

Addendum zu Analysis 2 aus dem Sommersemester 2004

Andreas Kriegl

March 25, 2004

Dies ist der aktuelle Stand meiner Vorlesung Analysis 2 im Sommersemester 2004. Ich werde diesen Text wie im vergangenen Semester laufend im WWW aktualisieren und am Ende der Vorlesung eine komplette Version bereit stellen.

Diejenigen Nummern, wo nichts dabei steht, habe ich wie im Skriptum Analysis 1 vorgetragen. Dort wo die Nummer ein vorgesetztes 'Ad' trägt, sind dies Ergänzungen die ich in der Vorlesung zum entsprechenden Teil im Skriptum gemacht habe. Dort wo eine Nummer ohne vorgesetztes 'Ad' steht, bin ich in der Vorlesung wesentlich vom Skriptum abgewichen und ersetzt der hier vorliegende Text die entsprechende Passage aus dem Skriptum.



Ich habe diejenigen Teile, die über die Vorlesung hinausgehen und für jene gedacht sind, die keine Angst habe zeitweilig ein wenig den Boden unter den Füßen zu verlieren und in höhere Sphären aufzusteigen, durch linkseitiges Symbol eines Hängegleiters gekennzeichnet.



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

Andreas Kriegl, Wien im März 2004

Ad Kapitel 5

Ad (5.1.1)

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 ein- und umschriebener Vielecken erreicht. Ganz ähnlich gehen wir nun vor, um die durch eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ 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 $a = t_0 \leq t_1 \leq \dots \leq t_{N_1} \leq t_N = b$ vermöge $I_1 := [t_0, t_1], \dots, I_i := [t_{i-1}, t_i], \dots, I_N := [t_{N_1}, t_N]$. Wir werden (logisch nicht völlig korrekt) zwischen der Interpretation der Zerlegung Z als $Z = \{I_1, \dots, I_N\}$ und als $Z = \{t_0, \dots, t_N\}$ je nach Bedarf hin- und herwechseln.

Hierher gehört die Definition von $O(f, Z)$, $U(f, Z)$, $O(f)$ und $U(f)$ aus dem Skriptum.

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.

Wir haben anstelle der Bezeichnung Riemann-integrierbar aus dem Skriptum den Namen Darboux-integrierbar (kurz D -integrierbar) verwendet, da wir die Äquivalenz dieser beiden Begriffe später zeigen wollen. Im Skriptum war dies ursprünglich nicht vorgesehen, und deshalb der gebräuchlichere Begriff der Riemann-Integrierbarkeit verwendet. In dieser Vorlesung könnten wir auch kurz von Integrierbarkeit sprechen, aber es gibt auch andere wichtige nicht äquivalente Begriffe (wie Lebesgue-integrierbar und Henstock-integrierbar).

Es ist $O(f) \geq U(f)$, 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 $U(f, Z_1) \leq U(f, Z) \leq O(f, Z) \leq O(f, Z_2)$ und somit $U(f) = \sup\{U(f, Z_1) : Z_1\} \leq O(f, Z_2)$ und schließlich $U(f) \leq \inf\{O(f, Z_2) : Z_2\} = O(f)$.

Es ist $O(f) = U(f)$ genau dann, wenn

$$\forall \varepsilon > 0 \exists Z_1, Z_2 \text{ Zerlegungen : } O(f, Z_1) - U(f, Z_2) < \varepsilon.$$

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 $0 \leq O(f) - U(f) \leq O(f, Z_2) - U(f, Z_2) < \varepsilon$ für gewisse Zerlegungen Z_1 und Z_2 und somit ist $O(f) = U(f)$. \square

Ad (5.1.2)

Ad (5.1.3)

Monotone Funktionen können natürlich in einzelnen Punkten unstetig sein, denn zwar existiert $\lim_{x \rightarrow a-} f(x)$, da f monoton und durch $f(a)$ beschränkt ist, und ebenso existiert $\lim_{x \rightarrow a+} f(x) \geq f(a) \geq \lim_{x \rightarrow a+} f(x)$, also kann dennoch 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. überabzählbar viele mindestens ε sind, und deren gesamte Sprunghöhe wäre somit $\geq \infty \cdot \varepsilon = \infty$, ein Widerspruch.

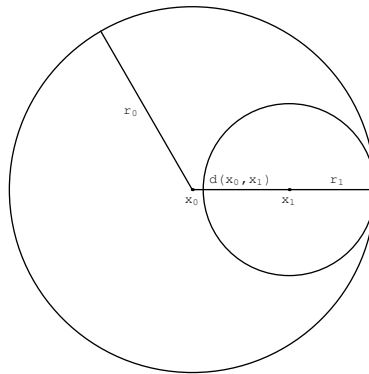
Ad (5.1.4)

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.

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:

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

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:

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

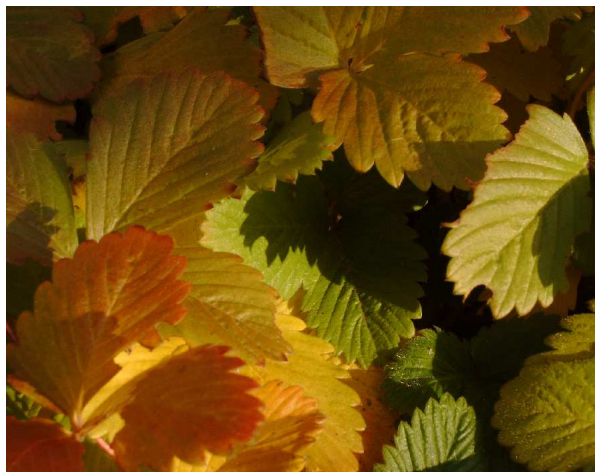
5.1.4a Überdeckungssatz von Lebesgue.

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.

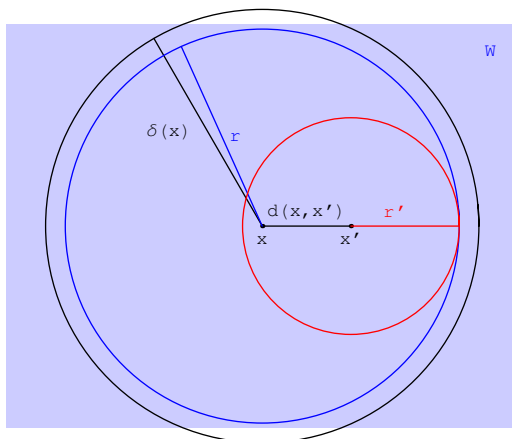


Beweis. Sei also \mathcal{W} eine Überdeckung von X mit offenen Teilmengen. Wegen $X = \bigcup \mathcal{W}$ können wir für jedes $x \in X$ ein $W \in \mathcal{W}$ wählen mit $x \in W$. Da $W \in \mathcal{W}$ offen ist existiert ein $\delta > 0$ mit $U_\delta(x) \subseteq W$. Es sei $\delta(x) := \sup\{0 < \delta \leq 1 : \exists W \in \mathcal{W} : U_\delta(x) \subseteq W\}$. 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 $\delta : X \rightarrow \mathbb{R}$ 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 also gezeigt, daß die Funktion $\delta : X \rightarrow (0, 1] \subseteq \mathbb{R}$ 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. \square

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

Folgerung.

Es sei $f : X \rightarrow Y$ 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 (5.5.1) 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 somit ist $d(f(x_i), f(x)) < \varepsilon$ für $i \in \{1, 2\}$, also $d(f(x_1), f(x_2)) < 2\varepsilon$. Dies ist die gleichmäßige Stetigkeit von f . \square

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 X .

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 Folge $(x_n) \in X$ 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 Folgenglieder enthält. Andererseits ist $\{U_x : x \in X\}$ eine Überdeckung von X mit offenen Mengen, also existiert eine endlich Teilüberdeckung $\{U_1, \dots, U_N\}$ und damit enthält $X = \bigcup_{i=1}^N U_i$ nur für endlich viele Folgenglieder, ein Widerspruch. \square

5.1.4 Folgerung. Lebesgue’sches Integrabilitätskriterium.

Eine beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ 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 wo $\Omega(f|_I)$ 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 $f : [a, b] \rightarrow \mathbb{R}$ 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. offen!) 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 $O(f, Z) - U(f, Z) < (2\|f\|_\infty + (b-a))\varepsilon$, d.h. f ist Darboux-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. \square

Ad (5.1.6)

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 $f \circ g$ 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 mißverständliche 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 $h \geq f$ und $h \geq g$.

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 $g(x) \neq 0$ 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. \square

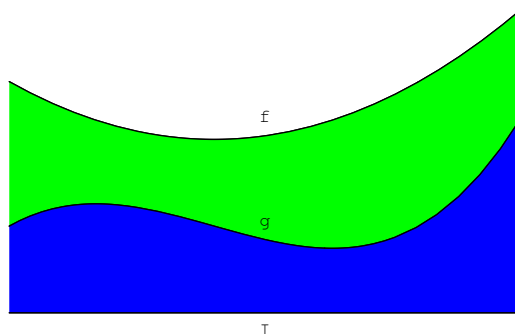
Beachte, daß diese Folgerung (ungleich der Kettenregel (4.1.14) und der Produktregel (4.1.15)) nur die Integrierbarkeit der Zusammensetzung $f \circ g$, der Summe $f + g$ und des Produkts $f \cdot g$, aber keine Formel für $\int_a^b f \circ g$ und $\int_a^b f \cdot g$ liefert. Für die Summe wollen wir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

5.1.7 Lemma. Integrieren ist linear.

Die Menge der D-integrierbaren Funktionen $f : I \rightarrow \mathbb{R}$ auf einem Intervall I bilden einen Vektorraum $R(I, \mathbb{R})$ und integrieren $\int_I : R(I, \mathbb{R}) \rightarrow \mathbb{R}$, $f \mapsto \int_I f$ ist linear, d.h. $\int_I (f + \lambda g) = \int_I f + \lambda \int_I g$ für alle $\lambda \in \mathbb{R}$ und all D-integrierbare Funktionen $f, g : I \rightarrow \mathbb{R}$.

Insbesondere bedeutet dies, daß wir die (orientierte) Fläche A zwischen zwei Funktionen f und g über einen 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 $f : I \rightarrow \mathbb{R}$ 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 $\int_I : R(I, \mathbb{R}) \rightarrow \mathbb{R}$. Sei vorerst $\lambda \geq 0$. Dann ist $U(\lambda f, Z) = \sum_J \inf(f(J)) |J| = \sum_J \lambda \inf(f(J)) |J| = \lambda \sum_J \inf(f(J)) |J| = \lambda U(f, Z)$ 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

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

also ist

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

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

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

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

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

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

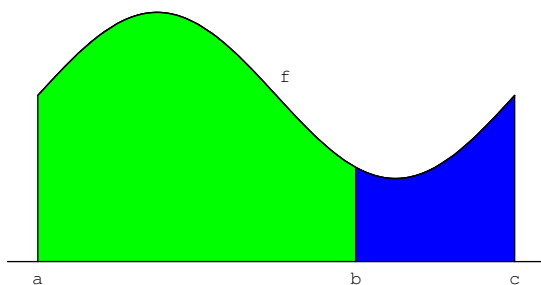
Fehlt noch die Homogenität für $\lambda < 0$. Diese genügt es für $\lambda = -1$ zu zeigen, was aus $0 = \int_I 0 = \int_I (f + (-f)) = \int_I f + \int_I (-f)$ folgt. \square

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. □

Ad (5.1.10)

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

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

Ad (5.1.11)

5.1.12 Proposition. Stetigkeit des Integrierens.

Es sei $f_n : [a, b] \rightarrow \mathbb{R}$ D -integrierbar und f_n konvergiere gegen f_∞ gleichmäßig. Dann ist auch f_∞ D -integrierbar und es gilt

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

Beachte, daß dies einerseits besagt, daß $R(I, \mathbb{R})$ ein abgeschlossener Teilraum des Raumes $B(I, \mathbb{R})$ aller beschränkten Funktionen $I := [a, b] \rightarrow \mathbb{R}$ ist (und somit nach (4.2.9) vollständig ist) und andererseits, daß $\int_I : R(I, \mathbb{R}) \rightarrow \mathbb{R}$ stetig ist (nach (3.1.3)).

Beweis. Nach dem Lebesgue'schen Integrierbarkeitskriterium (5.1.4) ist f_∞ Riemann-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 $|\int_a^b f - \int_a^b g| = |\int_a^b (f-g)| \leq \int_a^b |f-g| \leq |b-a| d_\infty(f, g)$ nach (4.1.10) und (5.1.11) ist das Integrieren \int_I Lipschitz und somit stetig, d.h. $\int_a^b f_n$ konvergiert gegen $\int_a^b f_\infty$ falls $d_\infty(f_n, f_\infty) \rightarrow 0$, also $f_n \rightarrow f_\infty$ gleichmäßig konvergiert. \square



5.1.13 Bemerkung.

Beachte, daß (5.1.2) 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 $f_n(x) \rightarrow 0$ für jedes x , also konvergiert $f_n \rightarrow 0$ 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 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 insbesondere 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 $f_n \leq f_{n+1}$. Für $x \in [a, b] \setminus \Delta(f)$ ist f stetig bei x , also existiert zu $\varepsilon > 0$ ein $\delta > 0$, sodaß $f(x') \geq f(x) - \varepsilon$ für alle x' mit $|x' - x| < \delta$. Für $(b-a)2^{-n-1} < \delta$ liegt x in einen $I \in Z_n$ (und ist nicht der rechte Randpunkt) mit $I \subseteq \{x' : |x' - x| < \delta\}$. Damit ist $\inf(f(I)) \geq f(x) - \varepsilon$, also $f_n(x) \geq f(x) - \varepsilon$. Somit konvergiert $f_n(x) \rightarrow f(x)$ 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 $f_n \leq f_{n+1}$ für alle n für welche f_n gegen eine D -integrierbare Funktion f_∞ fast überall konvergiert, auch $\int_I f_n \rightarrow \int_I f_\infty$ gilt. Indem wir $g_n := f_\infty - f_n$ betrachten, genügt es folgendes zu zeigen:



5.1.14 Theorem.

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 $g_n \rightarrow 0$ fast überall. Dann gilt $\int_I g_n \rightarrow 0$.

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 $\lim_{n \rightarrow \infty} g_n(x) = 0$ 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 $n' := \max\{n_{x_1}, \dots, n_{x_K}\}$. Dann ist $g_n < \varepsilon$ auf $\bigcup_{i=1}^N U_{x_i}$ für alle $n \geq n'$. Für offene Intervalle $J = (a', b')$ bezeichnen wir mit $\int_J f$ das Integral $\int_{a'}^{b'} f$. Somit ist

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

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

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 $\chi_{\mathbb{Q}}$.

Ad (5.2.2)

Wir haben im Beweis 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 aussagen: Wenn eine Abbildung $F : I \rightarrow \mathbb{R}$ 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

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

und somit

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

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 dies das unbestimmte Integral von f (obwohl es nicht immer durch Integration erhalten werden kann). Der Hauptsatz der Analysis besagt dann unter anderen, daß das bestimmte Integral D -integrierbarer Funktionen f wie folgt berechnet werden kann:

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

Beachte jedoch, daß die klassische und weit verbreitete Schreibweise $\int f(x) dx$ für den Funktionswert $(\int f)(x)$ – 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.

Ad (5.2.3)

Da Integrieren (fast) die Inverse zum Differenzieren ist und letzteres linear ist, gilt gleiches auch für das unbestimmte Integral von stetigen Funktionen. Für das bestimmte Integral haben wir die Linearität in (5.1.7) gezeigt. Daraus folgt die entsprechende Formel für das unbestimmte Integral $\int f : x \mapsto \int_a^x f$ integrierbarer Funktionen. Schließlich können wir dies aber sogar für Funktionen die eine Stammfunktion besitzen zeigen, denn wegen $(\int f + \lambda \int g)' = (\int f)' + \lambda (\int g)' = f + \lambda g$ ist $\int f + \lambda \int g$ Stammfunktion von $f + \lambda g$, also $\int (f + \lambda g) = \int f + \lambda \int g$.

Aus der Produktregel $(f \cdot g)' = f' \cdot g + f \cdot g'$ folgt durch Integration $\int f g' = f g - \int f' g$, oder allgemeiner aus $(f g - \int f' g)' = (f g)' - f' g = f' g + f g' - f' g = f g'$.

Die Substitutionsformel

$$\int (f \circ g) \cdot g' = \left(\int f \right) \circ g$$

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

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

Um die zweite Variante der Substitutionsformel 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

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

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.

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

Für das bestimmte Integral $\int_{-1}^1 \frac{2x}{1+x^2} dx$ erhalten wir somit

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

was wir allerdings auch ohne jegliche Rechnung hätten sehen können, denn $\text{id} : x \mapsto x$ 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 $\int_{-a}^a f = \int_{-a}^0 f + \int_0^a f = - \int_0^a f + \int_0^a f = 0$.

Wie geben nun noch geometrische Interpretationen der Substitutionsformel und der Formel für inverse Funktionen:

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

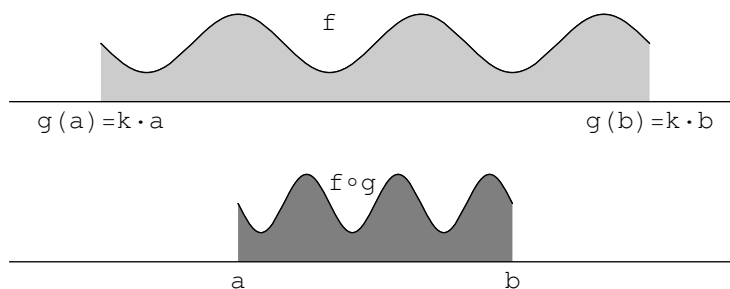
oder wenn wir $\alpha = g(a)$ und $\beta = g(b)$ setzen, also $a = g^{-1}\alpha$ und $b = g^{-1}(\beta)$, dann ist

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

Wenn insbesondere g linear (oder affin) ist, und somit g' konstant ist, also $g(x) = kx + d$ 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:



Die erste Fläche ist die um den Faktor k gestreckte zweite Fläche. Allgemein hängt der Streckungsfaktor $g'(x)$ eben vom Punkt x ab, was uns wegen $g(t_i) - g(t_{i-1}) = g'(\xi_i)(t_i - t_{i-1})$ nicht wundern darf.

Ad 5.2.7 Beispiel.

Es ist

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

und somit

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

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 $1/(x - x_0)^d$ abziehen können um dadurch eine einfachere rationale Funktion (die x_0 nicht mehr als d -fache Nullstelle besitzt) zu erhalten. Wegen $\lim_{x \rightarrow x_0} (x - x_0)^d \frac{P(x)}{Q(x)} =$

$\lim_{x \rightarrow x_0} \frac{P(x)}{Q_1(x)} = \frac{P(x_0)}{Q_1(x_0)}$, 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 $\text{grad}(R_1) < \text{grad}(Q_1) + g - 1$ wegen $\lim_{x \rightarrow \infty} \frac{P(x)}{Q(x)} = 0 = \lim_{x \rightarrow \infty} \frac{1}{(x - x_0)^g}$. Induktion nach der Anzahl der Nullstellen (also dem Grad von Q) liefert somit die Partialbruchzerlegung im Komplexen.



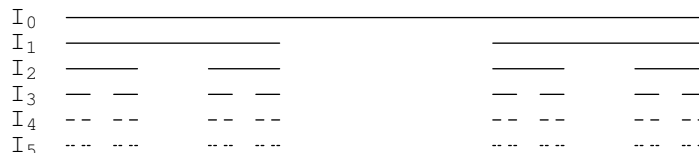
Nicht Nullmengen.

Kein Intervall $I := [a, b]$ mit $a < b$ ist eine 0-Menge, denn aus $I \subseteq \bigcup_{n \in \mathbb{N}} I_n$ folgt $|I| \leq \sum_{n=0}^{\infty} |I_n|$: 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 $Z_n := \{J \in Z : n = \min\{k : J \subseteq I_k\}\}$, dann ist $|I_n| \geq \sum_{J \in Z_n} |J|$. Und da Z die disjunkte Vereinigung der Z_n ist, ist $|I| = \sum_{J \in Z} |J| = \sum_{n=0}^{\infty} \sum_{J \in Z_n} |J| \leq \sum_{k=0}^{\infty} |I_k|$.



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 $(2/3)^k \rightarrow 0$ für $k \rightarrow \infty$.



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} aber überabzählbar viele.

Alle unendlichen Ternärzahlen die keinen 1'er in der Entwicklung enthalten liegen also in I_∞ , und das sind genausoviele wie $[0, 1]$, also überabzählbar viele.

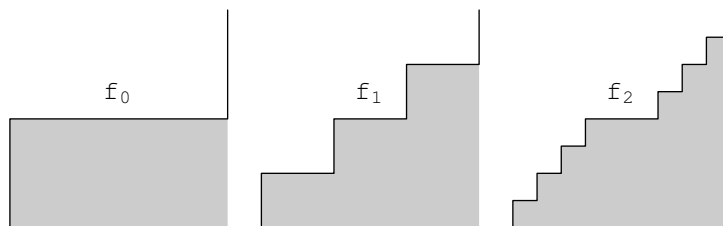
Umgekehrt, ist jede Ternärzahl die eine (möglicherweise abbrechende) Entwicklung ohne 1'er zuläßt in I_∞ , denn die rechten Endpunkte der entfernten Intervalle haben zwar nicht-abbrechende Entwicklungen, die genau einen 1'er enthalten welcher von lauter 2'ern gefolgt wird, also auch eine abbrechende Entwicklung der Form, wo dieser 1'er durch 2 ersetzt ist und die nachfolgenden 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 $f : [0, 1] \rightarrow [0, 1]$ beschreiben, die wir Cantor's Treppe nennen könnten. Dazu bezeichnen wir mit $f_n : [0, 1] \rightarrow [0, 1]$ 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.

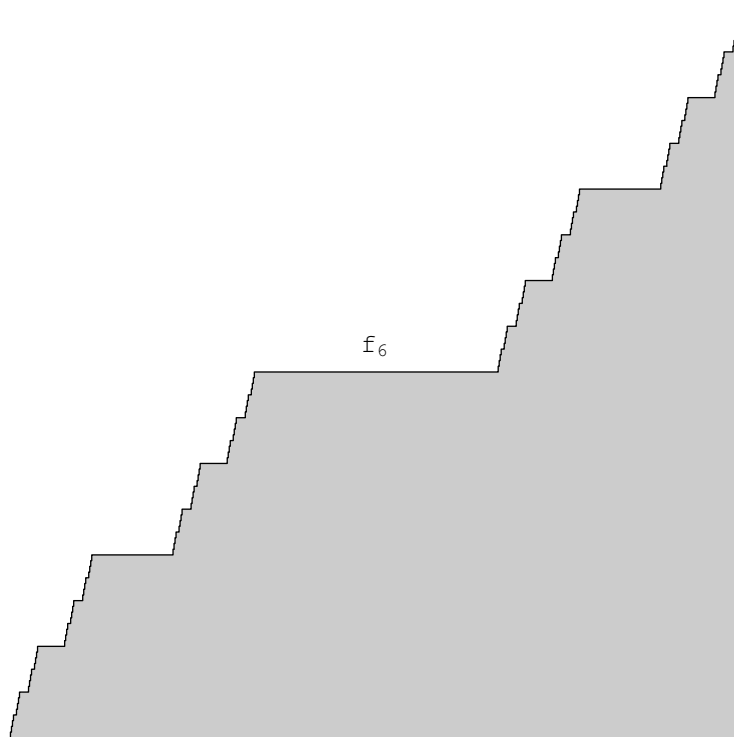
$$f_n(x) := \sum_{x \geq t \in \partial I_n} \frac{1}{2^{n+1}},$$

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



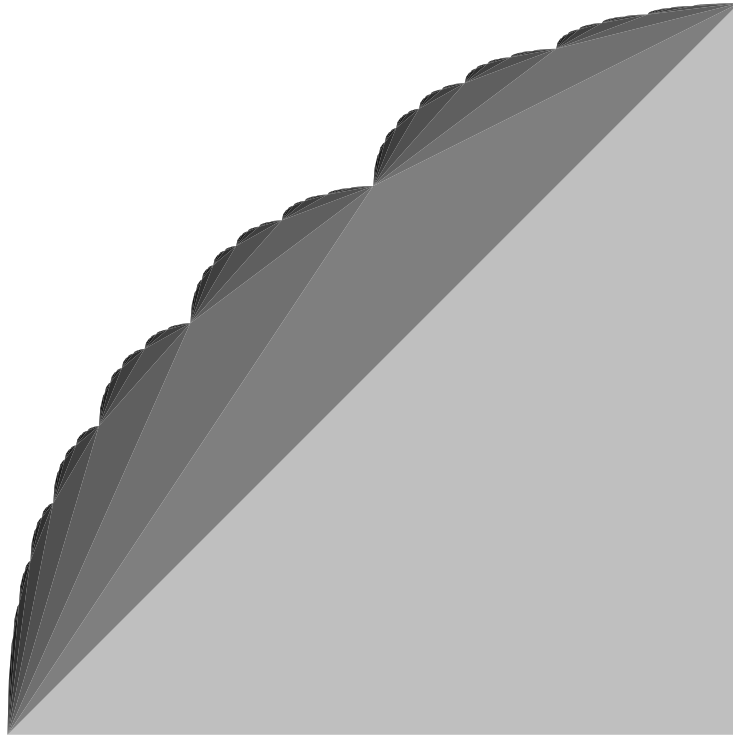
Es ist f_n monoton wachsend und konstant auf den Teilintervallen von $[0, 1] \setminus I_n$. Weiters konvergiert die Folge $(f_n)_n$ gleichmäßig, denn f_{n+1} unterscheidet sich von f_n (auf den Teilintervallen von I_n) um höchstens $1/2^{n+1}$. Somit ist die Grenzfunktion f_∞ stetig und ebenfalls monoton wachsend. Da f_∞ auf jedem Teilintervall von $[0, 1] \setminus I_\infty$ 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 existiert die Ableitung und ist 0.



Beispiel.

Wir zeigen nun mit einem weiteren Beispiel, daß stetige Bilder Nullmengen nicht Nullmengen zu sein brauchen. Dazu definieren wir für $0 < q < 1$ rekursiv eine monoton wachsende Folge strikt monotoner stückweise affiner stetiger Funktionen $h_n : [0, 1] \rightarrow [0, 1]$ 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 $h(x) := \lim_{n \rightarrow \infty} h_n(x)$. Dann ist $h : [0, 1] \rightarrow [0, 1]$ 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 $h(x) \leq h(\frac{k}{2^n}) = h_n(\frac{k}{2^n}) < h_n(y) \leq h(y)$.

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

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

Analog ist $h(\beta_{n+1}) - h(\alpha_{n+1}) = \frac{1-q}{2} (h(\beta_n) - h(\alpha_n))$ 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 $n \rightarrow \infty$, da die Faktoren vom Betrag kleiner als $\frac{1+q}{2} < 1$ sind.

Da monotone Funktionen nach einem Satz 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

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

Falls $h'(x) > 0$ ist, so konvergiert $\sum_{j=1}^n \log(1 + \varepsilon_j q) = \log(\prod_{j=1}^n (1 + \varepsilon_j q)) \rightarrow \log(h'(x))$ und somit $\log(1 + \varepsilon_n q) \rightarrow 0$. 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 einem 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 des Satzes von Saks ist $h(A)$ eine Nullmenge.

Da h streng monoton von $[0, 1]$ auf $[0, 1]$ ist, existiert die Umkehrfunktion $h^{-1} : [0, 1] \rightarrow [0, 1]$ und diese ist stetig und streng monoton, nach (3.4.3), also ebenfalls fast überall differenzierbar. Es ist $h^{-1}(h(A)) = A$ aber eine Menge vom Maß 1. D.h. das homöomorphe Bild einer 0-Menge muß keine 0-Menge mehr sein.

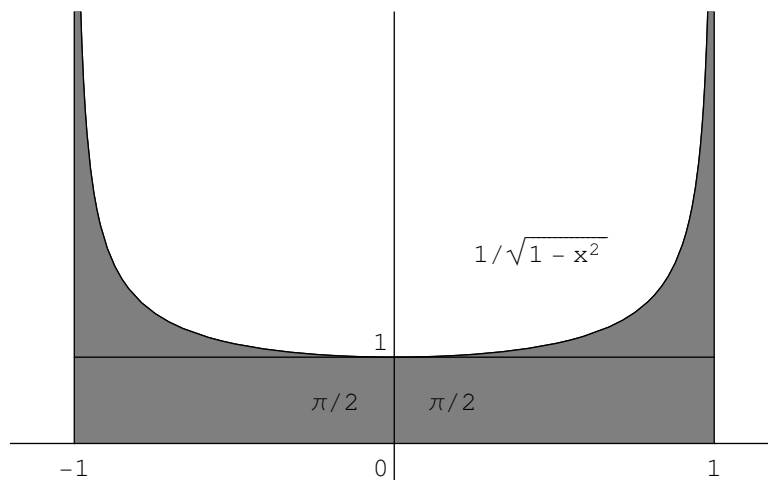
Ad 5.2.6

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 Kreissektors von $(1, 0)$ bis $(x, y) = (\cos(t), \sin(t))$ ist somit (via Substitution $s = \cos(r)$) durch

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

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.

Ad 5.3.2 Beispiel.



Es 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

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

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 könnten $\int_1^\infty \frac{1}{y\sqrt{y^2-1}} dy$ auch mit partieller Integration weiter umzuformen, indem wir $\frac{1}{y} = \ln'(y)$ oder $\frac{1}{\sqrt{y^2-1}} = \text{Arcosh}'(y)$ beachten:

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

$$\int \frac{1}{y\sqrt{y^2-1}} dy = \frac{\text{Arcosh}(y)}{y} - \int \frac{-\text{Arcosh}(y)}{y^2} dy.$$

Ad 5.3.2

Was stimmt nicht mit folgender Rechnung?

$$\begin{aligned}\int_{-1}^1 \frac{1}{\sqrt[3]{x^4}} dx &= \int_{-1}^1 x^{-\frac{4}{3}} dx = \frac{x^{-\frac{4}{3}+1}}{-\frac{4}{3}+1} \Big|_{x=-1}^1 \\ &= \frac{-3}{\sqrt[3]{x}} \Big|_{x=-1}^1 = -3 - (-3) = 6.\end{aligned}$$

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

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

5.3.3a Proposition.

Es seien X und Y metrische Räume, Y sei vollständig, $x_\infty \in X$ sei nicht isoliert in X und $f : X \setminus \{x_\infty\} \rightarrow Y$ eine Funktion.

Dann existiert $\lim_{x \rightarrow x_\infty} f(x)$ genau dann, wenn die Cauchy-Bedingung

$d(f(x), f(x')) \rightarrow 0$ für $x, x' \rightarrow x_\infty$

erfüllt ist, d.h.

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

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

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

2. Möglichkeit: Es ist $d(f(x), f(x')) \leq d(f(x), f(x_\infty)) + d(f(x_\infty), f(x')) \rightarrow 0$ für $x, x' \rightarrow x_\infty$.

(\Leftarrow) Für jede Folge $(x_n)_n$ in $X \setminus \{x_0\}$ ist also $f(x_n)$ eine Cauchy-Folge in Y und konvergiert somit. Wir zeigen nun, daß dieser Grenzwert $\lim_{n \rightarrow \infty} f(x_n)$ 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 $\eta := \lim_{n \rightarrow \infty} f(x_n'')$. Damit konvergieren aber auch die Teilfolgen $(f(x_n))_n$ und $(f(x_n'))_n$ beide gegen η , haben also den gleichen Grenzwert. Mit (3.1.3) folgt nun, daß $\lim_{x \rightarrow x_\infty} f(x)$ existiert und dieser Grenzwert ist. \square

Ad 5.3.3

Lemma (5.3.3) ist nun eine Folgerung aus (5.3.3a), denn die Cauchy-Bedingung für das uneigentliche Integral $\int_a^b f = \lim_{\beta \rightarrow b^-} \int_a^\beta f$ ist:

$$\int_a^{\beta'} f - \int_a^\beta f = \int_\beta^{\beta'} f \rightarrow 0 \text{ für } \beta, \beta' \rightarrow b. \quad \square$$

5.3.3b Majorantentest.

Es seien $f, g : [a, b] \rightarrow \mathbb{R}$ Funktionen mit $0 \leq f \leq g$ 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 $\int_a^b f \leq \int_a^b g$.

Beweis. Es sei $F(\beta) := \int_a^\beta f$ und $G(\beta) := \int_a^\beta g$. Wegen $g \geq f \geq 0$ sind F und G monoton wachsend und $F \leq G$. Also ist $F(\beta) \leq G(\beta) \leq \lim_{\beta \rightarrow b^+} G(\beta) = \int_a^b g < \infty$ und somit existiert $\int_a^b f = \lim_{\beta \rightarrow b} F(\beta) \leq \int_a^b g$. \square

Ad 5.3.5

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

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

$$\int_0^1 e^{-t} t^{x-1} dt \leq \int_0^1 t^{x-1} dt = \lim_a \rightarrow 0 + \frac{t^x}{x} \Big|_{t=a}^1 = \frac{1}{x} < \infty.$$

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

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

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

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

Zusammengefaßt existiert also

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

Ad 5.3.5

Ausführlicher lautet die Herleitung der Rekursionsvorschrift:

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

5.3.7 Bemerkung.

Wir versuchen nun $\Gamma(1/2)$ zu berechnen:

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

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



Ad 5.2.4 und 5.2.5

Wir haben in der Analysis 1 die (Ableitung der) Logarithmusfunktion \ln wie folgt erhalten:

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

Man könnte diesen Zugang auch umdrehen, d.h. den Logarithmus durch $\ln(x) := \int_1^x \frac{1}{t} dt$ definieren. Dann nachrechnen, daß $\ln(a \cdot b) = \ln(a) + \ln(b)$ erfüllt. Die Exponentialfunktion \exp als Umkehrfunktion von \ln definieren und allgemein $a^x := (e^{\ln(a)})^x := \exp(x \ln(a))$ und nun nachrechnen, daß $x := a^{1/n}$ die Gleichung $x^n = a$ löst.

5.5 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 $f : \mathbb{R} \rightarrow \mathbb{R}^m$, für Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$ oder gleich für Abbildungen $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ treiben.

5.5.1 Vektorräume von Funktionen.

Wir haben bereits an einigen Stellen gesehen, daß für das Verständnis von niedrig dimensional 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 $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ bei 0 benötigen wir die Konvergenz von $f(rx) \rightarrow f(0)$ für $r \rightarrow 0+$ gleichmäßig bzgl. $|x| = 1$, also von der Familie von Funktionen $f_r : y \mapsto f(ry)$ 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 in $B(Y, \mathbb{R})$.
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 $\int : R[a, b] \rightarrow \mathbb{R}$ bzgl. der gleichmäßigen Konvergenz.

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

1. Unser Hauptbeispiel ist der Raum \mathbb{K}^n aller endlichen Folgen (x_0, \dots, x_{n-1}) mit n Gliedern $x_1, \dots, x_n \in \mathbb{K}$, also aller Funktionen $x : \{0, \dots, n-1\} \rightarrow \mathbb{K}$.
2. Der Raum $B(X, \mathbb{K}) := \{f : X \rightarrow \mathbb{K} : \|f\|_\infty := \sup\{|f(x)| : x \in X\} < \infty\}$ aller beschränkten Funktionen $X \rightarrow \mathbb{K}$ auf einer Menge X .
3. Oder allgemeiner der Raum $B(X, \mathbb{K}^n) := \{f : X \rightarrow \mathbb{K}^n : \|f\|_\infty := \sup\{|f(x)| : x \in X\} < \infty\}$ aller beschränkten Abbildungen $X \rightarrow \mathbb{K}^n$ 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 $X \rightarrow \mathbb{K}$ auf einem metrischen Raum X .

6. Oder allgemeiner der Raum $C(X, \mathbb{K}^n)$ der stetigen Abbildungen $X \rightarrow \mathbb{K}^n$ auf einem metrischen Raum X .
7. Speziell der Raum $c := \{x \in \mathbb{K}^{\mathbb{N}} : \exists \lim_{n \rightarrow \infty} x_n\}$ aller konvergenten Folgen in \mathbb{K} , siehe (2.3.6).
8. Der Raum $\ell^1 := \{x \in \mathbb{K}^{\mathbb{N}} : \|x\|_1 := \sum_{n=0}^{\infty} |x_n| < \infty\}$ aller absolut konvergenten Reihen.
9. Der Raum $\ell^2 := \{x \in \mathbb{K}^{\mathbb{N}} : (\|x\|_2)^2 := \sum_{n=0}^{\infty} |x_n|^2 < \infty\}$ aller 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 $f : \mathbb{R} \rightarrow E$ benötigen wir auch eine Skalarmultiplikation um Ausdrücke der Form

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

betrachten zu können. Die Metrik sollte dann natürlich auch verträglich mit diesen Operationen sein.

5.5.2 Definition. Norm und normierter Raum.

Auf den Vektorräumen von Funktionen aus (5.5.1) haben wir Metriken der Form $d(f, g) := \|f - g\|$ betrachtet, wobei die Funktion $f \mapsto \|f\| \in \mathbb{R}$ folgende Eigenschaften hat:

- $\|f\| = 0 \Leftrightarrow f = 0$.
- $\|\lambda f\| = |\lambda| \cdot \|f\|$ für alle $\lambda \in \mathbb{K}$.
- $\|f + g\| \leq \|f\| + \|g\|$.

So eine Funktion $\|\cdot\|$ nennt man Norm, und Vektorräume mit Norm heißen normierte Räume. 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 Funktionen f 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 $|(|f|)|$ gelesen werden.

Für endliches X haben wir in (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

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

und ebenso für ℓ^2

$$\begin{aligned}\sum_{k=0}^n |x_k + y_k|^2 &\leq \left(\sqrt{\sum_{k=0}^n |x_k|^2} + \sqrt{\sum_{k=0}^n |y_k|^2} \right)^2 \leq (\|x\|_2 + \|y\|_2)^2 \\ \Rightarrow \|x + y\|_2^2 &= \sup \left\{ \sum_{k=0}^n |x_k + y_k|^2 : n \in \mathbb{N} \right\} \leq (\|x\|_2 + \|y\|_2)^2.\end{aligned}$$

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

Für die LeserInnen/HörerInnen denen dieser Schritt vom 1-dimensionalen auf das unendlich-dimensionale (für den Moment) zu groß ist und Banach-Räume zu abstrakt sind, 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.

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. 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 $x^n = (x_k^n)_{k \in \mathbb{N}}$ reeller Zahlen $x_k^n \in \mathbb{R}$. 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_k^n als $x_n(k)$ zu schreiben, und damit zum Ausdruck zu bringen, daß die Elemente von ℓ^1 auch als Funktionen $\mathbb{N} \rightarrow \mathbb{R}$ 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	x_0^0	x_1^0	x_2^0	x_3^0	...	x_k^0	...
x^1	x_0^1	x_1^1	x_2^1	x_3^1	...	x_k^1	...
x^2	x_0^2	x_1^2	x_2^2	x_3^2	...	x_k^2	...
x^3	x_0^3	x_1^3	x_2^3	x_3^3	...	x_k^3	...
\vdots	\vdots	\vdots	\vdots	\vdots		\vdots	
x^n	x_0^n	x_1^n	x_2^n	x_3^n	...	x_k^n	...
\vdots	\vdots	\vdots	\vdots	\vdots		\vdots	
\downarrow	\downarrow	\downarrow	\downarrow	\downarrow		\downarrow	
x^∞	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 $|x_k^n| \leq \|x^n\|_1$) und somit konvergent. Sei $x_k^\infty := \lim_{n \rightarrow \infty} x_k^n$. Weiters ist $\|x^n\|_1$ eine Cauchy-Folge (wegen $|\|x^n\|_1 - \|x^m\|_1| \leq \|x^n - x^m\|_1$) also beschränkt. Damit ist

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

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

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

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

Ganz analog kann man auch bei ℓ^2 vorgehen.

5.5.5 Lemma.

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 $\|\sum_{j=n}^{n+p} x_j\| \leq \sum_{j=n}^{n+p} \|x_j\|$, 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 $\sum_n \|\lambda_n b_n\| \leq \|\lambda\|_1 \cdot \sup\{\|b_n\| : n \in \mathbb{N}\}$. Also konvergiert diese Reihe nach (2).

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

$$\begin{aligned} & \forall k \exists n_k \forall i, j \geq n_k : \|x_i - x_j\| \leq \lambda_k \quad (\text{o.B.d.A. } n_k \leq n_{k+1}) \\ \Rightarrow & \left\| \frac{1}{\lambda_k} (x_{n_{k+1}} - x_{n_k}) \right\| \leq 1 \text{ für } n \leq k \\ \Rightarrow & \frac{1}{\lambda_k} y_k \text{ ist beschränkt, wobei } y_k := x_{n_{k+1}} - x_{n_k} \\ \stackrel{(3)}{\Rightarrow} & x_{n_j} = \sum_{k < j} \lambda_k \frac{1}{\lambda_k} y_k \text{ konvergiert.} \end{aligned}$$

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

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

d.h. $x_n \rightarrow x_\infty$. □

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 Lemma.

Sei X eine Menge und E ein Banach-Raum. Dann ist der Raum

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

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 $f : X \rightarrow E$ 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 die Punktevaluationen $\text{ev}_x : B(X, E) \rightarrow E$, $f \mapsto f(x)$ stetig (wegen $\|\text{ev}_x(f)\| = \|f(x)\| \leq \|f\|_\infty$) und linear sind (also gleichmäßig stetig sind), 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

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

für $n \geq n_0(\varepsilon)$ (und geeignet gewählten k). Also konvergiert $f_n \rightarrow f$ in der Supremums-Norm.

Weiters ist

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

also gehört f_∞ zu $B(X, E)$. □

Lemma.

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 . □

Folgerung.

Für kompaktes X ist auch der Teilraum $C(X) := C(X, \mathbb{K})$ von $B(X, \mathbb{K})$ ein Banach-Raum.

Beweis. Nach (5.5.6) ist $B(X, E)$ vollständig, wenn E es ist. Nach (4.2.8) ist $C(C, E)$ abgeschlossen in $B(X, E)$, also nach obigen Lemma selbst vollständig. □

5.5.7 Produkte von Räumen.

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

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n)$$

$$\lambda \cdot (x_1, \dots, x_n) := (\lambda x_1, \dots, \lambda x_n).$$

Als Norm können wir $\|(x_1, \dots, x_n)\|_\infty := \max\{\|x_i\| : i \in \{1, \dots, n\}\}$ 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) $\|(x_1, \dots, x_n)\|_2 := \left(\sum_{i=1}^n \|x_i\|^2\right)^{1/2}$ 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 $E_1 \times \dots \times E_n$ von Banach-Räumen ebenfalls vollständig ist.

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

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

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

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

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

5.5.8 Definition. Stetige lineare Abbildungen.

Um verschiedene Banach-Räume zueinander in Beziehung zu setzen benötigen wir (strukturerehaltende) 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 $f : \mathbb{R}^n \rightarrow F$ ist die Stetigkeit automatisch erfüllt, denn in (3.1.5) haben wir gezeigt, daß die Koordinaten-Projektionen $\text{pr}_j : \mathbb{R}^n \rightarrow \mathbb{R}, x = (x^1, \dots, x^n) \mapsto x^j$ stetig sind und wegen

$$f(x) = f\left(\sum_{j=1}^n x^j e_j\right) = \sum_{j=1}^n x^j f(e_j) = \sum_{j=1}^n \text{pr}_j(x) f(e_j),$$

ist f als Summe von Produkten stetiger Funktionen stetig.



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

$$\sum \left(\frac{1}{n} \chi_{\{1, \dots, n\}} \right) := \sum_{i=1}^{\infty} \frac{1}{n} \chi_{\{1, \dots, n\}}(i) = \sum_{i=1}^n \frac{1}{n} = 1 \neq 0 = \sum(0).$$

5.5.9 Lemma. Charakterisierung stetig linearer Abbildungen.

Es sei $f : E \rightarrow F$ eine lineare Abbildung zwischen normierten Räumen. Dann sind folgende Aussagen äquivalent:

1. f ist stetig;
2. Das Bild $\{f(x) : \|x\| \leq 1\}$ des Einheitsballs $\{x \in E : \|x\| \leq 1\}$ ist beschränkt;
3. f ist beschränkt auf allen beschränkten Mengen;

4. $\exists L \geq 0 \forall x \in E: \|f(x)\| \leq L\|x\|$;

5. f ist Lipschitz, d.h. $\exists L \geq 0 \forall x, y \in E: \|f(x) - f(y)\| \leq L \cdot \|x - y\|$;

Beweis. (1 \Rightarrow 2) Andernfalls gäbe es zu jedem $n \in \mathbb{N}$ ein $x_n \in E$ mit $\|x_n\| \leq 1$ und $\|f(x_n)\| \geq n$. Dann konvergiert $\frac{1}{n}x_n \rightarrow 0$ aber $\|f(\frac{1}{n}x_n)\| = 1$ und somit konvergiert $f(\frac{1}{n}x_n)$ nicht gegen $0 = f(0)$.

(2 \Rightarrow 3) Sei $B \subseteq E$ beschränkt, also existiert ein $r > 0$ mit $B \subseteq \{x : \|x\| \leq r\} = r \cdot \{y : \|y\| \leq 1\}$ (wobei wir $x = ry$ gesetzt haben). Damit ist aber auch $f(B) \subseteq f(r \cdot \{y : \|y\| \leq 1\}) = r \cdot f(\{y : \|y\| \leq 1\})$ beschränkt.

(3 \Rightarrow 2) Ist offensichtlich.

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

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

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

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

Bemerkung.



Ein wesentlicher Unterschied zwischen endlich-dimensionalen und unendlich-dimensionalen Banach-Räumen E besteht darin, daß der Einheitsball $\{x \in E : \|x\| \leq 1\}$ 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

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

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

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. [Kri02, 7.28].

5.5.10 Bemerkung.

Wir haben für lineare Abbildungen $f : E \rightarrow F$ 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

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

aller stetig linearen Abbildungen $f : E \rightarrow F$ 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 $\mathbb{K}^n \cong E$, $(x^1, \dots, x^n) \mapsto \sum_{i=1}^n x^i e_i$ und $\mathbb{K}^m \cong F$, $(y^1, \dots, y^m) \mapsto \sum_{j=1}^m y^j f_j$ jede lineare Abbildung $a : E \rightarrow F$ als lineare Abbildung $\tilde{a} : \mathbb{K}^n \rightarrow \mathbb{K}^m$ 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

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

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 bedeutet, sondern wir bloß Indizes wie schon in (5.5.4) oben geschrieben 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 Operator-Norm einer (der Einfachheit halber 2×2 -)Matrix zu bestimmen, wobei wir vorerst am $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ die Maximums-Norm verwenden. Dann ist

$$\begin{aligned} \left\| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \right\|_{\infty} &= \|(ax + by, cx + dy)\|_{\infty} \\ &= \max\{|ax + by|, |cx + dy|\} \\ &\leq \max\{|x|, |y|\} \cdot \max\{|a| + |b|, |c| + |d|\}, \end{aligned}$$

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

Wenn wir anstelle der Maximums-Norm die geometrischere Euklid'sche 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.

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.

5.5.11 Folgerung.

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 Operator-Norm

$$\begin{aligned}\|f\| &:= \sup\{\|f(x)\| : \|x\| \leq 1\} = \sup\{\|f(x)\| : \|x\| = 1\} = \sup\left\{\frac{\|f(x)\|}{\|x\|} : x \neq 0\right\} \\ &= \inf\{K : \|f(x)\| \leq K\|x\| \text{ für alle } x\}\end{aligned}$$

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 aller stetigen linearen Funktionale auf E .



Falls $f : E \rightarrow F$ ein stetig linearer Operator ist, so bezeichnen wir mit $f^* : F^* \rightarrow E^*$ 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 $a : F \times F \rightarrow F$ und die skalar-Multiplikation $m_\lambda : F \rightarrow F$, $x \mapsto \lambda x$ 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 $f, g : E \rightarrow F$ wieder stetig, also bildet $L(E, F)$ einen linearen Teilraum des Vektorraums aller linearer Abbildungen.

Daß obige Beschreibungen der Operator-Norm äquivalent sind, zeigt:

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

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

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

$$\begin{aligned}\forall x : \|(f + g)(x)\| &\leq \|f(x)\| + \|g(x)\| \leq (\|f\| + \|g\|) \|x\| \Rightarrow \|f + g\| \leq \|f\| + \|g\| \\ \forall x : \|(\lambda f)x\| &= |\lambda| \|f(x)\| \Rightarrow \|\lambda f\| = |\lambda| \|f\|\end{aligned}$$

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

5.5.12 Bemerkung.

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

In (5.5.7) hatten wir die Stetigkeit der Addition $a : E \times E \rightarrow E$ 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 $\|a(x_1, y_1)\| = \|x_1 + y_1\| \leq \|x_1\| + \|y_1\| \leq 2\|(x_1, y_1)\|$ gefolgt.

Ebenso hatten wir die Stetigkeit der Skalarmultiplikation $m : \mathbb{K} \times E \rightarrow E$ gezeigt. Diese ist aber nicht linear, sondern nur bilinear, d.h. in jedem Faktor getrennt linear, also 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.

Wir benötigen also eine Verallgemeinerung von (5.5.9) auf multi-lineare Abbildungen:

Proposition.

Es seien E_1, \dots, E_n, F normierte Räume und $f : E_1 \times \dots \times E_n \rightarrow F$ eine multi-lineare Abbildung.

Dann ist f genau dann stetig, wenn

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

ist.

Beweis. (\Rightarrow) 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.

(\Leftarrow) 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

$$\begin{aligned} \|f(x_1, y_1) - f(x_0, y_0)\| &\leq \|f(x_1 - x_0, y_1)\| + \|f(x_0, y_1 - y_0)\| \\ &\leq \|f\| \underbrace{\left(\|x_1 - x_0\| \|y_1\| + \|x_0\| \|y_1 - y_0\|\right)}_{\rightarrow 0} \end{aligned}$$

für $(x_1, y_1) \rightarrow (x_0, y_0)$. \square

5.5.13 Bemerkung.

Wir wollen nun bilineare Abbildungen $f : E_1 \times E_2 \rightarrow F$ als (lineare) Abbildungen auf E_1 mit Werten in einem Raum $L(E_2, F)$ von (linearen) Abbildungen von E_2 nach F auffassen.

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

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

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 $E_2 \rightarrow F$ eine Abbildung $\hat{g} : E_1 \times E_2 \rightarrow F$ 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 $E_1 \times E_2 \rightarrow F$ einerseits und dem Banach-Raum $L(E_1, L(E_2, F))$ der linearen stetigen Abbildungen $E_1 \rightarrow L(E_2, F)$ andererseits.

Lemma. Bilineare versus lineare Abbildungen.

Es bezeichne $L(E_1, E_2; F)$ den Vektorraum der bilinearen stetigen Abbildungen $E_1 \times E_2 \rightarrow F$ 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 $f \mapsto \check{f}, \hat{f} \leftarrow f$.

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

Das 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 $\circ : A \times A \rightarrow A$ die zusätzlich assoziativ ist (und eine Einheit $id \in A$ besitzt).

Achtung $\|f \circ g\| \neq \|f\| \cdot \|g\|$, z.B. $f(x, y) := (x, 0)$ und $g(x, y) := (0, y)$.

Beweis.

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

Es gilt allerdings bei $\|f \circ g\| \leq \|f\| \cdot \|g\|$ nicht immer Gleichheit, wie das Beispiel $g(x) := (x, 0)$ und $f(x, y) := y$ zeigt.

5.5.15 Definition .

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 , insbesondere ist also der Raum der $n \times n$ -Matrizen eine Banach-Algebra.

Lemma.

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.

Inbesondere 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). □

5.5.16 Folgerung.

Die Menge der invertierbaren Elemente einer Banach-Algebra ist offen und die Inversion $a \mapsto a^{-1}$ ist eine stetige Abbildung.

Ein Element $a \in A$ heißt bekanntlich invertierbar, wenn ein 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

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

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.

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 $\langle \cdot | \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, $(x, y) \mapsto \sum_{i=1}^n x^i y^i$ eine symmetrische Bilinearform ist. Diese Bilinearform erlaubt uns nämlich Winkel $\angle xy$ zwischen Vektoren $x, y \neq 0$ durch

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

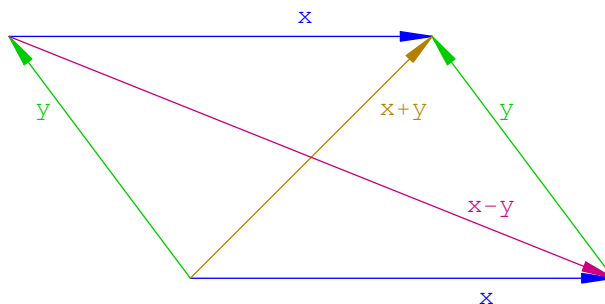
zu definieren, und insbesondere Normalität vermöge $x \perp y \Leftrightarrow \langle x|y \rangle = 0$.



Die selbe Konstruktion funktioniert 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 $\langle \cdot | \cdot \rangle : E \times E \rightarrow \mathbb{R}$ zuläßt, welche $\|x\|^2 = \langle x|x \rangle$ für alle $x \in E$ erfüllt. Man kann zeigen (siehe [Kri02, 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 $E \rightarrow \mathbb{R}$.



Man kann zeigen (siehe [Kri02, 6.7]), daß auch für Hilbert-Räume ein Isomorphismus $E \cong E^*$ besteht, der wie im euklidischen Fall für orthonormal Basen durch $E \ni x \mapsto \langle x | \cdot \rangle \in E^*$ gegeben ist, d.h. wir können lineare stetige Abbildungen $f : E \rightarrow \mathbb{R}$ durch jenen Vektor $x \perp \text{Ker}(f)$ beschreiben auf welchen $f(x) = 1$ ist.



5.5.18 Weitergehende Sätze für Banach-Räume aus der Funktionalanalysis.

Im \mathbb{R}^n haben wir viele Aussagen auf den 1-dimensionalen Fall zurückführen können, indem wir die Koordinaten $x^i \in \mathbb{R}$ der Vektoren $x = (x^1, \dots, x^n) \in \mathbb{R}^n$ betrachtet haben, siehe z.B. Aufgabe (2.11), (2.3.4), (2.4.5), (3.1.4), (4.1.1) und (5.3.8). Die Abbildungen $\text{pr} : \mathbb{R}^n \rightarrow \mathbb{R}, x \mapsto x^i$ die wir dabei verwendet haben sind linear (und somit stetig). Auch bei Funktionen-Räumen können wir Identitäten durch Auswerten mittels der stetigen linearen Funktionen $\text{ev}_x : \mathbb{R}^X \rightarrow \mathbb{R}, f \mapsto f(x)$ verifizieren. Allerdings können wir Konvergenz so nicht beschreiben, denn $\text{ev}_x(f_n) \rightarrow \text{ev}_x(f_\infty)$ 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 einem normierten Raum E ist. Für Hilbert-Räume haben wir das oben schon beantwortet. Die allgemeine Antwort liefert der Satz von Hahn-Banach (siehe [Kri02, 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 stetig lineare Abbildung $f : E \rightarrow \mathbb{R}$ mit $f(x_1) \neq f(x_2)$.

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 [Kri02, 5.4]) zeigen: Es sei $\mathcal{F} \subseteq L(E, F)$ eine punktweise beschränkte Menge stetig linearer Abbildungen auf einem Banach-Raum E , dann ist \mathcal{F} beschränkt im normierten Raum $L(E, F)$, d.h. $\{\|f\| : f \in \mathcal{F}\}$ ist beschränkt, insbesondere 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 [Kri02, 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 Folgenglieder 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 [Kri02, 5.7].

References

[Kri02] Andreas Kriegel. Funktionalanalysis 1. Skriptum, 2002. 32, 38, 39, 40