Kapitel 6
Mehrdimensionale Integrale

Auch im Mehrdimensionalen ist die Integralrechnung nach der Differentialrechnung das zweite Standbein der Analysis. Während allerdings für Funktionen einer Veränderlichen Differentiation und Integration einander entgegengesetzte Operationen sind, ist die Situation hier komplizierter: Die Ableitung einer reellwertigen Funktion mehrerer Veränderlicher ist schließlich keine reellwertige Funktion mehr, sondern eine vektorwertige; so etwas wie eine Stammfunktion kann daher nur für Funktionen in n Variablen mit Werten in $\mathbb{R}^n$ existieren, die sogenannten *Vektorfeldern.* Mit Vektorfeldern und Integralen darüber beschäftigt sich die Analysis III; hier in der Analysis II geht es um einen Integralbegriff, der die Flächeninterpretation des eindimensionalen *bestimmten* Integrals verallgemeinert. Anstelle eines Integrationsintervalls $[a,\, b]$ können (und müssen) wir im $\mathbb{R}^n$ natürlich sehr viel allgemeinere Mengen betrachten; um möglichst schnell zu Beispielen zu kommen, für die wir die bekannte eindimensionale Integrationstheorie verwenden können, beginnen wir aber mit Integralen über Quader. Der Zugang hier folgt im wesentlichen dem von ROLF WALTER in seiner *Einführung in die Analysis 2,* der wiederum Argumente von SERGE LANG verwendet, dem wohl produktivsten Autor mathematischer Bücher im zwanzigsten Jahrhundert.

§1: Integration stetiger Funktionen über Quader

Unter einen Quader im $\mathbb{R}^n$ wollen wir im folgenden, wenn nichts anderes gesagt wird, stets einen *abgeschlossenen achsenparallelen Quader* verstehen, also eine Menge der Form

$$Q = \prod_{i=1}^{n}[a_i,\, b_i] = \big\{(x_1, \dots, x_n) \bigm| a_i \le x_i \le b_i \text{ für alle } i\big\}\,,$$

wobei jeweils $a_i \le b_i$ sein soll. Als zugehörigen *offenen* Quader bezeichnen wir das entsprechende Produkt der offenen Intervalle, also

$$\mathring{Q} \underset{\text{def}}{=} \prod_{i=1}^{n}(a_i,\, b_i) = \big\{(x_1, \dots, x_n) \bigm| a_i < x_i < b_i \text{ für alle } i\big\}\,.$$

Im Eindimensionalen sind Quader natürlich einfach Intervalle; im Zweidimensionalen sind sie achsenparallele Rechtecke.

Das mehrdimensionale Integral, das wir in diesem Kapitel betrachten wollen, soll insbesondere auch zur Volumenbestimmung oder – besser – zur Volumen*definition* dienen: Für eine von krummen Flächen begrenzten Teilmenge wissen wir schließlich selbst im Dreidimensionalen nicht wirklich, wie wir ihr Volumen definieren sollen.

Dieses Problem stellte sich bereits den klassischen griechischen Mathematikern; auf sie geht auch der Ansatz zurück, auf dem jeder heute gebräuchliche Integralbegriff beruht: Wir können mit elementaren Methoden die Volumina vieler geradlinig begrenzter Flächen und Körper berechnen; durch *Exhaustion* oder *Ausschöpfung* konnten EUDOXOS (erste Hälfte des vierten vorchristlichen Jahrhunderts), ARCHIMEDES (287–212) und andere viele kompliziertere Flächen und Volumina zurückführen auf die bekannter Polygone und Körper.

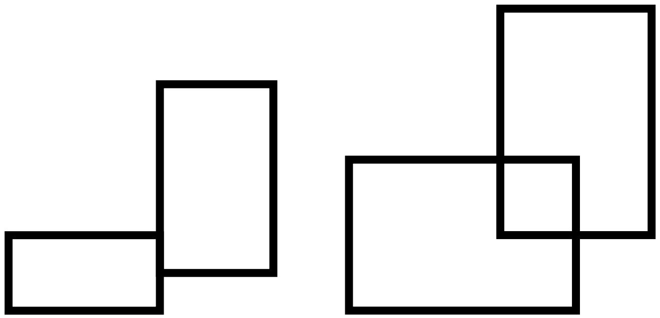
Am einfachsten sind Rechtecke und Quader; wir definieren allgemein das Volumen $\mu(Q)$ eines Quaders im $\mathbb{R}^n$ als Produkt seiner n Kantenlängen, d.h.

$$\mu(Q) = \mu(\mathring{Q}) = \prod_{i=1}^{n}(b_i - a_i) \quad \text{für} \quad Q = \prod_{i=1}^{n}[a_i,\, b_i]\,.$$

Falls irgendein a_i gleich dem zugehörigen b_i ist, verschwindet dieses Produkt; solche *entarteten* Quader haben also das Volumen Null. Das entspricht insofern unserem Alltagsgebrauch des Wortes Volumen, als wir auch da einem Rechteck zwar eine Fläche zuordnen; wenn wir aber aber vom (dreidimensionalen) Volumen reden, betrachten wir es als Quader der Dicke Null und damit auch mit Volumen Null.

In der Tradition von EUDOXOS und ARCHIMEDES wollen wir kompliziertere Volumina „ausschöpfen“ durch Vereinigungen von Quadern; dabei müssen wir natürlich darauf achten, daß sich diese nicht überschneiden, zumindest nicht so, daß der Durchschnitt ein positives Volumen hat.

Definition: Zwei Quader Q und Q' heißen *volumenfremd,* wenn die zugehörigen offenen Quader leeren Durchschnitt haben.



volumenfremd nicht volumenfremd

Als ersten Schritt in Richtung auf die Definition eines n-dimensionalen Integrals definieren wir nun Integrale stetiger Funktionen über Quader in einer zwar sehr willkürlichen, dafür aber unmittelbar zum Rechnen geeigneten Weise:

Definition: $f\colon D \to \mathbb{R}^n$ sei eine stetige Funktion und $Q = \prod_{i=1}^{n}[a_i,\, b_i]$ sei ein in D enthaltener Quader. Dann ist

$$\int_Q f = \int_{a_n}^{b_n} \left(\cdots \left(\int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x_1, \ldots, x_n)\, dx_1 \right) dx_2 \right) \cdots \right) dx_n$$

für $Q \neq \emptyset$ und $\int_\emptyset f = 0$.

Bei dieser Definition bezieht sich jede einzelne Integration nur auf *eine* Variable; wenn wir über dx_i integrieren, betrachten wir die übrigen Variablen x_j als festgehaltene Parameter. Beim innersten Integral über dx_1 halten wir also x_2 bis x_n fest und integrieren nur über x_1; das Ergebnis ist für jedes feste $(n-1)$-Tupel $(x_2, \ldots, x_n)$ eine reelle Zahl und insgesamt betrachtet eine Funktion, die nun nur noch von den Variablen x_2 bis x_n abhängt. Die zweite Integration eliminiert auch noch die Variable x_2 und so weiter, bis wir zuletzt nur noch eine Funktion von x_n integrieren und als Ergebnis eine reelle Zahl erhalten. Somit lassen sich bei allen n Integralen die in Kapitel 4 entwickelten Techniken anwenden.

Für konkret gegebene Funktionen f lassen sich die so definierten Integrale, wie wir bald anhand von Beispielen sehen werden, meist problemlos ausrechnen – zumindest, wenn wir die Stammfunktionen der jeweiligen Integranden angeben können. Ansonsten haben wir mit dieser Definition allerdings mindestens zwei Probleme:

Erstens können wir nicht sicher sein, daß das Integral überhaupt existiert: Wie wir in Kapitel 4 gesehen haben, gibt es durchaus Funktionen einer Veränderlichen, die *nicht* integrierbar sind. Zwar wissen wir, daß jede stückweise stetige Funktion einer Veränderlichen integrierbar ist, und wir haben f sogar als stetige Funktion vorausgesetzt, aber wir wissen nicht, ob nach der ersten Integration die entstehende Funktion von x_2 bis x_n noch stetig zumindest in x_2 ist.

Zweitens haben wir uns willkürlich darauf festgelegt, zuerst über x_1, dann überx_2 und so weiter zu integrieren. Wenn wir bei einer anderen Integrationsreihenfolge ein anderes Ergebnis bekommen, widerspricht dies allem, was wir von einem sinnvoll definierten Integral erwarten. Außerdem kann es auch aus praktischen Gründen gelegentlich sehr viel einfacher sein, die Integrationen in einer anderen Reihenfolge auszuführen.

Das erste Problem läßt sich glücklicherweise recht einfach lösen: Wie das folgende Lemma induktiv zeigt, sind für eine stetige Funktion f auch alle Zwischenintegranden stetig:

Lemma: $D \subseteq \mathbb{R}^n$ sei eine offene Menge, $[a,\, b] \subset \mathbb{R}$ ein abgeschlossenes Intervall und $\widetilde{D} \subseteq \mathbb{R}^{n-1}$ eine offene Menge derart, daß $[a,\, b] \times \widetilde{D}$ in D liegt. Für eine stetige Funktion $f\colon D \to \mathbb{R}$ definiert dann die Zuordnung

$$F(y) \underset{\text{def}}{=} \int_a^b f(x, y)\, dx$$

eine stetige Funktion $F\colon \widetilde{D} \to \mathbb{R}$.

Beweis: Wegen der Stetigkeit von f insbesondere auch in x integrieren wir für jedes feste $y \in \widetilde{D}$ über eine stetige Funktion; das Integral existiert also. Um die Stetigkeit von F zu beweisen, benutzen wir die

Charakterisierung durch Folgen: F ist genau dann stetig in einem Punkt $y_0 \in \widetilde{D}$, wenn für jede gegen y_0 konvergierende Folge $(y_k)_{k\in\mathbb{N}}$ auch die Folge der Funktionswerte $F(y_k)$ gegen $F(y_0)$ konvergiert, wenn es also zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so daß $|F(y_k) - F(y_0)| < \varepsilon$ ist für alle $k \geq N$.

Wegen der Linearität der (eindimensionalen) Integration und der Monotonieregel ist

$$\begin{aligned} |F(y_k) - F(y_0)| &= \left| \int_a^b f(x, y_k)\,dx - \int_a^b f(x, y_0)\,dx \right| \\ &= \left| \int_a^b \big(f(x, y_k) - f(x, y_0) \big)\,dx \right| \\ &\leq \int_a^b |f(x, y_k) - f(x, y_0)|\,dx\,, \end{aligned}$$

und das können wir weiter abschätzen, falls wir eine handhabbare obere Schranke für den Integranden finden. Dazu verhilft uns ein Kompaktheitsargument:

Wir beginnen mit der Menge $Y \subset \mathbb{R}^{n-1}$ bestehend aus y_0 und den y_k mit $k \in \mathbb{N}$. Diese ist kompakt, denn jede offene Überdeckung $\mathfrak{U}$ von Y enthält insbesondere eine offene Menge U, die y_0 enthält. Damit enthält U auch für irgendein $\varepsilon > 0$ alle $y \in \mathbb{R}^{n-1}$ mit $\|y - y_0\|_\infty < \varepsilon$, also gibt es wegen der Konvergenz der Folge $(y_k)_{k\in\mathbb{N}}$ gegen y_0 ein $N \in \mathbb{N}$, so daß $y_k \in U$ für alle $k \geq N$. Damit enthält U alle y_k mit höchstens endlich vielen Ausnahmen. Für jede dieser Ausnahmen gibt es (mindestens) eine offene Menge $V \in \mathfrak{U}$, in der sie enthalten ist; also gibt es eine endliche Teilüberdeckung von $\mathfrak{U}$.

Somit ist Y nach dem Satz von HEINE-BOREL abgeschlossen und beschränkt, und damit gilt natürlich dasselbe auch für $[a,\,b] \times Y$, so daß, wieder nach dem Satz von HEINE-BOREL, auch diese Menge kompakt ist. Nach dem letzten Lemma aus Kapitel 5, §4*a)* ist die stetige Funktion f daher auf der Teilmenge $[a,\,b] \times Y$ gleichmäßig stetig, es gibt also

zu jedem $\varepsilon > 0$ ein $\delta > 0$, so daß $|f(x, y) - f(x^*, y^*)| < \varepsilon$ ist für alle $(x, y), (x^*, y^*) \in D$ mit

$$\|(x, y) - (x^*, y^*)\|_\infty = \max\big\{ |x - x^*|\,, \|y - y^*\|_\infty \big\} < \delta\,.$$

Dies gilt insbesondere für alle Punkte (x, y_0) und (x, y_k) aus $[a,\,b] \times Y$ mit $\|y_0 - y_k\|_\infty < \delta$, und wegen der Konvergenz der Folge $(y_k)_{k\in\mathbb{N}}$ gegen y_0 gibt es ein $N \in \mathbb{N}$, so daß dies für alle $k \geq N$ erfüllt ist. Für diese k ist daher auch $|f(x, y_0) - f(x, y_k)| < \varepsilon$ für alle $x \in [a,\,b]$.

Um nun die Konvergenz der Folge der $F(y_k)$ gegen $F(y_0)$ zu beweisen, betrachten wir irgendein $\varepsilon > 0$. Wie wir gerade gesehen haben, gibt es dazu ein $N \in \mathbb{N}$, so daß

$$|f(x, y_0) - f(x, y_k)| < \frac{\varepsilon}{b - a} \quad \text{für alle } k \geq N \text{ und alle } x \in [a,\,b]\,.$$

Damit ist nach der ersten Abschätzung dieses Beweises für $k \geq N$ auch

$$|F(y_k) - F(y_0)| \leq \int_a^b \frac{\varepsilon}{b - a}\,dx = \varepsilon\,,$$

die Folge konvergiert also gegen $F(y_0)$ und F ist stetig in y_0. Da $y_0 \in \widetilde{D}$ beliebig war, ist F stetig auf $\widetilde{D}$. ■

Auch für unserer zweites Problem brauchen wir als ersten Schritt zu einer Lösung ein Lemma nach Art des gerade bewiesenen, jetzt allerdings für differenzierbare Funktionen, wobei wir uns auf zunächst auf $\mathbb{R}^2$ beschränken wollen:

Lemma: $I \subseteq \mathbb{R}$ sei ein Intervall, $D \subseteq \mathbb{R}^2$ sei eine offene Menge, die $[a,\,b] \times I$ enthalte, und $f\colon D \to \mathbb{R}$ erfülle die beiden Bedingungen

1.) Für jedes $y \in [c,\,d]$ ist die Funktion $x \mapsto f(x, y)$ stetig auf $[a,\,b]$
2.) Die partielle Ableitung f_y existiert in jedem Punkt (x, y) aus $[a,\,b] \times [c,\,d]$ und ist dort stetig. Dann definiert die Vorschrift

$$F(y) \underset{\text{def}}{=} \int_a^b f(x, y)\,dx$$

eine in jedem Punkt $x \in [a, b]$ stetig differenzierbare Funktion F und

$$F'(y) = \int_a^b f_y(x, y)\,dx\,.$$

Beweis: Wir beschränken uns zunächst auf den Fall, daß $I = [c, d]$ ein kompaktes Intervall ist und zeigen, daß der Differenzenquotient für $h \to 0$ gegen das rechtsstehende Integral konvergiert, d.h.

$$\lim_{h\to 0} \frac{F(y_0 + h) - F(y_0)}{h} = \int_a^b f_y(x, y_0)\,dx\,.$$

für alle $y_0 \in [c, d]$. Liegt auch $y_0 + h$ in $[c, d]$, können wir den Zähler schreiben als

$$\begin{aligned} F(y_0 + h) - F(y_0) &= \int_a^b f(x, y_0 + h)\,dx - \int_a^b f(x, y_0)\,dx \\ &= \int_a^b \big(f(x, y_0 + h) - f(x, y_0)\big)\,dx\,. \end{aligned}$$

Da f bezüglich y eine Stammfunktion von f_y ist, können wir den Integranden weiter ausrechnen als

$$f(x, y_0 + h) - f(x, y_0) = \int_{y_0}^{y_0+h} f_y(x, y)\,dy\,,$$

und nach dem Mittelwertsatz der Integralrechnung aus Kapitel 4, §*a3)* gibt es ein $\theta \in [0, 1]$, so daß dieses Integral gleich $hf_y(x, y_0+\theta h)$ ist. Somit ist

$$f(x, y_0 + h) - f(x, y_0) = hf_y(x, y_0 + \theta h)\,,$$

also

$$\frac{F(y_0 + h) - F(y_0)}{h} = \int_a^b f(x, y_0 + \theta h)\,dx\,.$$

Wir müssen zeigen, daß wir beim Grenzübergang $h \to 0$ auch rechts einfach $h = 0$ setzen können. Dazu verwenden wir die Kompaktheit von $[a, b]\times[c, d]$: Nach Voraussetzung ist f_y dort stetig, also wegen der Kompaktheit nach dem letzten Lemma aus Kapitel 5, §4*a)* sogar gleichmäßig stetig. Daher gibt es zu jedem $\eta > 0$ ein $\delta > 0$, so daß

$$\left|f_y(x, y) - f_y(x^*, y^*)\right| < \eta \quad \text{falls } \left\|(x, y) - (x^*, y^*)\right\|_\infty < \delta\,.$$

Insbesondere ist also $|f(x, y_0 + \theta h) - f(x, y_0)| < \eta$, falls $|\theta h| < \delta$, also erst recht, wenn $|h| < \delta$ ist. Somit ist

$$\begin{aligned} &\left|\frac{F(y + h) - F(y)}{h} - \int_a^b f_y(x, y)\,dx\right| \\ &= \left|\int_a^b f_y(x, y_0 + \theta h)\,dx - \int_a^b f_y(x, y)\,dx\right| \\ &\le \int_a^b \left|f_y(x, y_0 + \theta h) - f_y(x, y_0)\right|\,dx < \eta\cdot(b - a) \end{aligned}$$

für alle h mit $|h| < \delta$.

Zum Abschluß des Beweises für kompakte Intervalle betrachten wir irgendein $\varepsilon > 0$ und setzen $\eta = \varepsilon/(b - a)$. Für $|h| < \delta$ ist dann

$$\left|\frac{F(y + h) - F(y)}{h} - \int_a^b f_y(x, y)\,dx\right| < \eta\cdot(b - a) = \varepsilon\,,$$

das heißt

$$\lim_{h\to 0} \frac{F(y + h) - F(y)}{h} = \int_a^b f_y(x, y)\,dx\,.$$

Damit ist die Differenzierbarkeit von F und auch die behauptete Formel für F' gezeigt.

Bleibt noch der Fall, daß I kein kompaktes Intervall ist. Dann liegt trotzdem jedes $y \in I$ in einem kompakten Teilintervall; da die Behauptung

für alle Punkte aus diesem Teilintervall richtig ist, gilt sie insbesondere für y, also, da y beliebig war, für alle Punkte aus I.
■

Damit können wir nun zeigen, daß das oben definierte Integral einer stetigen Funktion über einen Quader unabhängig ist von der Integrationsreihenfolge. Wir betrachten zunächst den Fall einer stetigen Funktion zweier Variablen; hier ist die Behauptung die folgende schwache Form des Satzes von FUBINI:

Satz: $f\colon D \to \mathbb{R}$ sei eine stetige Funktion auf der offenen Teilmenge $D \subseteq \mathbb{R}^2$. Für ein Rechteck $[a, b] \times [c, d] \subset D$ ist dann

$$\int_a^b \left(\int_c^d f(x,y)\,dy \right) dx = \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy\,.$$

Beweis: Wir betrachten die beiden Funktionen

$$F(x) = \int_a^x \left(\int_c^d f(\xi,y)\,dy \right) d\xi \quad \text{und} \quad G(x) = \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy\,;$$

die Behauptung ist dann äquivalent zur Gleichung $F(b) = G(b)$. Nach dem Hauptsatz der Differential- und Integralrechnung und den beiden vorigen Lemmata ist

$$F'(x) = \int_c^d f(x,y)\,dx = G'(x)\,,$$

F und G unterscheiden sich also höchstens um eine additive Konstante. Da $F(a) = G(a) = 0$ ist, muß diese aber verschwinden, d.h. $F(x) = G(x)$ für alle x und damit insbesondere für $x = b$.
■

Wie wir in den Übungen sehen werden, kann sich der Aufwand für die Berechnung der beiden Seiten der Gleichung im Satz von FUBINI deutlich unterscheiden; der Satz hat also durchaus auch praktische Bedeutung. Um ihn auf Funktionen von $n > 2$ Variablen zu verallgemeinern, beachten wir, daß sich jede Permutation π der n Variablen $x_1, \ldots, x_n$ als Produkt von Transpositionen schreiben läßt, d.h. also als

Produkt von Vertauschungen zweier Variablen x_i und x_j. Tatsächlich können wir π sogar als Produkt von Vertauschungen benachbarter Variablen schreiben: Bezeichnet nämlich $(i\ j)$ für $i < j$ die Vertauschung von x_i und x_j, so ist die Hintereinanderausführung $(j\ j+1)(i\ j)(j\ j+1)$ gleich $(i\ j+1)$. Somit folgt induktiv, daß sich die Transposition $(i\ j)$ mit $i < j$ schreiben läßt als

$$(j-1\ j)(j-2\ j)\cdots(i+1\ i+2)(i\ i+1)(i+1\ i+2)\cdots(j-2\ j-1)(j-1\ j)\,.$$

Auf die Vertauschung zweier benachbarter Variablen können wir aber den gerade bewiesenen Satz anwenden.

Daher können wir in der Formel

$$\int_Q f = \int_{a_n}^{b_n} \left(\cdots \left(\int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x_1, \ldots, x_n)\,dx_1 \right) dx_2 \right) \cdots \right) dx_n$$

die Reihenfolge der n Integrationen beliebig permutieren.

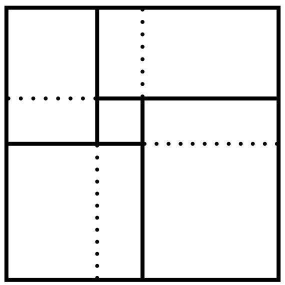
Der italienische Mathematiker GUIDO FUBINI (1879–1943) arbeitete zunächst auf dem Gebiet der Differentialgeometrie, interessierte sich dann aber immer mehr für analytische Themen wie Differentialgleichungen und Funktionen mehrerer komplexer Veränderlicher. 1901 wurde er Professor in Catania auf Sizilien, später in Genua und ab 1908 in Turin, wo er blieb, bis er 1939 trotz seiner angegriffenen Gesundheit wegen des italienischen Faschismus nach USA emigrierte und ans Institute for Advanced Study in Princeton wechselte. Der hier zitierte Satz ist zwar sein bekanntestes, aber ganz sicher nicht sein bedeutendstes Ergebnis.

Damit sind beide der gleich nach der Definition erwähnten Probleme gelöst und wir können unbesorgt mit den so definierten Integral rechnen. Wir wollen als nächstes einige zwar ziemlich offensichtliche, aber dennoch zu beweisende weitere Eigenschaften zusammenstellen:

Zerlegungsregel: Läßt sich der Quader Q als Vereinigung der paarweise volumenfremden Quader $Q_1, \ldots, Q_r$ schreiben, so ist

$$\int_Q f = \int_{Q_1} f + \cdots + \int_{Q_r} f\,.$$

Zum *Beweis* schreiben wir
$$Q=\prod_{i=1}^n[a_i,\,b_i]\quad\text{und}\quad Q_j=\prod_{i=1}^n[a_i^{(j)},\,b_i^{(j)}]\,.$$
Im Falle $n=1$ haben wir dann einfach ein Zerlegung des Intervalls $[a_1,\,b_1]$ in Teilintervalle; durch Umordnung der Q_j können wir erreichen, daß
$$a_1=a_1^{(1)}\le b_1^{(1)}=a_1^{(2)}\le b_1^2=a_1^{(3)}\le\cdots b_1^{(r-1)}=a_1^{(r)}\le b_1^{(r)}=b_1$$
ist, und wie wir aus Kapitel 4 wissen, ist
$$\int_{a_1}^{b_1}f(x_1)\,dx_1=\sum_{j=1}^r\int_{a_1^{(j)}}^{b_1^j}f(x_1)\,dx_1\,.$$
Leider können wir diese Zerlegung für $n>1$ nicht einfach durch die Definition durchziehen, denn unsere Quaderzerlegung muß nicht von festen Zerlegungen der Intervalle $[a_i,\,b_i]$ induziert sein. Wir können sie allerdings so verfeinern, daß dies der Fall ist, indem wir die Quader Q_j noch weiter zerlegen: Für jedes i betrachten wir die Menge der sämtlichen a_i^j und b_i^j und ordnen sie der Größe nach zu einer Folge
$$a_i=c_{i1}<c_{i2}<\ldots<c_{ir_i}=b_i\,,$$
wobei wir mehrfach vorkommende Zahlen nur einfach berücksichtigen. Dann betrachten wir die (in der Zeichnung gestrichelten) $\prod_{i=1}^n r_i$ Quader der Form
$$\prod_{i=1}^r[c_{ij_i},\,c_{i,j_i+1}]\,.$$



Auch sie bilden eine Zerlegung des Quaders Q durch paarweise volumenfremde Teilquader, und für jeden der Quader Q_j bildet die Teilmenge jener Quader, die in Q_j liegen, eine Zerlegung von Q_j, die ebenfalls von dieser speziellen Form ist. Es genügt daher, wenn wir zeigen, daß die Behauptung für diese spezielle Zerlegung von Q gilt, und da folgt sie induktiv aus der Definition von $\int_Q f$ und der obigen Formel für den Fall $n=1$. ■

Monotonieregel: Ist $f(x)\le g(x)$ für alle $x\in Q$, so ist $\int_Q f\le\int_Q g$. Insbesondere ist $\left|\int_Q f\right|\le\int_Q|f|$.

Beweis: Für den eindimensionalen Fall kennen wir die Monotonieregel aus Kapitel 4; die Verallgemeinerung auf n-dimensionale Quader folgt aus der Definition von $\int_Q f$ sofort durch vollständige Induktion. Die Aussage über den Betrag wiederum folgt daraus wegen der Ungleichung $-|x|\le x\le|x|$. ■

Aus den beiden vorigen Regeln folgt sofort

Korollar: $f\colon D\to\mathbb{R}$ sei stetig auf der offenen Menge $D\subseteq\mathbb{R}^n$ und die Quader $Q,Q_1,\ldots,Q_r$ seien Teilmengen von D.

a) Ist $Q\subseteq Q_1\cup\cdots\cup Q_r$, so ist
$$\int_Q f\le\int_{Q_1}f+\cdots+\int_{Q_r}f\,.$$
b) Ist $Q_1\cup\cdots\cup Q_r\subseteq Q$ und sind die Q_i paarweise volumenfremd, so ist
$$\int_{Q_1}f+\cdots+\int_{Q_r}f\le\int_Q f\,.$$
Beweis: a) Durch Unterteilung können wir annehmen, daß Q die Vereinigung gewisser volumenfremder Q_i ist; dann ist $\int_Q f$ die Summe der entsprechenden $\int_{Q_i}f$. Die restlichen $\int_{Q_j}f$ sind wegen der Nichtnegativität von f allesamt größer oder gleich null.

b) Hier können wir durch Unterteilung annehmen, daß Q die volumenfremde Vereinigung der Q_i zusammen mit einigen weiteren Quadern $Q_j\subset Q$ ist. Da $\int_{Q_j}f\ge0$ für alle j, folgt auch hier die Behauptung. ■

Linearität: Für zwei stetige Funktionen $f, g: D \to \mathbb{R}$ und $a, b \in \mathbb{R}$ gilt für einen Quader $Q \subset D \subseteq \mathbb{R}^n$

$$\int_Q (af + bg) = a \int_Q f + b \int_Q g \,.$$

Der *Beweis* folgt induktiv aus der entsprechenden Aussage im Eindimensionalen. ∎

Lemma: Für jedes $c \in \mathbb{R}$ ist $\int_Q c = c \cdot \mu(Q)$.

Beweis: Auch dies folgt induktiv aus dem entsprechenden Lemma für den eindimensionalen Fall

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

∎

Damit haben wir Integrale über Quader recht gut verstanden; in vielen Anwendungen müssen wir aber über allgemeinere Teilmengen von $\mathbb{R}^n$ integrieren. Als ersten Zwischenschritt in Richtung auf entsprechende Integrale betrachten wir stetige Funktionen, die außerhalb eines Quaders verschwinden:

Definition: *a)* Für eine Teilmenge $X \subseteq \mathbb{R}^n$ ist der *Abschluß* $\overline{X}$ von X die kleinste abgeschlossene Teilmenge $Y \subseteq \mathbb{R}^n$, die X enthält, d.h. der Durchschnitt aller abgeschlossener Teilmengen $Z \subseteq \mathbb{R}^n$, die X enthalten.
b) $f: D \to \mathbb{R}$ sei eine stetige Funktion auf der offenen Menge $D \subseteq \mathbb{R}^n$. Der *Träger* von f ist der Abschluß der Menge aller $x \in D$, für die $f(x)$ nicht verschwindet, d.h. die Menge $\operatorname{Tr}(f) = \overline{\{x \in D \mid f(x) \neq 0\}}$.
c) Eine stetige Funktion $f: D \to \mathbb{R}$ heißt *Funktion mit kompaktem Träger,* falls $\operatorname{Tr}(f)$ eine kompakte Teilmenge von $\mathbb{R}^n$ ist.
d) $C^0(D, \mathbb{R})$ ist die Menge aller stetiger Funktionen $D \to \mathbb{R}$ und $K^0(D, \mathbb{R})$ die Teilmenge der Funktionen mit kompaktem Träger.

$C^0(D, \mathbb{R})$ ist ein Vektorraum, da jede Linearkombination stetiger Funktionen wieder stetig ist, und $K^0(D, \mathbb{R})$ ist ein Untervektorraum, denn der Träger einer Linearkombination zweier Funktionen f, g ist enthalten

in der Vereinigung der Träger von f und von g. Bei der Definition eines möglichst allgemeinen n-dimensionalen Integrals werden die Funktionen mit kompaktem Träger im folgenden die gleiche Rolle spielen wie die Treppenfunktionen bei der Definition des RIEMANN-Integrals für Funktionen einer Veränderlichen.

Da jede kompakte Teilmenge von $\mathbb{R}^n$ beschränkt ist, liegt sie insbesondere in einem Quader; für eine Funktion $f: \mathbb{R}^n \to \mathbb{R}$ mit kompaktem Träger gibt es also einen Quader $Q \subseteq \mathbb{R}^n$, so daß $f(x) = 0$ für alle $x \not\in Q$. Da f nach Definition stetig auf ganz $\mathbb{R}^n$ und damit insbesondere auch auf Q ist, existiert das Integral über Q und wir *definieren*

$$\int_{R^n} f = \int_Q f \,.$$

Diese Definition ist nur sinnvoll, wenn wir zeigen können, daß sie nicht vom Quader Q abhängt; das ist aber einfach, denn sind $Q, \widetilde{Q}$ zwei Quader, die den Träger von f enthalten, so liegt der Träger auch im Durchschnitt $Q \cap \widetilde{Q}$. Wie im Beweis der Zerlegungsregel können wir Quader $Q_1, \ldots, Q_r$ finden derart, daß sowohl Q als auch $\widetilde{Q}$ als auch der Durchschnitt $Q \cap \widetilde{Q}$ Vereinigung gewisser Q_i sind. Alle Q_i, die nicht in $Q \cap \widetilde{Q}$ liegen, sind disjunkt zum Träger von f; nach der Zerlegungsregel sind sowohl $\int_Q f$ als auch $\int_{\widetilde{Q}} f$ als auch $\int_{Q \cap \widetilde{Q}} f$ gleich der Summe über diejenigen $\int_{Q_i} f$, für die Q_i im Durchschnitt $Q \cap \widetilde{Q}$ liegt.

Für die Konstruktion allgemeinerer Integrale wollen wir als erstes Integrale allgemeinerer als der bisher betrachteten Funktionen über $\mathbb{R}^n$ definieren; danach wollen wir auch über allgemeinere Teilmengen von $\mathbb{R}^n$ integrieren.

Für beides ist es wichtig, $K^0(\mathbb{R}^n, \mathbb{R})$ zu einem *normierten* Vektorraum zu machen. Um zu sehen, daß nicht auf jedem normierten Vektorraum alle Normen, so wie auf $\mathbb{R}^n$, äquivalent sein müssen, definieren wir gleich zwei Normen: Die erste Norm, die L^1-Norm ist definiert als

$$\|f\|_1 = \int_{\mathbb{R}^n} f\,;$$

nach Definition einer Funktion mit kompaktem Träger existiert dieses Integral für jede Funktion mit kompaktem Träger.

Als zweite Norm auf $K^0(\mathbb{R}^n, \mathbb{R})$ betrachten wir die Supremumsnorm

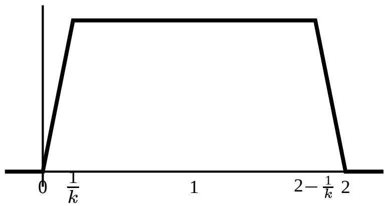
$$\|f\|_\infty = \sup\big\{|f(x)| \;\big|\; x \in \mathbb{R}^n\big\}\,.$$

Für Funktionen mit kompaktem Träger ist das Supremum des Betrags über $\mathbb{R}^n$ gleich dem Supremum von $|f(x)|$ über den kompakten Träger, denn auch $|f(x)|$ ist eine stetige Funktion ist. Damit ist dieses Supremum nach dem Satz vom Maximum gleich dem Funktionswert an einem Punkt im Träger, also eine reelle Zahl.

Wir bezeichnen den normierten Vektorraum $K^0(\mathbb{R}^n, \mathbb{R})$ mit der L^1-Norm aus $K_1^0(\mathbb{R}^n, \mathbb{R})$; mit der Supremumsnorm bezeichnen wir ihn als $K_\infty^0(\mathbb{R}^n, \mathbb{R})$. Bezüglich beider Normen ist der Vektorraum nicht vollständig:

Betrachten wir etwa die Folge der Funktionen $f_k\colon \mathbb{R} \to \mathbb{R}$ mit

$$f_k(x) = \begin{cases} 0 & \text{falls } x \notin (0,\ 2) \\ kx & \text{falls } 0 \le x \le \frac{1}{k} \\ 1 & \text{falls } \frac{1}{k} \le x \le 2 - \frac{1}{k} \\ k(2-x) & \text{falls } 2 - \frac{1}{k} \le x \le 2 \end{cases}\,.$$



Sie konvergiert punktweise gegen die Funktion $f\colon \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 0 & \text{falls } x \notin (0,\ 2) \\ 1 & \text{falls } x \in (0,\ 2) \end{cases}\,,$$

denn für jedes $x \in (0,\ 2)$ gibt es ein $N \in \mathbb{N}$, so daß $\frac{1}{k} \le x \le 2 - \frac{1}{k}$ für alle $k \ge N$. Diese Funktion ist unstetig an den Sprungstellen bei $x = 0$ und $x = 2$, liegt also insbesondere nicht in $K^0(\mathbb{R}, \mathbb{R})$. Trotzdem ist die Folge der f_k eine CAUCHY-Folge, denn für $\ell \ge k$ stimmen die Funktionen f_k und f_l überein außer in den beiden Intervallen $(0,\ \frac{1}{k})$ und $(2 - \frac{1}{k},\ 2)$. Diese haben jeweils die Länge $\frac{1}{k}$, und da alle Funktionswerte

in $[0,\ 1]$ liegen, ist dort $|f_\ell(x) - f_k(x)| \le 1$. Daher ist

$$\|f_\ell - f_k\|_1 = \int_{\mathbb{R}} |f_\ell - f_k| \le \frac{2}{k}\,.$$

Wählen wir also zu vorgegebenem $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit $2/N < \varepsilon$, so ist $\|f_\ell - f_k\| < \varepsilon$ für alle $k, \ell \ge N$, die Folge $(f_k)_{k\in\mathbb{N}}$ ist also eine CAUCHY-Folge bezüglich der L^1-Norm, konvergiert aber nicht gegen eine Funktion aus $K^0(\mathbb{R}, \mathbb{R})$.

Bezüglich der Supremumsnorm ist $(f_k)_{k\in\mathbb{N}}$ keine CAUCHY-Folge, denn für $k \le \ell$ ist $f_\ell\big(\frac{1}{\ell}\big) - f_k\big(\frac{1}{\ell}\big) = 1 - \frac{k}{\ell}$, und das ist für $\ell \ge 2k$ stets größer als $\frac{1}{2}$.

Es gibt allerdings auch CAUCHY-Folgen bezüglich der Supremumsnorm, die nicht gegen ein Element von $K^0(\mathbb{R}, \mathbb{R})$ konvergieren. Als Beispiel betrachten wir die Funktionen $g(x) = e^{-x^2}$ und

$$g_k(x) = \begin{cases} e^{-x^2} & \text{falls } |x| \le k \\ e^{-k^2}\big(|k+1| - |x|\big) & \text{falls } k \le |x| \le k+1 \\ 0 & \text{sonst} \end{cases}\,.$$

Offensichtlich liegen alle g_k in $K^0(\mathbb{R}, \mathbb{R})$, denn sie sind stetig und der Träger $[-k-1, k+1]$ ist kompakt. Für $\ell \ge k$ ist $|g_\ell(x) - g_k(x)| \le e^{-k^2}$ für alle $x \in \mathbb{R}$, also ist $\|g_k - g_\ell\| < e^{-N^2}$ für alle $k, \ell \ge N$ und wie haben eine CAUCHY-Folge bezüglich der Supremumsnorm. Da die Grenzfunktion g für kein $x \in \mathbb{R}$ verschwindet, hat g aber natürlich keinen kompakten Träger.

Man könnte die Unvollständigkeit von $K^0(\mathbb{R}, \mathbb{R})$ bezüglich beider Normen als Nachteil sehen, tatsächlich aber das genaue Gegenteil der Fall: Die Funktionen, die uns wirklich interessieren, haben schließlich nur in den seltensten Fällen kompakte Träger, und zumindest gelegentlich müssen wir auch Funktionen mit Sprungstellen betrachten, beispielsweise bei Kostenfunktionen mit Rabattstaffeln. Wie wir sehen werden, lassen sich fast alle solche Funktionen als Grenzwerte von CAUCHY-Folgen von Funktionen mit kompaktem Träger bezüglich der L^1-Norm realisieren, die zur Vollständigkeit von $K_1^0(\mathbb{R}, \mathbb{R})$ „fehlenden“ Funktionen sind also genau die Funktionen, die uns interessieren.