$$

Mittels folgendem Trick können wir Gleichungen der Form $x''(t) = f(x(t))$ lösen: Sei dazu F eine Stammfunktion von f und somit ist

$$(x' \cdot x')'(t) = 2x'(t)x''(t) = 2x'(t)F'(x(t)) = (2F \circ x)'(t), \text{ also } x'(t)^2 \qquad = 2F(x(t)) + C,$$

wobei C aus der Anfangsbedingung $x'(0)$ bestimmt werden kann. Weiters kann $x(t)$ durch Lösen der Differentialgleichung 1-ter Ordnung

$$x'(t) = \pm\sqrt{C + 2F(x(t))}$$

im Prinzip mittels Separation der Variablen bestimmt werden:

$$\int \frac{1}{\sqrt{C + 2F(x)}}\, dx = \pm \int dt = \pm t + C_1.$$

In unseren Fall ist $f(x) := -\frac{\gamma M}{x^2}$ und somit $F(x) = \frac{\gamma M}{x}$, also

$$x'(t)^2 = C + \frac{2\gamma M}{x} \text{ mit } C = v_0^2 - \frac{2\gamma M}{x_0},$$

wobei x_0 Anfangshöhe und v_0 Anfangsgeschwindigkeit ist.

Falls $C = 0$ ist, also $v_0 = \sqrt{\frac{2\gamma M}{x_0}}$ die Fluchtgeschwindigkeit ist, und somit x' nur für $x \to \infty$ verschwindet und folglich $x(t) \to \infty$ für $t \to \infty$ gilt, liefert Separation der Variablen:

$$\int \sqrt{x}\, dx = \pm \int \sqrt{2\gamma M}\, dt,$$

also $x^{3/2} = \pm\frac{3}{2}(\sqrt{2\gamma M}\, t + C_1)$ und somit

$$x(t) = \left(\pm 3\sqrt{\frac{\gamma M}{2}}\, t + C_2\right)^{2/3} = \left(\pm 3\sqrt{\frac{\gamma M}{2}}\, t + x_0^{3/2}\right)^{2/3}.$$

Andernfalls versuchen wir zuerst die Konstanten zu vereinfachen: Dazu substituieren wir $y(t) := \frac{1}{\lambda}x(t)$ mit $\lambda > 0$ und erhalten somit

$$y'(t)^2 = \frac{1}{\lambda^2}\, x'(t)^2 = \frac{1}{\lambda^2}\left(C + \frac{2\gamma M}{\lambda\, y(t)}\right) = \frac{C}{\lambda^2} + \frac{2\gamma M}{\lambda^3\, y(t)}$$

Wählen wir also $\lambda := \sqrt[3]{2\gamma M}$ so lautet unsere Differentialgleichung

$$y'(t)^2 = C_1 + \frac{1}{y(t)} \text{ mit } C_1 = \frac{C}{(2\gamma M)^{2/3}}$$

Mittels Separation der Variablen formen wir die Differentialgleichung um in

$$\int \sqrt{\frac{y}{1 + C_1\, y}}\, dy = \pm \int dt = \pm t + C_2.$$

Substituieren wir nun $z := \sqrt{\frac{y}{1+C_1\, y}} > 0$, also

$$y = \frac{z^2}{1 - C_1 z^2} \text{ und } dy = \frac{2z}{(1 - C_1 z^2)^2}\, dz,$$

so erhalten wir mittels partieller Integration

$$\begin{aligned}
\int \sqrt{\frac{y}{1 + C_1\, y}}\, dy &= \int \frac{2z^2}{(1 - C_1 z^2)^2}\, dz \\
&= \int \frac{z}{C_1}\, \frac{2C_1 z}{(1 - C_1 z^2)^2}\, dz \\
&= \frac{z}{C_1}\, \frac{1}{1 - C_1 z^2} - \frac{1}{C_1} \int \frac{1}{1 - C_1 z^2}\, dz
\end{aligned}$$

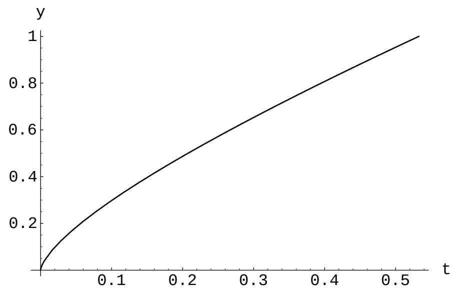
Und je nach Vorzeichen von C_1 ist

$$\begin{aligned}
\int \frac{1}{1-C_1 z^2}\,dz &= \frac{1}{2}\int \frac{1}{1-\sqrt{C_1}z} + \frac{1}{1+\sqrt{C_1}z}\,dz \\
&= \frac{1}{2\sqrt{C_1}}\Big(-\ln(1-\sqrt{C_1}z) + \ln(1+\sqrt{C_1}z)\Big) \\
&= \frac{1}{2\sqrt{C_1}}\Big(-\ln(1-\sqrt{C_1}z) + \ln(1+\sqrt{C_1}z)\Big) \\
&= \frac{1}{2\sqrt{C_1}}\ln\left(\frac{1+\sqrt{C_1}z}{1-\sqrt{C_1}z}\right) \qquad \text{für } C_1 > 0 \\
\text{und} \qquad \int \frac{1}{1-C_1 z^2}\,dz &= \frac{1}{\sqrt{-C_1}}\int \frac{\sqrt{-C_1}}{1+(\sqrt{-C_1}z)^2}\,dz \\
&= \frac{1}{\sqrt{-C_1}}\arctan(\sqrt{-C_1}z) \qquad \text{für } C_1 < 0.
\end{aligned}$$

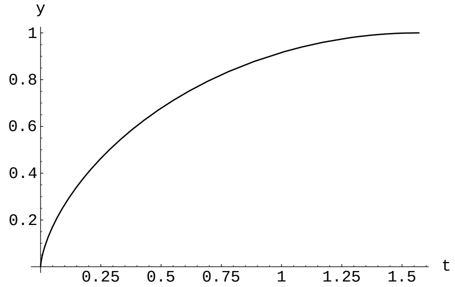
Rücksubstitution von $z = \sqrt{\frac{y}{1+C_1\,y}}$ ergibt somit

$$\begin{aligned}
\pm t + C_2 &= \frac{z}{C_1}\,\frac{1}{1-C_1 z^2} - \frac{1}{C_1}\int \frac{1}{1-C_1 z^2}\,dz \\
&= \frac{z}{C_1(1-C_1 z^2)} - \frac{1}{2C_1^{3/2}}\ln\left(\frac{1+\sqrt{C_1}z}{1-\sqrt{C_1}z}\right) \\
&= \frac{2\sqrt{C_1\,y\,(1+C_1\,y)} - \ln(1+2\,C_1\,y+2\sqrt{C_1\,y\,(1+C_1\,y)})}{2\,C_1^{3/2}} \\
\text{bzw.} \qquad &= \frac{z}{C_1(1-C_1 z^2)} - \frac{1}{C_1\sqrt{-C_1}}\arctan(\sqrt{-C_1}z) \\
&= \frac{\sqrt{-C_1\,y\,(1+C_1\,y)} - \arctan(\sqrt{\frac{-C_1\,y}{1+C_1\,y}})}{C_1\sqrt{-C_1}} \\
&= \frac{\sqrt{-C_1\,y\,(1+C_1\,y)} - \arcsin(\sqrt{-C_1\,y})}{C_1\sqrt{-C_1}}
\end{aligned}$$

Es gelingt uns allerdings nicht mehr die Umkehrfunktion diese Ausdruckes für $\pm t + C_2$ explizit zu berechnen, aber zumindest könne wir damit exakt berechnen zu welchen Zeitpunkt t eine vorgegebene Höhe x erreicht wird (oder auch nicht).
Für $C_1 = 1$ erhalten wir



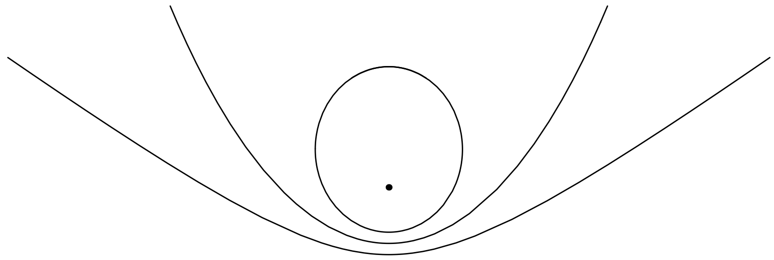
für $C_1 = -1$ hingegen



Wenn wir nun mit $x(t) \in \mathbb{R}^3$ die Koordinaten eines Objektes im Raums bezeichnen, wobei wir den Ursprung in den Sonnenmittelpunkt gelegt haben, so gilt für die Kraft welche die Sonne auf das Objekt der Masse m ausübt entsprechend $F = -\frac{\gamma M m}{\|x\|^2} \frac{x}{\|x\|}$ und die Differentialgleichung welche die Bahn des Objekts beschreibt ist durch

$$x''(t) = -\frac{\gamma M m}{\|x\|^3}\, x$$

gegeben. Wenn wir dies numerisch lösen, dann sehen wir, daß je nach Anfangsbedingungen (Anfangsort $x(0)$ und Anfangsgeschwindigkeit $x'(0)$) die Bahnen durch Ellipsen (Kreise), Parabeln und Hyperbeln beschrieben werden.



Beachte, daß die entsprechende Differentialgleichung bereits 6 Variablen aufweist. Stellt man für mehrere Körper (z.B. die 9 Planeten unseres Sonnensystems) ein ganz analog gebildetes Gleichungssystem auf, so involviert dieses bereits $9 \cdot 6 = 54$ Variablen. Wir sehen also, daß es keine akademische Spielerei ist Vektorräume der Dimension größer als 3 zu behandeln.

6.2.11. Newton-Verfahren

Um eine Nullstelle der (nicht-linearen) Gleichung $f(x) = 0$ zu erhalten beginnen wir mit irgend einem Näherungswert x_0, betrachten die affine Funktion $x \mapsto f'(x_0)(x - x_0) + f(x_0)$ die f nahe x_0 am besten approximiert, bestimmen ihre Nullstelle x_1, d.h. lösen $f'(x_0)(x_1 - x_0) + f(x_0) = 0$ und erhalten als neue Näherungsnullstelle von f den Punkt $x_1 := x_0 - f'(x_0)^{-1} f(x_0)$. Das Newton-Verfahren besteht nun darin auf rekursive Weise $x_{n+1} := g(x_n) := x_n - f'(x_n)^{-1} f(x_n)$ zu definieren und zu hoffen, daß $x_\infty := \lim_{n\to\infty} x_n$ existiert, weil dann $x_\infty = g(x_\infty)$, also $f(x_\infty) = 0$ ist (Verwende dazu die stetige Differenzierbarkeit von f und damit die Stetigkeit von g). Da aber die Bestimmung der Inversen $f'(x_0)^{-1}$ einer Matrix $f'(x_0)$ sehr aufwendig ist, modifizieren wir dieses Verfahren derart, daß wir Anstelle von $f'(x_n)$ immer das fixe $f'(x_0)$ verwenden. Wir betrachten also $g(x) := x - f'(x_0)^{-1} \cdot f(x)$. Damit g eine Kontraktion wird benötigen wir ein $q < 1$ mit

$$q \geq \|g'(z)\| = \|\,\mathrm{id} - f'(x_0)^{-1} \circ f'(z)\| = \|f'(x_0)^{-1} \circ (f'(x_0) - f'(z))\|.$$

Für z hinreichend nahe an x_0 (sagen wir $\|z - x_0\| \leq r$) ist dies erreichbar. Für $x, y \in B$ ist dann nach dem Mittelwertsatz $\|g(x) - g(y)\| \leq \sup\{\|g'(\xi)\|\,\|x - y\| : \xi \in B\} \leq q\,\|x - y\|$.

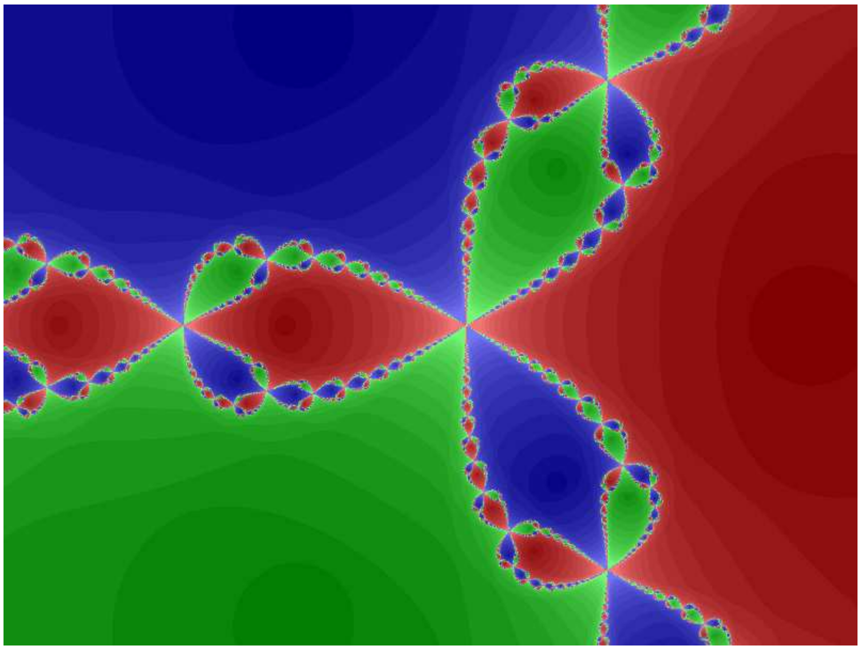
Wir müssen aber auch zeigen, daß g diese Menge $B := \{z : \|z - x_0\| \leq r\}$ invariant läßt: Sei also $z \in B$, dann ist

$$\|g(z) - x_0\| \leq \|g(z) - g(x_0)\| + \|g(x_0) - x_0\| \leq q\,\|z - x_0\| + \|g(x_0) - x_0\| \leq q\,r + \|g(x_0) - x_0\| \leq r,$$

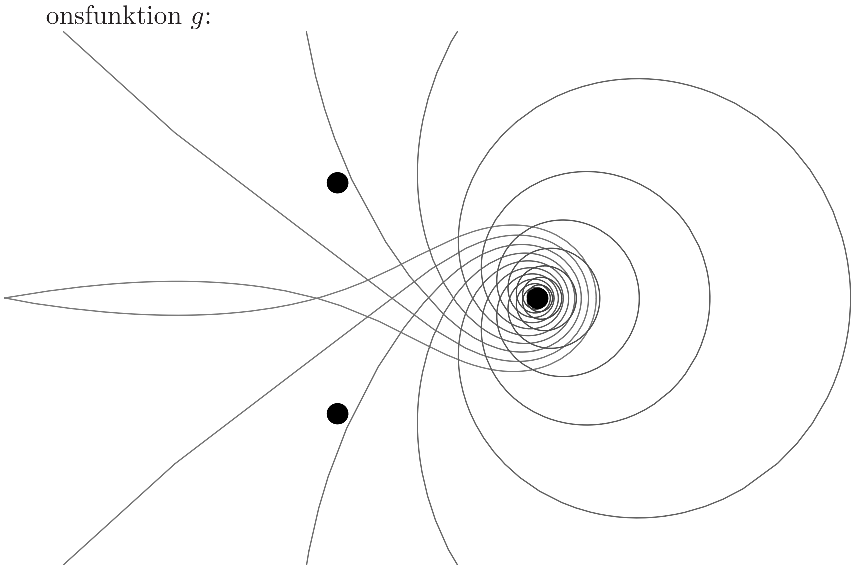
falls $\|g(x_0) - x_0\| \leq r(1 - q)$ ist. Anwendung des Banach'schen Fixpunktsatzes (3.4.17) liefert somit $\exists\, x_\infty := \lim_{n\to\infty} g^n(x_0)$ und x_∞ ist der eindeutige Fixpunkt von g, also $x_\infty = g(x_\infty) = x_\infty - f'(x_0)^{-1} f(x_\infty)$, d.h. x_∞ ist die eindeutige Nullstelle von f:

Vereinfachtes Newton-Verfahren.
Es sei $f : E \to F$ stetig differenzierbar und $x_0 \in E$. Es sei $f'(x_0)$ invertierbar und es existiere ein $r > 0$ und ein $q < 1$ mit $\|f'(x_0)^{-1} \cdot f(x_0)\| \le (1-q)r$ und $\|f'(x_0)^{-1}\| \cdot \|f'(x_0) - f'(x)\| \le q$ für alle x mit $\|x - x_0\| \le r$. Dann konvergiert die vereinfachte Newton-Folge $x_{n+1} := x_n - f'(x_0)^{-1} \cdot f(x_n)$ gegen eine Nullstelle $\xi \in B := \{z : \|z - x_0\| \le r\}$. □

6.2.12 Beispiel. Julia-Menge.
Die Menge der Startwerte, für welche die Newton-Folge gegen eine fixe Nullstelle konvergiert, ist im allgemeinen schwer beschreibbar. Betrachten wir z.B. das Polynom $p(z) := z^3 - 1$ als Abbildung $p : \mathbb{C} \to \mathbb{C}$. Dieses hat 3 komplexe Nullstellen $1, -\frac{1}{2} \pm i\,\frac{\sqrt{3}}{2}$. Wir können es auch als Abbildung $p : \mathbb{R}^2 \to \mathbb{R}^2$ auffassen. Die Ableitung von p ist gegeben durch (Multiplikation mit) $3z^2$ und die Newton-Rekursion somit durch $g : z \mapsto z - \frac{p(z)}{p'(z)} = z - \frac{z^3-1}{3z^2} = \frac{2z^3+1}{3z^2}$. Färben wir die Punkte der Ebene in 3 Farben entsprechend den 3 Nullstellen gegen welche die Newton-Folge mit diesen Startwerten konvergiert so erhalten wir folgendes fraktales Bild, eine Julia-Menge:



Die folgende Zeichnung zeigt die Bilder konzentrischer Kreise um den Fixpunkt 1 unter der Rekursionsfunktion g:



Wir können hier die Konvergenz für Startwerte in der Nähe der Nullstellen direkt verifizieren. Betrachten wir dazu o.B.d.A. (wegen Rotationssymmetrie) die Nullstelle 1. Es ist $g(z)-1=(z-1)^2\frac{1+2z}{3z^2}$. Wegen $\lim_{z\to 1}\frac{1+2z}{3z^2}=1$ können wir ein $0<\delta<1$ finden, sodaß $|\frac{1+2z}{3z^2}|<\frac{1}{\delta}$ für alle $|z-1|\le\delta$ (z.B. $\delta=\frac{2}{5}$, denn $|\frac{1+2z}{3z^2}|\le\frac{1+2|z|}{3|z|^2}\le\frac{1+2(1-\delta)}{3(1-\delta)^2}\le\frac{1}{\delta}$). Somit ist $q:=\delta\,\max\{|\frac{1+2z}{3z^2}|:|z-1|\le\delta\}<1$) und $|g(z)-1|<\delta^2\frac{1}{\delta}=\delta$ sowie $|g(z)-1|=|z-1|^2\,\left|\frac{1+2z}{3z^2}\right|\le|z-1|\,q$, also konvergiert $z_n\to 1$ für $|z_0-1|\le\delta$.

Es ist $g'(z)=\frac{2}{3}\left(1-\frac{1}{z^3}\right)$. Für $z_0^3=1$ ist $\lim_{z\to z_0}\frac{1}{z^3}=1$ und somit existiert ein $\delta>0$, s.d. $|g'(z)|\le\frac{2}{3}$ für alle z mit $|z-z_0|\le\delta$. Also ist g auf diesen Scheiben B eine Kontraktion und wegen $|g(z)-z_0|=|g(z)-g(z_0)|\le\frac{2}{3}|z-z_0|$ ist z_0 der eindeutige Fixpunkt in B.

6.3 Höhere Ableitungen

6.3.1 Definition. Zweite Ableitung.
Wir wollen nun die 2. Ableitung einer Funktion $f : E \supseteq U \to F$ definieren. Für $x \in U$ ist die Ableitung von f bei x eine lineare Abbildung $f'(x) : E \to F$ (also $f'(x) \in L(E,F)$) und bei variierendem $x \in U$ ist für differenzierbares f die Ableitung eine Abbildung $f' : E \supseteq U \to L(E,F)$. Falls dies im Punkte $x \in U$ seinerseits differenzierbar ist, also $(f')'(x) : E \to L(E,F)$ als lineare Abbildung existiert, so können wir $(f')'(x)$ auch als bilineare Abbildung $E \times E \to F$ auffassen (siehe (5.5.13)) und mit $f''(x) \in L(E,E;F)$ bezeichnen. Falls $f' : E \supseteq U \to L(E,F)$ auf ganz U differenzierbar ist, so nennen wir f ZWEIMAL DIFFERENZIERBAR. Die Ableitung $(f')'$ von f' ist dann eine Abbildung $(f')' : E \supseteq U \to L(E, L(E,F))$ und vermöge der Identifikation von $L(E_1, L(E_2,F))$ mit $L(E_1,E_2;F)$ aus (5.5.13) erhalten wir eine Abbildung $f'' : E \supseteq U \to L(E,E;F)$ die sogenannte ZWEITE ABLEITUNG von f gegeben durch

$$f''(x)(v,w) := (f')'(x)(v)(w) \text{ für } x \in U \text{ und } v,w \in E.$$

Versuchen wir nun die zweite Ableitung – so sie existiert – durch Richtungsableitungen auszudrücken:

$$\begin{aligned}
f''(x)(v,w) &= (f')'(x)(v)(w) = \mathrm{ev}_w(d_v(f')(x)) = \mathrm{ev}_w\left(\lim_{t\to 0+}\frac{f'(x+tv)-f'(x)}{t}\right)\\
&= \lim_{t\to 0+}\mathrm{ev}_w\left(\frac{f'(x+tv)-f'(x)}{t}\right) = \lim_{t\to 0+}\frac{f'(x+tv)(w)-f'(x)(w)}{t}\\
&= \lim_{t\to 0+}\frac{d_wf(x+tv)-d_wf(x)}{t} = d_v(d_wf)(x) =: d_vd_w\, f\,(x),
\end{aligned}$$

also kann $f''(x)(v,w)$ dadurch berechnet werden, daß wir hintereinander f zuerst in Richtung w und das Ergebnis dann in Richtung v an der Stelle x differenzieren.

Da $f''(x)$ bilinear ist können wir die zweite Ableitung für Abbildungen $f : \mathbb{R}^n \supseteq U \to F$ wie folgt ausdrücken:

$$\begin{aligned}
f''(x)(v,w) &= f''(x)\Big(\sum_{i=1}^n v_i\, e_i, \sum_{j=1}^n w_j\, e_j\Big) = \sum_{i=1}^n v_i\, f''(x)\Big(e_i, \sum_{j=1}^n w_j\, e_j\Big)\\
&= \sum_{i=1}^n\sum_{j=1}^n v_i\, w_j\, f''(x)(e_i,e_j) = \sum_{i=1}^n\sum_{j=1}^n v_i\, w_j\, \underbrace{d_{e_i}\, d_{e_j}\, f(x)}_{\partial_i\partial_j f(x)}.
\end{aligned}$$

6.3.9 Beispiel. Zweite Ableitung.
Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(x,y) := x^2 \cdot \sin(y)$. Dann ist

$$\begin{aligned}
\partial_1 f(x,y) &= 2\,x \cdot \sin(y)\\
\partial_1\partial_1 f(x,y) &= 2 \cdot \sin(y)\\
\partial_2\partial_1 f(x,y) &= 2\,x \cdot \cos(y)\\
\partial_2 f(x,y) &= x^2 \cdot \cos(y)\\
\partial_1\partial_2 f(x,y) &= 2\,x \cdot \cos(y)\\
\partial_2\partial_2 f(x,y) &= -x^2 \cdot \sin(y)
\end{aligned}$$

Die Ableitung $f'(z)$ mit $z := (x,y)$ wirkt durch

$$f'(z)(u_1,u_2) = \partial_1 f(z)\, u_1 + \partial_2 f(z)\, u_2 = \Big(\partial_1 f(z),\, \partial_2 f(z)\Big) \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$$

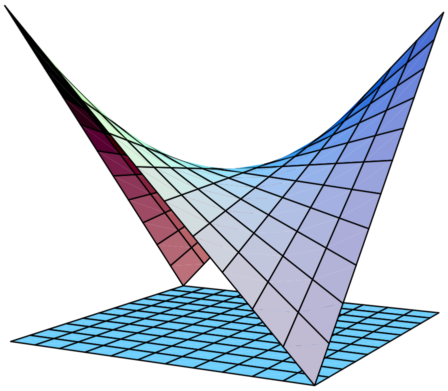
Für die zweite Ableitung $f''(z)$ erhalten wir

$$\begin{aligned} f''(z)(u_1,u_2;v_1,v_2) &= (f')'(z)(u_1,u_2)(v_1,v_2) = (\partial_1 f'(z)\,u_1 + \partial_2 f'(z)\,u_2)(v_1,v_2) \\ &= \partial_1\partial_1 f(z)\,u_1\,v_1 + \partial_1\partial_2 f(z)\,u_1\,v_2 + \partial_2\partial_1 f(z)\,u_2\,v_1 + \partial_2\partial_2 f(z)\,u_2\,v_2 \\ &= (u_1,u_2)\cdot\begin{pmatrix}\partial_1\partial_1 f(z) & \partial_1\partial_2 f(z)\\ \partial_2\partial_1 f(z) & \partial_2\partial_2 f(z)\end{pmatrix}\cdot\begin{pmatrix}v_1\\ v_2\end{pmatrix}\end{aligned}$$

Die Matrixdarstellung der bilinear-Form $f''(z)$ ist somit die sogenannte Hesse'sche Matrix

$$[f''(z)] = \begin{pmatrix}\partial_1\partial_1 f(z) & \partial_1\partial_2 f(z)\\ \partial_2\partial_1 f(z) & \partial_2\partial_2 f(z)\end{pmatrix}$$

Wer sich wundert, daß man für die zweite Ableitung auch gemischte partielle Ableitungen wie $\partial_1\partial_2 f$ berücksichtigen muß, betrachte ein hyperbolisches Paraboloid, welches in Koordinaten z.B. durch $f(x,y) := x\cdot y$ gegeben ist,



Bei festem x beschreibt f eine Gerade $y \mapsto x\,y$ mit Anstieg x und analog für y. Die iterierten Ableitungen $\partial_1\partial_1 f$ und $\partial_2\partial_2 f$ sind folglich überall 0. Wäre hingegen die zweite Ableitung $f''=0$, dann wäre f' nach (5.6.16) konstant und somit nach (6.1.23a)

$$f(z) = f(0) + \int_0^1 \frac{d}{dt} f(tz)\,dt = f(0) + \int_0^1 f'(tz)(z)\,dt = 0 + \int_0^1 f'(0)(z)\,dt = f'(0)(z),$$

also f linear, was offensichtlich nicht der Fall ist da f 2-homogen ist (d.h. $f(t\,z) = t^2\,f(z)$). Wenn man f längs Geraden $t \mapsto z + t\,w$ mit Richtungsvektoren $w \notin \{e_1,e_2\}$ betrachtet, so ist

$$f(z) = f(z_1 + t\,w_1, z_2 + t\,w_2) = (z_1 + t\,w_1)\cdot(z_2 + t\,w_2) = z_1\,z_2 + (z_1\,w_2 + z_2\,w_1)\,t + w_1\,w_2\,t^2$$

ein Polynom 2-ten Grades, also eine Parabel, und die Richtungsableitung

$$\begin{aligned}\partial_w\partial_w f(z) &= \frac{d}{dt}\Big|_{t=0} \partial_w f(z+t\,w) = \frac{d}{dt}\Big|_{t=0}\frac{d}{ds}\Big|_{s=0} f(z+t\,w+s\,w) \\ &= \frac{d}{dt}\Big|_{t=0}\frac{d}{ds}\Big|_{s=t} f(z+t\,w) = \left(\frac{d}{dt}\right)^2\Big|_{t=0} f(z+s\,w) = 2\,w_1\,w_2 \neq 0.\end{aligned}$$

Siehe dazu auch (6.3.9a).

6.3.1a Definition. Höhere Ableitungen.
Rekursiv nennen wir f $n+1$-mal (stetig) differenzierbar, falls f' n-mal (stetig) differenzierbar ist. Wir schreiben dafür $f \in D^{n+1}$ ($f \in C^{n+1}$). Schließlich heißt f glatt (kurz $f \in C^\infty$), wenn f beliebig oft differenzierbar ist, d.h. D^n oder auch C^n für alle $n \in \mathbb{N}$ ist.

Unter der $(n+1)$-ten Ableitung einer $(n+1)$-mal differenzierbaren Abbildung $f : E \supseteq U \to F$ versteht man

$$f^{(n+1)} : E \supseteq U \xrightarrow{(f')^{(n)}} L(E,\dots,E;L(E,F)) \cong L(E,\dots,E,E;F).$$

Beachte die Schreibweise $f^{(n)}$ für höhere Ableitungen, denn $f', f'', f''', \ldots$ läßt sich schlecht allgemein aufschreiben, darum setzen wir $f^{(0)} := f$, $f^{(1)} := f'$, $f^{(2)} := f''$,....

Wegen $f'(x)(v) = d_v f(x)$ können wir $f^{(n)}(x)(v_1, \ldots, v_n)$ wie folgt durch iterierte Richtungsableitungen berechnen:

$$f^{(n)}(x)(v_1, \ldots, v_n) = d_{v_1}(d_{v_2}(\ldots(d_{v_n} f)\ldots))(x) =: d_{v_1}\, d_{v_2}\, \ldots\, d_{v_n}\, f\,(x).$$

Beachte den Vorteil dieser rekursiven Definition im Gegensatz zu der auf den ersten Blick vielleicht einleuchtenderen Definition "$f \in D^{n+1} \Leftrightarrow f \in D^n$ und $f^{(n)} \in D$" für welche wir gleichzeitig $f^{(n)}$ rekursiv mitdefinieren müssen.

6.3.3 Lemma. Grundlegendes für höhere Differenzierbarkeit.

(2) Die Kettenregel (6.1.5) ist gültig für C^p und D^p für $p \in \{1, 2, \ldots, \infty\}$.

(1) Jede stetige multilineare Abbildung ist glatt.

(3) Eine Abbildung $f = (f^1, \ldots, f^n)$ in ein Produkt $F = F_1 \times \ldots \times F_n$ ist genau dann C^p bzw. D^p wenn es jede Komponente f^i für $1 \le i \le n$ ist.

Beweis. (3) Es gilt: $f \in D^{p+1} :\Longleftrightarrow f = (f^1, \ldots, f^n) \in D$ und $(g_1, \ldots, g_n) = g := f' = ((f^1)', \ldots, (f^n)') \in D^p \overset{\text{(6.1.7) und I.A.}}{\Longleftrightarrow} \forall i$: $f^i \in D$ und $g_i = (f^i)' \in D^p \Longleftrightarrow: \forall i$: $f^i \in D^{p+1}$.

Wegen $f^{(p)}(x) = ((f^1)^{(p)}(x), \ldots, (f^n)^{(p)}(x))$ gilt: $f \in C^p \Leftrightarrow \forall i : f^i \in C^p$.

(1) Dies gilt offensichtlich für affine und für bilineare Abbildungen, denn die Ableitung einer bilineare Abbildung ist nach (6.1.4) und Aufgabe (6.3) linear und, die einer linearen Abbildung ist nach (6.1.3.1) konstant, und somit alle höheren Ableitungen gleich 0. Da jedoch die Ableitung eine 3-linearen Abbildung nicht bilinear ist, können wir in der allgemeinen Situation so nicht vorgehen, sondern zeigen zuerst die Kettenregel:

(2) Wir beginnen mit dem Fall, wo die äußere Funktion a linear und stetig ist und die innere Funktion $f \in D^p$ ist.
Für $p = 1$ wissen wir, daß $(a \circ f)' = a_* \circ f'$ und da $a_* : L(E, F_1) \to L(E, F_2)$ stetig (mit Norm $\|a_*\| = \sup\{\|a_*(T)\| : \|T\| \le 1\} = \sup\{\|a \circ T\| : \|T\| \le 1\} \le \|a\|$) und linear ist, können wir Induktion anwenden um $a \circ f \in D^p$ für $f \in D^p$ zu erhalten.

Der allgemeine Fall einer nicht-linearen äußeren Funktion f folgt mittels Induktion nach der Differentiationsordnung p. Für $(p = 1)$ ist nach (6.1.5) auch $f \circ g \in D$ und $(f \circ g)'(x) = f'(g(x)) \circ g'(x)$, also $(f \circ g)' = \operatorname{komp} \circ (f' \circ g, g')$.
Nun zum Induktionsschritt auf $(p + 1)$: Sei dazu $f, g \in D^{p+1}$, also $f, g \in D$ und $f', g' \in D^p$. Insbesonders ist $g \in D^p$ (siehe (6.3.4)) und somit nach Induktionsannahme $f' \circ g \in D^p$. Nach (3) ist damit $(f' \circ g, g') \in D^p$ und da komp bilinear und stetig ist nach (5.5.13), folgt komp $\in D^p$ wegen (1) und somit ist $(f \circ g)' = \operatorname{komp} \circ (f' \circ g, g') \in D^p$ nach Induktionsannahme, also $f \circ g \in D^{p+1}$.

Für C^p folgt (2) nun ebenfalls durch Induktion, denn komp $\in C^\infty$.

(1) zeigen wir nun für n-lineare stetige Abbildungen $T : E_1 \times \ldots \times E_n \to F$ ebenfalls mittels Induktion nach der Differentiationsordnung p.
$(p = 0)$ ist trivial.
$(p = 1)$ haben wir (für $n = 2$) bereits in (6.1.4) gezeigt, wobei

$$\begin{aligned} T'(x_1, \ldots, x_n)(v_1, \ldots, v_n) &= \sum_i \partial_i T(x_1, \ldots, x_n)(v_i) \\ &= T(v_1, x_2, \ldots, x_n) + \cdots + T(x_1, \ldots, x_{n-1}, v_n). \end{aligned}$$

$(p+1)$ Der j-te Summand obiger Formel für T' ist durch folgende Zusammensetzung gegeben:

$$\begin{tikzcd} E_1 \times \ldots \times E_n \arrow[r] \arrow[d, "\mathrm{pr}_{1,\ldots,\widehat{i},\ldots,n}"] & L(E_1 \times \ldots \times E_n, F) \\ E_1 \times \ldots \times \widehat{E_i} \times \ldots \times E_n \arrow[r, "T^\vee"'] & L(E_i, F) \arrow[u, "\mathrm{pr}_i^*"'] \end{tikzcd}$$

Dabei bedeutet $\widehat{E_i}$, daß dieser Faktor fehlt. Der untere Pfeil ist $n-1$-linear und stetig und die vertikalen Pfeilen sind lineare und stetig, daher sind alle D^p nach Induktionsannahme für (1). Nach der Induktionsannahme für (2) ist die Zusammensetzung D^p. Also ist auch $T' \in D^p$ (da Summieren stetig und linear ist) und daher ist $T \in D^{p+1}$. □

Aus dem Spezialfall der Kettenregel (2) folgt, daß die Differenzierbarkeitsklasse zweier Abbildungen gleich ist, falls diese bis auf einen linearen Isomorphismus α der Bildräume übereinstimmen, und wir wollen die Bezeichnung $f \simeq g$ benützen, wenn $f = \alpha \circ g$ ist für irgendeinen natürlichen linearen Isomorphismus zwischen den Bild-Räumen.

6.3.4 Lemma. Differenzierbarkeit in Schritten.
Es sei $0 \le p < \infty$ und $0 \le q \le \infty$ und $f : E \supseteq U \to F$. Dann gilt:

$$f \text{ ist } D^{p+q} \Leftrightarrow f \text{ ist } D^p \text{ und } f^{(p)} \text{ ist } D^q.$$
$$f \text{ ist } C^{p+q} \Leftrightarrow f \text{ ist } C^p \text{ und } f^{(p)} \text{ ist } C^q.$$

Beweis. Wir zeigen dies durch Induktion nach p. Für $p = 0$ ist nichts zu zeigen.
$(p+1)$ $(\Rightarrow)$

$$\begin{aligned} & f \in D^{p+q+1} \\ \Rightarrow & f \in D,\ g := f' \in D^{p+q} && \text{nach Definition} \\ \Rightarrow & f \in D,\ f' = g \in D^p,\ g^{(p)} \in D^q && \text{nach Induktions-Voraussetzung} \\ \Rightarrow & f \in D^{p+1},\ f^{(p+1)} \simeq (f')^{(p)} \in D^q && \text{nach Definition.} \end{aligned}$$

$(\Leftarrow)$

$$\begin{aligned} & f \in D^{p+1},\ (f')^{(p)} \simeq f^{(p+1)} \in D^q \\ \Rightarrow & f \in D,\ g := f' \in D^p,\ g^{(p)} \in D^q && \text{nach Definition} \\ \Rightarrow & f \in D,\ f' = g \in D^{p+q} && \text{nach Induktions-Voraussetzung} \\ \Rightarrow & f \in D^{1+p+q} && \text{nach Definition.} \end{aligned}$$

Für stetige Differenzierbarkeit haben wir nur zu zeigen, daß $f \in C^{p+q} \Rightarrow f \in C^p$. O.B.d.A. sei $q > 0$. Dann ist $f \in D^{p+1}$ wegen $(\Rightarrow)$ und daher $f \in D^p$ und $f^{(p)} \in D$. Somit ist $f^{(p)} \in C$, d.h. $f \in C^p$. □

6.3.5. Höhere Differenzierbarkeit via partielle Ableitungen
Wir wollen als nächstes (6.1.24) auf höhere Differenzierbarkeit verallgemeinern, d.h. aus der Existenz der iterierten partiellen Ableitungen $\partial_{i_1} \ldots \partial_{i_p} f := \partial_{i_1}(\ldots(\partial_{i_p} f))$ und einer geeigneten Stetigkeitsbedingung auf $f \in C^p$ schließen.

Proposition.
Es ist $f : E = \mathbb{R}^n \supseteq U \to F$ genau dann C^p, wenn alle ITERIERTEN PARTIELLEN ABLEITUNGEN *$\partial_{i_1} \ldots \partial_{i_p} f : U \to F$ mit $i_1, \ldots, i_p \in \{1, \ldots, n\}$ existieren und stetig sind. Weiters haben wir dann*

$$f^{(p)}(x)(v_1, \ldots, v_p) = \sum_{i_1=1}^{n} \cdots \sum_{i_p=1}^{n} \partial_{i_1} \cdots \partial_{i_p} f(x) \cdot v_1^{i_1} \cdot \ldots \cdot v_p^{i_p},$$

wobei v_i^j die j-te Koordinate von v_i bezeichnet.

Beweis. Wir zeigen dies mittels Induktion und (6.1.24): $f \in C^{p+1} \xLeftrightarrow{(6.3.4)} f \in C^1$ und $f' \in C^p \xLeftrightarrow{(6.1.24)} \forall i$: $\exists \partial_i f \in C$ und $f' : x \mapsto (\partial_1 f(x), \dots, \partial_n f(x)) \in F \times \dots \times F \cong L(\mathbb{R}^n, F)$ ist $C^p \xLeftrightarrow{(6.3.3.3)} \forall i :\ \exists \partial_i f \in C^p \xLeftrightarrow{\text{I.A.}} \forall i_0, \dots, i_p\ \ \exists \partial_{i_0} \dots \partial_{i_p} f \in C$, und wegen der Multilinearität von $f^{(p)}(x)$ ist

$$\begin{aligned} f^{(p)}(x)(v_1, \dots, v_p) &= f^{(p)}(x)\Big(\sum_{i_1} v_1^{i_1} e_{i_1}, \dots, \sum_{i_p} v_p^{i_p} e_{i_p}\Big) \\ &= \sum_{i_1} \cdots \sum_{i_p} v_1^{i_1} \dots v_p^{i_p} \cdot f^{(p)}(x)(e_{i_1}, \dots, e_{i_p}) \\ &= \sum_{i_1, \dots, i_p} v_1^{i_1} \dots v_p^{i_p} \cdot \partial_{i_1} \dots \partial_{i_p} f(x). \qquad \square \end{aligned}$$

6.3.6 Differenzierbarkeit in Räume multilinearer Abbildungen via Richtungsableitungen. Man kann auch (6.1.23) auf höhere Differenzierbarkeit verallgemeinern, d.h. aus der Existenz der iterierten Richtungsableitungen $\partial_{v^1} \dots \partial_{v^p} f$ und einer geeigneten Stetigkeitsbedingung auf $f \in C^p$ schließen, siehe (6.3.8). Bereits im Fall $p = 2$ müssen wir dabei zumindest die Existenz der Richtungsableitungen von $f' : E \supseteq U \to L(E, F)$ zeigen, haben aber nur die Richtungsableitungen $\partial_v(\mathrm{ev}_w \circ f') = \partial_v \partial_w f$ zur Verfügung. Wir brauchen also ein Resultat der Form (6.3.7). Wir zeigen statt dessen aber gleich die folgende Kombination von (6.3.7) und (6.3.8):

Theorem.
Eine Abbildung $g : E \supseteq U \to L(E_1, \dots, E_q; F)$ ist genau dann C^p, wenn $\forall x \in U$, $\forall v^1, \dots, v^p \in E$ $\forall w^1 \in E_1, \dots, w^q \in E_q$, $\exists \partial_{v^1} \dots \partial_{v^p}(\mathrm{ev}_{w^1, \dots, w^q} \circ g)(x) \in F$ und

$$\begin{aligned} &x \mapsto \Big((v^1, \dots, v^p; w^1, \dots, w^q) \mapsto \partial_{v^1} \dots \partial_{v^p}(\mathrm{ev}_{w^1, \dots, w^q} \circ g)(x)\Big) \\ &U \to L(E, \dots, E, E_1, \dots, E_q; F) \end{aligned}$$

ist wohldefiniert und stetig. Weiters ist dann

$$\partial_{v^1} \dots \partial_{v^p}(\mathrm{ev}_{w^1, \dots, w^q} \circ g)(x) = g^{(p)}(x)(v^1, \dots, v^p)(w^1, \dots, w^q)$$

Beweis. Wir machen Induktion nach p. Für $(p = 0)$ ist nichts zu zeigen.

$(p = 1)$ $(\Downarrow)$ Da $\mathrm{ev}_{w^1, \dots, w^q} : L(E_1, \dots, E_q; F) \to F \in L$ ist, ist $\mathrm{ev}_{w^1, \dots, w^q} \circ g$ C^p und $(\mathrm{ev}_{w^1, \dots, w^q} \circ g)'(x)(v) = g'(x)(v)(w^1, \dots, w^q)$. Nach (6.1.23) existiert weiters $\partial_{v^1}(\mathrm{ev}_{w^1, \dots, w^q} \circ g)(x) = (\mathrm{ev}_{w^1, \dots, w^q} \circ g)'(x)(v^1)$.

$(\Uparrow)$ Nach (6.1.23) genug zu zeigen, daß $\partial_v g$ existiert und $x \mapsto (v \mapsto \partial_v g(x))$ wohldefiniert und stetig $U \to L(E, E_1, \dots, E_q; F)$ ist. Beachte, daß

$$\begin{aligned} \frac{g(x + tv) - g(x)}{t}(w^1, \dots, w^q) &= \frac{\mathrm{ev}_{w^1, \dots, w^p}(g(x + tv)) - \mathrm{ev}_{w^1, \dots, w^p}(g(x))}{t} \\ &= \int_0^1 \partial_v \Big(\mathrm{ev}_{w^1, \dots, w^p} \circ g\Big)(x + t\tau v)\, d\tau \\ &\to \partial_v \Big(\mathrm{ev}_{w^1, \dots, w^p} \circ g\Big)(x) \\ &\qquad \text{für } t \searrow 0 \text{ glm. in } \|w^1\| \le 1, \dots, \|w^q\| \le 1, \end{aligned}$$

da

$$x \mapsto ((v, w^1, \dots, w^q) \mapsto \partial_v(\mathrm{ev}_{w^1, \dots, w^q} \circ g)(x)$$

nach Voraussetzung stetig ist. Daher existiert $\partial_v g(x)$ und

$$\partial_v g(x)(w^1, \dots, w^q) := \mathrm{ev}_{w^1, \dots, w^q}(\partial_v g(x)) = \partial_v \Big(\mathrm{ev}_{w^1, \dots, w^p} \circ g\Big)(x),$$

ist nach Voraussetzung multilinear und beschränkt in v, w^1,..., w^q und stetig in x.

$(p+1)$ $(\Downarrow)$

$$
\begin{aligned}
g \in C^{p+1} \Rightarrow & g \in D,\ g' \simeq f : U \to L(E, F_1, \dots, F_q; F) \in C^p \qquad \text{nach Definition}\\
& \mathrm{ev}_{v^{p+1},w^1,\dots,w^q} \circ f := \mathrm{ev}_{w^1,\dots,w^q} \circ \mathrm{ev}_{v^{p+1}} \circ g'\\
\Rightarrow & \forall v^{p+1} \in E, w^1 \in E_1, \dots, w^q \in E_q\ \forall v^1, \dots, v^p \in E\\
& \exists\, \partial_{v^1} \dots \partial_{v^p} (\mathrm{ev}_{v^{p+1},w^1,\dots,w^q} \circ f)(x) \overset{\mathrm{IA}}{=}\\
& = f^{(p)}(x)(v^1, \dots, v^p)(v^{p+1}, w^1, \dots, w^q)\\
& = (\mathrm{ev}_{v^{p+1},w^1,\dots,w^q} \circ f)^{(p)}(x)(v^1, \dots, v^p) \quad \text{Def. von } f\\
& = (\mathrm{ev}_{w^1,\dots,w^q} \circ \mathrm{ev}_{v^{p+1}} \circ g')^{(p)}(x)(v^1, \dots, v^p) \quad \mathrm{ev} \in L,\ \mathrm{ev} \circ g' \in C^p\\
& = \mathrm{ev}_{w^1,\dots,w^q} \Big(g^{(p+1)}(x)(v^1, \dots, v^p, v^{p+1}) \Big) \quad \text{Def. von } g^{(p+1)}
\end{aligned}
$$

$(\Uparrow)$

$$
\begin{aligned}
& \exists\, \partial_{v^1} \dots \partial_{v^{p+1}} (\mathrm{ev}_{w^1,\dots,w^q} \circ g)(x)\\
\Rightarrow & \exists\, \partial_{v^{p+1}} (\mathrm{ev}_{w^1,\dots,w^q} \circ g)(x) \quad \text{nach Def.}\\
\Rightarrow & g \in C^1,\ \partial_{v^{p+1}}((\mathrm{ev}_{w^1,\dots,w^q} \circ g)(x)) =\\
& = g'(x)(v^{p+1})(w^1, \dots, w^q) = \partial_{v^{p+1}} g(x)(w^1, \dots, w^q) \quad \text{nach (p=1)}
\end{aligned}
$$

Es sei $g' \simeq: f : U \to L(E, E_1, \dots, E_p; F)$

$$
\begin{aligned}
\Rightarrow & \exists\, \partial_{v^1} \dots \partial_{v^p} (\mathrm{ev}_{v^{p+1},w^1,\dots,w^q} \circ f)\\
& = \partial_{v^1} \dots \partial_{v^p} (\mathrm{ev}_{w^1,\dots,w^q} \circ \mathrm{ev}_{v^{p+1}} \circ g') \quad \text{Def. von } f\\
& = \partial_{v^1} \dots \partial_{v^p} (\mathrm{ev}_{v^{p+1}} \circ (\mathrm{ev}_{w^1,\dots,w^q})_* \circ g')\\
& = \partial_{v^1} \dots \partial_{v^p} (\mathrm{ev}_{v^{p+1}} \circ (\mathrm{ev}_{w^1,\dots,w^q} \circ g)')\\
& = \partial_{v^1} \dots \partial_{v^{p+1}} (\mathrm{ev}_{w^1,\dots,w^q} \circ g)
\end{aligned}
$$

Nach Induktionshypothese folgt nun

$$
\begin{aligned}
\Rightarrow & f \simeq g' \in C^p,\ \partial_{v^1} \dots \partial_{v^p} (\mathrm{ev}_{v^{p+1},w^1,\dots,w^q} \circ f)(x) =\\
& = f^{(p)}(x)(v^1, \dots, v^p)(v^{p+1}, w^1, \dots, w^q)\\
& = g^{(p+1)}(x)(v^1, \dots, v^{p+1})(w^1, \dots, w^q) \quad \square
\end{aligned}
$$

6.3.7 Folgerung. Differenzierbarkeit via Richtungsableitungen.
$f : E \supseteq U \to F$ ist C^p $\Leftrightarrow$ $\partial_{v^1} \dots \partial_{v^p} f$ existiert für alle $v^j \in E$ und $x \mapsto ((v_1, \dots, v_p) \mapsto \partial_{v^1} \dots \partial_{v^p} f(x)$ ist wohldefiniert und stetig als Abbildung

$$E \supseteq U \to L(E, \dots, E; F).$$

Weiters haben wir dann

$$f^{(p)}(x)(v^1, \dots, v^p) = \partial_{v^1} \dots \partial_{v^p} f(x).$$

Für endlich dimensionale E genügt die Stetigkeit von $\partial_{v^1,\dots,v^p} f$ zu verlangen, denn eine Abbildung nach $L(E, \dots, E; F)$ ist genau dann stetig, wenn alle Zusammensetzungen mit $\mathrm{ev}_{w^1,\dots,w^p} : L(E, \dots, E; F) \to$

F es sind. Die multi-Linearität von $\partial_{v^1,\dots,v^p} f(x)$ bzgl. $v^1,\dots,v^p$ folgt dann wie in [13, 163.1]. Und analog wie in [13, 163.1] genügt sogar die Existenz und Stetigkeit der iterierten partiellen Ableitungen $\partial_{i^1}\dots\partial_{i_p} f$ für $f \in C^p$.

6.3.8 Folgerung. Diff.barkeit in Räume multilin. Abbildungen via Punktevaluationen. *Es sei $g : E \supseteq U \to L(E_1,\dots,E_q;F)$. Dann ist g C^p vorausgesetzt*

$$\mathrm{ev}_{v^1,\dots,v^q} \circ g : E \supseteq U \to L(E_1,\dots,E_q;F) \to F$$

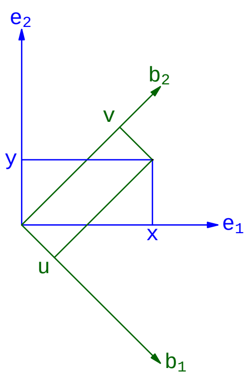
ist D^p, und

$$\begin{aligned}&x \mapsto ((w^1,\dots,w^p;v^1,\dots,v^q) \mapsto (\mathrm{ev}_{v^1,\dots,v^q} \circ g)^{(p)}(x)(w^1,\dots,w^p)\\ &U \to L(E,\dots,E,E_1,\dots,E_q;F)\end{aligned}$$

ist wohldefiniert und stetig. □

6.3.9a Beispiel. Wirkung einer linearen Transformation auf die zweite Ableitung. Wir können allgemein durchrechnen, wie sich die Ableitungen ändern, wenn wir eine lineare Koordinatentransformation a auf f anwenden, d.h. $\tilde f := f \circ a$ mit stetig linearen a setzen:

$$\begin{aligned}
(f\circ a)'(z) &= f'(a(z)) \circ a'(z) = f'(a(z)) \circ a = a^*(f'(a(z)))\\
&= (a^* \circ f' \circ a)(z) \Rightarrow\\
(f\circ a)'(z)(h) &= f'(a(z))(a(h)) \Rightarrow\\
((f\circ a)')'(z) &= (a^* \circ f' \circ a)'(z) = (a^*)'\Big(f'(a(z))\Big) \circ (f')'(a(z)) \circ a'(z)\\
&= a^* \circ (f')'(a(z)) \circ a \qquad \text{da } a^* \text{ und } a \text{ linear ist}\\
(f\circ a)''(z)(h,k) &= ((f\circ a)')'(z)(h)(k) = \Big(a^* \circ (f')'(a(z)) \circ a\Big)(h)(k)\\
&= \bigg(a^*\Big((f')'\big(a(z)\big)\big(a(h)\big)\Big)\bigg)(k)\\
&= \bigg(\Big((f')'\big(a(z)\big)\big(a(h)\big)\Big) \circ a\bigg)(k)\\
&= \bigg(\Big((f')'\big(a(z)\big)\big(a(h)\big)\Big)\big(a(k)\big)\bigg)\\
&= f''\big(a(z)\big)\big(a(h),a(k)\big)
\end{aligned}$$



Sei insbesonders die Standardbasis $\mathcal{E}$ gegeben durch e_1 und e_2, und eine dazu um sagen wir -45^o gedrehte Basis $\mathcal{B}$ mit Richtungsvektoren b_1 und b_2 gegeben. Anstelle Punkte z in der Ebene durch die Standardkoordinaten $z = x\,e_1 + y\,e_2$ darzustellen verwenden wir $z = u\,b_1 + v\,b_2$, also die Basisdarstellungen $[z]_{\mathcal{E}} = (x,y)$ und $[z]_{\mathcal{B}} = (u,v)$. Offensichtlich ist $b_1 = \frac{1}{\sqrt2}(e_1 - e_2)$ und $b_2 = \frac{1}{\sqrt2}(e_1+e_2)$.

Sei r die Drehung in der Ebene um 45^o. Diese bildet die neuen Basisvektoren auf die Standardbasisvektoren ab, d.h. $r(b_i) = e_i$. In der linearen Algebra lernt man, daß die Matrixdarstellungen dieser Basiswechselabbildung wie folgt beschrieben werden können:

$$[r]_{\mathcal{E},\mathcal{E}} = [r]_{\mathcal{B},\mathcal{B}} = [\mathrm{id}]_{\mathcal{E},\mathcal{B}}.$$

In der Tat sind die Spalten $(r_i^j)_{j=1,\dots,n}$ von $[r]_{\mathcal{B},\mathcal{B}} = (r_i^j)_{i=1,\dots,n;j=1\dots,n}$ gerade die Komponenten der Bildvektoren $r(b_i) = e_i$ bzgl. der Basis $\mathcal{B}$, d.h. $e_i = \sum_{j=1}^n r_i^j\, b_j$, in unserem Beispiel also

$$e_1 = \frac{1}{\sqrt2}(b_1+b_2), \quad e_2 = \frac{1}{\sqrt2}(b_2 - b_1) \Rightarrow [r]_{\mathcal{B},\mathcal{B}} = \begin{pmatrix} r_1^1 & r_2^1\\ r_1^2 & r_2^2\end{pmatrix} = \frac{1}{\sqrt2}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}.$$

Ebenso sind die Spalten von $[\mathrm{id}]_{\mathcal{E},\mathcal{B}}$ die Komponten von e_i bzg. $\mathcal{B}$ und schließlich sind jene von $[r]_{\mathcal{E},\mathcal{E}}$ die Komponenten von $r(e_i)$ bzg. e_i, also gegeben durch:

$$r(e_i) = r\Big(\sum_{j=1}^n r_i^j\, b_j\Big) = \sum_{j=1}^n r_i^j\, r(b_j) = \sum_{j=1}^n r_i^j\, e_j.$$

Für die entsprechenden Koordinaten bedeutet dies:

$$\begin{aligned} u\, b_1 + v\, b_2 = z = x\, e_1 + y\, e_2 &= x\,(r_1^1\, b_1 + r_1^2\, b_2) + y\,(r_2^1\, b_1 + r_2^2\, b_2) \\ &= (r_1^1\, x + r_2^1\, y)\, b_1 + (r_1^2\, x + r_2^2\, y)\, b_2, \end{aligned}$$

also

$$u = r_1^1\, x + r_2^1\, y, \quad v = r_1^2\, x + r_2^2\, y \Rightarrow \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} r_1^1 & r_2^1 \\ r_1^2 & r_2^2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = [r]_{\mathcal{E},\mathcal{E}} \cdot \begin{pmatrix} x \\ y \end{pmatrix}.$$

In (6.3.9) haben wir eine Funktion f in den Standard-Koordinaten durch $f(x,y) := x \cdot y$ definiert. Wenn wir diese Funktion nun in den neuen Koordinaten (u,v) beschreiben, also $\tilde f(u,v) := f(x,y) = f(a(u,v))$ mit

$$[a] = [r^{-1}] = \frac{1}{r_1^1 r_2^2 - r_1^2 r_2^1} \begin{pmatrix} r_2^2 & -r_2^1 \\ -r_1^2 & r_1^1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix},$$

so erhalten wir

$$\tilde f(u,v) = f\Big(\tfrac{1}{\sqrt{2}}(u+v), \tfrac{1}{\sqrt{2}}(v-u)\Big) = \frac{v^2 - u^2}{2}.$$

Zusammengefaßt:

$$\begin{aligned} b_1 &= \tfrac{1}{\sqrt{2}}(e_1 - e_2) & e_1 &= \tfrac{1}{\sqrt{2}}(b_1 + b_2) \\ b_2 &= \tfrac{1}{\sqrt{2}}(e_1 + e_2) & e_2 &= \tfrac{1}{\sqrt{2}}(b_2 - b_1) \\ u &= \tfrac{1}{\sqrt{2}}(x - y) & x &= \tfrac{1}{\sqrt{2}}(u + v) \\ v &= \tfrac{1}{\sqrt{2}}(x + y) & y &= \tfrac{1}{\sqrt{2}}(v - u) \end{aligned}$$

Dieses Drehen der Koordinatenrichtungen können wir ebensogut als Drehung des Graphs der Funktion in die umgekehrte Richtung interpretieren. Insbesonders ist

$$\begin{aligned} \partial_2\partial_2 \tilde f(u,v) &= \partial_2\partial_2 (f\circ a)(u,v) = (f\circ a)''(u,v)(e_2,e_2) = f''\big(a(u,v)\big)\big(a(e_2), a(e_2)\big) \\ &= f''(x,y)\left(\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) \\ &= \partial_1\partial_1 f(x,y)\tfrac{1}{\sqrt{2}}\tfrac{1}{\sqrt{2}} + \partial_1\partial_2 f(x,y)\tfrac{1}{\sqrt{2}}\tfrac{1}{\sqrt{2}} + \partial_2\partial_1 f(x,y)\tfrac{1}{\sqrt{2}}\tfrac{1}{\sqrt{2}} + \partial_2\partial_2 f(x,y)\tfrac{1}{\sqrt{2}}\tfrac{1}{\sqrt{2}} \\ &= 0 + \tfrac{1}{2} + \tfrac{1}{2} + 0 = 1. \end{aligned}$$

Man beachte, daß in diesem Beispiel $\partial_1\partial_2 f(z) = \partial_2\partial_1 f(z)$ gilt. Daß dies ziemlich allgemein der Fall ist, zeigt folgender

6.3.10 Satz von Schwarz. Vertauschbarkeit von Ableitungen.
Es sei $f : E \supseteq U \to F$ und $u, v \in E$ gegeben. Wir nehmen an, daß $d_u f(x)$, $d_v f(x)$ und $d_u d_v f(x)$ existieren und stetig in $x \in U$ sind. Dann existiert $d_v d_u f(x)$ und stimmt überein mit $d_u d_v f(x)$.

Beweis. Es ist

$$\begin{aligned} (d_u d_v f)(x) &= \lim_{t\to 0+} \frac{d_v f(x+tu) - d_v f(x)}{t} \\ &= \lim_{t\to 0+}\lim_{s\to 0+} \left(\frac{f(x+tu+sv) - f(x+tu)}{s} - \frac{f(x+sv) - f(x)}{s}\right)\cdot \frac{1}{t} \\ &= \lim_{t\to 0+}\lim_{s\to 0+} \varphi(t,s), \end{aligned}$$

wobei die Funktion $\varphi : \mathbb{R}^2 \supseteq \{(t,s) : x + t\,v + s\,w \in U\} \to F$ gegeben ist durch

$$\begin{aligned}\varphi(t,s) &:= \left(\frac{f(x+tu+sv)-f(x+tu)}{s} - \frac{f(x+sv)-f(x)}{s}\right)\cdot\frac{1}{t}\\ &\xlongequal{\text{(6.1.23a), } d_v \text{ stetig}} \left(\int_0^1 d_vf(x+tu+\sigma sv)\,d\sigma - \int_0^1 d_vf(x+\sigma sv)\,d\sigma\right)\cdot\frac{1}{t}\\ &= \int_0^1 \frac{d_vf(x+tu+\sigma sv)-d_vf(x+\sigma sv)}{t}\,d\sigma\\ &\xlongequal{\text{(6.1.23a), } d_ud_v \text{ stetig}} \int_0^1\int_0^1 d_ud_vf(x+\tau tu+\sigma sv)\,d\tau\,d\sigma.\end{aligned}$$

Es ist φ stetig bei $(0,0)$, denn für für $t,s \to 0+$ konvergiert $d_ud_vf(x+\tau tu+\sigma sv) \to d_ud_vf(x)$ gleichmäßig für $\tau,\sigma \in [0,1]$ wegen der Stetigkeit von d_ud_vf, und wegen (5.6.14) folgt die Konvergenz von $\varphi(t,s) \to d_ud_vf(x)$.

Für $s \neq 0$ existiert wegen der Existenz von d_uf

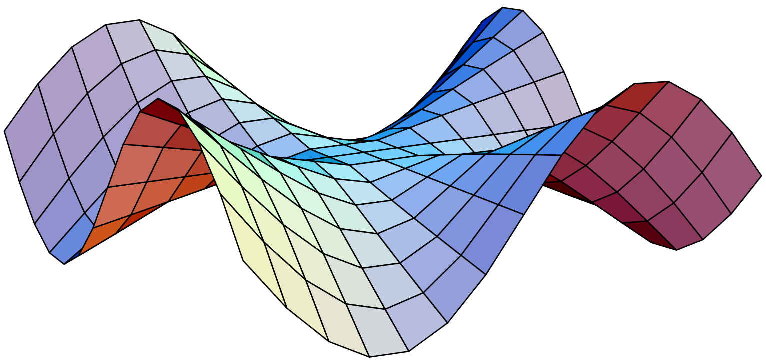
$$\lim_{t\to 0+}\varphi(t,s) = \lim_{t\to 0+}\frac{f(x+tu+sv)-f(x+tu)-f(x+sv)+f(x)}{t\,s} = \frac{d_uf(x+sv)-d_uf(x)}{s}.$$

Nach dem ersten Lemma in (3.2.8) folgt

$$\exists\, d_vd_uf(x) := \lim_{s\to 0+}\lim_{t\to 0+}\varphi(t,s) = \lim_{s\to 0+,t\to 0+}\varphi(t,s) = d_ud_vf(x) \quad \square$$

Beispiel. Nicht-Symmetrie der zweiten Ableitung.
Betrachte

$$f(x,y) := \begin{cases} 0 & \text{für } (x,y)=0\\ \frac{xy(x^2-y^2)}{x^2+y^2} = \frac{r^2\sin(2\varphi)(\cos(\varphi)^2-\sin(\varphi)^2)}{2} = r^2\frac{\sin(4\varphi)}{2} & \text{andernfalls.}\end{cases}$$



Dann existiert $\partial_i\partial_j$ für alle i,j, aber $\partial_1\partial_2 f(0,0) \neq \partial_2\partial_1 f(0,0)$.
In der Tat $\partial_1 f(0,y) = -y$, $\partial_2 f(x,0) = x$ und daher erhalten wir $\partial_2\partial_1 f(0,0) = -1 \neq 1 = \partial_1\partial_2 f(0,0)$.

Folgerung. Symmetrie der Ableitungen.
Es sei $f : E \supseteq U \to F$ C^p. Dann ist $f^{(p)}(x)$ symmetrisch für alle $x \in U$.

Beweis. Wir zeigen dies mittels Induktion nach p. Für $(p=2)$ ist das (6.3.10).

Für $(p+1)$ haben wir

$$f^{(p+1)}(x)(v_1,\dots,v_{p+1}) := (f')^{(p)}(x)(v_1,\dots,v_p)(v_{p+1}).$$

Da $g := f' : U \to L(E,F)$ C^p ist, erhalten wir nach Induktions-Hypothese, daß der Ausdruck $g^{(p)}(x)(v_1,\dots,v_p)$ symmetrisch in $(v_1,\dots,v_p)$ ist und

$$g^{(p)}(x)(v_1,\dots,v_p) = (g')^{(p-1)}(x)(v_1,\dots,v_{p-1})(v_p).$$

Also ist

$$f^{(p+1)}(x)(v_1,\dots,v_{p+1}) = g^{(p)}(x)(v_1,\dots,v_p)(v_{p+1})$$

symmetrisch in $(v_1,\dots,v_p)$ und

$$\begin{aligned} f^{(p+1)}(x)(v_1,\dots,v_{p+1}) &= (g')^{(p-1)}(x)(v_1,\dots,v_{p-1})(v_p)(v_{p+1}) \\ &= ((f')')^{(p-1)}(x)(v_1,\dots,v_{p-1})(v_p)(v_{p+1}) \\ &= (f'')^{(p-1)}(x)(v_1,\dots,v_{p-1})(v_p,v_{p+1}) \end{aligned}$$

ist symmetrisch in (v_p, v_{p+1}) da $f''(x)$ symmetrisch ist. Somit folgt das Resultat, da die Permutationen von $\{1,\dots,p+1\}$ erzeugt werden durch jene von $\{1,\dots,p\}$ und die Transposition $(p,p+1)$, siehe (1.3.27). □

Bemerkung. Koordinatendarstellung der höheren Ableitungen.
Es sei $f : \mathbb{R}^n \supseteq U \to F$ p-mal stetig differenzierbar. Dann ist nach (6.3.5)

$$f^{(p)}(x)(v_1,\dots,v_p) = d_{v_1}\dots d_{v_p} f(x) = \sum_{i_1,\dots,i_p} v_1^{i_1}\dots v_p^{i_p}\cdot \partial_{i_1}\dots\partial_{i_p} f(x) \text{ für } v_i = (v_i^1,\dots,v_i^n)\in\mathbb{R}^n.$$

Wenn wir nun die Symmetrie der p-ten Ableitung verwenden, so können wir die partiellen Ableitungen nach Größe des Index sortieren und brauchen nur den Koeffizienten von $(\partial_1)^{p_1}\dots(\partial_n)^{p_n}$ bestimmen, wobei $p_1+\dots+p_n = p$ und alle $p_i \geq 0$. Dieser ist gerade

$$\sum_{i_1,\dots,i_p} v_1^{i_1}\dots v_p^{i_p},$$

wobei die Summe über all jene Indizes $i_1,\dots,i_p$ läuft, von denen gerade p_1 viele 1 sind, p_2 viele 2 sind, …, und p_n viele n sind. Ist insbesonders $v_1 = \dots = v_p = v = (v^1,\dots,v^n) \in \mathbb{R}^n$, so ist der entsprechende Koeffizient $v_1^{i_1}\dots v_n^{i_n} = (v^1)^{p_1}\dots(v^n)^{p_n}$ und es gibt genau $\frac{p!}{p_1!\dots p_n!}$ viele solche Summanden (Anzahl der Möglichkeiten die $p = p_1+\dots+p_n$ vielen partiellen Ableitungen auf die Kästchen mit gegebener Koordinatenrichtung zu verteilen, vergleiche dies mit (1.5.3)), d.h.

$$f^{(p)}(x)(v,\dots,v) = \sum_{p_1+\dots+p_n=p} \frac{p!}{p_1!\dots p_n!}\,(v^1)^{p_1}\dots(v^n)^{p_n}\,(\partial_1)^{p_1}\dots(\partial_n)^{p_n}\,f(x).$$

6.3.11 Taylor'sche Formel.
Es sei $f : E \supseteq U \to F$ C^p und die Verbindunsgstrecke $\overline{x(x+h)} \subseteq U$. Dann ist

$$f(x+h) = \sum_{0\le j<p} \frac{f^{(j)}(x)(h,\dots,h)}{j!} + \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!}\, f^{(p)}(x+th)(h,\dots,h)\,dt.$$

Beachte, daß wir im Gegensatz zum Taylor'schen Lehrsatz (4.2.2) nun nicht das Restglied in der Form von Lagrange beschrieben sondern durch ein explizites Integral. Man schreibt auch gerne h^n anstelle von $(h,\dots,h)$ um die Ähnlichkeit mit der Taylor-Formel im 1-dimensionalen perfekt zu machen:

$$f(x+h) = \sum_{0\le j<p} \frac{f^{(j)}(x)\cdot h^j}{j!} + \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!}\, f^{(p)}(x+th)\cdot h^p\,dt.$$

Für $E = \mathbb{R}^n$ lautet die Taylor-Formel somit in Koordinaten

$$f(x+h) = \sum_{p_1+\dots+p_n<p} \frac{1}{p_1!\dots p_n!} \partial_1{}^{p_1}\dots\partial_n{}^{p_n} f(x)\cdot h_1{}^{p_1}\dots h_n{}^{p_n} + \\ + \int_0^1 (1-t)^{p-1}\, p \sum_{p_1+\dots+p_n=p} \frac{1}{p_1!\dots p_n!} \partial_1{}^{p_1}\dots\partial_n{}^{p_n} f(x+t\,h)\cdot h_1{}^{p_1}\dots h_n{}^{p_n}\,dt.$$

Beweis. O.B.d.A. dürfen wir annehmen, daß $E = \mathbb{R}$ und U ein offenes Intervall ist, welches $x = 0$ und $h = 1$ enthält. Indem wir $c(t) := f(x + t\,h)$ setzen erhalten wir nach iterierter partieller Integration

$$\begin{aligned} c(1) &= c(0) + \int_0^1 1\, c'(t)\,dt \\ &= c(0) + \Big[(t-1)\,c'(t)\Big]_{t=0}^1 - \int_0^1 (t-1)\,c''(t)\,dt \\ &= c(0) + c'(0) - \Big[\frac{(t-1)^2}{2}\,c''(t)\Big]_{t=0}^1 + \int_0^1 \frac{(t-1)^2}{2}\,c'''(t)\,dt \\ &\ \ \vdots \\ &= c(0) + c'(0) + \dots + \frac{1}{(p-1)!}\,c^{(p-1)}(0) + \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!}\,c^{(p)}(t)\,dt \quad \square \end{aligned}$$

6.3.12 Lemma. Vollständigkeit von C^p.
Es sei $f_j \in C^p$ so, daß $f_j^{(k)}$ gleichmäßig konvergiert auf beschränkten Mengen gegen eine Funktion f^k für jedes $k \le p$. Dann ist $f := f^0$ C^p und $f^{(k)} = f^k$ für alle $k \le p$.

Beweis. Es genügt dies für $p = 1$ zu zeigen. Die allgemeine Aussage folgt dann mittels Induktion nach p.
Beachte zuerst, daß wir aus der gleichmäßigen Konvergenz auf beschränkten Mengen auf die Stetigkeit der Grenzfunktion schließen können:
In der Tat sei $x \in E$ und U sei eine Umgebung von $f(x)$. Dann

$$f(y) - f(x) = f(y) - f_n(y) + f_n(y) - f_n(x) + f_n(x) - f(x) \in \frac{1}{3}U + \frac{1}{3}U + \frac{1}{3}U = U,$$

vorausgesetzt wir wählen eine beschränkte Umgebung B und n so groß, daß $f_n(y) - f(y) \in \frac{1}{3}U$ für alle $y \in B$ und wählen eine Umgebung $V \subseteq B$ von x so daß $f_n(y) - f_n(x) \in \frac{1}{3}U$ für alle $y \in V$.
Da f_n C^1 ist, haben wir

$$\frac{f_n(x+tv) - f_n(x)}{t} = \int_0^1 f_n'(x + t\tau v)(v)\,d\tau.$$

Die linke Seite konvergiert (punktweise) gegen

$$\frac{f(x+tv) - f(x)}{t}$$

und die rechte gegen

$$\int_0^1 f^1(x + t\tau v)(v)\,d\tau, \text{ da } f_n' \to f^1 \text{ gleichmäßig auf beschränkten Mengen.}$$

Also ist

$$\frac{f(x+tv) - f(x)}{t} = \int_0^1 f^1(x + t\tau v)(v)\,d\tau.$$

Da f^1 stetig ist (als Limes der stetigen f_n') erhalten wir, daß die rechte Seite für $t \to 0$ gegen $\int_0^1 f^1(x+t\tau v)(v)\,d\tau = f^1(x)(v)$ konvergiert. Also existiert $d_v f(x)$ und $f^1 : x \mapsto (v \mapsto d_v f(x))$ ist als Limes der stetigen Abbildungen f_n' eine wohldefinierte stetige Abbildung $E \to L(E,F)$. Daraus folgt, daß f differenzierbar ist und $f' = f^1$ nach (6.1.2). □

6.3.13 Lemma. Glattheit von Potenzreihen in Banach-Algebren.
Es sei $\sum_k \lambda_k t^k$ eine Potenzreihe mit skalaren Koeffizienten λ_k, welche auf $\{t \in \mathbb{R} : |t| < \varepsilon\}$ konvergiert. Es sei A eine Banach-Algebra. Dann ist $a \mapsto \sum_k \lambda_k a^k$ eine wohldefinierte C^∞-Abbildung $\{a \in A : \|a\| < \varepsilon\} \to A$.

Es ist also $a \mapsto \exp(a)$, $a \mapsto \sin(a)$ und $a \mapsto \cos(a)$ unendlich oft differenzierbar auch für Elemente a einer Banach-Algebra A wie z.B. $A = L(E)$. Für $A = C(I,\mathbb{R})$ folgt dies auch induktiv aus (6.1.20).

Beweis. Es sei $f_k(a) := \lambda_k a^k$. Dann ist f_k ein k-homogenes Polynom und daher C^∞. Beachte, daß $f_k'(a)(v) = \lambda_k(v \cdot a^{k-1} + \cdots + a^{k-1} \cdot v)$. Die Reihe $\sum_k f_k^{(p)}$ konvergiert gleichmäßig auf abgeschlossen Bällen in $\{a \in A : \|a\| < \varepsilon\}$, in der Tat haben wir, daß $f_k^{(p)}(a)(v_1,\dots,v_p)$ λ_k-mal die Summe von allen Produkten von k Faktoren ist, wobei p Faktoren $v_1,\dots,v_p$ sind und alle anderen sind a. Jeder solcher Faktor definiert eine stetige Abbildung $A \to L(A,\dots,A;A)$, welche beschränkt ist durch $\|a\|^{k-p}$. Es gibt $k(k-1)\dots(k-p+1)$ viele solche Faktoren, daher ist $\|f_k^{(p)}(a)\| \le k(k-1)\dots(k-p+1)\|a\|^{k-p}$. Da

$$\Big(\frac{d}{dt}\Big)^p\Big(\sum_{k\ge 0}\lambda_k t^k\Big) = \sum_{k\ge p}\lambda_k\, k(k-1)\dots(k-p)t^{k-p}$$

auf kompakten Intervallen in $\{t : |t| < \varepsilon\}$ konvergiert, folgt die Konvergenz von $\sum_k f_k^{(p)}$. Somit folgt das Resultat mittels (6.3.12). □

6.3.14 Theorem. Glatte Lösungen von Gleichungen.
Folgende Sätze gelten auch für n-mal stetig differenzierbare und auch für glatte Abbildungen:

(1) Satz (6.2.2) über inverse Funktionen;
Genauer: Ist die Gleichung C^n so auch die C^1-Lösung.

(2) Satz (6.2.4) über implizite Funktionen ;
Genauer: Ist die Gleichung C^n so auch die C^1-Lösung.

(3) Die Inversion $\mathrm{inv} : \mathrm{Inv}(A) \to A$ von $A := L(E,E)$ (oder einer Banach-Algebra) ist C^n, vgl. (5.5.16) und (6.2.6).

(4) Aus $f : E \supseteq U \to F$ C^n folgt $f_* : C(I,E) \supseteq C(I,U) \to C(I,F)$ C^n, vgl. (6.1.18) und (6.1.20).

(5) Existenz und Eindeutigkeitssatz (6.2.8) für Differentialgleichungen;
Genauer: Ist die Gleichung C^n so auch die C^1-Lösung.

Beweis. Wir wenden die Methode von Baron Münchhausen an und ziehen uns mittels Induktion nach n an den eigenen Füßen aus dem Sumpf.

Den Induktionsanfang ($n = 1$) haben wir bereits in (6.2.2), (6.2.4), (6.2.6), (6.1.20) und (6.2.8) erledigt. Nun der Induktionsschritt auf $(n+1)$:

(1) Für die Inverse $g := f^{-1}$ von f gilt $g'(y) = \mathrm{inv}(f'(g(y)))$ nach (6.2.2), d.h.

$$g' = \mathrm{inv} \circ f' \circ g.$$

Nach (3) ist inv C^n, nach Voraussetzung ist f' C^n und nach (2) ist g C^n, also ist g' C^n nach der Kettenregel (6.3.3.2), d.h. g ist C^{n+1}.

(2) Die Lösung g der impliziten Gleichung $f(x, g(x, z)) = z$ ist nach Induktionsanfang C^1 mit

$$\partial_1 g(x, z) = -\partial_2 f(x, g(x, z))^{-1} \cdot \partial_1 f(x, g(x, z))$$
$$\partial_2 g(x, z) = -\partial_2 f(x, g(x, z))^{-1}.$$

Die rechte Seite ist C^n nach Induktionsvoraussetzung für (3) und (2) und (6.3.3.1). Also ist g C^{n+1} nach (6.1.24).

(3) Dies ist ein spezial-Fall von (2), siehe dem Beweis von (6.2.6).

(4) Beachte, daß wir nach (6.1.20) für $n = 1$ erhalten haben, daß $(f_*)' = \beta \circ (f')_*$, wobei $\beta : C(I, L(E, F)) \to L(C(I, E), C(I, F))$ stetig und linear ist. Also folgt mittels Induktion, daß $(f')_*$ C^n ist und damit nach (6.3.3) auch $(f_*)'$, also f_* C^{n+1} ist.

(5) Im Beweis für C^1 in (6.2.8), haben wir ein implizite Lösung $\bar{x} : \mathbb{R} \times P \to C^1_0$ erhalten. Diese ist nun nach (4) und (2) auch C^{n+1}, also ist $x_p(t) := \bar{x}_{\varepsilon t/t_0, p}(t_0/\varepsilon)$ im Beweis von (6.2.8) in (t, p) C^{n+1} und der Definitionsbereich hat sich nicht verkleinert.

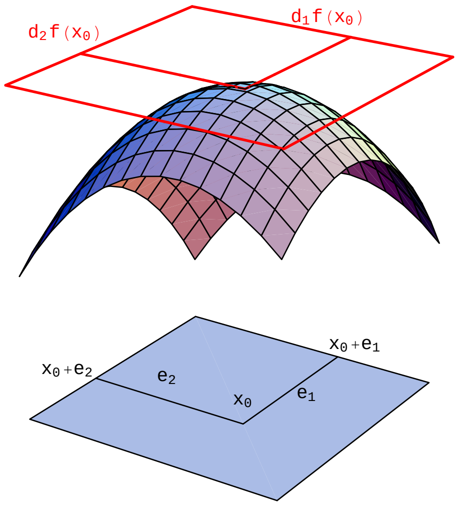
Für zeitabhängige Differentialgleichungen mit einer allgemeinen Anfangsbedingung folgt das nun ebenso wie im C^1-Fall. □

6.4 Lokale Extrema

6.4.1 Definition. Lokales Minimum und Maximum.
Es sei $f : E \supseteq U \to \mathbb{R}$ und $\xi \in U$. Dann hat f ein LOKALES MINIMUM (resp. LOKALES MAXIMUM) an der Stelle ξ, wenn ein $\delta > 0$ existiert, so daß für all $x \in U$ mit $|x - \xi| < \delta$ die Ungleichung $f(x) \geq f(\xi)$ (resp. $f(x) \leq f(\xi)$) gilt. Gilt sogar die strikte Ungleichung $f(x) > f(\xi)$ (resp. $f(x) < f(\xi)$) für all jene $x \neq \xi$, so spricht man von einem LOKALEN STRIKTEN MINIMUM (resp. LOKALES STRIKTES MAXIMUM).

6.4.2 Lemma. Notwendige Bedingung für Extremum.
Es sei $f : E \supseteq U \to \mathbb{R}$ differenzierbar bei $\xi \in U$ und habe ein lokales EXTREMUM *(d.h. Maximum oder Minimum) bei ξ.*

Dann ist ξ ein KRITISCHER PUNKT *von f, d.h. $f'(\xi) = 0$.*



Beweis. Für $v \in E$ besitzt auch $t \mapsto f(\xi + tv)$ ein lokales Extremum bei $t = 0$ und somit z.B. im Falle eines lokalen Minimums $f(\xi + tv) \geq f(\xi)$ für alle t und damit

$$\frac{f(\xi + tv) - f(\xi)}{t} \begin{cases} \geq 0 & \text{für } t > 0 \\ \leq 0 & \text{für } t < 0 \end{cases}$$

also ist $f'(\xi)(v) = 0 = \lim_{t \to 0} \frac{f(\xi+t\,v)-f(\xi)}{t} = 0$. □

6.4.3 Proposition. Hinreichende Bedingung für lokales Extremum.
Es sei $f : E \supseteq U \to \mathbb{R}$ C^2, $f'(\xi) = 0$ und die symmetrische Form $f''(\xi) : E \times E \to \mathbb{R}$ sei POSITIV DEFINIT, *d.h. $\exists\,\delta > 0$ $\forall v \neq 0$: $f''(\xi)(v, v) \geq \delta \|v\|^2$. Dann hat f ein striktes lokales Minimum bei ξ.*

Im Endlich-dimensionalen ist eine bilineare symmetrische Form b genau dann positiv definit, wenn $\forall h : b(h, h) > 0$: Denn die Einheitssphäre $\{h : \|h\| = 1\}$ ist kompakt und somit existiert $\delta := \min\{b(h, h) : \|h\| = 1\} > 0$ und für $0 \neq v \in E$ gilt $v = \|v\|h$, wobei $h := \frac{1}{\|v\|}v$ und damit $b(v, v) = b(\|v\|h, \|v\|h) = \|v\|^2 b(h, h) \geq \|v\|^2 \delta$.

Im Unendlich-dimensionalen genügt die obige schwächere Bedingung nicht: Betrachte zum Beispiel $E := c_0$ und das Polynom

$$f(x) := \sum_n \frac{1}{2^n} x_n^2 \Big(1 - (n x_n)^2\Big) = \sum_n \frac{1}{2^n} x_n^2 + \sum_n \frac{n^2}{2^n} x_n^4$$

vom Grad 4: Um die Wohldefiniertheit und Stetigkeit der quadratischen Form $x \mapsto \sum_n \frac{1}{2^n} x_n^2$ und jener vierter Ordnung $x \mapsto \sum_n \frac{n^2}{2^n} x_n^4$ zu zeigen verwenden wir die Hölder'sche Ungleichung

$$\|a \cdot b\|_1 := \sum_n |a_n b_n| \leq \sum_n |a_n| \sup\{|b_n| : n \in \mathbb{N}\} = \|a\|_1 \cdot \|b\|_\infty.$$

Es ist nun aber $f(2e_n/n) < 0$, $f(0) = 0$, $f'(0) = 0$ und $f''(0)(h,h) = \sum_n \frac{2}{2^n} x_n > 0$ für $h \neq 0$. Beachte, daß wegen $\ell^p \subseteq c_0$ dasselbe auch für ℓ^p stimmt.

In der linearen Algebra zeigt man, daß eine symmetrische bilinear-Form $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ genau dann positiv definit ist, wenn es die zugehörige quadratische Matrix $(b_{i,j})_{i,j=1}^n$ mit $b_{i,j} := b(e_i, e_j)$ ist, d.h. die Determinante der Hauptminore $(b_{i,j})_{i,j=1}^k$ positiv ist für jedes $1 \leq k \leq n$. Äquivalent dazu ist auch, daß alle Eigenwerte positiv sind.

Eine symmetrische bilinear-Form b heißt negativ definit, wenn $-b$ positiv definit ist, d.h. $b(v,v) < 0$ für alle $v \neq 0$ ist. Dies ist also genau dann der Fall, wenn alle Eigenwerte negativ, bzw. die Hauptminoren alternierendes Vorzeichen haben und zwar beginnend mit $-$ für die 1×1-Minore.

Beweis. O.B.d.A. sei $\xi = 0$. Nach Taylor's Theorem (6.3.11) für $p = 2$ ist

$$f(x) = f(0) + f'(0) \cdot x + \int_0^1 (1-t)\, f''(tx)(x,x)\, dt$$

Die Menge der stetigen positiv definiten symmetrischen bilinear-Formen ist offen in allen symmetrischen bilinear-Formen, denn wenn $b_0(x,x) \geq \delta \|x\|^2$ ist und $\|b - b_0\| < \delta$ ist, so ist

$$b(x,x) = b_0(x,x) + \Big(b(x,x) - b_0(x,x)\Big) \geq b_0(x,x) - |(b-b_0)(x,x)| \geq \overbrace{(\delta - \|b - b_0\|)}^{>0} \|x\|^2$$

Da $x \mapsto f''(x)$ stetig ist und $f''(0)$ positiv definit ist existiert ein $\delta > 0$ und ein $\varepsilon > 0$, sodaß alle $\|x\| < \varepsilon$ in U liegen und $f''(x)(h,h) \geq \frac{\delta}{2}\|h\|^2$ für alle $h \in E$ gilt. Somit ist für diese x

$$\int_0^1 (1-t)\, f''(tx)(x,x)\, dt \geq \int_0^1 (1-t) \|x\|^2 \frac{\delta}{2}\, dt \geq \|x\|^2 \frac{\delta}{4},$$

daher ist $f(x) > f(0)$ für alle $0 \neq \|x\| < \varepsilon$. □

Bemerkung. Nicht-degenerierte bilinear-Formen.
Eine stetige symmetrische bilinear-Form $b : E \times E \to \mathbb{R}$ die $b(x,x) \geq 0$ erfüllt, ist genau dann positiv definit ist, wenn $b^\vee : E \to E^* := L(E, \mathbb{R})$ ein Isomorphismus ist.

Wir können $b(x,y)$ als $b(x,y) = (Bx)(y)$ schreiben, wobei $B := \check{b} : E \to L(E,\mathbb{R}) = E^*$. In der linearen Algebra lernt man, daß der Annihilator $(\operatorname{Bild} B)_o := \{y \in E : B(x)y = 0\}$ genau der Kern von $B^t : E \to L(E,\mathbb{R})$ definiert durch $B^t(x)(y) := B(y)(x)$ ist. Da aber b symmetrisch ist, ist $B = B^t$ und wenn b positiv definit ist, so ist B injektiv und damit $(\operatorname{Bild} B)_o = \{0\}$ also $\operatorname{Bild} B$ dicht in $L(E,\mathbb{R})$. Andererseits folgt aus $b(x,x) \geq \delta \|x\|^2$, daß $\|B(x)\| \geq \frac{\|b(x,x)\|}{\|x\|} \geq \delta \|x\|$, und deshalb ist das Bild von B abgeschlossen, also B surjektiv.

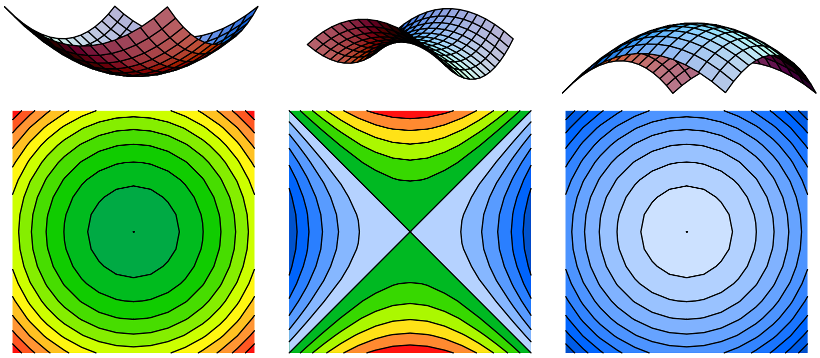
Umgekehrt sei $B : E \to E \cong L(E,\mathbb{R})$ bijektiv. Dann ist 0 nicht im Spektrum $\sigma(B) := \{\lambda \in \mathbb{C} : B - \lambda \operatorname{id}$ ist nicht bijektiv $\}$ von B. Mittels Spektral-Theorie (siehe [17, 9.27]) folgt für die Quadrat-Wurzel $C := \sqrt{B}$ von B, daß $0 \notin \sigma(C) = \{\sqrt{\lambda} : \lambda \in \sigma(B)\}$, d.h. C ist bijektiv. Also ist

$$b(x,x) = \langle B(x)|x\rangle = \langle C^2(x)|x\rangle = \langle C(x)|C(x)\rangle = \|C(x)\|^2 \geq (\|C^{-1}\|^{-1}\|x\|)^2,$$

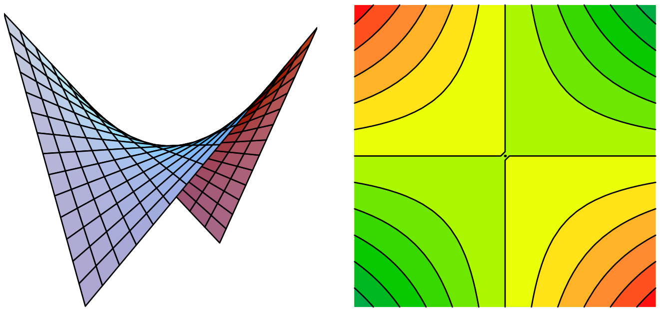
da $\|x\| = \|C^{-1}Cx\| \leq \|C^{-1}\|\|C(x)\|$.

6.4.4 Beispiele von Extremalpunkten.
Die typischen QUADRATISCHEN (NICHT DEGENERIERTEN) FORMEN mit 2 positiven, resp. einem positiven und einem negativen Eigenwert, resp. zwei negativen Eigenwerten sind:

$x^2 + y^2$	$x^2 - y^2$	$-x^2 - y^2$
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
positiv definit	indefinit	negativ definit

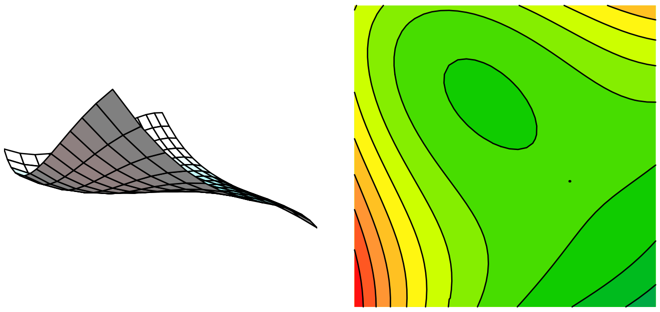


1. $f(x, y) := xy$ hat nur $(0, 0)$ als kritischen Punkt. Dieser ist ein SATTELPUNKT.



   In der Tat ist $f'(x, y) = (y, x) = 0 \Leftrightarrow x = 0 = y$ und die Eigenwerte λ von
   $$f''(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
   sind durch $0 = \det\Big(\lambda \,\mathrm{id} - f''(0,0)\Big) = \lambda^2 - 1$, also $\lambda = \pm 1$ gegeben.

2. $f(x, y) := x^3 + y^3 - 3xy$ hat nur $(0, 0)$ und $(1, 1)$ als kritische Punkte. Der erste ist ein Sattelpunkt (setze $y = 0$) und der zweite ein lokales striktes Minimum. Beachte, daß die Küstenlinie $\{(x, y) : f(x, y) = 0\}$ gerade das kartesische Blatt ist.

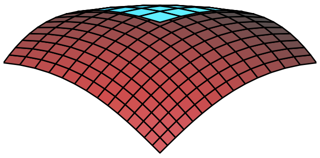
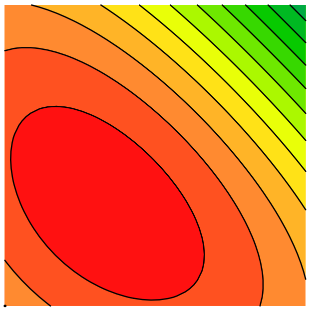


In der Tat ist $f'(x,y) = 3(x^2 - y, y^2 - x) = 0 \Leftrightarrow (x,y) = (0,0)$ oder $(x,y) = (1,1)$. Die Eigenwerte λ von

$$f''(x,y) = 3\begin{pmatrix} 2x & -1 \\ -1 & 2y \end{pmatrix}$$

sind für $(x,y) = (0,0)$ ähnlich wie zuvor ± 3 und für $(x,y) = (1,1)$ durch $0 = \det(\lambda\,\mathrm{id} - f''(1,1)) = (\lambda - 6)^2 - 3^2$, also $\lambda = 6 \pm 3$ gegeben.

3. $f(x,y) := \sin(x) + \sin(y) + \sin(x+y)$ für $0 \le x, y \le \frac{\pi}{2}$.

Der einzige kritische Punkt im Inneren ist $(\frac{\pi}{3}, \frac{\pi}{3})$, ein globales Maximum. Am Rand sind $(0,0)$ und $(\frac{\pi}{2}, \frac{\pi}{2})$ lokale Minima (und der erste Punkt ein globales) und $(\frac{\pi}{2}, \frac{\pi}{4})$ und $(\frac{\pi}{4}, \frac{\pi}{2})$ lokale Maxima am Rand aber nicht von f am ganzen Quadrat:

Hier ist $f'(x,y) = (\cos(x) + \cos(x+y), \cos(y) + \cos(x+y)) = (0,0)$ genau dann wenn $\cos(x) = \cos(y)$, also $x = y$, und $0 = \cos(x) + \cos(2x) = \cos(x) + 2\cos^2(x) - 1$, also $\cos(x) = -\frac{1}{4} \pm \frac{3}{4}$, d.h. $x = \arccos(\frac{1}{2}) = \frac{\pi}{3}$. In diesem Punkt ist

$$f''(x,y) = \begin{pmatrix} -\sin(x) - \sin(x+y) & -\sin(x+y) \\ -\sin(x+y) & -\sin(x) - \sin(x+y) \end{pmatrix} = \begin{pmatrix} -\sqrt{3} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\sqrt{3} \end{pmatrix}$$

mit Eigenwerten $-\frac{3\sqrt{3}}{2}$ und $-\frac{\sqrt{3}}{2}$, also liegt ein Maximum vor.

Bleibt noch das Verhalten von f am Rand zu studieren. Es ist $x \mapsto f(x,0) = 2\sin(x)$ streng monoton wachsend auf $[0, \pi/2]$ von 0 auf 2. Weiters hat $y \mapsto f(\pi/2, y) = 1 + \sin(y) + \sin(\pi/2 + y) = 1 + \sin(y) + \cos(y) = 1 + \sqrt{2}\sin(\pi/4 + y)$ ein Maximum von $1 + \sqrt{2}$ bei $\pi/4$ und fällt auf 2 gegen die Randpunkte $y \in \{0, \pi/2\}$. Allerdings ist dies kein Maximum von f, denn $\partial_1 f(\pi/2, \pi/4) = \cos(\pi/2) + \cos(\pi/2 + \pi/4) = -\frac{\sqrt{2}}{2}$ als fällt f in Richtung e_1 und wächst in Richtung $-e_1$. Gleiches gilt aus Symmetriegründen für die anderen beiden Kanten.

6.4.5 Bemerkung. Gradient.

Da es nur Sinn macht für skalar-wertige Funktionen $f : E \supseteq U \to \mathbb{R}$ von lokalen Extrema zu sprechen, wollen wir für diese eine andere Interpretation der Ableitung $f'(x) \in L(E, \mathbb{R})$ geben. Wenn $E = \mathbb{R}^n$ ist, dann ist die Jacobi-Matrix

$$[f'(x)] = \Big(\partial_1 f(x), \dots, \partial_n f(x)\Big)$$

und $f'(x)(v)$ ist durch Matrix-Multiplikation gegeben:

$$f'(x)(v) = \Big(\partial_1 f(x), \dots, \partial_n f(x)\Big) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

wobei $v_1, \dots, v_n$ die Koordinaten von v sind. Wir müssen also den Vektor v als Spaltenvektor interpretieren und $f'(x)$ als Zeilenvektor der gleichen Länge. Wenn wir die Matrix $[f'(x)]$ transponieren, also

$f'(x)$ ebenfalls als Spalten-Vektor in $\mathbb{R}^n$ auffassen, dann nennen wir diesen den Gradienten von f bei x und schreiben $\operatorname{grad} f(x)$ dafür. Es gilt somit

$$f'(x)(v) = \Big(\partial_1 f(x), \dots, \partial_n f(x)\Big) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^n \partial_i f(x) \cdot v_i = \Big\langle \operatorname{grad} f(x) \Bigm| v \Big\rangle.$$

Leider ist das Transponieren von Matrizen, und insbesonders der Isomorphismus $E \cong L(E,\mathbb{R})$ der dadurch gegeben ist, daß man eine Basis in E wählt und diese auf die duale Basis in $L(E,\mathbb{R})$ abbildet, wesentlich von der Basis abhängig. Die Gleichung $f'(x)(v) = \langle \operatorname{grad} f(x) | v\rangle$ hingegen benutzt keine Koordinaten (von v und von $\operatorname{grad} f(x)$) mehr, macht also als Definition von $\operatorname{grad} f(x)$ auf Vektorräumen ohne vorgegebene Basis, nur versehen mit einem skalaren Produkt, Sinn.

6.4.6 Lemma. Geometrische Bedeutung des Gradienten.
Es sei E ein Hilbert-Raum und $f : E \supseteq U \to \mathbb{R}$ differenzierbar $x \in U$ mit $f'(x) \neq 0$. Dann ist die Richtung $v \in E$ mit $\|v\| = 1$ für welche $d_v f(x)$ maximal ist durch

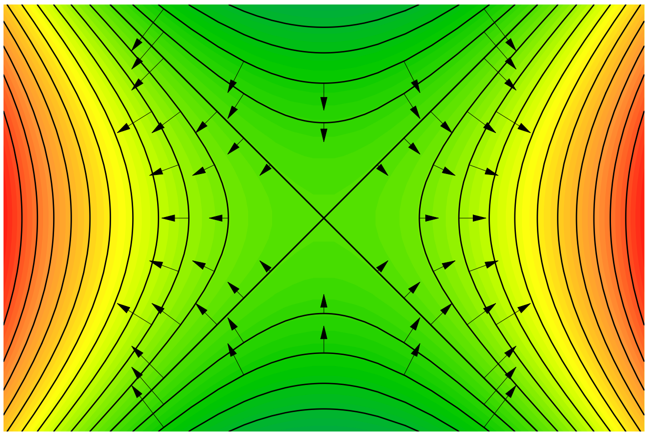
$$\operatorname{grad} f(x)/\|\operatorname{grad} f(x)\|$$

gegeben, wobei der Gradient von *f* bei *x definiert ist durch $\langle \operatorname{grad} f(x), v\rangle := f'(x)(v)$ vermöge dem Isomorphismus von E mit E^* welcher durch $w \mapsto \langle w|\_\rangle$ gegeben ist. Es zeigt also $\operatorname{grad} f(x) \in E$ in die Richtung des größten Anstiegs von f.*

Jede Höhenlinie *von f, d.h. differenzierbare Kurve $c : \mathbb{R} \supseteq I \to U$ mit konstanten $f \circ c$, steht in jedem Punkt $c(t)$ normal auf $\operatorname{grad} f(c(t))$.*

Für $E = \mathbb{R}^2$ ist die zu einem Vektor $v = (v_1, v_2)$ gegebene normale Richtung bis auf Vielfache durch $v^\perp := (-v_2, v_1)$ gegeben, also erhalten wir eine Parametrisierung c der Höhenschichtlinien als Lösung der Differentialgleichung

$$c'(t) = \Big(\operatorname{grad} f(c(t))\Big)^\perp.$$



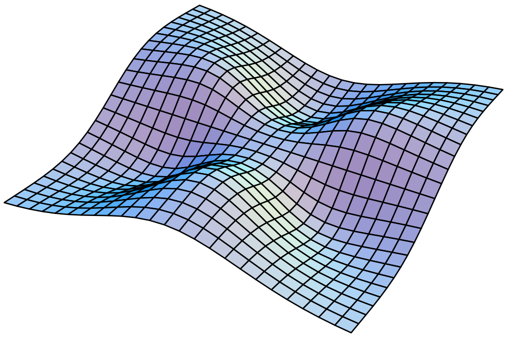
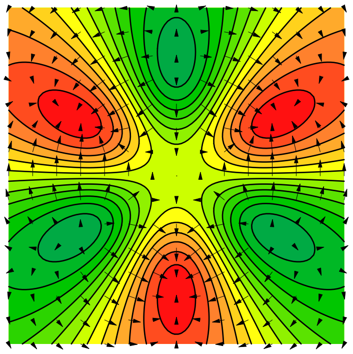
Beweis. In der Tat ist wegen der Cauchy-Schwarz Ungleichung (2.2.3) (die auch in Hilberträumen gilt, siehe [17, 6.6])

$$d_v f(x) = f'(x)(v) = \langle \operatorname{grad} f(x)|v\rangle = \|\operatorname{grad} f(x)\| \, \|v\| \, \cos(\measuredangle(\operatorname{grad} f(x), v)).$$

Dies wird also maximal, wenn $\cos(\measuredangle(\operatorname{grad} f(x), v)) = 1$ ist, d.h. $\measuredangle(\operatorname{grad} f(x), v) = 0$ ist, also v in die selbe Richtung wie $\operatorname{grad} f(x)$ zeigt.

Der zweite Teil der Aussage folgt nun so: Da $f \circ c$ konstant ist folgt durch Differenzieren $0 = (f \circ c)'(t) = f'\big(c(t)\big)\big(c'(t)\big) = \big\langle \operatorname{grad} f(c(t)) \bigm| c'(t)\big\rangle$, also ist $c'(t) \perp \operatorname{grad} f(c(t))$. $\square$

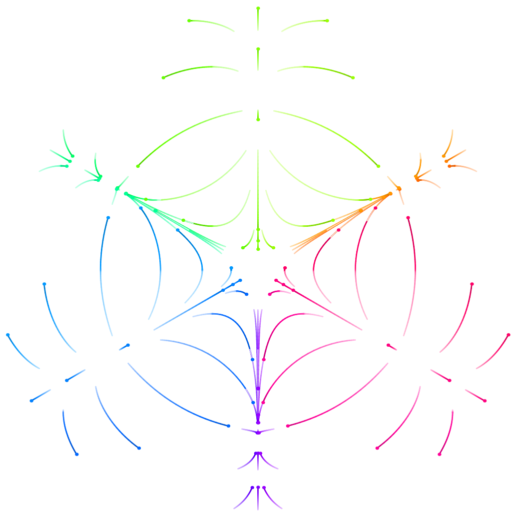
Bemerkung. Gradientenfeld.
Es sei eine C^1 Funktion $f : E \supseteq U \to \mathbb{R}$ vorgegeben. Dann können wir das Vektorfeld $\operatorname{grad} f : E \supseteq U \to E$ betrachten, das sogenannte GRADIENTENFELD von f.

Die Lösungen der zugehörigen gewöhnlichen Differentialgleichung

$$x'(t) = \operatorname{grad} f\,(x(t))$$

beschreiben dann die Wege, die Punkte nehmen, die denn stärksten Anstieg folgen. Wenn man f als Luftdruck interpretiert, so beschreiben die Lösungen dieser Differentialgleichung die (abgesehen von der Corioliskraft) durch die Druckverteilung f entstehenden Winde.



6.4.7 Morse-Palais-Lemma. In krit. Punkten beschreibt die 2.Ableitung die Funktion.
Es sei $f : E \to \mathbb{R}$ C^{p+2} mit $p \geq 1$, E ein Hilbert-Raum. Wir nehmen an, daß $f(0) = 0$ und 0 ein NICHT-DEGENERIERTER KRITISCHER PUNKT *ist, d.h. $f'(0) = 0$ und $f''(0)^\vee : E \to E^*$ ist ein Isomorphismus. Dann existiert ein C^p-Diffeomorphismus ψ, sodaß für x nahe 0 folgende Darstellung von f gilt:*

$$f(x) = \tfrac{1}{2} f''(0)\big(\psi(x), \psi(x)\big).$$

Dieses Lemma besagt also, daß bei nicht degenerierten kritischen Punkten eine Funktion im wesentlichen (d.h. bis auf ψ) wie ihre 2-te Ableitung aussieht.

Beweis. Wir haben

$$f(x) = f(0) + f'(0)\cdot x + \left(\int_0^1 (1-t) f''(tx)\,dt\right)(x,x) =: 0 + \rho(x)(x,x).$$

Beachte, daß $\rho : E \to L_{\mathrm{sym}}(E,E;\mathbb{R}) := \{f \in L(E,E;\mathbb{R}) : f \text{ ist symmetrisch}\}$ C^p ist. Indem wir der Isomorphismus

$$L_{\mathrm{sym}}(E,E;\mathbb{R}) \cong L_{\mathrm{sym}}(E,E) := \{a \in L(E,E) : a^* = a\}$$

benutzen, erhalten wir eine C^p-Abbildung $R : E \to L_{sym}(E,E)$, $\langle R(x)v|w\rangle := \rho(x)(v,w)$. Wir haben, daß $R(0)$ invertierbar ist, da $f''(0)^\vee$ invertierbar vorausgesetzt ist. Daher ist auch $R(x)$ invertierbar für kleine x. Wir machen den Ansatz $\psi(x) := C(x)\cdot x$, wo $C : E \to GL(E) := \{A \in L(E,E) : A \text{ ist invertierbar}\}$, d.h. wir wollen

$$\begin{aligned}\langle R(x)x|x\rangle = \rho(x)(x,x) = f(x) &\overset{z.z.}{=} \frac{f''(0)}{2}\big(\psi(x),\psi(x)\big) = \rho(0)\big(\psi(x),\psi(x)\big)\\ &= \Big\langle R(0)C(x)x \;\Big|\; C(x)x\Big\rangle = \Big\langle C(x)^*R(0)C(x)x \;\Big|\; x\Big\rangle.\end{aligned}$$

Dafür genügt

$$C(x)^*R(0)C(x) = R(x)$$

zu zeigen. Es sei $B(x) := R(0)^{-1}R(x)$. Da $\sqrt{1+t} = \sum_{k=0}^\infty \binom{1/2}{k} t^k$ für $|t|<1$ gilt, ist $C(x) := \sqrt{B(x)}$ wohldefiniert und C^p nach (6.3.13). Da $R(0)$ und $R(x)$ selbstadjungiert (=symmetrisch) sind erhalten wir

$$B(x)^* = R(x)R(0)^{-1}$$

und folglich

$$B(x)^*R(0) = R(x) = R(0)B(x).$$

Da $C(x)$ eine Potenzreihe in $1-B(x)$ ist, erhalten wir aus der Stetigkeit von $(\_)^*$, daß $C(x)^*$ die selbe Potenzreihe in $1-B(x)^*$ ist. Die letzte Gleichung gilt auch für irgendeine Potenz von $B(x)$ anstelle von $B(x)$ und daher für irgendeine Polynom in $1-B(x)$ oder sogar für Potenzreihen in $1-B(x)$, insbesonders also für $C(x)$, d.h.

$$C(x)^*R(0)C(x) = R(0)C(x)^2 = R(0)B(x) = R(x).$$

Aus

$$\psi'(x)(v) = C'(x)(v)\cdot x + C(x)\cdot v$$

erhalten wir für $x=0$ unter Benützung von $C(0)=1$, mittels dem Satz (6.2.2) über inverse Funktionen, daß ψ ein lokaler Diffeomorphismus ist. □

Folgerung. Zerlegung in positiv- und negativ-definiten Teil.
Es sei $f : E \to \mathbb{R}$ C^{p+2} mit $p \geq 1$ so, daß 0 ein nicht-degenerierter kritischer Punkt ist. Dann haben wir bis auf einen lokalen C^p-Diffeomorphismus eine orthogonale Zerlegung von $E = E_+ \oplus E_-$, so daß

$$f(x_+ + x_-) = \langle x_+|x_+\rangle - \langle x_-|x_-\rangle = \|x_+\|^2 - \|x_-\|^2.$$

Dies zeigt, daß die Beispiele in (6.4.4) typisch waren.

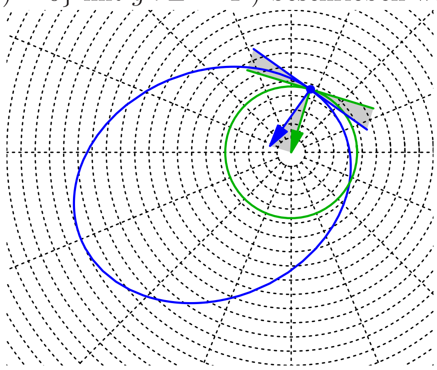
Beweis. Für endlich dimensionale Räume existiert nach dem Trägheitssatz von Silvester aus der linearen Algebra eine Orthonormalbasis bestehend aus Eigenvektoren zu Eigenwerten $\lambda \in \mathbb{R}$ für die symmetrische lineare Abbildung b, welche durch $\langle b(v)|w\rangle := f''(0)(v,w)/2$ gegeben ist. Da 0 nicht-degenerierter kritischer Punkt ist, ist 0 kein Eigenwert und wenn wir die Basis Vektoren durch die Wurzel aus dem Absolutbetrag des zugehörigen Eigenvektoren dividieren, so erhalten wir für die resultierenden Eigenvektoren v somit $f''(0)(v,v)/2 = \langle b(v)|v\rangle = \lambda\|v\|^2 = \lambda/\sqrt{|\lambda|}^2 = \mathrm{sgn}(\lambda) \in \{-1,1\}$ und damit die gewünschte Zerlegung.

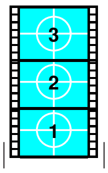
Für unendlich-dimensionale Räume können wir wie folgt vorgehen: Nach [17, 10.13] können wir $A = f''(0)$ in seinem positiven Teil A_+ und seinem negativen Teil A_- zerlegen, welche beide positiv definit sind nach [17, 11.19]. Dazu beachte, daß $0 \notin \sigma(A) \subseteq \mathbb{R}$, und daher $0 \notin \sigma(A_\pm) := \sigma(A) \cap \mathbb{R}_\pm$. □

6.4.8 Bemerkung. Extremalpunkte unter Nebenbedingung.
Sehr oft wird die Menge X in der wir eine Extremalstelle von f suchen nicht unbedingt eine offene

Teilmenge eines Banach-Raumes E sein. Falls wir X lokal durch eine Abbildung $h : H \to X$ parametrisieren können, d.h. zu vorgegebenen $x_0 \in X$ eine offenen Menge $V \subseteq E$ mit $x_0 \in V$ und offene Menge U eines Banach-Raumes H mit $0 \in U$ sowie eine surjektive Abbildung $h : U \to V \cap X$ finden können, dann ist x_0 genau dann ein lokales Extremum von $f|_X : X \to \mathbb{R}$ in X, wenn $h_0 \in h^{-1}(x_0)$ ein solches von $f \circ h : H \supseteq U \to V \cap X \to \mathbb{R}$ ist. Somit ist in dieser Situation das Problem wieder auf Funktionen zurückgespielt, die auf offenen Teilmengen von Banach-Räumen definiert sind.

Bisweilen werden wir aber eine Parametrisierung nicht vorgegeben haben, man denke z.B. an den Kreis bzw. die Sphäre $X := \{x : \|x\| = 1\}$. Wir wollen trotzdem versuchen die lokalen Extrema von $f : X \to \mathbb{R}$ zu lokalisieren und zwar ohne eine Parametrisierung von X explizit benutzen zu müssen. Die Idee dabei ist bereits für eine Kurve $X \subseteq \mathbb{R}^2$ zu erkennen. Ein Punkt $x_0 \in X$ ist höchstens dann extremal für eine Funktion f, wenn die Kurve die entsprechende Höhenschichtlinie $\{x : f(x) = f(x_0)\}$ im Punkt x_0 berührt. Die Tangente an die Höhenschichtlinie steht normal auf den Gradienten $\operatorname{grad} f\,(x_0)$. Wenn also X durch eine Gleichung $X := \{x : g(x) = 0\}$ gegeben ist, dann muß ebenfalls die Tangente an X in x_0 normal auf den Gradienten $\operatorname{grad} g(x_0)$ stehen, und somit $\operatorname{grad} f\,(x)^\perp = \operatorname{grad} g(x)^\perp$ gelten, d.h. $\operatorname{grad} f\,(x_0)$ proportional zu $\operatorname{grad} g(x_0)$ sein, d.h. ein $\lambda \in \mathbb{R}$ existieren mit $f'(x_0) = \lambda \cdot g'(x_0)$. Einen entsprechenden allgemeineren Satz für X die durch mehrere Gleichungen (also eine Vektor-wertige Gleichung $X = \{x : g(x) = 0\}$ mit $g : E \to F$) beschrieben werden ist der folgende:





6.4.9 Methode der Lagrange Multiplikatoren für krit. Punkte unter Nebenbedingung. *Es sei $f : E \supseteq U \to \mathbb{R}$ C^1 und E ein Hilbert-Raum. Weiters sei $g : E \supseteq U \to F$ C^1. Falls $x \in U$ ein lokales Extremum von f unter der* NEBENBEDINGUNG *$g(x) = 0$ ist und falls $g'(x)$ surjektiv ist, so ist $f'(x) = \lambda \circ g'(x)$ für ein $\lambda \in L(F, \mathbb{R})$ (dieses heißt* LAGRANGE-MULTIPLIKATOR*).*

Beweis. O.B.d.A. sei $x = 0$. Wir wollen die 0-Stellenmenge $g^{-1}(0)$ von g parametrisieren. Um dazu den Satz (6.2.4) über implizite Funktionen anwenden zu können benötigen wir eine (möglichst gut an das Problem angepaßte) Zerlegung $E \cong E_1 \times E_2$. Als E_1 bietet sich der Kern von $g'(0)$ als vermeintlicher Tangentialraum an $g^{-1}(0)$ an. Es sei E_2 ein Komplementärraum zu E_1 also z.B. das orthogonale Komplement $E_2 := E_1^\perp$. Dann ist $E_1 \times E_2 \cong E$ vermöge $(x_1, x_2) \mapsto x_1 + x_2$ und wir bezeichnen mit $\tilde g : E_1 \times E_2 \to F$ die Abbildung $\tilde g(x_1, x_2) := g(x_1 + x_2)$ und analog $\tilde f(x_1, x_2) := f(x_1 + x_2)$.

Offensichtlich ist $(x_1, x_2) = (0, 0)$ ein lokales Extremum von $\tilde f$ unter der Nebenbedingung $\tilde g(x_1, x_2) = 0$. Es ist $\partial_2 \tilde g(0, 0)$ injektiv, denn $\partial_2 \tilde g(0, 0) = g'(0)|_{E_2} \in L(E_2, F)$ und aus $0 = \partial_2 \tilde g(0, 0)(v_2) = g'(0)(v_2)$ folgt $v_2 \in E_1 \cap E_2 = \{0\}$. Weiters ist $\partial_2 \tilde g(0, 0)$ surjektiv, denn da $g'(0)$ surjektiv ist, existiert zu $w \in F$ ein $v \in E$ mit $w = g'(0)(v)$. Sei $v = v_1 + v_2$ mit $v_i \in E_i$, dann ist $w = g'(0)(v) = g'(0)(v_1) + g'(0)(v_2) = 0 + \partial_2 \tilde g(0, 0)(v_2)$. Nach dem Satz (6.2.4) über implizite Funktionen existiert somit lokal eine C^1-Abbildung $h : E_1 \to E_2$ mit $h(0) = 0$ und $\tilde g(x_1, h(x_1)) = 0$ für alle $x_1 \in E_1$ nahe 0 ($g^{-1}(0)$ läßt sich also lokal durch den Graphen der Abbildung $h : E_1 \to E_2$ parametrisieren).

Somit erhalten wir

$$
\begin{aligned}
0 = \tilde g \circ (\mathrm{id}, h) \quad &\Rightarrow \quad 0 = (\tilde g \circ (\mathrm{id}, h))'(0) = \partial_1 \tilde g(0, 0) + \partial_2 \tilde g(0, 0) \circ h'(0) \\
&\Rightarrow \quad h'(0) = -\partial_2 \tilde g(0, 0)^{-1} \circ \partial_1 \tilde g(0, 0)
\end{aligned}
$$

Da 0 ein lokales Extremum von $\tilde f \circ (\mathrm{id}, h)$ ist nach (6.4.2)

$$\begin{aligned} & 0 = (\tilde f \circ (\mathrm{id}, h))'(0) = \partial_1 \tilde f(0,0) + \partial_2 \tilde f(0,0) \circ h'(0) \\ \Rightarrow \quad & \partial_1 \tilde f(0,0) = -\partial_2 \tilde f(0,0) \circ h'(0) = \underbrace{\partial_2 \tilde f(0,0) \circ \partial_2 \tilde g(0,0)^{-1}}_{=:\lambda \in L(F,\mathbb{R})} \circ \partial_1 \tilde g(0,0) \end{aligned}$$

Also ist

$$\begin{aligned} & \partial_1 \tilde f(0,0) = \lambda \circ \partial_1 \tilde g(0,0) \qquad \text{und} \\ & \partial_2 \tilde f(0,0) = \partial_2 \tilde f(0,0) \circ \partial_2 \tilde g(0,0)^{-1} \circ \partial_2 \tilde g(0,0) = \lambda \circ \partial_2 \tilde g(0,0) \end{aligned}$$

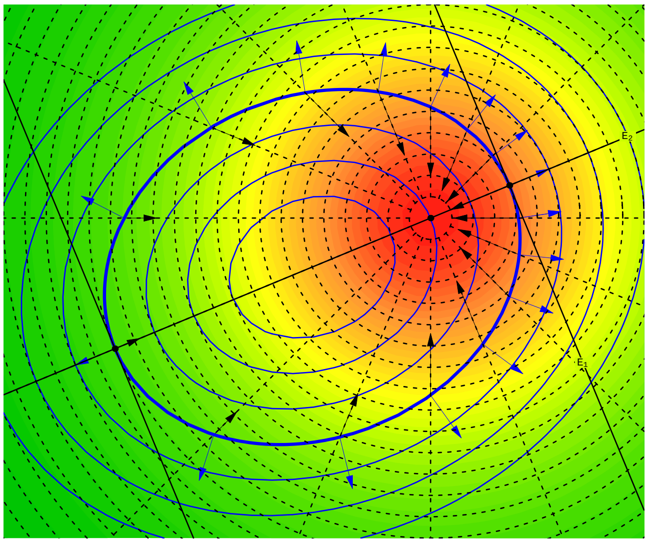
Insgesamt also

$$\tilde f'(0,0) = \lambda \circ \tilde g'(0,0)$$

und damit

$$f'(0) = \lambda \circ g'(0),$$

da $(x_1, x_2) \mapsto x_1 + x_2$ ein linearer Isomorphismus $E_1 \times E_2 \to E$ ist der f und g in $\tilde f$ und $\tilde g$ übersetzt. $\square$



Beachte, daß dieser Beweis auch funktioniert, wenn E ein Banach-Raum ist und $\ker g'(0)$ komplementiert in E ist.
Wenn $F = \mathbb{R}^n$ ist und somit $g = (g_1, \dots, g_n)$, dann ist λ durch $(\lambda_i)_i \in \mathbb{R}^n$ gegeben und somit muß ein relatives lokales Extremum x folgendes erfüllen

$$f'(x) = \sum_{i=1}^n \lambda_i \, g_i'(x) \text{ und } g(x) = 0.$$

Ist zusätzlich $E = \mathbb{R}^m$ so bedeutet dies folgendes Gleichungssystem von $n+m$ Gleichungen in den $n+m$ vielen Variablen $(x, \lambda) = (x_1, \dots, x_m; \lambda_1, \dots, \lambda_n)$:

$$\begin{aligned} \partial_j f(x) &= \sum_{i=1}^n \lambda_i \, \partial_j g_i(x) \quad \text{für } j = 1, \dots, m, \\ g_i(x) &= 0 \quad \text{für } i = 1, \dots, n. \end{aligned}$$

Da der letzte Satz nur eine notwendige Bedingung liefert, benötigt man zusätzliche Argumente um den Nachweis eines lokalen Extremums zu liefern. Das kann z.B. sein, daß die Nebenbedingung eine kompakte Menge beschreibt und die stetige Funktion f somit ein Maximum und ein Minimum besitzen muß. Einen anderen Existenznachweis hat man in folgender Situation:

6.4.10 Lemma. Distanz zweier Mengen.
Es seien A eine abgeschlossene und K eine kompakte Teilmenge eines $\mathbb{R}^n$. Dann ist die Distanz $d(A,K) := \inf\{\|a-x\| : a \in A, x \in K\}$ ein Minimum, d.h. es existieren $a_0 \in A$ und $x_0 \in K$ mit $\|a_0 - x_0\| \le \|a - x\|$ für alle $a \in A$ und $x \in K$.

Beweis. Die Abbildung $d_A : K \to \mathbb{R}$, $x \mapsto d(A,x) := \inf\{\|a-x\| : a \in A\}$ ist stetig, da $d(A,x) - d(A,x') \le \|x - x'\|$, und besitzt somit auf der kompakten Menge K ein Minimum, d.h. ein $x_0 \in K$ mit $d(A,x_0) = \inf\{d(A,x) : x \in K\} = d(A,K)$. Es sei $r > d(A,x_0)$ und $A_r := A \cap \{a : \|a - x_0\| \le r\}$, dann ist $d(A,x_0) = d(A_r,x_0)$ und A_r beschränkt und abgeschlossen also kompakt. Folglich besitzt auch die stetige Funktion $a \mapsto \|a - x_0\|$ ein Minimum auf A_r, d.h. es existiert ein $a_0 \in A_r \subseteq A$ mit $\|a_0 - x_0\| = d(A_r, x_0) = d(A,x_0) = d(A,K)$. □

Bemerkung.
Falls $K \subseteq E$ kompakt und $A \subseteq E$ abgeschlossen in einem Banach-Raum E sind, so muß dies nicht mehr gelten. Sei z.B. $K = \{0\}$ und $A := \{x \in \ell^2 : \|(\lambda_k x^k)_k\| = 1\}$ für eine Folge $\lambda_k \nearrow 1$ für $k \to \infty$. Dann ist $d(A,K) = \inf\{\|x\| : \|(\lambda_k x^k)_k\| = 1\} = 1$, da $\frac{1}{\lambda_k} e_k \in A$ und $1 = \|(\lambda_k x^k)_k\|^2 = \sum_k (\lambda_k x^k)^2 < \sum_k (x^k)^2 = \|x\|^2$. Insbesonders existiert kein $a \in A$ mit $\|a\| = 1$.

6.4.11 Beispiele von Extremalwerten unter Nebenbedingungen.

1. Gesucht sei jener Punkt der Ebene $z = x + y$ der von $(1,0,0)$ minimalen Abstand hat. Hier ist $f(x,y,z) := \sqrt{(x-1)^2 + y^2 + z^2}$ oder besser $f(x,y,z) := (x-1)^2 + y^2 + z^2$ und $g(x,y,z) := x + y - z$. Es ist somit folgendes Gleichungssystem in $(x,y,z;\lambda)$ zu lösen:

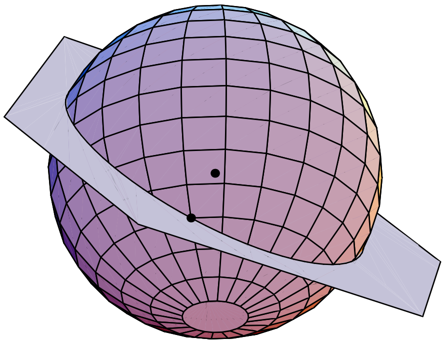
$$\begin{aligned} 2(x-1) &= \partial_1 f(x,y,z) = \lambda\, \partial_1 g(x,y,z) = \lambda \\ 2y &= \partial_2 f(x,y,z) = \lambda\, \partial_2 g(x,y,z) = \lambda \\ 2z &= \partial_3 f(x,y,z) = \lambda\, \partial_3 g(x,y,z) = -\lambda \\ x + y - z &= g(x,y,z) = 0 \end{aligned}$$

   Die einzige Lösung ist $\frac{\lambda}{2} = x - 1 = y = -z = -\frac{1}{3}$ und somit das Minimum bei $(\frac{2}{3}, -\frac{1}{3}, \frac{1}{3})$. Dies kann auch durch normal-Projektion von $(1,0,0)$ auf die Ebene erhalten werden.

2. Gesucht sind die Extremalwerte von $f(x,y,z) := 5x + y - 3z$ auf dem Durchschnitt der Ebene $x + y + z = 0$ mit der Sphäre $x^2 + y^2 + z^2 = 1$. Die Abbildung g ist in diesem Beispiel $g = (g_1, g_2) : \mathbb{R}^3 \to \mathbb{R}^2$ mit $g_1(x,y,z) := x + y + z$ und $g_2(x,y,z) := x^2 + y^2 + z^2 - 1$. Es ist somit folgendes nicht-lineare Gleichungssystem in $(x,y,z;\lambda_1,\lambda_2)$ zu lösen:

$$\begin{aligned} 5 &= \partial_1 f(x,y,z) = \lambda_1\, \partial_1 g_1(x,y,z) + \lambda_2\, \partial_1 g_2(x,y,z) = \lambda_1\, 1 + \lambda_2\, 2x \\ 1 &= \partial_2 f(x,y,z) = \lambda_1\, \partial_2 g_1(x,y,z) + \lambda_2\, \partial_2 g_2(x,y,z) = \lambda_1\, 1 + \lambda_2\, 2y \\ -3 &= \partial_3 f(x,y,z) = \lambda_1\, \partial_3 g_1(x,y,z) + \lambda_2\, \partial_3 g_2(x,y,z) = \lambda_1\, 1 + \lambda_2\, 2z \\ 0 &= g_1(x,y,z) = x + y + z \\ 0 &= g_2(x,y,z) = x^2 + y^2 + z^2 - 1 \end{aligned}$$

   Die einzigen zwei Lösungen sind $(\pm\frac{1}{\sqrt{2}}, 0, \mp\frac{1}{\sqrt{2}}; 1, \pm 2\sqrt{2})$. Einsetzen in f zeigt, daß $(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}})$ die Stelle eines lokalen (und somit auch globalen) Maximums und $(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$ die eines globalen Minimums ist. Genauer, da die Sphäre und somit auch der Schnitt mit der Ebene kompakt ist und f stetig ist, muß nach (3.3.6) ein Maximum und ein Minimum von f unter dieser Nebenbedingung existieren. Nach (6.4.9) sind die einzigen möglichen Punkte dafür aber die gerade berechneten, und da f auf dem ersten Punkt einen größeren Wert als am zweiten hat, muß der erste die Stelle des Maximums und der zweite jene des Miniums sein.

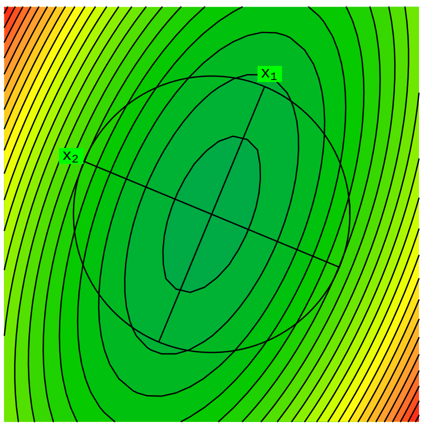


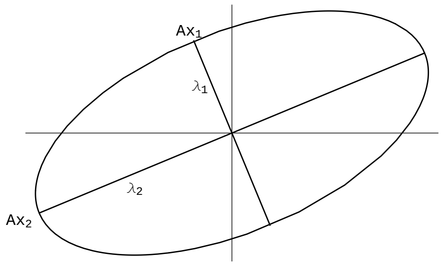
3. Sei $A \in L(E,E)$ gegeben. Gesucht seien die Extremalwerte von $\|Ax\|$ (oder äquivalent von $\|Ax\|^2 = \langle Ax|Ax\rangle$) unter der Nebenbedingung $\|x\| = 1$ (äquivalent von $\|x\|^2 = \langle x|x\rangle$). Die Methode der Lagrange Multiplikatoren liefert:

$$2\langle A^t Ax|v\rangle = \langle Ax|Av\rangle + \langle Av|Ax\rangle = \lambda(\langle x|v\rangle + \langle v|x\rangle) = 2\,\langle \lambda\, x|v\rangle \qquad \forall v$$
$$\langle x|x\rangle = 1$$

Somit ist $A^t Ax = \lambda\, x$ und $\|x\| = 1$ und folglich x ein normierter Eigenvektor von $A^t A$ zum Eigenwert λ. Ist $\dim(E) = 2$, dann liefert die beiden Eigenwerte λ das Maximum und das Minimum von $\|Ax\|^2$ und die Bilder $A(v)$ der zugehörigen normierten Eigenvektoren v sind gerade die HALBACHSEN der durch $\{A(x) : \|x\| = 1\}$ beschriebenen Ellipse. Die Längen der Halbachsen sind $\|A(x)\| = \sqrt{\langle A(x)|A(x)\rangle} = \sqrt{\lambda\langle x|x\rangle} = \sqrt{\lambda}\|x\| = \sqrt{\lambda}$.

Andererseits können wir auch die QUADRIK $X := \{x : \|Ax\| = 1\}$ untersuchen. Die Punkte $x \in X$ mit extremalen Abstand $\|x\|$ zu 0 sind wie folgt zu erhalten. Wir setzen $v := \frac{1}{\|x\|}x$ und erhalten $\|v\| = 1$ und $\|A(v)\| = \frac{\|Ax\|}{\|x\|} = \frac{1}{\|x\|}$. Wir suchen also unter allen v mit $\|v\| = 1$ jene, für welche $\|A(v)\|$ extremal wird, dies sind nach obigen genau die normierten Eigenvektoren v von $A^t A$ und die zugehörigen Eigenwerte sind $\lambda = \frac{1}{|x|^2}$. Die Punkte $x \in X$ mit extremalen Abstand zu 0 sind also gerade $\frac{1}{\sqrt{\lambda}}v$ mit Norm $\frac{1}{\sqrt{\lambda}}$. Die Eigenvektoren v beschreiben also die Richtungen der HAUPTACHSEN der Quadrik (d.h. im $\mathbb{R}^2$ der Ellipse oder Hyperbel) X.





6.5 Stammfunktionen & Kurven-Integrale

6.5.1 Totale Differentialgleichungen

Wir versuchen nun den Hauptsatz (5.6.17) der Analysis für Kurven auf Abbildungen $g : E \supseteq U \to F$ zu verallgemeinern. Für Kurven besagte dieser grob gesprochen, daß Integrieren beinahe invers zum Ableiten $g \mapsto g'$ von Kurven $g : \mathbb{R} \supseteq [a,b] \to F$ ist. Wir wollen also allgemeiner eine Umkehrung zum Differenzieren $(g : E \supseteq U \to F) \mapsto (f := g' : E \supseteq U \to L(E,F))$ finden, d.h. zu gegeben f eine sogenannte Stammfunktion g finden, welche $g'(x) = f(x)$ für alle x erfüllt. Wir dürfen hier aber nicht erwarten, daß Stetigkeitsbedingungen an f ausreichen um wie im 1-dimensionalen die Existenz einer Stammfunktion zu gewährleisten, denn selbst wenn f C^1 ist, also die Stammfunktion C^2 wäre, dann muß $(g')' = f'$ symmetrisch sein, also für f die algebraische INTEGRABILITÄTSBEDINGUNG

$$f'(x)(v)(w) = f'(x)(w)(v) \text{ für alle } x \in U \text{ und } v, w \in E$$

erfüllt sein um eine Chance auf eine Stammfunktion zu besitzen.

Für 1-dimensionales E haben wir das Problem der Stammfunktion in (6.2) verallgemeinert und gewöhnliche Differentialgleichungen $g'(t) = f(t, g(t))$ betrachtet. Für mehrdimensionales E werden wir gleich diese allgemeinere Situation betrachten. Da nun $g'(x) \in L(E,F)$ liegt, muß die die Differentialgleichung definierende Funktion eine Abbildung $f : E \times F \to L(E,F)$ sein.

Satz von Frobenius für totale Differentialgleichungen.
Es sei $f : E \times F \to L(E,F)$ C^p für irgendein $1 \le p \le \infty$. Dann existiert lokal eine (eindeutig bestimmte) differenzierbare Lösung $g_{x_0,y_0} : E \to F$ der TOTALEN DIFFERENTIALGLEICHUNG

$$g'_{x_0,y_0}(x) = f(x, g_{x_0,y_0}(x)) \text{ mit } g_{x_0,y_0}(x_0) = y_0$$

dann und nur dann, wenn die INTEGRABILITÄTSBEDINGUNG*, d.h.daß $f'(z)(v_1, f(z)v_1)v_2$ symmetrisch in v_1, v_2 für alle $z \in E \times F$ ist, erfüllt ist.*
Weiters ist die Abbildung $(x_0, y_0, x) \mapsto g_{x_0,y_0}(x)$ C^p.

Bemerkung. Totale Differentialgleichung im endlich-dimensionalen.
Für $E := \mathbb{R}^n$ sei $f_i(z) := f(z)(e_i)$, dann ist die totale Differentialgleichung $g'(x) = f(x, g(x))$ gleichbedeutend mit folgendem SYSTEM PARTIELLER DIFFERENTIALGLEICHUNGEN

$$\partial_i g(x) = g'(x)(e_i) = f(x, g(x))(e_i) = f_i(x, g(x)) \text{ für } i \in \{1, \dots, n\}.$$

Beweis.

($\Rightarrow$) Für $z_0 = (x_0, y_0) \in E \times F$ sei g eine lokale Lösung obiger Differentialgleichung mit Anfangsbedingung $g(x_0) = y_0$. Dann ist $g' = f \circ (\mathrm{id}, g)$ stetig und nach der Kettenregel (6.3.3.2) ist g C^2 mit

$$\begin{aligned} g''(x_0)(v_1, v_2) &= (g')'(x_0)(v_1)(v_2) = \mathrm{ev}_{v_2}\Big((g')'(x_0)(v_1)\Big) = \mathrm{ev}_{v_2}\Big((f \circ (\mathrm{id}, g))'(x_0)(v_1)\Big) \\ &= \mathrm{ev}_{v_2}\bigg(f'(x_0, g(x_0))\Big(v_1, g'(x_0)(v_1)\Big)\bigg) = f'(z_0)\Big(v_1, f(z_0)(v_1)\Big)(v_2) \end{aligned}$$

Der letzte Ausdruck ist also symmetrisch in v_1 und v_2, weil $g''(x_0)$ es nach dem Satz (6.3.10) von Schwarz ist.

($\Leftarrow$) Es sei $(x_0, y_0) \in E \times F$ fix. Wir versuchen die totale Differentialgleichung auf eine gewöhnliche zurückzuführen indem wir zuerst nur untersuchen was bei x_0 in Richtung $v \in E$ passiert.
Dazu nehmen vorerst an, daß eine lokale Lösung g der totalen Differentialgleichung mit Anfangswert $g(x_0) = y_0$ existiert und setzen $\bar{g}(t, v) := g(x_0 + tv)$. Dann gilt

$$\begin{aligned} \frac{\partial}{\partial t}\bar{g}(t, v) &= g'(x_0 + tv) \cdot v = f(x_0 + tv, g(x_0 + tv)) \cdot v = f(x_0 + tv, \bar{g}(t, v)) \cdot v, \\ \bar{g}(0, v) &= g(x_0) = y_0. \end{aligned}$$

Dies ist eine gewöhnliche Differentialgleichung für $\bar g$ welche nach (6.2.8) lokal (d.h. für $|t| \le \varepsilon$, $\|v\| \le \varepsilon$ mit einem gewissen $\varepsilon > 0$) eine eindeutig bestimmte Lösung $\bar g$ besitzt, welche von (t, v, x_0, y_0) C^p abhängt. Daraus sollten wir eine Lösung g der totalen Differentialgleichung vermöge $g(x) := \bar g(t, v)$ erhalten, wenn wir $tv := x - x_0$ setzen. Naheliegend wäre $t = 1$ zu wählen, aber solange existiert womöglich die Lösung $\bar g$ nicht, darum wählen wir $t := \varepsilon$ und somit $v := \frac{x-x_0}{\varepsilon}$, also setzen wir $g(x) := \bar g(\varepsilon, \frac{x-x_0}{\varepsilon})$ für $\|x - x_0\| \le \varepsilon^2$ und müssen nun $g'(x)(w)$ und dazu insbesonders $\partial_2 \bar g$ berechnen. Die Idee dabei ist, daß

$$\begin{aligned}
\partial_2\bar g(t,v)(w) &= \frac{\partial}{\partial s}\Big|_{s=0} \bar g(t, v+sw) = \frac{\partial}{\partial s}\Big|_{s=0} g(x_0 + t(v+sw)) \\
&= g'(x_0+tv)(tw) = f(x_0+tv, g(x_0+tv))(tw) \\
&= f(x_0+tv, \bar g(t,v))(tw)
\end{aligned}$$

gelten sollte. Also definieren wir $k : \mathbb{R} \to F$ durch

$$k(t) := \partial_2\bar g(t,v)(w) - f(x_0+tv, \bar g(t,v))(tw).$$

Dann ist $k(0) = \partial_2\bar g(0,v)(w) - f(x_0+0v, \bar g(0,v))(0w) = 0$ und somit ergibt sich – wobei wir der Übersichtlichkeit halber nach Anwendung der Kettenregel (6.1.5) das Argument (t,v) bei $\bar g$ und bei deren Ableitungen sowie das Argument $(x_0+tv, \bar g(t,v))$ bei f und bei seinen Ableitungen weglassen –

$$\begin{aligned}
\frac{d}{dt}k(t) &= \frac{\partial}{\partial t}\Big(\partial_2\bar g(t,v)(w) - f(x_0+tv,\bar g(t,v))(tw)\Big) \\
&\overset{(6.3.10)}{=\!=} \partial_2 \underbrace{\frac{\partial}{\partial t}\bar g(t,v)}_{f(x_0+tv,\bar g(t,v))\cdot v}(w) - \Big(\partial_1 f\cdot v\cdot tw + \partial_2 f\cdot \underbrace{\frac{\partial}{\partial t}\bar g}_{f\cdot v}\cdot tw + f\cdot w\Big) \\
&\overset{(6.1.5)}{=\!=} \Big(\partial_1 f\cdot tw\cdot v + \partial_2 f\cdot(\partial_2\bar g\cdot w)\cdot v + f\cdot w\Big) - \Big(f'\cdot(v, f\cdot v)\cdot tw + f\cdot w\Big) \\
&\overset{\text{Int.Bed.}}{=\!=} \partial_1 f\cdot tw\cdot v + \partial_2 f\cdot(\partial_2\bar g\cdot w)\cdot v - f'\cdot(tw, f\cdot tw)\cdot v \\
&\overset{(6.1.10)}{=\!=} \partial_2 f\cdot(\partial_2\bar g\cdot w - f\cdot tw)\cdot v = \partial_2 f\cdot k(t)\cdot v.
\end{aligned}$$

Da letzteres eine lineare Differentialgleichung (mit nicht konstanten Koeffizienten) und $k(0) = 0$ ist, folgt $k = 0$. Folglich ist für $t := \varepsilon$ und $v := \frac{x-x_0}{\varepsilon}$

$$\begin{aligned}
g'(x)(w) &= \partial_2\bar g(\varepsilon, \tfrac{x-x_0}{\varepsilon})(\tfrac{1}{\varepsilon}w) = \partial_2\bar g(t,v)(\tfrac{1}{\varepsilon}w) = f(x_0+tv, \bar g(t,v))\,(t\tfrac{1}{\varepsilon}w) = f\Big(x, \bar g(\varepsilon, \tfrac{x-x_0}{\varepsilon})\Big)(w) \\
&= f(x, g(x))(w). \quad \square
\end{aligned}$$

Wir betrachten nun den Spezialfall, wo $f : E \times F \to L(E,F)$ nur vom ersten Faktor abhängt:

6.5.2 Folgerung. Lokale Stammfunktion.
Es sei $f : E \supseteq U \to L(E,F)$ C^1. Dann existiert genau dann lokal um jeden Punkt $x_0 \in U$ ein $g : E \to F$ mit $g(x_0) = 0$ und $g'(x) = f(x)$ für alle x nahe x_0, wenn die Integrabilitätsbedingung $f'(x)(v)(w) = f'(x)(w)(v)$ für alle $x \in U$ und $v, w \in E$ erfüllt ist.

Was uns an diesem Resultat noch stört, ist, daß wir die Stammfunktion nur lokal erhalten. Im 1-dimensionalen war die Stammfunktion am selben Intervall wie f definiert.

6.5.3 Definition. 1-Form.
Unter einer (F-WERTIGEN) 1-FORM (auf $U \subseteq E$) versteht man eine Abbildung $f : E \supseteq U \to L(E,F)$. Ausführlicher spricht man auch von einer DIFFERENTIALFORM DER ORDNUNG 1.

Sie heißt EXAKT, falls sie eine Stammfunktion auf ganz U besitzt, d.h. eine differenzierbare Abbildung $g : U \to F$, mit $g' = f$.

Eine differenzierbare 1-Form f heißt GESCHLOSSEN, wenn die Integrabilitätsbedingung

$$f'(x)(v)(w) = f'(x)(w)(v) \text{ für alle } x \in U \text{ und } v, w \in E$$

erfüllt ist.

Diese Namen kommen aus der dualen Situation, die man in der Homologietheorie (einen Teilgebiet der algebraischen Topologie) behandelt. Siehe [22].

Falls $F = \mathbb{R}$ ist und E ein Hilbert-Raum, dann haben wir den Isomorphismus $L(E, \mathbb{R}) =: E^* \cong E$ und können 1-Formen auch als Abbildungen $f : E \supseteq U \to E^* \cong E$ auffassen. SKALARWERTIGE 1-FORMEN sind also für Hilbert-Räume nur eine andere Schreibweise für Vektorfelder.

Wenn $g : E \supseteq U \to F$ differenzierbar ist, dann schreibt man auch dg für die exakte 1-Form die durch die Ableitung $g' : E \supseteq U \to L(E, F)$ gegeben ist, diese heißt in der klassischen Literatur öfters auch TOTALES DIFFERENTIAL von f. Für $E = \mathbb{R}^n$ und die lineare Koordinatenprojektion $\mathrm{pr}_i : x = (x_1, \dots, x_n) \mapsto x_i$ erhält man insbesonders die exakte 1-Form $d\,\mathrm{pr}_i$ für die man üblicherweise dx_i schreibt, und die an jeder Stelle durch $v = (v_1, \dots, v_n) \mapsto v_i$ gegeben ist.

Eine allgemeine 1-Form $f : E \supseteq U \to L(E, F)$ kann für $E = \mathbb{R}^n$ nun wie folgt geschrieben werden:

$$\begin{aligned} f(x)(v) &= f(x)\Big(\sum_{i=1}^n v_i\, e_i\Big) = \sum_{i=1}^n \underbrace{v_i}_{dx_i(x)(v)} \cdot \underbrace{f(x)(e_i)}_{=:f_i(x)} \\ &= \sum_{i=1}^n dx_i(x)(v) \cdot f_i(x) = \sum_{i=1}^n f_i(x) \cdot dx_i(x)(v) = \Big(\sum_{i=1}^n f_i(x) \cdot dx_i(x)\Big)(v), \end{aligned}$$

also

$$f(x) = \sum_{i=1}^n f_i(x) \cdot dx_i(x) = \Big(\sum_{i=1}^n f_i \cdot dx_i\Big)(x)$$

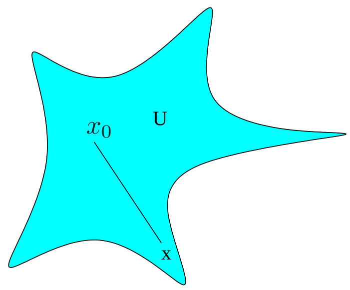
und schließlich $f = \sum_{i=1}^n f_i\, dx_i$, d.h. f wird vollständig durch die Komponenten $f_i : E \supseteq U \to F$ für $i = 1, \dots, n$ beschrieben.

Für $F = \mathbb{R}$ haben wir also so etwas wie eine Basis $dx_1, \dots, dx_n$ des Raums aller skalar-wertigen 1-Formen $\mathbb{R}^n \supseteq U \to L(\mathbb{R}^n, \mathbb{R})$ gefunden. Die Komponenten $f_1, \dots, f_n$ einer beliebigen 1-Form f sind allerdings nicht Zahlen sondern Funktionen. Eine Abel'sche Gruppe (wie z.B. die Menge der 1-Formen mit punktweiser Addition) zusammen mit einer Multiplikation mit Elementen eines Rings R (hier der Menge aller reell-wertigen Funktionen) welche die üblichen Vektorraum-Axiome erfüllt (abgesehen davon, daß R kein Körper zu sein braucht) heißt R-MODUL.

Ist $f = dg = g'$, dann ist $f_i(x) = g'(x)(e_i) = \partial_i g(x) = \frac{\partial g(x)}{\partial x_i}$, also gerade die i-te partielle Ableitung von g und

$$dg(x) = \sum_i \frac{\partial g(x)}{\partial x_i}\, dx_i, \quad \text{oder kurz} \quad dg = \sum_i \partial_i g \cdot dx_i.$$

6.5.4 Theorem. Exakte 1-Formen.
Es sei $f : E \supseteq U \to L(E, F)$ eine 1-Form und C^1.
Falls f exakt ist, so ist f geschlossen.
Ist U zusätzlich sternförmig und offen, dann gilt die Umkehrung.



Eine Menge $U \subseteq E$ heißt STERNFÖRMIG, wenn ein Punkt $x_0 \in U$ existiert, s.d. für alle $x \in U$ die Verbindungsstrecke $\overline{x_0 x} = \{x_0 + t(x - x_0) : 0 \le t \le 1\}$ in U enthalten ist.

Beweis. ($\Rightarrow$) folgt sofort aus dem Satz (6.3.10) von Schwarz.

(⇐) Der Satz (6.5.1) von Frobenius liefert uns nur eine lokale Stammfunktion. Der Grund dafür war, daß im Beweis $\bar g$ eine lokale Lösung einer gewöhnlichen Differentialgleichung war. Da im vorliegenden Fall f aber nicht vom Faktor F abhängt, ist die entsprechende Differentialgleichung:

$$\frac{\partial}{\partial t}\bar g(t,v) = f(x_0+tv)(v)$$

aus der $\bar g$ als

$$\bar g(t,v) = \bar g(0,v) + \int_0^1 f(x_0+tsv)(v)\,ds \qquad \text{für } t \text{ mit } \overline{x_0, x_0+tv} \subseteq U$$

berechnet werden kann. Da U sternförmig bzgl. x_0 ist, ist

$$g(x) = \bar g(1, x-x_0) = y_0 + \int_0^1 f(x_0+s(x-x_0))(x-x_0)\,ds.$$

die gesuchte Stammfunktion. □

Zur Übung kann man auch direkt unter Verwendung der Integrabilitätsbedingung nachrechnen, daß durch $g(x) := y_0 + \int_0^1 f(x_0+t(x-x_0))\,(x-x_0)\,dt$ wirklich eine Stammfunktion g von f gegeben ist:

$$\begin{aligned}
g(x) &= y_0 + \int_0^1 f(x_0+t(x-x_0))(x-x_0)\,dt \quad \Rightarrow\\
g(x+s\,v) &= y_0 + \int_0^1 f(x_0+t(x-x_0+s\,v))(x-x_0+s\,v)\,dt \quad \Rightarrow\\
g'(x)(v) &= \frac{\partial}{\partial s}\Big|_{s=0} g(x+s\,v) = \int_0^1 \frac{\partial}{\partial s}\Big|_{s=0} f(x_0+t(x-x_0+s\,v))\cdot(x-x_0+s\,v)\,dt\\
&= \int_0^1 f'(x_0+t(x-x_0))\cdot(t\,v)\cdot(x-x_0) + f(x_0+t(x-x_0))\cdot v\,dt\\
&\overset{\text{Int.Bed.}}{=\!=} \int_0^1 \Big(t\,f'(x_0+t(x-x_0))\cdot(x-x_0) + f(x_0+t(x-x_0))\Big)\cdot v\,dt\\
&= \int_0^1 \frac{\partial}{\partial t}\Big(t\,f(x_0+t(x-x_0))\Big)\cdot v\,dt = \int_0^1 \frac{\partial}{\partial t}\Big(t\,f(x_0+t(x-x_0))\Big)\,dt\cdot v\\
&= \Big[t\,f(x_0+t(x-x_0))\Big]_{t=0}^1 \cdot v = f(x)\cdot v.
\end{aligned}$$

Auch dieses Resultat ist nicht völlig befriedigend, denn die Sternförmigkeit von U scheint eine zu restriktive Bedingung zu sein. Es sollte doch genügen, wenn wir von $x_0 \in U$ zu einen beliebigen $x \in U$ zwar nicht längs einer Gerade in U gelangen können aber doch zumindest längs einer Kurve c in U. Die obige Formel der Stammfunktion g können wir mittels der Parametrisierung $c: t \mapsto x_0 + t(x-x_0)$ auch als

$$g(x) = y_0 + \int_0^1 f(\underbrace{x_0+t(x-x_0)}_{=:c(t)})\,\underbrace{(x-x_0)}_{c'(t)}\,dt = y_0 + \int_0^1 f(c(t))\cdot c'(t)\,dt$$

schreiben, und letzteres macht auch für stetiges f und stetig differenzierbare Kurven c einen Sinn.

6.5.5 Proposition. Stammfunktion via Kurvenintegral.
Es sei $g: E \supseteq U \to F$ C^1. Dann ist

$$\int_c g' := \int_a^b (g'\circ c)\cdot c' = g(c(b)) - g(c(a)) \text{ für alle } C^1\text{-Kurven } c:[a,b]\to U.$$

Beweis. Nach Kettenregel (6.1.5) und Hauptsatz (5.6.17) gilt:

$$\int_a^b (g'\circ c)(t)\cdot c'(t)\,dt = \int_a^b (g\circ c)' = g(c(b)) - g(c(a)) \quad \square$$

Dies führt dazu auch für allgemeine 1-Formen $f : E \supseteq U \to L(E,F)$ das Integral längs C^1-Kurven $c : [a,b] \to U$ wie folgt zu definieren:

6.5.6 Definition. Kurvenintegral längs C^1-Kurven.
Es sei $c : [a,b] \to E$ C^1 und $f : E \supseteq c([a,b]) \to L(E,F)$ eine stetige 1-Form. Dann ist das KURVENINTEGRAL $\int_c f$ wie folgt definiert:

$$\int_c f := \int_a^b (f \circ c) \cdot c' = \int_a^b f(c(t)) \cdot c'(t)\, dt.$$

6.5.7 Proposition. Reparametrisierungsinvarianz des Kurvenintegrals.
Das Kurvenintegral ist Reparametrisierungs-invariant, d.h. für $c : [a,b] \to E$ C^1 und $h : [\bar a, \bar b] \to [a,b]$ C^1 mit $h(\bar a) = a$ und $h(\bar b) = b$ ist für alle stetigen 1-Formen $f : E \supseteq c([a,b]) \to L(E,F)$:

$$\int_c f = \int_{c\circ h} f.$$

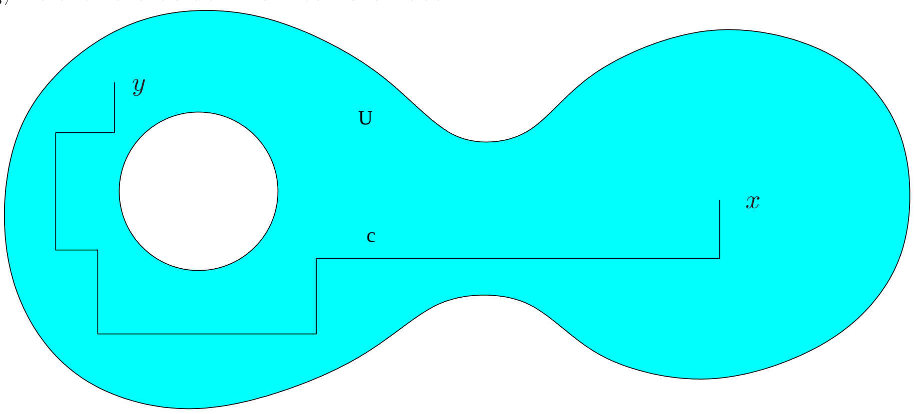
Beweis.

$$\int_{c\circ h} f = \int_{\bar a}^{\bar b} f\big((c\circ h)(t)\big) \cdot (c\circ h)'(t)\, dt = \int_{\bar a}^{\bar b} f\big(c(h(t))\big) \cdot c'\big(h(t)\big)\, h'(t)\, dt = \int_{h(\bar a)}^{h(\bar b)} f(c(s)) \cdot c'(s)\, ds \qquad \square$$

Eine 1-Form f transformiert sich unter Parametrisierungswechsel $c \mapsto c \circ h$ also gerade so, daß das Kurvenintegral gleich bleibt.

6.5.8 Theorem. Exaktheit via Wegunabhängigkeit.
Es sei $f : E \supseteq U \to L(E,F)$ eine stetige 1-Form auf einer zusammenhängenden offenen Menge U. Dann ist f genau dann exakt, wenn ihr Kurvenintegral WEGUNABHÄNGIG *ist, d.h. nur von den Endpunkten der Kurve und nicht ihrem Aussehen dazwischen abhängt. Als Stammfunktion können wir $g(z) := \int_c f$ definieren, wobei $c : [0,1] \to U$ eine C^1-Kurve von z_0 nach z ist. Weiters ist die Stammfunktion bis auf eine additive Konstante eindeutig.*

Definition. Zusammenhängende Menge.
Eine offene Menge $U \subseteq E$ heißt ZUSAMMENHÄNGEND $:\Leftrightarrow$ zu je zwei Punkten $x, y \in U$ existiert ein Polygonzug, welcher die beiden Punkte verbindet.



Wenn $c_i : [t_{i-1}, t_i] \to E$ Kurven sind, die an den Endpunkten übereinstimmen (d.h. $c_i(t_i) = c_{i+1}(t_i)$), so definieren wir die Kurve $c := c_1 \odot \cdots \odot c_n : [t_0, t_n] \to E$ durch $c|_{[t_{i-1},t_i]} := c_i$. Ein POLYGONZUG ist also $c_1 \odot \cdots \odot c_n$, wobei die c_i die affinen Verbindungsstücke der Ecken sind.

Die Segmente $\overline{x_{i-1}, x_i}$ des Polygonzugs können durch $c_i : t \mapsto x_{i-1} + t(x_i - x_{i-1})$ C^∞-parametrisiert werden. Die Ableitungen an der Ecken stimmen aber nicht überein, also liefert $c_1 \odot \cdots \odot c_n$ keine differenzierbare Parametrisierung. Aber wir können die c_i vermöge $h : [0,1] \to [0,1]$, $h(t) = t^2(3 - 2t)$ umparametrisieren. Wegen $h(0) = 0$ und $h(1) = 1$ können wir $c := (c_1 \circ h) \odot \cdots \odot (c_n \circ h)$ bilden. Wegen $h'(0) = 0 = h'(1)$ verschwindet $(c_i \circ h)'$ an den Randpunkten und somit ist c C^1. Also können wir den Polygonzug C^1 parametrisieren. Es geht sogar C^∞, wenn wir h so wählen, daß alle Ableitungen am Rand des Intervall verschwinden. Mittels (6.5.7) macht also das Kurvenintegral längs stückweiser C^1-Kurven einen Sinn, indem wir dieses als Summe der Kurvenintegrale über die C^1-Teile definieren oder so umparametrisieren, daß die stückweise differenzierbare Kurve C^1 wird.

Beweis von (6.5.8). ($\Rightarrow$) Wenn f exakt ist, so ist das Kurvenintegral wegunabhängig nach (6.5.5).

Eindeutigkeit: Es seien g_i für $i \in \{0, 1\}$ zwei Stammfunktionen. Dann ist die Differenz $g = g_1 - g_0$ differenzierbar mit Ableitung $g' = 0$. Sei x_0 und x_1 in U beliebig und $c : [a, b] \to U$ eine C^1-Kurve in U, die x_0 mit x_1 verbindet. Dann ist $g(x_1) - g(x_0) = [(g \circ c)(t)]_{t=a}^b = \int_a^b (g \circ c)' = \int_a^b g'(c(t))(c'(t))\, dt = 0$.

($\Leftarrow$) Umgekehrt sei $x \in U$ beliebig und $c_x : [0, 1] \to U$ eine C^1-Kurve die x_0 mit x verbindet. Dann setzen wir $g(x) := \int_{c_x} f$. Für $v \in E$ sei weiters $c_{x,v} : t \mapsto x + tv$. Dann ist $c_{x,v}([0, 1]) \subseteq U$ für alle kleinen v und

$$g(x + v) = \int_{c_x \odot c_{x,v}} = \int_{c_x} f + \int_{c_{x,v}} f = g(x) + \int_0^1 f(x + tv)(v)\, dt.$$

Folglich ist

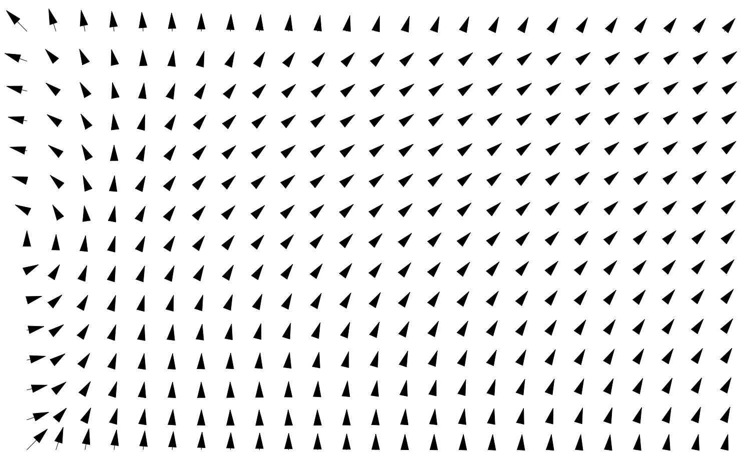
$$\|g(x + v) - g(x) - f(x)(v)\| = \left\| \int_0^1 f(x + tv)(v) - f(x)(v)\, dt \right\| \leq \|v\| \underbrace{\max\{\|f(x + tv) - f(x)\| : t \in [0, 1]\}}_{\to 0 \text{ für } v \to 0},$$

also ist $g'(x)(v) = f(x)(v)$. □

6.5.8a Beispiel. Berechnung einer Stammfunktion.
Es sei ein Vektorfeld $(f_1, f_2) : \mathbb{R}^2 \supseteq U \to \mathbb{R}^2$ durch

$$f_1(x, y) := -\frac{\tan(y)}{x^2} + 2xy + x^2$$
$$f_2(x, y) := \frac{1}{x \cos(y)^2} + x^2 + y^2$$

gegeben.



Gefragt ist, ob dieses KONSERVATIV, d.h. ein Gradientenfeld ist. Falls ja, so ist eine Funktion (das POTENTIAL DES VEKTORFELDS) gesucht, die dieses als Gradienten besitzt.

Wir übersetzen dieses Problem zuerst in die entsprechende Aussage über eine 1-Form $f : \mathbb{R}^2 \supseteq U \to L(\mathbb{R}^2, \mathbb{R}) \cong \mathbb{R}^2$. Der dabei verwendete Isomorphismus ist in Koordinaten durch Transponieren gegeben, d.h.

$$[f(z)] := (f_1(z), f_2(z)) = \begin{pmatrix} f_1(z) \\ f_2(z) \end{pmatrix}^t$$

bzw. via des inneren Produkt $\langle \_|\_ \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ als

$$f(z)(w) = \Big\langle \begin{pmatrix} f_1(z) \\ f_2(z) \end{pmatrix} \Big| w \Big\rangle \text{ für } z \in U \subseteq \mathbb{R}^2 \text{ und } w \in \mathbb{R}^2.$$

In der Basis (dx, dy) der 1-Formen ist also

$$f(x, y) = f_1(x, y)\, dx + f_2(x, y)\, dy.$$

Der maximale Definitionsbereich ist $U := \{(x, y) : x \neq 0, \cos(y) \neq 0\}$. Der größte zusammenhängende Definitionsbereich der z.B. den Punkt $(1, 0)$ enthält ist $U_0 := \{(x, y) : x > 0,\ -\frac{\pi}{2} < y < \frac{\pi}{2}\}$. Es ist f offensichtlich C^1, denn die Komponenten f_i von f sind (partiell) differenzierbar und ihre partiellen Ableitungen $\partial_i f_j : U \to \mathbb{R}$ sind stetig.

Die Integrabilitätsbedingung $f'(z)(v)(w) = f'(z)(w)(v)$ für alle $z \in U$ und $v, w \in \mathbb{R}^2$ können wir äquivalent auf der Standardbasis austesten, also ist $f'(z)(e_i)(e_j) = f'(z)(e_j)(e_i)$ für alle i und j (mit $i < j$) zu überprüfen. Dabei ist $f'(z)(e_i)(e_j) = \mathrm{ev}_{e_j}(f'(z)(e_i)) = (\mathrm{ev}_{e_j} \circ f)'(z)(e_i) = f_j{}'(z)(e_i) = \partial_i f_j(z)$. In unserem Beispiel ist Integrabilitätsbedingung somit die $\mathbb{R}$-wertige gültige Gleichung

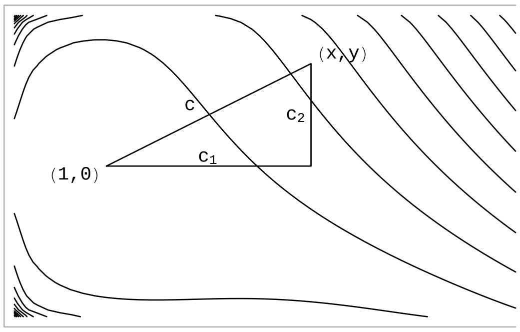
$$\partial_2 f_1(x, y) = -\frac{1}{x^2 \cos(y)^2} + 2x = \partial_1 f_2(x, y).$$

Nach (6.5.4) existiert somit auf der konvexen Menge U eine Stammfunktion g und nach (6.5.8) ist diese bis auf eine additive Konstante eindeutig bestimmt und kann mittels des Kurvenintegrals berechnet werden.

Wir können einen beliebigen Punkt $(x, y) \in U_0$ mit $(1, 0)$ durch die affine C^1-Kurve $c : [0, 1] \to U_0$ mit $c(t) := (1 + t(x - 1), ty)$ und Ableitung $c'(t) = (x - 1, y)$ verbinden. Also ist

$$\begin{aligned} g(x, y) &:= \int_c f = \int_0^1 f(c(t)) \cdot c'(t)\, dt = \int_0^1 f_1(1 + t(x-1), ty)\,(x-1) + f_2(1 + t(x-1), ty)\, y\, dt \\ &= \int_0^1 \Big(-\frac{\tan(ty)}{(1 + t(x-1))^2} + 2(1 + t(x-1))ty + (1 + t(x-1))^2 \Big)(x-1) \\ &\qquad + \Big(\frac{1}{(1 + t(x-1)) \cos(ty)^2} + (1 + t(x-1))^2 + (ty)^2 \Big) y\, dt \end{aligned}$$

eine Stammfunktion.



Wir können zur Vermeidung dieses doch recht häßlichen Integrals auch anders vorgehen und Achsenparallele Wege verwenden:

$$\begin{aligned} c_1(\bar{x}) &:= (\bar{x}, 0) \quad \bar{x} \in [1, x], \quad c_1'(\bar{x}) = (1, 0) \\ c_2(\bar{y}) &:= (x, \bar{y}) \quad \bar{y} \in [0, y], \quad c_2'(\bar{y}) = (0, 1). \end{aligned}$$

Dann ist

$$
\begin{aligned}
g(x,y) &:= \int_{c_1} f + \int_{c_2} f = \int_1^x f_1(\bar{x},0)\,d\bar{x} + \int_0^y f_2(x,\bar{y})\,d\bar{y} \\
&= \int_1^x \Big(-\frac{\tan(0)}{\bar{x}^2} + 2\bar{x}0 + \bar{x}^2\Big)\,d\bar{x} + \int_0^y \Big(\frac{1}{x\cos(\bar{y})^2} + x^2 + \bar{y}^2\Big)\,d\bar{y} \\
&= \Big[\frac{\bar{x}^3}{3}\Big]_{\bar{x}=1}^{x} + \Big[\frac{1}{x}\tan(\bar{y}) + x^2\bar{y} + \frac{\bar{y}^3}{3}\Big]_{\bar{y}=0}^{y} = \frac{x^3}{3} - \frac{1}{3} + \frac{\tan(y)}{x} + x^2 y + \frac{y^3}{3} \\
&= \frac{\tan(y)}{x} + x^2 y + \frac{x^3}{3} + \frac{y^3}{3} - \frac{1}{3}
\end{aligned}
$$

Ebenso können wir folgende Achsen-parallele Wege verwenden:

$$
\begin{aligned}
c_2(\bar{y}) &:= (1,\bar{y}) \quad \bar{y} \in [0,y], \quad c_2'(\bar{y}) = (0,1) \\
c_1(\bar{x}) &:= (\bar{x},y) \quad \bar{x} \in [1,x], \quad c_1'(\bar{x}) = (1,0).
\end{aligned}
$$

Dann ist

$$
\begin{aligned}
g(x,y) &:= \int_{c_2} f + \int_{c_1} f = \int_0^y f_2(1,\bar{y})\,d\bar{y} + \int_1^x f_1(\bar{x},y)\,d\bar{x} \\
&= \int_0^y \Big(\frac{1}{\cos(\bar{y})^2} + 1 + \bar{y}^2\Big)\,d\bar{y} + \int_1^x \Big(-\frac{\tan(y)}{\bar{x}^2} + 2\bar{x}y + \bar{x}^2\Big)\,d\bar{x} \\
&= \Big[\tan(\bar{y}) + \bar{y} + \frac{\bar{y}^3}{3}\Big]_{\bar{y}=0}^{y} + \Big[\frac{\tan(y)}{\bar{x}} + \bar{x}^2 y + \frac{\bar{x}^3}{3}\Big]_{\bar{x}=1}^{x} \\
&= \tan(y) + y + \frac{y^3}{3} - 0 + \frac{\tan(y)}{x} + x^2 y + \frac{x^3}{3} - \tan(y) - y - \frac{1}{3} \\
&= \frac{\tan(y)}{x} + x^2 y + \frac{x^3}{3} + \frac{y^3}{3} - \frac{1}{3}
\end{aligned}
$$

6.5.9 Bemerkung. Riemann-Stieltjes-Integral längs nicht differenzierbarer Kurven.
Falls $f : E \supseteq U \to L(E,F)$ stetig und $c : [a,b] \to U$ C^1 ist, so haben wir das Kurvenintegral als

$$\int_c f = \int_a^b f(c(t))(c'(t))\,dt$$

definiert. Die entsprechenden Riemann-Summen sind

$$S\big((f\circ c)\cdot c', Z, \xi\big) = \sum_{j=1}^n f\big(c(\xi_j)\big)\big(c'(\xi_j)\big)(t_j - t_{j-1}).$$

Wegen $\frac{c(t_j)-c(t_{j-1})}{t_j-t_{j-1}} \approx c'(\xi_j)$ können wir diese ihrerseits durch die sogenannten Riemann-Stieltjes-Summen

$$S_c(f,Z,\xi) := \sum_{j=1}^n f\big(c(\xi_j)\big)\big(c(t_j) - c(t_{j-1})\big)$$

approximieren. Folglich definieren wir das Kurvenintegral (oder Riemann-Stieltjes-Integral) als Limes der Riemann-Stieltjes-Summen

$$\int_c f := \lim_{|Z|\to 0} S_c(f,Z,\xi),$$

sofern dieser existiert.

Wir wollen nun zeigen, daß dieses Integral unter der schwächeren Bedingung daß f stetig ist und c endliche Länge hat existiert.

6.5.10 Definition. Rektifizierbare Kurven.
Eine stetige Kurve $c : [a, b] \to F$ heißt genau dann REKTIFIZIERBAR (oder auch von BESCHRÄNKTER VARIATION oder von endlicher Länge), wenn

$$V_a^b(c) = V(c) := \sup\Big\{V(c, Z) : Z \text{ ist Zerlegung von } [a, b]\Big\} < \infty$$

$$\text{wobei } V(c, Z) := \sum_{j=1}^{n} \|c(t_j) - c(t_{j-1})\| \text{ für Zerlegungen } Z = (a = t_0 < \cdots < t_n = b)$$

Dabei heißt $V(c)$ auch die LÄNGE (oder VARIATION) von c, vgl. dies mit Aufgabe (5.23).

6.5.11 Satz. Länge differenzierbarer Kurven.
Es sei $c : [a, b] \to F$ C^1. Dann ist c rektifizierbar, $V_a^b(c) = \int_a^b \|c'(t)\|\, dt$ und $\frac{d}{dt} V_a^t(c) = \|c'(t)\|$.

Vgl. dies mit Aufgabe (5.24).

Beweis. Wegen der Dreiecksungleichung für Integrale (5.6.14.2) ist

$$\|c(t_i) - c(t_{i-1})\| \le \int_{t_{i-1}}^{t_i} \|c'\|$$

und somit $V(c) \le \int \|c'\|$. Umgekehrt ist für $h > 0$

$$\Big\|\frac{c(t+h) - c(t)}{h}\Big\| \le \frac{1}{h} V_t^{t+h}(c) \le \frac{1}{h} \int_t^{t+h} \|c'\| \to \|c'(t)\| \quad \square$$

Um die Existenz des Kurvenintegrals zu zeigen benötigen wir folgendes Lemma

6.5.12 Lemma. Differenz zweier Riemann-Stieltjes-Summen.
Es sei $c : [a, b] \to E$ rektifizierbar und $f : E \supseteq c([a, b]) \to L(E, F)$ beschränkt. Weiters sei Z' eine Verfeinerung der Zerlegung Z von $[a, b]$. Dann ist

$$\Big\|S_c(f \circ c, Z, \xi) - S_c(f \circ c, Z', \xi')\Big\| \le \Omega_Z(f \circ c) \cdot V(c),$$

wobei $\Omega_Z(f) := \sup\Big\{\|f(c(t)) - f(c(s))\| : t, s \in I \in Z\Big\}$.

Beweis. O.B.d.A. sei $Z = \{a, b\}$ und $Z' = \{a = t_0 < \cdots < t_n = b\}$. Wegen $c(b) - c(a) = \sum_i c(t_i) - c(t_{i-1})$ ist

$$\begin{aligned}\Big\|f\big(c(\xi)\big)\big(c(b) - c(a)\big) - \sum_i f\big(c(\xi_i)\big)\,\big(c(t_i) - c(t_{i-1})\big)\Big\| &\le \sum_i \|f(c(\xi)) - f(c(\xi_i))\|\, \|c(t_i) - c(t_{i-1})\| \\ &\le \sup_i \|f(c(\xi)) - f(c(\xi_i))\| \sum_i \|c(t_i) - c(t_{i-1})\| \\ &\le \Omega(f \circ c)\, V(c) \quad \square\end{aligned}$$

6.5.13 Satz. Integral längs rektifizierbarer Kurven.
Es sei $c : [a, b] \to E$ rektifizierbar, $f : E \supseteq c([a, b]) \to L(E, F)$ stetig. Dann existiert das Kurvenintegral

$$\int_c f := \lim_{|Z| \to 0} S_c(f \circ c, Z, \xi).$$

Beweis. Es sei $Z := Z_1 \cup Z_2$. Dann ist

$$\begin{aligned}
&\|S_c(f\circ c, Z_1, \xi_1) - S_c(f\circ c, Z_2, \xi_2)\| \le \\
&\qquad \le \|S_c(f\circ c, Z_1, \xi_1) - S_c(f\circ c, Z, \xi)\| + \|S_c(f\circ c, Z, \xi) - S_c(f\circ c, Z_2, \xi_2)\| \\
&\qquad \le \Omega_{Z_1}(f\circ c)\cdot V(c) + \Omega_{Z_2}(f\circ c)\cdot V(c) < \varepsilon, \text{ falls } \Omega_{Z_i}(f\circ c) \le \frac{\varepsilon}{2V(c)}. \quad \square
\end{aligned}$$

Daß das in (6.5.9) bzw. (6.5.13) definierte Kurvenintegral mit jenem in (6,5.6) definierten Kurvenintegral für C^1-Kurven übereinstimmt, zeigt folgender Satz.

6.5.14 Satz. Kurvenintegral längs differenzierbarer Kurven.
Es sei $c : [a, b] \to E$ C^1 und $f : E \supseteq c([a, b]) \to L(E, F)$ C. Dann ist

$$\int_c f = \int_a^b f(c(t))(c'(t))\, dt.$$

Beweis.

$$\begin{aligned}
\Big\|S_c(f, Z, \xi) - \int_a^b f(c(t))\,(c'(t))\,dt\Big\| &= \Big\|\sum_i f(c(\xi_i))\,(c(t_i) - c(t_{i-1})) - \int_a^b f(c(t))\,(c'(t))\,dt\Big\| \\
&\le \sum_i \Big\|\int_{t_{i-1}}^{t_i} \Big(f(c(\xi_i)) - f(c(t))\Big)\,(c'(t))\,dt\Big\| \\
&\le \sum_i \int_{t_{i-1}}^{t_i} \|f(c(\xi_i)) - f(c(t))\|\,\|c'(t)\|\,dt \\
&\le \Omega_Z(f\circ c)\,V_a^b(c) \le \varepsilon\,(b-a)\,\|c'\|_\infty,
\end{aligned}$$

falls $|Z|$ so klein ist, daß $\Omega_Z(f\circ c) < \varepsilon$ wegen der glm. Stetigkeit von $f\circ c$. $\square$

6.5.15 Bemerkung. Arbeit.
In der Physik ist Arbeit grob gesprochen Kraft mal Weg. So einfach kann man das aber nur machen, wenn die Kraft längs des Weges konstant ist. Da Kräfte nicht skalare Größen sind, sondern auch eine Richtung haben, also Vektoren oder genauer gesagt Vektorfelder $f : E \supseteq U \to E$ sind, und bei der zu leistenden Arbeit im Kraftfeld f nur die Tangentialkomponente $\frac{1}{|c'(t)|}\langle f(c(t))|c'(t)\rangle$ an den Weg $c : [a, b] \to E$ zum Tragen kommt, sollte die Arbeit näherungsweise durch

$$A \approx \sum_{i=1}^n \frac{1}{\|c'(\xi_i)\|}\,\Big\langle f(c(\xi_i))\,\Big|\,c'(\xi_i)\Big\rangle\,\|c(t_i) - c(t_{i-1})\|$$

gegeben sein, wobei wir die Zerlegung $Z = (a = t_0 < \cdots < t_n = b)$ fein genug wählen, so daß $\frac{1}{|c'(t)|}\langle f(c(t))|c'(t)\rangle$ für $t_{i-1} \le t \le t_i$ annäherungsweise konstant gleich dem Wert an irgend einer Zwischenstelle $t_{i-1} \le \xi_i \le t_i$ ist. Wegen $c'(\xi_i) \approx \frac{c(t_i)-c(t_{i-1})}{t_i - t_{i-1}}$ ist dies annäherungsweise eine Riemann-Summe

$$A \approx \sum_i \Big\langle f(c(\xi_i))\,\Big|\,c'(\xi_i)\Big\rangle\,(t_i - t_{i-1})$$

des Kurvenintegrals der 1-Form $\tilde f : x \mapsto (v \mapsto \langle f(x)|v\rangle)$ die zum Vektorfeld f gehört. Man definiert folglich die Arbeit, die längs einer Kurve $c : [a, b] \to U \subseteq E$ in einem Kraftfeld $f : U \to E$ zu leisten ist, als

$$A := \int_c \tilde f = \int_a^b \langle f(c(t))|c'(t)\rangle\,dt.$$

Ein Kraftfeld heißt konservativ, wenn die Arbeit in diesem Kraftfeld nicht vom zurückgelegten Weg sondern nur von dessen Endpunkten abhängt. Nach (6.5.8) ist dies genau dann der Fall, wenn die zugehörige 1-Form exakt ist, also das Kraftfeld ein Gradientenfeld ist. Die Stammfunktion wird (bis auf ein Vorzeichen) in der Physik dann als Potential zum Vektorfeld bezeichnet.

6.6 Komplexe Analysis

6.6.1 Komplexe lineare Algebra und komplexe Funktionalanalysis.
Unter einem KOMPLEXEN BANACH-RAUM versteht man einen Vektorraum über dem Grundkörper $\mathbb{C}$ der komplexen Zahlen, versehen mit einer vollständigen Norm $\|\_\|$ die zusätzlich $\|\lambda \cdot x\| = |\lambda| \cdot \|x\|$ für alle $\lambda \in \mathbb{C}$ erfüllt. Jeder komplexe Banach-Raum ist natürlich insbesonders ein reeller Banach-Raum, wenn wir die Skalarmultiplikation nur für Skalare in $\mathbb{R} \subseteq \mathbb{C}$ betrachten. In diesem Kapitel seien alle Banach-Räume komplexe Banach-Räume.

Eine $\mathbb{C}$-LINEARE ABBILDUNG $f : E \to F$ zwischen komplexen Banach-Räumen ist eine Abbildung, die additiv ist (d.h. $f(x + y) = f(x) + f(y)$ erfüllt) und mit der Skalarmultiplikation vertauscht (d.h. $f(\lambda \cdot x) = \lambda \cdot f(x)$ für alle $\lambda \in \mathbb{C}$ erfüllt). Dies ist genau dann der Fall wenn dies für alle $\lambda \in \mathbb{R}$ gilt (d.h. sie $\mathbb{R}$-linear ist) und zusätzlich $f(i \cdot x) = i \cdot f(x)$ gilt. Falls $E = \mathbb{C}$ ist, so ist das genau dann der Fall, wenn $f(\lambda) = f(\lambda \cdot 1) = \lambda \cdot f(1)$ ist, d.h. f die Skalarmultiplikation mit dem Vektor $f(1) \in F$ ist. Insbesonders sind Abbildungen $f : \mathbb{C} \to \mathbb{C}$ genau dann $\mathbb{C}$-linear, wenn ihre Matrixdarstellung

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

(als $\mathbb{R}$-lineare Abbildung) bezüglich der reellen standard-Basis $(e_1, e_2) \subseteq \mathbb{R}^2 \cong \mathbb{C}$ folgende beide Bedingungen erfüllt: $a = d$ und $b = -c$.
In der Tat genügt es die $\mathbb{R}$-lineare Gleichung $f(i \cdot x) = i \cdot f(x)$ für $x \in \{e_1, e_2\}$ auszutesten, und wegen $i^2 = -1$ genügt $x = e_1$ zu setzen. Dann ist

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \left(i\, e_1\right) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot e_2 = \begin{pmatrix} b \\ d \end{pmatrix} \quad \text{und} \quad i \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot e_1 = i \begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} -c \\ a \end{pmatrix}$$

In diesem Fall ist die Wirkung von f auf $u + iv$ gegeben durch

$$\begin{pmatrix} a & -c \\ c & a \end{pmatrix} \cdot \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} au - cv \\ cu + av \end{pmatrix},$$

dies entspricht der Multiplikation $(a + ic) \cdot (u + iv) = (au - cv) + i(av + cu)$.

6.6.2 Proposition. Reell- versus komplex-differenzierbar.
Eine Abbildung $f : \mathbb{C} \supseteq U \to F$ mit Werten in einen komplexen Banach-Raum F ist bei $z \in U$ genau dann (reell-)differenzierbar mit $\mathbb{C}$-linearer Ableitung $f'(z) : \mathbb{C} \to F$, wenn $\lim_{\mathbb{C}\ni w\to 0} \frac{f(z+w)-f(z)}{w} \in F$ existiert. Dieser Limes ist dann gerade $f'(z)(1)$ und heißt KOMPLEXE ABLEITUNG *von f an der Stelle z und wird mit $\frac{df}{dz}(z)$ oder auch wieder mit $f'(z)$ bezeichnet.*

Folglich nennt man allgemein eine Abbildung $f : E \supseteq U \to F$ $\mathbb{C}$-DIFFERENZIERBAR oder auch HOLOMORPH, wenn sie als Abbildung zwischen den zugehörigen reellen Banach-Räumen differenzierbar ist und die Ableitung $f'(z) : E \to F$ $\mathbb{C}$-linear für alle $z \in U$ ist. Der Einfachheit halber wollen wir zusätzlich verlangen, daß f' stetig, also f C^1 ist. Goursat konnte zeigen, daß diese Bedingung für $E = F = \mathbb{C}$ überflüssig ist. Wir gehen darauf in (6.6.7) etwas genauer ein.

Beweis. ($\Rightarrow$) Wenn $f'(z)$ $\mathbb{C}$-linear ist, dann gilt $f'(z)(w) = f'(z)(w \cdot 1) = w \cdot f'(z)(1)$ und somit ist

$$0 = \lim_{w\to 0} \frac{\|f(z + w) - f(z) - f'(z)(w)\|}{\|w\|} = \lim_{w\to 0} \left\| \frac{f(z + w) - f(z)}{w} - f'(z)(1) \right\|,$$

d.h. es existiert $\frac{df}{dz}(z) := \lim_{w\to 0} \frac{f(z+w)-f(z)}{w}$ und ist $f'(z)(1)$.

($\Leftarrow$) Es sei $\frac{df}{dz}(z) := \lim_{w\to 0} \frac{f(z+w)-f(z)}{w}$ und $\ell(w) := w \cdot \frac{df}{dz}(z)$. Dann gilt:

$$\lim_{w\to 0} \frac{\|f(z+w)-f(z)-\ell(w)\|}{\|w\|} = \lim_{w\to 0} \left\| \frac{f(z+w)-f(z)}{w} - \frac{df}{dz}(z) \right\| = 0,$$

d.h. f ist differenzierbar mit Ableitung $f'(z) = \ell$ und $\frac{df}{dz}(z) = \ell(1) = f'(z)(1)$. □

Folgerung. Cauchy-Riemann-Differentialgleichung.
Eine komplex-wertige Abbildung $f = f^1 + i\, f^2 : \mathbb{C} \supseteq U \to \mathbb{C}$ ist genau dann holomorph, wenn sie C^1 ist und die Cauchy-Riemann'schen Differentialgleichungen

$$\partial_1 f^1 = \partial_2 f^2 \quad \textit{und} \quad \partial_2 f^1 = -\partial_1 f^2$$

erfüllt sind. □

Bevor wir nun das umgekehrte Problem angehen, nämlich aus der komplexen Ableitung $g := f'$ die Stammfunktion f zu rekonstruieren, machen wir folgende Bemerkung:

6.6.3 Lemma. $\mathbb{C}$-Differenzierbarkeit vererbt sich auf die Ableitung.
Es sei $f : \mathbb{C} \supseteq U \to F$ C^2 und $\mathbb{C}$-differenzierbar, dann ist $\frac{df}{dz} : U \to F$ ebenfalls $\mathbb{C}$-differenzierbar.

Wir werden in (6.6.12) zeigen, daß $\mathbb{C}$-differenzierbare Abbildungen C^∞-sind, also sind alle ihre Ableitungen $\mathbb{C}$-differenzierbar.

Beweis. Wir müssen zeigen, daß $\left(\frac{df}{dz}\right)'(z)$ $\mathbb{C}$-linear ist. Sei also $\lambda \in \mathbb{C}$. Dann ist

$$\begin{aligned}
\left(\frac{df}{dz}\right)'(z)(\lambda\cdot v) &= (f')'(z)(\lambda\cdot v)(1) = f''(z)(\lambda\cdot v, 1) \\
&\overset{(6.3.10)}{=\!=} f''(z)(1, \lambda\cdot v) = (f')'(z)(1)(\lambda\cdot v) = \lambda\cdot (f')'(z)(1)(v) \\
&\overset{(6.3.10)}{=\!=} \lambda\cdot (f')'(z)(v)(1) = \lambda\cdot \left(\frac{df}{dz}\right)'(z)(v). \quad \square
\end{aligned}$$

6.6.4 Satz über gewöhnliche komplexe Differentialgleichung.
Es sei $f : \mathbb{C}\times E \supseteq U \to E$ C^1 und in beiden Variablen (getrennt) $\mathbb{C}$-differenzierbar. Dann existiert lokal eine $\mathbb{C}$-differenzierbare Funktion $g : \mathbb{C} \to E$ mit $g'(x) = f(x, g(x))$ und $g(0) = 0$.

Beachte, daß, wenn f zusätzlich $\mathbb{C}$-differenzierbar von weiteren Parametern abhängt, so gilt dies auch für die Lösung und diese hängt auch $\mathbb{C}$-differenzierbar von den Anfangsdaten ab.

Beweis. Wir wenden den Satz (6.5.1) von Frobenius auf die Funktion $\tilde f : \mathbb{R}^2 \times E \supseteq U \to L(\mathbb{R}^2, E)$, $z \mapsto (\lambda \mapsto \lambda\cdot f(z))$ an. Die Integrabilitätsbedingung ist erfüllt, denn

$$\begin{aligned}
w\cdot f'(x,g(x))\,(v, v\cdot f(x,g(x))) &= w\cdot \Big(\partial_1 f(x,g(x))\,(v) + \partial_2 f(x,g(x))\,(v\cdot f(x,g(x)))\Big) \\
&= w\cdot v\cdot \Big(\partial_1 f(x,g(x))\,(1) + \partial_2 f(x,g(x))\,(f(x,g(x)))\Big)
\end{aligned}$$

ist symmetrisch in v, w. Also existiert nach (6.5.1) eine lokale C^1 Lösung $g : \mathbb{R}^2 \to E$ von

$$g'(x) = \tilde f(x, g(x)).$$

Da $\tilde f(z)$ nach Konstruktion $\mathbb{C}$-linear ist, ist g $\mathbb{C}$-differenzierbar. □

6.6.5 Folgerung. Komplexe Stammfunktion.
Es sei $f : \mathbb{C} \supseteq U \to F$ $\mathbb{C}$-differenzierbar auf der sternförmigen Menge U, dann existiert eine $\mathbb{C}$-differenzierbare Stammfunktion g von f.

Beweis. Dazu müssen wir nur die Differentialgleichung $g'(z) = f(z)$ lösen, diese hat nach (6.6.4) eine $\mathbb{C}$-differenzierbare Lösung. □

6.6.6. Komplexes Kurvenintegral.
Zu $f : \mathbb{C} \supseteq X \to F$ definieren wir eine 1-Form $\tilde{f} : \mathbb{C} \supseteq X \to F \cong L_{\mathbb{C}}(\mathbb{C}, F) \subseteq L(\mathbb{R}^2, F)$ durch $\tilde{f}(z)(\lambda) := \lambda \cdot f(z)$. Dabei bezeichnet $L_{\mathbb{C}}(\mathbb{C}, F)$ die Menge der $\mathbb{C}$-linearen (stetigen) Abbildungen $\mathbb{C} \to F$. Nach (6.6.2) hat die exakte 1-Form dg jeder $\mathbb{C}$-differenzierbaren Abbildung $g : \mathbb{C} \supseteq U \to F$ Werte in $L_{\mathbb{C}}(\mathbb{C}, F)$ und ist auf $\lambda \in \mathbb{C}$ durch Multiplikation mit der komplexen Ableitung $g'(z)$ gegeben. Insbesonders haben wir die komplexe 1-Form $dz : \mathbb{C} \to \mathbb{C} \cong L_{\mathbb{C}}(\mathbb{C}, \mathbb{C}) \subseteq L(\mathbb{R}^2, \mathbb{R}^2)$ die wir als Ableitung der Identität $z \mapsto z$ erhalten, sowie $dx, dy : \mathbb{C} \cong \mathbb{R}^2 \to L(\mathbb{R}^2, \mathbb{R}) \subseteq L(\mathbb{R}^2, \mathbb{C}) \cong L(\mathbb{R}^2, \mathbb{R}^2)$, die Ableitungen der Koordinaten-Projektionen $z = x + i\,y \mapsto x$ und $z = x + i\,y \mapsto y$ auf Real- und Imaginärteil. Wegen $z = x + i\,y$ gilt für die Ableitungen $dz = dx + i\,dy$ aufgefaßt als Abbildungen $\mathbb{C} \to L(\mathbb{R}^2, \mathbb{R}^2)$. Mit Hilfe dieser ausgezeichneten 1-Formen können wir die zu $f : \mathbb{C} \supseteq X \to F$ gehörende 1-Form als $\tilde{f}(z) = f(z)\,dz$ schreiben. Für $F = \mathbb{C}$ können wir f in Real- und Imaginärteil als $f(z) = u(z) + i\,v(z)$ zerlegen, und erhalten

$$\tilde{f}(z) = f(z) \cdot dz = \Big(u(z) + i\,v(z)\Big) \cdot \Big(dx + i\,dy\Big) = \Big(u(z)\,dx - v(z)\,dy\Big) + i\Big(v(z)\,dx + u(z)\,dy\Big),$$

oder kurz,

$$\tilde{f} = (u\,dx - v\,dy) + i(v\,dx + u\,dy).$$

Dies zeigt auch die folgende Rechnung mit $w = x + i\,y$

$$\begin{aligned}\tilde{f}(z)(w) &= (x + iy) \cdot (u(z) + iv(z)) = x \cdot \Big(u(z) + i\,v(z)\Big) + y\Big(-v(z) + i\,u(z)\Big)\\ &= \Big(x \cdot u(z) - y \cdot v(z)\Big) + i\Big(x \cdot v(z) + y \cdot u(z)\Big).\end{aligned}$$

Satz.
Es sei $c : [a, b] \to \mathbb{C}$ C^1 und $f : \mathbb{C} \supseteq c([a, b]) \to F$ stetig. Dann existiert das komplexe Kurvenintegral

$$\int_c f := \int_c \tilde{f} = \int_c f(z)\,dz := \int_a^b f(c(t))\,c'(t)\,dt.$$

Ist $F = \mathbb{C}$ und $f = u + i\,v$ die Zerlegung in Real- und Imaginärteil, so ist

$$\int_c f = \int_c (u\,dx - v\,dy) \;+\; i\int_c (v\,dx + u\,dy).$$

Weiters ist

$$\Big\|\int_c f\Big\| \le \int_a^b \|c'(t)\|\,dt \cdot \sup\Big\{\|f(c(t))\| : t \in [a, b]\Big\} = L(c) \cdot \|f \circ c\|_\infty \le L(c) \cdot \|f\|_\infty.$$

Beweis. Die Existenz folgt aus (6.5.6).
Für $F = \mathbb{C}$ und $f = u + i\,v$ ist wegen $dz = dx + i\,dy$ und $\tilde{f}(z) = f(z) \cdot dz = (u\,dx - v\,dy) + i\,(v\,dx + u\,dy)$

$$\int_c f(z)\,dz = \int_c \Big((u\,dx - v\,dy) + i\,(v\,dx + u\,dy)\Big) = \int_c (u\,dx - v\,dy) + i\,\int_c (v\,dx + u\,dy).$$

Die Normabschätzung erhalten wir aus (5.6.14) wie folgt:

$$\Big\|\int_c f\Big\| = \Big\|\int_a^b f(c(t))\,(c'(t))\,dt\Big\| \le \int_a^b \|c'(t)\|\,dt \cdot \sup\{\|f(c(t))\| : t \in [a, b]\} = L(c) \cdot \|f \circ c\|_\infty \le L(c) \cdot \|f\|_\infty.$$ □

6.6.6a Beispiel. Kurvenintegral der Potenzfunktionen.
Für $f(z) := (z - z_0)^m$ mit $m \in \mathbb{Z}$ und der Parametrisierung $c(t) := z_0 + r\,e^{it}$ für $t \in [0, 2\pi]$ des Kreises mit Radius r und Mittelpunkt z_0 ist

$$\int_c f = \begin{cases} 0 & \text{für } m \neq -1 \\ 2\pi i & \text{für } m = -1 \end{cases} .$$

In der Tat ist

$$\begin{aligned}
\int_c f &= \int_0^{2\pi} (r\,e^{it})^m\, r\,i\,e^{it}\,dt \\
&= \int_0^{2\pi} r^{m+1}\,i\,e^{it(1+m)}\,dt = \int_0^{2\pi} r^{m+1}\,i\,\Big(\cos(t(m+1)) + i\sin(t(m+1))\Big) \\
&= \begin{cases} \int_0^{2\pi} i\,dt = 2\pi i & \text{für } m = -1 \\ i r^{m+1} \int_0^{2\pi(m+1)} \frac{\cos(s)}{m+1}\,ds - r^{m+1}\int_0^{2\pi(m+1)} \frac{\sin(s)}{m+1}\,ds = 0 & \text{sonst.} \end{cases}
\end{aligned}$$

Für $m = -1$ und $z_0 = 0$ bedeutet dies insbesonders, daß

$$\int_c \frac{1}{z}\,dz = 2\,\pi\,i.$$

Zerlegung in Real- und Imaginärteil der $\mathbb{C}$-wertigen Form $\frac{1}{z}\,dz$ liefert

$$\frac{1}{z}\,dz = \frac{1}{x + i\,y}\,d(x + i\,y) = \frac{x - i\,y}{x^2 + y^2}\,(dx + i\,dy) = \frac{x\,dx + y\,dy}{x^2 + y^2} + i\,\frac{x\,dy - y\,dx}{x^2 + y^2}.$$

Der Realteil ist eine exakte 1-Form auf $\mathbb{R}^2 \setminus \{0\}$, denn eine Stammfunktion ist durch $(x, y) \mapsto \ln(\sqrt{x^2 + y^2})$ gegeben. Der Imaginärteil ist auch geschlossen, denn

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

Auf sternförmigen Mengen $U \subseteq \mathbb{R}^2 \setminus \{0\}$ ist dieser also exakt nach (6.5.4). Eine Stammfunktion ist für $x \neq 0$ z.B. durch $(x, y) \mapsto \arctan(\frac{y}{x})$ und für $y \neq 0$ durch $(x, y) \mapsto \operatorname{arccot}(\frac{x}{y})$ gegeben, denn

$$\begin{aligned}
\frac{\partial}{\partial x}\left(\arctan\left(\frac{y}{x}\right)\right) &= \frac{1}{1 + \left(\frac{y}{x}\right)^2}\left(-\frac{y}{x^2}\right) = -\frac{y}{x^2 + y^2} \\
\frac{\partial}{\partial y}\left(\arctan\left(\frac{y}{x}\right)\right) &= \frac{1}{1 + \left(\frac{y}{x}\right)^2}\,\frac{1}{x} = \frac{x}{x^2 + y^2}.
\end{aligned}$$

Wenn wir Real- und Imaginärteil kombinieren, so erhalten wir als lokale Stammfunktion die Umkehrfunktion ln der Exponentialfunktion $\exp : \mathbb{C} \to \mathbb{C}$. Da das Kurvenintegral von $\frac{1}{z}\,dz$ längs des Einheitskreises gerade $2\,\pi\,i$ ist, ist es nicht egal, ob wir den Wert der Stammfunktion an der Stelle -1 durch Integration längs des oberen Halbkreises oder des unteren berechnen. Um Wohldefiniertheit der Umkehrfunktion zu erreichen, verwendet man folglich als Definitionsbereich die geschlitzte Ebene $\mathbb{C} \setminus \{t \in \mathbb{R} : t \leq 0\}$, oder man verwendet als Definitionsbereich eine sogenannte Riemann'sche Fläche indem man sich $\mathbb{C} \setminus \{0\}$ als Grundriß einer "Wendeltreppe" interpretiert und somit -1 in jedem "Stockwerk" $k \in \mathbb{Z}$ vorkommt und ln dort den Wert $2\,k\,\pi\,i$ hat.

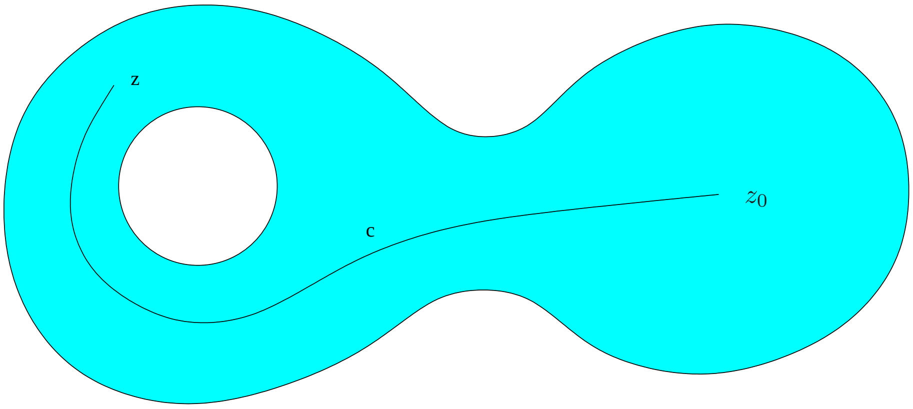
6.6.7 Cauchy'scher Integralsatz. $\mathbb{C}$-diff.bare Fkt. liefern lokal wegunabh. Kurvenintegrale.
Es sei f $\mathbb{C}$-differenzierbar auf einer sternförmigen offenen Menge $U \subseteq \mathbb{C}$. Dann ist $\int_c f$ wegunabhängig, oder anders ausgedrückt, falls c geschlossen ist, d.h. an den Randpunkten des Definitionsbereichs den selben Wert hat, so ist $\int_c f = 0$.

Beweis. Da f $\mathbb{C}$-differenzierbar ist, so erfüllt die zugehörige 1-Form $\tilde{f}$ die Integrabilitätsbedingung, denn $\tilde{f}'(z)(v)(w) = \frac{df}{dz}(z) \cdot v \cdot w$. Also folgt aus (6.5.4), daß $\tilde{f}$ exakt ist und aus (6.5.5), daß das Kurvenintegral wegunabhängig ist.

Ist $c : [a,b] \to U$ geschlossen, so ist $\int_c f = \int_{c_1} f + \int_{c_2} f$, wobei $c_1 := c|_{[a,\frac{a+b}{2}]}$ und $c_2 := c|_{[\frac{a+b}{2},b]}$, und $\int_{c_2} f = -\int_{c_1} f$, denn c_1 und der umgekehrt durchlaufene Weg c_2 haben die gleichen Endpunkte, also das selbe Kurvenintegral. □

Beachte, daß wir in diesen Beweis verwendet haben, daß die Ableitung von f stetig ist. Goursat konnte zeigen, daß diese Voraussetzung im Cauchy'schen Integralsatz nicht nötig ist, siehe dazu jedes Buch über KOMPLEXE ANALYSIS (= FUNKTIONENTHEORIE), z.B. [20, IV.3.13].

6.6.8 Satz von Morera. Wegunabh. komplexe 1-Formen haben $\mathbb{C}$-diff.bare Stammfkt.
Es sei $f : \mathbb{C} \supseteq U \to \mathbb{C}$ stetig, U zusammenhängend und $\int_c f$ wegunabhängig. Dann ist $g(z) := \int_c f$, wobei c eine Kurve ist, die z_0 mit z verbindet, eine $\mathbb{C}$-differenzierbare Stammfunktion. Für jede Stammfunktion g von f gilt: $\int_c f = g(c(b)) - g(c(a))$.

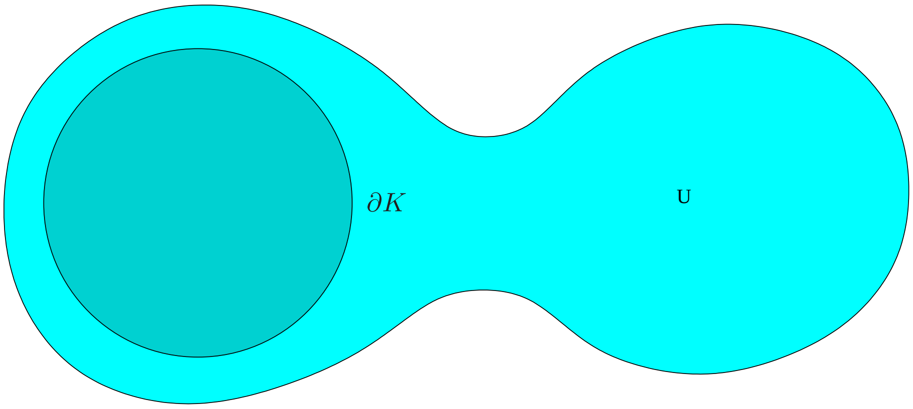


Beweis. Aus (6.5.8) folgt die Existenz der Stammfunktion, und da deren Ableitung f $\mathbb{C}$-linear ist, ist sie $\mathbb{C}$-differenzierbar nach (6.6.1). □

6.6.10 Cauchy'sche Integralformel. $\mathbb{C}$-diff.bare Fkt. sind Mittelwert ihrer Randwerte.
Es sei f $\mathbb{C}$-differenzierbar auf der offenen Menge U und K eine Kreisscheibe in U. Dann ist

$$f(z) = \frac{1}{2\pi i}\int_{\partial K} \frac{f(\zeta)}{\zeta - z}\,d\zeta \text{ für alle } z \text{ im Inneren von } K,$$

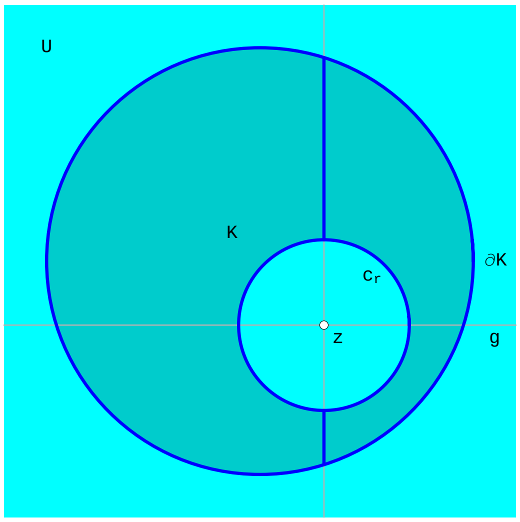
wobei ∂K den 1-mal positiv orientiert durchlaufenen Rand von K bezeichnet.



Beweis. Nach (6.6.6a) ist

$$\int_{\partial K} \frac{f(\zeta)-f(z)}{\zeta - z}\,d\zeta = \int_{\partial K} \frac{f(\zeta)}{\zeta - z}\,d\zeta - \int_{\partial K} \frac{f(z)}{\zeta - z}\,d\zeta = \int_{\partial K} \frac{f(\zeta)}{\zeta - z}\,d\zeta - f(z)\underbrace{\int_{\partial K} \frac{1}{\zeta - z}\,d\zeta}_{2\pi i}.$$

Um das erste Integral abzuschätzen verwenden wir, daß $\zeta \mapsto \frac{f(\zeta)-f(z)}{\zeta - z}$ $\mathbb{C}$-differenzierbar auf $U \setminus \{z\}$ ist, und somit auf allen sternförmigen Teilmengen wegunabhängig ist. Wenn wir somit eine Gerade g durch z legen, so ist K vermindert um einen der beiden Halbstrahlen sternförmig bezüglich jedem Punkt in K am anderen Halbstrahl. Wenn wir die orthogonal liegende Gerade durch z legen, so können wir das Kurvenintegral längs der enstprechenden Abschnitte von ∂K durch jenes längs der entsprechenden Abschnitte eines kleinen Kreises c_r um z in K plus jenen der beiden Abschnitte der orthogonalen Gerade zwischen den beiden Kreisen. Da diese Geradenstücke insgesamt also jeweils zweimal auftreten und zwar in beiden Richtungen durchlaufen, heben sich die dazugehörigen Kurvenintegrale auf und somit ist



$$\left\| \int_{\partial K} \frac{f(\zeta)-f(z)}{\zeta - z}\, d\zeta \right\| = \left\| \int_{c_r} \frac{f(\zeta)-f(z)}{\zeta - z}\, d\zeta \right\| \le L(c_r) \sup \left\{ \left\| \frac{f(x + r\, e^{i\,t}) - f(x)}{r\, e^{i\,t}} \right\| : t \in [0, 2\pi] \right\}$$
$$= 2\,\pi \sup \left\{ \| f(x + r\, e^{i\,t}) - f(x) \| : t \in [0, 2\pi] \right\} \to 0 \text{ für } r \to 0+$$

Da das ganz linke Integral aber nicht von r abhängt muß dieses 0 sein. □

6.6.11 Satz. Stetigkeit des Kurvenintegrals.
Es sei c C^1, $f_n : c(I) \to \mathbb{C}$ stetig und konvergiere glm. gegen f, dann gilt:

$$\int_c f_n \to \int_c f.$$

Beweis. Wegen $\int_c f_n = \int_a^b f(c(t))\, c'(t)\, dt$ folgt dies aus den entsprechenden Satz (5.6.14.6) für Riemann-Integrale. □

6.6.12 Entwicklungssatz & Cauchy'sche Ableitungsformeln.

Es sei f $\mathbb{C}$-differenzierbar, dann ist f am Inneren der größten Kreisscheibe K in U um einen gegebenen Punkt in eine Potenzreihe entwickelbar. Jede $\mathbb{C}$-differenzierbare Funktion ist C^∞ und

$$\frac{d^n f}{dz^n}(z) = \frac{n!}{2\pi i} \int_{\partial K} \frac{f(\zeta)}{(\zeta - z)^{n+1}}\, dz,$$

wobei ∂K wieder den 1-mal positiv orientiert durchlaufenen Rand von K bezeichnet.

Beweis. Nach (6.1.22) kann Differenzieren und Integrieren vertauscht werden, also folgt mittels Induktion aus (6.6.10)

$$\frac{d^n f}{dz^n}(z) == \left(\frac{d}{dz}\right)^n \frac{1}{2\pi i} \int_{\partial K} \frac{f(\zeta)}{\zeta - z}\, d\zeta \overset{(6.1.22)}{=\!=} \frac{1}{2\pi i} \int_{\partial K} \left(\frac{\partial}{\partial z}\right)^n \frac{f(\zeta)}{\zeta - z}\, d\zeta = \frac{n!}{2\pi i} \int_{\partial K} \frac{f(\zeta)}{(\zeta - z)^{n+1}}\, dz.$$

Aus (6.6.11) folgt wegen der glm. Konvergenz von

$$\frac{1}{\zeta - z} = \frac{1}{(\zeta - z_0) - (z - z_0)} = \frac{1}{(\zeta - z_0)\left(1 - \frac{z - z_0}{\zeta - z_0}\right)} = \frac{1}{\zeta - z_0} \sum_k \left(\frac{z - z_0}{\zeta - z_0}\right)^k$$

für $|z-z_0| < |\zeta - z_0|$ und der Cauchy'schen Integralformel (6.6.10) folgende Darstellung als lokal gleichmäßig konvergente Potenzreihe:

$$f(z) = \frac{1}{2\pi i}\int_c \frac{f(\zeta)}{\zeta - z}\,d\zeta = \frac{1}{2\pi i}\int_c \sum_k \frac{f(\zeta)}{(\zeta-z_0)^{k+1}}(z-z_0)^k\,d\zeta = \sum_k (z-z_0)^k \frac{1}{2\pi i}\int_c \frac{f(\zeta)}{(\zeta-z_0)^{k+1}}\,d\zeta$$

Und wegen (4.2.14) (leicht auf $\mathbb{C}$-Ableitungen verallgemeinert) dürfen wir diese Potenzreihe gliedweise differenzieren und erhalten somit nochmals die Existenz der Ableitungen von f und die entsprechende Formel als Kurvenintegral für sie. □

6.6.13 Satz von Morera. Komplexe Fkt. mit wegunabh. Integral sind $\mathbb{C}$-differenzierbar.
Ist f stetig mit wegunabhängigen Integral auf einer zusammenhängenden offenen Menge, so ist f $\mathbb{C}$-differenzierbar.

Beweis. Nach (6.6.8) existiert eine $\mathbb{C}$-differenzierbare g Stammfunktion, nach (6.6.12) ist g C^∞ und wegen (6.6.3) ist $f = g'$ $\mathbb{C}$-differenzierbar. □

6.6.14 Satz. Holomorphie via Potenzreihen.
Es sei $I \subseteq \mathbb{R}$ offen, dann ist $f : I \to \mathbb{R}$ genau dann in eine Potenzreihe lokal um $z_0 \in I$ entwickelbar, wenn es eine $\mathbb{C}$-differenzierbare Fortsetzung auf eine Umgebung von z_0 in $\mathbb{C}$ gibt.

Beweis. ($\Leftarrow$) nach (6.6.12).

($\Rightarrow$) Nach (4.2.5) konvergiert die Potenzreihe auch für komplexe Argumente. Die Summenfunktion ist dann wegen (4.2.14) $\mathbb{C}$-differenzierbar. □

6.6.15 Identitätssatz für komplex-differenzierbare Funktionen.
Falls zwei $\mathbb{C}$-differenzierbare Funktionen auf einer zusammenhängenden Menge auf einer in der Menge konvergenten Folge übereinstimmen, so sind sie ident.

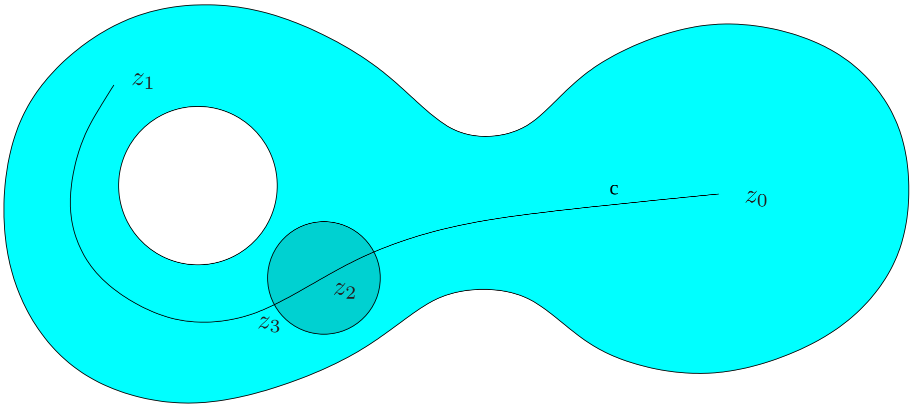
Beweis. Es sei $z_n \to z_\infty$ in U mit $f(z_n) = g(z_n)$ für alle n. Angenommen es gäbe ein $z \in U$ mit $f(z) \neq g(z)$. Wir wählen eine C^1-Kurve $c : [0,1] \to U$ mit $c(0) = z_\infty$ und $c(1) = z$. Dann ist $A := \{t \in [0,1] : f(c(t)) \neq g(c(t))\} \neq \emptyset$ nach unten beschränkt. Sei $\alpha := \inf(A)$. Es ist $\alpha > 0$, denn nach (6.6.12) und (4.2.18a) ist $f = g$ lokal um z_∞ also $f(c(t)) = g(c(t))$ für alle t nahe 0. Für alle $t < \alpha$ ist $f(c(t)) = g(c(t))$ und somit wieder nach (6.6.12) und (4.2.18a) $f(c(t)) = g(c(t))$ für alle t nahe α, ein Widerspruch zur Konstruktion von α. □

6.6.16 Maximumprinzip. $\mathbb{C}$-differenzierbare Funktionen haben ihr Maximum am Rand.
Es sei f auf einer zusammenhängenden offenen Menge $\mathbb{C}$-differenzierbar und nicht konstant, dann besitzt $|f|$ kein Maximum.

Beweis. Es genügt zu zeigen, daß eine $\mathbb{C}$-differenzierbare Funktion $f : \mathbb{C} \supseteq U \to \mathbb{C}$, die ein Maximum z_0 besitzt, konstant ist. Sei also $|f(z)| \leq |f(z_0)|$ für alle $z \in U$. Wir zeigen zuerst, daß $|f(z)| = |f(z_0)|$. Dazu nehmen wir indirekt an, daß ein z_1 existiert mit $|f(z_1)| < |f(z_0)|$. Wir verbinden z_0 und z_1 mit einem Polygonzug c, und nehmen t_0 maximal, s.d. $|f(c(t_0))| = |f(z_0)|$. Sei $z_2 := c(t_0)$ Dann existiert beliebig nahe ein z_3 mit $\varepsilon := |f(z_2)| - |f(z_3)| > 0$. Sei K der Kreis mit Mittelpunkt z_2 und Radius $r := |z_3 - z_2|$. Dann ist $|f(z)| < |f(z_2)| - \varepsilon/2$ für z nahe z_3 (auf einem Bogen der Länge $\delta > 0$) und somit ist

$$|f(z_2)| = \frac{1}{2\pi}\Big|\int_{\partial K} \frac{f(\zeta)}{\zeta - z_2}\,d\zeta\Big| \leq \frac{1}{2\pi r}\Big(\delta(|f(z_2)| - \varepsilon/2) + (2\pi r - \delta)|f(z_0)|\Big) = |f(z_0)| - \frac{\delta\varepsilon}{4\pi r} < |f(z_0)|,$$

ein Widerspruch.



Da f nicht konstant (gleich 0) ist, ist $c := |f(z_0)| > 0$. Sei $f = u + i\,v$ und somit $c^2 := |f(z_0)|^2 = |f(z)|^2 = u(z)^2 + v(z)^2 > 0$. Dann ist

$$\partial_1 u \cdot u + \partial_1 v \cdot v = 0$$
$$\partial_2 u \cdot u + \partial_2 v \cdot v = 0$$

Da f nicht konstant ist gibt es nach (6.5.5) ein z_1 mit $f'(z_1) \neq 0$ und somit

$$0 \neq \det f'(z_1) = \partial_1 u(z_1) \cdot \partial_2 v(z_1) - \partial_2 u(z_1) \cdot \partial_1 v(z_1).$$

Folglich hätte obiges lineares Gleichungssystem an der Stelle z_1 die eindeutige Lösung $u(z_1) = v(z_1) = 0$. Wegen $u(z_1)^2 + v(z_1)^2 = c > 0$ ist dies unmöglich. □

6.6.18 Satz von Liouville. Beschränkte globale $\mathbb{C}$-differenzierbare Fkt. sind konstant.
Jede beschränkte $\mathbb{C}$-differenzierbare Funktion auf ganz $\mathbb{C}$ ist konstant.

Beweis. Aus $\|f\|_\infty < \infty$ und der Cauchy'schen Formel (6.6.12) für einen Kreis um z mit Radius r folgt $|f'(z)| \leq \frac{\|f\|_\infty}{r}$. Mit $r \to \infty$ folgt $f' = 0$ und somit ist f nach (6.5.5) konstant. □

Als weitere Folgerung können wir einen zweiten Beweis des Fundamentalsatzes der Algebra liefern:

2. Beweis von (3.4.8). Es sei p ein nicht-konstantes Polynom. Angenommen p hat keine Nullstelle. Dann ist $\frac{1}{p} : \mathbb{C} \to \mathbb{C}$ $\mathbb{C}$-differenzierbar und wegen (3.4.5) ist $\lim_{|z|\to\infty} \frac{1}{p(z)} = 0$, also ist $\frac{1}{p}$ beschränkt und nach (6.6.18) konstant, ein Widerspruch. □

Literatur

[1] Herbert Amann and Joachim Escher. *Analysis II.* Grundstudium Mathematik. Birkhäuser, Basel - Boston - Berlin, 1999.

[2] Herbert Amann and Joachim Escher. *Analysis III.* Grundstudium Mathematik. Birkhäuser, Basel - Boston - Berlin, 2001.

[3] Herbert Amann and Joachim Escher. *Analysis I.* Grundstudium Mathematik. Birkhäuser, Basel - Boston - Berlin, 2002.

[4] Christian Blatter. *Analysis 1*, volume 151 of *Heidelberger Taschenbücher.* Springer, Berline - Heidelberg - New York, 1974.

[5] Christian Blatter. *Analysis 2*, volume 152 of *Heidelberger Taschenbücher.* Springer, Berline - Heidelberg - New York, 1974.

[6] Christian Blatter. *Analysis 3*, volume 153 of *Heidelberger Taschenbücher.* Springer, Berline - Heidelberg - New York, 1974.

[7] Richard Courant. *Vorlesungen über Differential- und Integralrechnung 1.* Springer, Berlin - Heidelberg - New York, 4 edition, 1971.

[8] Richard Courant. *Vorlesungen über Differential- und Integralrechnung 2.* Springer, Berlin - Heidelberg - New York, 4 edition, 1971.

[9] Jean Dieudonné. *Foundations of modern analysis, I.* Academic Press, New York – London, 1960.

[10] Michal Elconin. ??? *Acta Mathematica*, 68:71–107, 1937.

[11] Ernst Henze. *Einführung in die Maßtheorie*, volume 505 of *B.I.Hochschultaschenbücher.* Bibliographisches Institut, Mannheim, 1971. 7

[12] Harro Heuser. *Lehrbuch der Analysis, Teil 1.* Mathematische Leitfäden. Teubner, Stuttgart, 1980.

[13] Harro Heuser. *Lehrbuch der Analysis, Teil 2.* Mathematische Leitfäden. Teubner, Stuttgart, 1981. 35, 120

[14] Andreas Kriegl. *Differentialgeometrie.* Universität Wien, 1996. 63

[15] Andreas Kriegl. *Topologie 1.* Universität Wien, 2000. 61

[16] Andreas Kriegl. *Funktionalanalysis 1.* Universität Wien, 2002. 15, 42, 48, 49

[17] Andreas Kriegl. *Funktionalanalysis 2.* Universität Wien, 2002. 128, 131, 133

[18] Andreas Kriegl, Mark Losik, and Peter W. Michor. Choosing roots of polynomials smoothly, II. *Israel J.M.*, 139:183–188, 2004. 93

[19] John C. Oxtoby. *Maß und Kategorie.* Hochschultext. Springer, Berline - Heidelberg - New York, 1971. 11

[20] Ernst Peschl. *Funktionentheorie*, volume 131 of *B.I.Hochschultaschenbücher.* Bibliographisches Institut, Mannheim, 1967. 152

[21] Lang Serge. *Differentiable Manifolds.* Addison-Wesley, 1962.

[22] Ralph Stöcker and Heiner Zieschang. *Algebraische Topologie.* Mathematische Leitfäden. Teubner, Stuttgart, 1988. 63, 140

[23] Lee Peng Yee and Rudolf Výborný. *The integral: An Easy Approach after Kurzweil and Henstock*, volume 14 of *Australian Mathematical Society Lecture Series.* Cambridge University Press, Cambridge, 2000. 18

Liste der Symbole

$B(X,E)$	Raum der beschränkten Abbildungen, Seite 40
$B(X,\mathbb{K})$	Raum der beschränkten Funktionen, Seite 36
$B(X,\mathbb{K}^n)$	Raum der beschränkten vektorwertigen Funktionen, Seite 36
$C(X,E)$	Raum der stetigen Abbildungen, Seite 40
$C(X,\mathbb{K})$	Raum der stetigen Funktionen, Seite 36
$C(X,\mathbb{K}^n)$	Raum der stetigen vektorwertigen Funktionen, Seite 36
$C^1(I,\mathbb{R})$	Raum der stetig differenzierbaren reell-wertigen Funktionen auf einem Intervall I, Seite 17
$L(E,F)$	Raum der stetig linearen Abbildungen, Seite 43
$L(E_1,E_2;F)$	Raum der bilinearen stetigen Abbildungen $E_1 \times E_2 \to F$, Seite 46
$R(I,\mathbb{R})$	Raum aller $\mathbb{R}$-wertigen integrierbaren Funktionen auf einem Intervall I, Seite 12
$R([a,b],\mathbb{R})$	Raum der integrierbaren Funktionen, Seite 36
ℓ^1	Raum der absolut konvergenten Reihen, Seite 36
ℓ^2	Raum der quadratisch konvergenten Reihen, Seite 36
ℓ^∞	Raum der beschränkten Folgen, Seite 36
c	Raum der konvergenten Folgen, Seite 36
C^1	stetig differenzierbar, Seite 77
C^∞	beliebig oft (stetig) differenzierbar, Seite 115
C^p	p-fach stetig differenzierbar, Seite 115
D^p	p-fach differenzierbar, Seite 115
$E^* := L(E,\mathbb{K})$	Raum der stetigen linearen Funktionale, Seite 42
$E_1 \times \ldots \times E_n$	Produkt normierter Räume, Seite 40
$\operatorname{Inv}(A)$	Menge der invertierbaren Elemente (einer Banach-Algebra A), Seite 47
$L(E,F_1 \times F_2) \cong L(E,F_1) \times L(E,F_2)$	Komponentenzerlegung im Wertebereich, Seite 72
$L(E_1,E_2;F) \cong L(E_1,L(E_2,F))$	Exponentialgesetz für multilineare Abbildungen, Seite 46
$L(E_1,\ldots,E_n;L(E_{n+1},F)) \cong L(E_1,\ldots,E_{n+1};F)$	Exponentialgesetz multilinearer Abbildungen, Seite 116
$L(E_1 \times E_2,F) \cong L(E_1,F) \times L(E_2,F)$	Zerlegung nach den Argumenten, Seite 74
$L(E_1 \times \ldots \times E_m, F_1 \times \ldots \times F_n) \cong \prod_{j=1}^n \prod_{i=1}^m L(E_i,F_j)$	Komponentenzerlegung linearer Abbildungen, Seite 76
$L(\mathbb{R},F) \cong F$	Geraden via Richtungsvektor, Seite 66
$L_{\mathrm{sym}}(E,E;\mathbb{R}) \cong L_{\mathrm{sym}}(E,E)$	Exponentialgesetz symmetrischer multilinearer Abbildungen, Seite 133
$\mathbb{R} \times \mathbb{R}^{m-1} \cong \mathbb{R}^m$	Assoziativität des Produkts, Seite 72
$C(X \times I,E) \cong C(X,C(I,E))$	Exponentialgesetz für stetige Abbildungen und kompaktes I, Seite 60

$C(I, F_1 \times F_2) \cong C(I, F_1) \times C(I, F_2)$	Funktionenräume stetiger Abbildungen sind Produkterhaltend, Seite 82
$E \cong E^*$	Isomorphismus eines Hilbert-Raums mit seinem Dualraum, Seite 48
$\ln$	Natürliche Logarithmusfunktion zur Basis e, Seite 23
$\log_a$	Logarithmusfunktion zur Basis a, Seite 24
a^x	Potenz mit Basis a, Seite 24
$\sinh$	Sinushyperbolicus Funktion, Seite 24
$\cosh$	Cosinushyperbolicus Funktion, Seite 24
$\tanh$	Tangenshyperbolicus Funktion, Seite 24
Arsinh	Areasinushyperbolicus, Umkehrfunktion von sinh, Seite 26
Arcosh	Areacosinushyperbolicus, Umkehrfunktion von cosh, Seite 26
Artanh	Areatangenshyperbolicus, Umkehrfunktion von tanh, Seite 26
$\|{-}\|$	Operatornorm für bilineare Abbildungen, Seite 46
$\|{-}\|$	Operatornorm für lineare Abbildungen, Seite 42
$\|{-}\|_1$	1-Norm, Seite 36
$\|{-}\|_2$	2-Norm, Seite 36
$\|{-}\|_\infty$	Supremums oder ∞-Norm, Seite 40
$\mathrm{ev} : (f, x) \mapsto f(x)$	Evaluationsabbildung, Seite 61
$\mathrm{ev}_x : f \mapsto f(x)$	Punktevaluation bei x, Seite 40
inj_j	lineare Injektionsabbildung, Seite 74
ins_j	lineare Insertionsabbildung auf den j-ten Faktor, Seite 72
ins_j	lineare Insertionsabbildung, Seite 74
ins_j^x	affine Insertionsabbildung, Seite 75
$\mathrm{inv} : a \mapsto a^{-1}$	Inversionsabbildung (einer Banach-Algebra A), Seite 47
$\mathrm{komp} : (f, g) \mapsto f \circ g$	Kompositionsabbildung, Seite 77
pr_j	Projektion auf j-te Koordinate bzw. j-ten Faktor, Seite 41
pr_j	Projektion auf die j-te Koordinate bzw. den j-ten Faktor, Seite 40
Γ	Euler'sche Gammafunktion, Seite 34
∂_i	i-te partielle Ableitung, Seite 67
∂_i	i-te partielle Ableitung, Seite 75
$\frac{\partial}{\partial x}$	partielle Ableitung nach x, Seite 67
dx_i	Ableitung der i-ten Koordinatenprojektion, Seite 140
$dz = dx + i\,dy$	Basiselement der komplexen 1-Formen, Seite 150
$O(f)$	Infimum der Obersummen einer Funktion f, Seite 1

$O(f, Z)$	Obersumme der Funktion f bzgl. der Zerlegung Z, Seite 1
$S(f, Z, \xi)$	Riemann-Summe, Seite 52
$S_c(f, Z, \xi)$	Riemann-Stieltjes Summe einer 1-Form bzgl. einer Kurve und einer punktierten Zerlegung, Seite 145
$U(f)$	Supremum der Untersummen einer Funktion f, Seite 1
$U(f, Z)$	Untersumme der Funktion f bzgl. der Zerlegung Z, Seite 1
$[f]$	Matrixdarstellung einer linearen Abbildung f zwischen Produkten, Seite 76
$[f]_{B,C}$	Matrixdarstellung einer linearen Abbildung a bzgl. Basen B und C, Seite 44
$[x]_B$	Komponentendarstellung eines Vektors x bzgl. einer Basis B, Seite 43
$\Omega_f(Z)$	Oszillation bzgl. einer Zerlegung Z, Seite 54
$\operatorname{grad} f\ (x)$	Gradient von f an der Stelle x, Seite 131
$\int f$	(Familie der) Stammfunktion(en) von f, Seite 18
$\int f$	Stammfunktion und unbestimmtes Integral von f, Seite 17
$\int_I f = \int_a^b f$	Riemann-Integral, Seite 53
$\int_I f = \int_a^b f(x)\,dx$	Darboux-Integral von f über das Intervall I, Seite 2
$\int_a^{+\infty} f$	uneigentliches Integral, Seite 30
$\int_c f$	Kurvenintegral von f längs c, Seite 142
$\int_c f = \int_c f(z)\,dz$	komplexes Kurvenintegral, Seite 150
$\partial_{i_1} \dots \partial_{i_p} f$	iterierte partielle Ableitung von f, Seite 117
$\frac{df}{dz}(z)$	komplexe Ableitung von f an der Stelle z, Seite 148
$\operatorname{grad} f$	Gradient von f, Seite 131
$\operatorname{grad} f$	Gradientenfeld von f, Seite 132
$d_v f$	Richtungsableitung von f in Richtung v, Seite 64
$dg = g'$	exakte 1-Form, Seite 140
f'	Ableitung der Abbildung f, Seite 66
f''	Zweite Ableitung von f, Seite 114
$f'(x_0)$	Ableitung von f an der Stelle x_0, Seite 66
$f^{(n)}$	n-te Ableitung von f, Seite 116
$\check{f} : x \mapsto (y \mapsto f(x, y))$	Assoziierte funktionenwertige Abbildung, Seite 45
$\hat{g} : (x, y) \mapsto g(x)(y)$	Assoziierte Abbildung am Produkt, Seite 45
$f^* : g \mapsto g \circ f$	adjungierter Operator, Seite 42
$f_* : g \mapsto f \circ g$	Kovariant assoziierte Abbildung, Seite 61
$f_1 \times f_2$	Produkt zweier Abbildungen $f_1 : X_1 \to Y_1$ und $f_2 : X_2 \to Y_2$, Seite 33
$V(c)$	Variation von c, Seite 146

$V(c, Z)$	Variation von c bzgl. Zerlegung Z, Seite 146
$c_1 \odot \cdots \odot c_n$	Aneinanderhängung von Kurven, Seite 142
$\lvert Z \rvert$	Feinheitsmaß einer Zerlegung Z, Seite 53
$\bigcup \mathcal{W}$	Vereinigung einer Menge $\mathcal{W}$ von (vielen) Mengen, Seite 3
∂K	(parametrisierte) Rand einer Menge, einer Kreisscheibe, Seite 152
$d(A, B) := \inf\{d(a, b) : a \in A, b \in B\}$	Distanz zweier Teilmengen eines metrischen Raums, Seite 136
$h^n = (h, \ldots, h)$	n-Tupel gleicher Vektoren, Seite 123

Index

1-Form, 139
  exakte, 139
  geschlossene, 140
  skalarwertige, 140
  vektorwertige, 139
1-dimensionale inhomogene lineare Differentialgleichung mit nicht-konstanten Koeffizienten, 105
1-dimensionale homogene lineare Differentialgleichung mit nicht-konstanten Koeffizienten, 105

Abbildung
  holomorphe, 148
  komplex-lineare, 148
Ableitung
  iterierte partielle, 117
  komplexe, 148
  partielle, 75
  zweite, 114
Ableitung einer Abbildung, 66
  höhere, 115
Ableitung einer Kurve, 50
Ableitung von f bei x_0, 66
absolut summierbare Folge, 36
absolut-konvergente Reihe, 39
adjungierter Operator f^*, 42
affine Insertionsabbildung, 75
algebraischen Topologie, 63
äquidistante Zerlegung, 2
Arbeit bei konstanter Kraft, 147
Arbeit in einem Kraftfeld, 147
Areacosinushyperbolicus, 26
Areasinushyperbolicus, 26
Areatangenshyperbolicus, 26
äußere Maß einer Teilmenge von $\mathbb{R}$, 7

Banach-Algebra, 46
Banach-Raum, 38
  komplexer, 148
beschränkte Variation einer Kurve, 146
bestimmte Integral der Umkehrfunktion, 22
bestimmtes Darboux-Integral, 2
bilinear-Form
  negativ definite symmetrische, 128
  positiv definite symmetrische, 127
biquadratische Gleichung, 87

$\mathbb{C}$-differenzierbare Abbildung, 148
$\mathbb{C}$-lineare Abbildung, 148
Cantor'sche Diskontinuum, 7
Cauchy'sche Ableitungsformeln, 153
Cauchy'sche Integralformel, 152
Cauchy'scher Integralsatz, 151
Cauchy-Kriterium für R-Integrierbarkeit, 53
Cauchy-Kriterium für uneigentliche Integrale, 33
Cauchy-Riemann'schen Differentialgleichungen, 149
Cosinushyperbolicus, 24

D-integrierbar, 1
Darboux-integrierbar, 1
Darboux-Integral
  bestimmtes, 2
Definitheit jeder Norm, 37
Diffeomorphismus
  lokaler, 85
Differentialform
  exakte, 139
  geschlossene, 140
  skalarwertige, 140
Differentialform der Ordnung 1
  vektorwertige, 139
Differentialgeometrie, 63
Differentialgleichung
  partielle, 138
  totale, 138
Differentialgleichung n-ter Ordnung, gewöhnliche, 107
Differentialgleichung 1-ter Ordnung
  gewöhnliche, 99
  zeitunabhängige, 95
Differentialgleichung mit nicht-konstanten Koeffizienten
  lineare inhomogene, 105
Differentialgleichungen
  System partieller, 138
differenzierbar
  stetig, 77
differenzierbar bei x_0, 66
differenzierbare Abbildung, 66
  mehrmals, 115
  mehrmals stetig, 115
  zweimal, 114
Differenzierbarkeit einer Kurve, 50
Dirichlet'sche Sprungfunktion, 6
Dreiecksungleichung jeder Norm, 37

einfach geschlossene Kurve, 63
Euler'sche Gammafunktion, 34
exakte 1-Form, 139
Exponentialfunktion, 24
Exponentialgesetz für stetige Funktionen, 60
Extremum einer skalarwertigen Funktion, 127

fast überall, 5
Feinheitsmaß einer Zerlegung, 53
Fixpunkt der Differentialgleichung, 105
Funktionenräume, 36
Funktionentheorie, 152

Gammafunktion
  Euler'sche, 34
generischen Fall, 105
geschlossene 1-Form, 140
gewöhnliche Differentialgleichung n-ter Ordnung, 107
gewöhnliche Differentialgleichung 1-ter Ordnung, 99
  zeitunabhängige, 95
gewöhnliche zeitunabhängige Differentialgleichung 1-ter Ordnung, 95
glatte Abbildung, 115
Gradient einer skalarwertigen Funktion, 131
Gradient einer skalarwertigen Funktion am $\mathbb{R}^n$, 131
Gradientenfeld, 132

Höhenlinie, 131
höhere Ableitung einer Abbildung, 115
Hölder'sche Ungleichung, 128
Halbachse, 137
Hauptachse, 137
Hauptsatz der Analysis für Kurven, 59
Hauptsatz der Differential und Integralrechnung, 17
Hesse'sche Matrix, 115
Hilbert-Raum, 47
holomorphe Abbildung, 148
homöomorphes Bild, 11
homogene lineare 1-dimensionale Differentialgleichung mit konstanten Koeffizienten, 102
homogene lineare mehrdimensionale Differentialgleichungen mit konstanten Koeffizienten, 102

Identitätssatz für komplex-differenzierbare Funktionen, 154
Injektion, 74
Insertion, 74
Insertionsabbildung
  affine, 75
Integrabilitätsbedingung für Existenz einer Stammfunktion, 138
Integrabilitätsbedingung für Lösbarkeit der totalen Differentialgleichung, 138
Integral, 2
  bestimmtes Darboux-, 2
  Darboux-, 2
  Kurven-, 142
  Kurzweil-, 18
  Riemann-, 53
  uneigentliches, 30
Integral einer skalarwertigen Funktion
  unbestimmtes, 16, 18
Integraltest für Reihenkonvergenz, 33
Integration
  partielle, 19
integrierbare skalarwertige Funktion, 1
  Kurzweil-, 18
Inversion einer Banach-Algebra, 47
invertierbares Element, 47
Isobaren, 89
Isothermen, 89
iterierte partielle Ableitung, 117

Jacobi-Matrix, 76
Jordan'sche Kurvensatz, 63
Jordan'sche Normalform, 105
Julia-Menge, 112

kartesische Blatt, 90
Kettenregel, 71
Kettenregel für lineare äußere Abbildung, 50
Knotentheorie, 63
komplexe Ableitung, 148
komplexe Analysis, 152
komplexer Banach-Raum, 148
Komponentendarstellung einer linearen Abbildung zwischen Produkten, 76
konservatives Kraftfeld, 147
konservatives Vektorfeld, 143
  Potential eines, 147
konvexe Teilmenge, 48, 58
Kraftfeld
  konservatives, 147
kritischer Punkt
  nicht-degenerierter, 132
kritischer Punkt einer Abbildung, 127
Kugelkoordinaten, 88
Kurve
  einfach geschlossene, 63
Kurvenintegral, 142
  wegunabhängiges, 142
Kurvenintegral längs nicht differenzierbarer Kurven, 145
Kurzweil-Integral, 18
Kurzweil-integrierbare skalarwertige Funktion, 18

Länge einer Kurve, 146
Lagrange-Multiplikator, 134
Lebesgue'schen Integrationstheorie, 15
Lebesgue'sches Integrabilitätskriterium, 5
Lebesgue-Integrabilitätskriterium, 55
Lebesgue-Nullmenge, 6

Lebesgue-Zahl einer Überdeckung, 4
Leibniz'sche Produktregel, 73
lineare Differentialgleichung mit konstanten Koeffizienten
  homogene, 102
  homogene mehrdimensionale, 102
lineare Differentialgleichung mit nicht-konstanten Koeffizienten
  homogene, 105
  inhomogene, 105
lineares Funktional
  stetiges, 42
Lipschitz Abbildung, 41
Lipschitz-Abbildung, 100
Logarithmus
  natürlicher, 23
lokaler C^1-Diffeomorphismus, 85
lokales Maximum, 127
  striktes, 127
lokales Minimum, 127
  striktes, 127
lokales striktes Maximum, 127
lokales striktes Minimum, 127

Mannigfaltigkeiten, 63
Matrixdarstellung der bilinear-Form, 115
Matrixdarstellung einer linearen Abbildung zwischen Produkten, 76
Maximum
  lokales, 127
  lokales striktes, 127
Maximumprinzip, 154
Mehrfachnullstelle eines Polynoms, 92
mehrmals differenzierbare Abbildung, 115
mehrmals stetig differenzierbare Abbildung, 115
Methode der Lagrange Multiplikatoren, 134
Minimum
  lokales, 127
  lokales striktes, 127
Mittelwertsatz der Differentialrechnung, 60
Mittelwertsatz der Integralrechnung, 14
Mittelwertsatz für integrierbare Kurven, 58
Modul, 140
Morse-Palais-Lemma, 132

natürlicher Logarithmus, 23
Nebenbedingung, 134
negativ definite symmetrische Form, 128
Newton'sche Kraftgesetz, 108
Newton-Verfahren, 111
nicht-degenerierter kritischer Punkt, 132
Niveaulinien, 89
normierte Räume, 37

Obersumme bzgl. einer Zerlegung, 1
offene Teilmenge, 3
offene Umgebung eines Punktes, 85
Operator
  adjungierter, 42
Operatornorm, 42
ortsunabhängiges Vektorfeld, 99
Oszillation einer Funktion bzgl. einer Zerlegung, 54

Parameterabhängigkeit von Differentialgleichungen, 98
Partialbruchzerlegung im Komplexen, 27
Partialbruchzerlegung im Reellen, 28
partielle Ableitung, 67, 75
  iterierte, 117
partielle Integration, 19
Polarkoordinaten, 87
Polygonzug, 142
positiv definit, 127
Positive Homogenität jeder Norm, 37
Potential eines konservativen Vektorfelds, 143, 147
Prinzip der gleichmäßigen Beschränktheit, 48
punktierte Zerlegung, 53

quadratisch konvergenten Reihen, 36
quadratisch summierbare Folge, 36
quadratische Form
  nicht degenerierte, 129
quadratische nicht degenerierte Form, 129
Quadrieren, 86
Quadrik, 137
Quelle einer Differentialgleichung, 106

R-integrierbar, 53
R-Modul, 140
rektifizierbare Kurve, 146
Richtungsableitung, 64
Riemann-Integral, 53
Riemann-integrierbar, 53
Riemann-Stieltjes-Summe, 145
Riemann-Stieltjes-Integral, 145
Riemann-Summe, 52, 53

Sattelpunkt einer skalarwertigen Funktion, 129
Sattelpunkte einer Differentialgleichung, 106
Satz über implizite Funktionen, 89
Satz über inverse Funktionen, 85
Satz über offene Abbildungen, 49
Satz von Banach-Steinhaus, 48
Satz von Frobenius, 138
Satz von Hahn-Banach, 48
Satz von Liouville, 155
Satz von Morera, 152, 154
Satz von Picard-Lindelöf, 100
Satz von Schönflies, 63
Satz von Schwarz, 121

Separation der Variablen, 97
Sinushyperbolicus, 24
skalarwertige 1-Form, 140
skalarwertige Differentialform, 140
Stammfunktion einer skalarwertigen Funktion, 17
stationäre Lösung einer Differentialgleichung, 105
sternförmige Teilmenge, 140
stetig differenzierbar, 77
stetig differenzierbare Abbildung
    mehrmals, 115
stetig lineares Funktional, 42
striktes Maximum
    lokales, 127
striktes Minimum
    lokales, 127
Substitutionsformel für bestimmte Integrale, 21
Substitutionsformel für Kurven, 60
Substitutionsformel für unbestimmte Integrale, 20
symmetrische Form
    negativ definite, 128
    positiv definite, 127
System partieller Differentialgleichungen, 138

Tangenshyperbolicus, 24
Taylor'sche Formel, 123
Teufelstreppe, 8
Topologie
    algebraische, 63
Torus, 93
totale Differentialgleichung, 138
totales Differential einer Abbildung, 140
Trägheitssatz von Silvester, 133

Überdeckung, 3
    Lebesgue-Zahl einer, 4
    Verfeinerung einer, 4
    Vitali-, 15
Überdeckungssatz von Heine-Borel, 5
Überdeckungssatz von Lebesgue, 4
Umkehrfunktion
    bestimmte Integral der, 22
    unbestimmte Integral der, 22
unbestimmte Integral der Umkehrfunktion, 22
unbestimmtes Integral einer skalarwertigen Funktion, 16
unbestimmtes Integral einer skalarwertigen Funktion, 18
uneigentliches Integral, 30
Untersumme bzgl. einer Zerlegung, 1

Van der Waerden's Funktion, 8
Variation der Konstanten, 103, 105
Variation einer Kurve, 146
    beschränkte, 146
Vektorfeld
    konservatives, 143
    ortsunabhängiges, 99
    Potential eines konservativen, 143, 147
    zeitabhängiges, 99
    zeitunabhängiges, 99
vektorwertige 1-Form, 139
vektorwertige Differentialform der Ordnung 1, 139
Vereinfachtes Newton-Verfahren, 112
Vereinigung einer Menge von Mengen, 3
Verfeinerung einer Überdeckung, 4
Vergleichstest für uneigentliche Integrale, 32
Vitali-Überdeckung, 15

Wegunabhängigkeit eines Kurvenintegrals, 142

zeitabhängiges Vektorfeld, 99
zeitunabhängiges Vektorfeld, 99
Zerlegung
    äquidistante, 2
    Obersumme bzgl. einer, 1
    punktierte, 53
    Untersumme bzgl. einer, 1
Zerlegung eines Intervalls, 1
zusammenhängende Menge, 142
zweimal differenzierbare Abbildung, 114
zweite Ableitung, 114
Zwischenvektor einer Zerlegung, 52, 53