Scriptum zur Vorlesung
Analysis C

Prof. W. Hoffmann

1 Der Brouwersche Fixpunktsatz

In Analysis I wird folgender Fixpunktsatz bewiesen:

Satz 0 *Jede stetige Abbildung f eines kompakten Intervalls $[a_0, a_1]$ in sich selbst hat einen Fixpunkt.*

Zum Beweis betrachtet man die Funktion $g(x) = f(x) - x$. Wegen $f(a_0) \in [a_0, a_1]$ ist $f(a_0) \geq a_0$, also $g(a_0) \geq 0$. Analog zeigt man $g(a_1) \leq 0$. Nach dem Zwischenwertsatz existiert ein $c \in [a_0, a_1]$ mit $g(c) = 0$, also $f(c) = c$.

Im n-dimensionalen Fall funktioniert der Trick mit der Funktion g nicht. Wir wollen versuchen, den Beweis umzuformulieren, in der Hoffnung, ihn zu verallgemeinern. Es sei

$$C_0 = \{x \in [a_0, a_1] \mid f(x) \geq x\}, \qquad C_1 = \{x \in [a_0, a_1] \mid f(x) \leq x\}.$$

Jeder Punkt aus $C_0 \cap C_1$ ist ein Fixpunkt von f.

Lemma 1 *Ist das kompakte Intervall $[a_0, a_1]$ die Vereinigung von abgeschlossenen Teilmengen C_0 und C_1 mit $a_0 \in C_0$ und $a_1 \in C_1$, so ist $C_0 \cap C_1$ nicht leer.*

Beweis. Es sei $c = \sup C_0$. Da C_0 abgeschlossen ist, gehört c zu C_0. Ist $c = a_1$, so sind wir fertig, andernfalls ist das nichtleere halboffene Intervall $(c, a_1]$ in $[a_0, a_1] \setminus C_0$ enthalten, also in C_1, und aus der Abgeschlossenheit von C_1 folgt $c \in C_1$. $\square$

Definition 1 *Ein metrischer (oder topologischer) Raum X heißt Hausdorff-zusammenhängend, wenn es keine Zerlegung von X in zwei disjunkte abgeschlossene Teilmengen gibt.*

Bemerkung. Eine Teilmenge von $\mathbb{R}$ ist genau dann Hausdorff-zusammenhängend, wenn sie ein Intervall ist.

Ist nämlich ein Intervall die Vereinigung der disjunkten abgeschlossenen Teilmengen A_0 und A_1, so nehme man $a_0 \in A_0$, $a_1 \in A_1$ und betrachte $C_0 = A_0 \cap [a_0, a_1]$ und $C_1 = A_1 \cap [a_0, a_1]$ (falls $a_0 \leq a_1$, ansonsten vertausche man die Rollen). Ist hingegen $A \subset \mathbb{R}$ kein Intervall, so gibt es a_0, $a_1 \in A$ und $x \notin A$ mit $a_0 < x < a_1$, und dann ist A die Vereinigung der disjunkten abgeschlossenen Teilmengen $A_0 = (-\infty, x] \cap A$ und $A_1 = [x, \infty) \cap A$.

Der Satz 0 verallgemeinert sich nicht auf beliebige zusammenhängende Teilmengen von $\mathbb{R}^n$ (siehe Aufgabe 4). Wir brauchen die stärkere Eigenschaft der Konvexität.

Definition 2 *Eine Teilmenge K eines Vektorraums heißt konvex, wenn für beliebige Elemente x, $y \in K$ auch die Verbindungsstrecke*

$$\{\lambda x + (1 - \lambda) y \mid 0 \leq \lambda \leq 1\}$$

in K liegt.

Satz 1 (Brouwer) *Es sei K eine nichtleere kompakte konvexe Teilmenge eines endlichdimensionalen reellen Vektorraums und $f : K \to K$ eine stetige Abbildung. Dann besitzt f einen Fixpunkt.*

Für den Beweis brauchen wir einige Vorbereitungen. Was sind eigentlich die einfachsten konvexen Menge in $\mathbb{R}^n$?

Definition 3 *Ist X eine Teilmenge in einem Vektorraum, so bezeichnen wir die kleinste konvexe Menge, die X enthält, als konvexe Hülle von X, kurz* $\operatorname{conv} X$.

Bemerkung: Gemeint ist, dass $\operatorname{conv} X$ in jeder konvexen Menge enthalten ist, welche X enthält. Die konvexe Hülle existiert: Sie ist der Durchschnitt aller konvexen Mengen, welche X enthalten.

Lemma 2 *Die konvexe Hülle einer endlichen Menge $A = \{a_0, \ldots, a_n\}$ besteht aus allen Punkten der Form*

$$x = \lambda_0 a_0 + \cdots + \lambda_n a_n,$$

wobei

$$\lambda_i \geq 0 \quad \forall i, \qquad \lambda_0 + \cdots + \lambda_n = 1.$$

Beweis. Es sei S die Menge aller Punkte der angegebenen Form. Dann ist S konvex, denn ist ein weiterer Punkt $x' \in S$ gegeben, also

$$x' = \lambda'_0 a_0 + \cdots + \lambda'_n a_n \quad \text{mit} \quad \lambda'_i \geq 0, \quad \lambda'_0 + \cdots + \lambda'_n = 1,$$

so ist für $0 \leq \lambda \leq 1$ auch $\lambda x + (1-\lambda)x' \in S$. Offensichtlich ist $A \subseteq S$.

Es bleibt zu zeigen, dass für eine beliebige konvexe Menge $K \supseteq A$ gilt $K \supseteq S$. Für $n = 0$ ist nichts zu beweisen. Angenommen, die Behauptung gilt bereits für n Punkte. Ist nun $x \in S$, so ist entweder $\lambda_0 = 1$, also $x = a_0 \in A \subseteq K$, oder aber $\lambda_0 < 1$, und wir können schreiben

$$x = \lambda_0 a_0 + (1-\lambda_0)y,$$

wobei

$$y = \mu_1 a_1 + \cdots + \mu_n a_n \quad \text{mit} \quad \mu_i = \frac{\lambda_i}{1-\lambda_0}.$$

Dann ist $\mu_i \geq 0$ und $\mu_1 + \cdots + \mu_n = 1$, also nach Induktionsvoraussetzung $y \in K$ und nach Definition der Konvexität $x \in K$. □

Bemerkung. Die konvexe Hülle einer endlichen Menge $\{a_0, \ldots, a_n\}$ ist als stetiges Bild der kompakten Menge

$$\{(\lambda_0, \ldots, \lambda_n) \in [0, \infty)^{n+1} \mid \lambda_0 + \cdots + \lambda_n = 1\}$$

ebenfalls kompakt.

Definition 4 *Die Punkte $a_0, \ldots, a_n$ heißen affin unabhängig, wenn für beliebige Zahlen $\mu_0, \ldots, \mu_n$ mit den Eigenschaften*

$$\mu_0 a_0 + \cdots + \mu_n a_n = 0 \quad \textit{und} \quad \mu_0 + \cdots + \mu_n = 0$$

gilt, dass alle μ_i gleich Null sein müssen. In diesem Fall heißt die konvexe Hülle

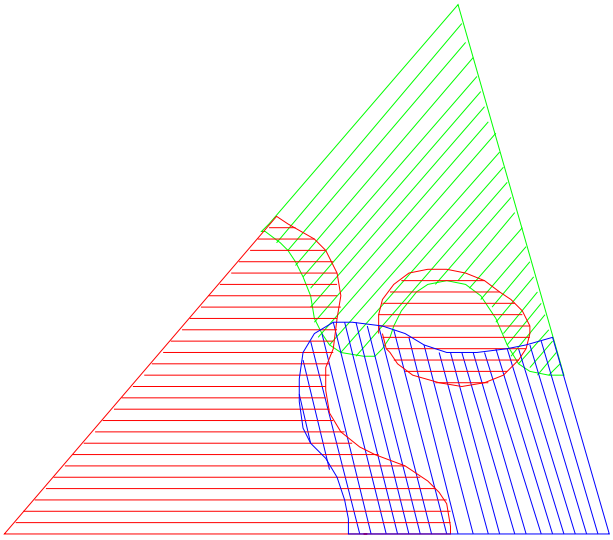
$$S = \operatorname{conv}\{a_0, \ldots, a_n\}$$

das Simplex mit den Ecken $a_0, \ldots, a_n$. Ist $B \subseteq \{a_0, \ldots, a_n\}$, so nennt man $\operatorname{conv} B$ *eine Seite von S. Ein Simplex mit $n+1$ Ecken heißt n-Simplex.*

Ein Punkt $x \in S$ lässt sich als Schwerpunkt (griechisch: Barykentron) von Massenpunkten in den Ecken interpretieren, wobei im Punkt a_i eine Masse der Größe λ_i konzentriert ist. Für jeden Punkt $x \in S$ sind die Zahlen λ_i in Lemma 2 eindeutig bestimmt und heißen baryzentrische Koordinaten von x bezüglich $(a_0, \ldots, a_n)$. Sie lassen sich durch die Koordinaten von x bezüglich gewisser linear unabhängiger Vektoren ausdrücken und hängen darum stetig von x ab.

Eine Teilmenge einer affin unabhängigen Menge ist ebenfalls affin unabhängig, darum ist jede Seite eines Simplexes wieder ein Simplex.

Satz 2 (Knaster-Kuratowski-Mazurkiewicz) *Es sei S ein Simplex mit den Ecken $a_0, \ldots, a_n$, und es seien abgeschlossene Teilmengen $C_0, \ldots, C_n$ von S gegeben. Wir nehmen an, dass eine beliebige Seite von S von den Mengen C_i, für die a_i eine Ecke dieser Seite ist, überdeckt wird. Dann ist der Durchschnitt $C_0 \cap \cdots \cap C_n$ nicht leer.*



Bemerkung. Jede Ecke von S ist eine Seite, also wird im Satz insbesondere $a_i \in C_i$ vorausgesetzt. Auch S ist eine Seite von sich selbst, also wird auch vorausgesetzt, dass $S = C_0 \cup \cdots \cup C_n$. Der Spezialfall $n = 1$ ist natürlich Lemma 1.

Ist $x \in S$ mit baryzentrischen Koordinaten λ_i, so nennt man die Menge der Ecken a_i mit $\lambda_i \neq 0$ den Träger von x. Man kann die Voraussetzung von Satz 2 auch so formulieren: Für einen beliebigen Punkt $x \in S$ gibt es eine Ecke a_i im Träger von x mit der Eigenschaft $x \in C_i$.

Beweis des Brouwerschen Fixpunktsatzes für ein beliebiges Simplex S unter Benutzung von Satz 2. Wir bezeichnen die baryzentrischen Koordinaten eines Punktes $x \in S$ mit $\lambda_i(x)$, und für eine gegebene stetige Abbildung $f : S \to S$ setzen wir

$$C_i = \{x \in S \mid \lambda_i(f(x)) \leq \lambda_i(x)\}.$$

Wegen der Stetigkeit von f und λ_i sind diese Mengen abgeschlossen. Für einen beliebigen Punkt ist die Summe der baryzentrischen Koordinaten, die zu seinem Träger gehören, gleich 1, also können nicht alle von ihnen bei der Anwendung von f wachsen. Es muss also ein a_i im Träger von x geben, so dass $\lambda_i(f(x)) \leq \lambda_i(x)$, d. h. $x \in C_i$.

Nach Satz 2 existiert somit ein $c \in S$, so dass keine baryzentrische Koordinate von c bei Anwendung von f wächst. Da aber die Summe gleich 1 bleiben muss, kann auch keine Koordinate fallen, d. h.

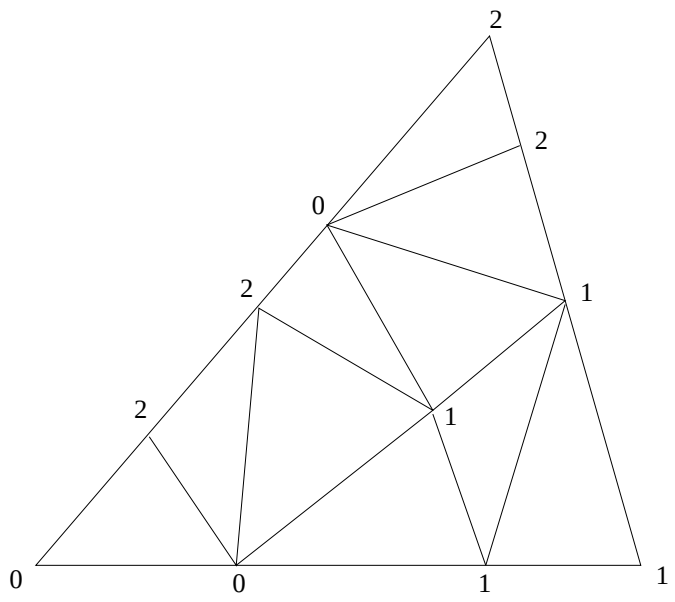
$$\lambda_i(f(c)) = \lambda_i(c) \quad \forall i$$

und somit $f(c) = c$. $\square$

Definition 5 *Ein Simplizialkomplex*[1] *in einem reellen Vektorraum ist eine endliche Menge $\mathcal{T}$ von Simplexen $T \subset \mathbb{R}^n$ mit folgenden Eigenschaften:*

1. *Jede Seite eines Simplexes $T \in \mathcal{T}$ gehört zu $\mathcal{T}$.*
2. *Der Durchschnitt zweier Simplexe T_1, $T_2 \in \mathcal{T}$ ist leer oder eine gemeinsame Seite von T_1 und T_2.*

Die 0-Simplexe heißen auch Ecken von $\mathcal{T}$. Ein Simplizialkomplex $\mathcal{T}$ heißt Unterteilung des Simplexes S, wenn die Vereinigung aller Simplexe $T \in \mathcal{T}$ gleich S ist. Allgemeiner heißt ein Simplizialkomplex $\mathcal{U}$ Unterteilung des Simplizialkomplexes $\mathcal{T}$, wenn jedes Simplex von $\mathcal{U}$ in einem Simplex von $\mathcal{T}$ enthalten ist und jedes Simplex von $\mathcal{T}$ eine Vereinigung von Simplexen von $\mathcal{U}$ ist.



Definition 6 *Es sei $\mathcal{T}$ eine Unterteilung des Simplexes $S = \operatorname{conv}\{a_0, \ldots, a_n\}$. Eine zulässige Färbung von $\mathcal{T}$ ist eine Abbildung („Färbung“), die jeder Ecke*

[1] genauer: endlicher Simplizialkomplex

b von $\mathcal{T}$ einen Index i so zuordnet, dass die Ecke a_i von S im Träger von b liegt. Ein n-Simplex von $\mathcal{T}$ heißt bunt, wenn seine $n+1$ Ecken mit sämtlichen Indizes $0, \ldots, n$ gefärbt sind.

Lemma 3 (Sperner) *Die Anzahl der bunten Simplexe einer zulässigen Färbung ist ungerade.*

Beweis durch Induktion nach n. Für $n = 0$ besteht T nur aus dem 0-Simplex S, und dieser ist bunt. Wir betrachten also den Fall $n > 0$ und halten eine $(n-1)$-dimensionale Seite R von S fest. Ein $(n-1)$-Simplex von $\mathcal{T}$ heiße R-bunt, wenn seine Ecken mit sämtlichen Indizes markiert sind, die bei den Ecken von R vorkommen. (In diesem Sinne bedeutet „bunt“ nichts anderes als „S-bunt“.)

Angenommen, das n-Simplex $T \in \mathcal{T}$ hat die R-bunte Seite V. Ist die Ecke von T, die nicht zu V gehört, mit dem in R fehlenden Index gefärbt, so ist T bunt und hat keine weitere R-bunte Seite. Andernfalls ist T nicht bunt und hat genau zwei R-bunte Seiten. Wir sehen, dass ein beliebiges n-Simplex $T \in \mathcal{T}$ keine, eine oder zwei R-bunte Seiten hat, wobei der mittlere Fall genau dann eintritt, wenn T selbst bunt ist. Bezeichnen wir also die Anzahl der bunten n-Simplexe von $\mathcal{T}$ mit k und für jedes n-Simplex $T \in \mathcal{T}$ die Anzahl der R-bunten Seiten von T mit $k_R(T)$, so folgt

$$k \equiv \sum_T k_R(T) \pmod 2,$$

wobei über alle n-Simplexe $T \in \mathcal{T}$ summiert wird.

Bezeichnen wir für jedes $(n-1)$-Simplex $V \in \mathcal{T}$ die Anzahl der n-Simplexe $T \in \mathcal{T}$, die V als Seite haben, mit $l(V)$, so gilt

$$\sum_T k_R(T) = \sum_V l(V),$$

wobei rechts über alle R-bunten $(n-1)$ Simplexe $V \in \mathcal{T}$ summiert wird. Nun ist $l(V)$ gleich 1 oder 2, je nachdem, ob V auf dem Rand oder im Inneren von S liegt.[2] Bezeichnen wir die Anzahl der R-bunten $(n-1)$-Simplexe von $\mathcal{T}$ auf dem Rand von S mit l_R, so folgt

$$\sum_V l(V) \equiv l_R \pmod 2.$$

Es bleibt zu zeigen, dass l_R ungerade ist.

[2] Genaugenommen müssten wir das beweisen.

Liegt ein $(n-1)$-Simplex V auf dem Rand von S, so liegt es in genau einer $(n-1)$-dimensionalen Seite von S, und die Vereinigung der Träger aller Elemente von V ist die Eckenmenge dieser Seite. Wir wollen zeigen, dass für ein R-buntes V diese Seite gleich R ist. Ist a_i eine Ecke von R, so hat das R-bunte Simplex V eine Ecke b mit der Färbung i, und wegen der Zulässigkeit der Färbung liegt a_i im Träger von b. Somit muss V tatsächlich in der Seite R liegen, und l_R ist die Anzahl der R-bunten Simplexe von $\mathcal{T}$ in R.

Sämtliche Simplexe von T, die in R liegen, bilden eine Unterteilung von R, und die Einschränkung der Färbung auf diese Unterteilung ist zulässig. Nach Induktionsvoraussetzung ist l_R ungerade. □

Beweis von Satz 2. Es sei $\mathcal{T}$ eine beliebige Unterteilung von S. Ist b eine Ecke von $\mathcal{T}$, so gibt es nach Voraussetzung eine Ecke a_i von S im Träger von b, so dass $b \in C_i$. Färben wir b mit solch einem Index i, so erhalten wir eine zulässige Färbung von $\mathcal{T}$. Nach Lemma 3 existiert ein buntes Simplex, und wenn wir seine Ecken entsprechend der Färbung mit $b_0, \ldots, b_n$ bezeichnen, so gilt $b_i \in C_i$ für jedes i.

Nun sei $\mathcal{T}_m$ eine Folge von Unterteilungen von S, so dass der maximale Durchmesser δ_m der Simplexe von $\mathcal{T}_m$ für $m \to \infty$ gegen Null konvergiert.[3] Dann erhalten wir Folgen $b_0^{(m)} \in C_0, \ldots, b_n^{(m)} \in C_n$. Da S kompakt ist, können wir durch Übergang zu einer Teilfolge annehmen, dass diese Folgen konvergieren. Wegen $\lim_{m\to\infty} \delta_m = 0$ konvergieren sie gegen denselben Punkt c, und wegen der Abgeschlossenheit von C_i ist $c \in C_i$ für alle i. □

Damit ist der Satz von Knaster-Kuratowski-Mazurkiewicz und somit auch der Brouwersche Fixpunktsatz für Simplexe bewiesen.

Definition 7 *Es sei X ein metrischer Raum und $Y \subseteq X$. Eine Retraktion von X auf Y ist eine stetige Abbildung $r : X \to Y$ mit der Eigenschaft $r|_Y = \mathrm{id}_Y$. Die Teilmenge Y des metrischen Raumes X heißt Retrakt von X, wenn es eine Retraktion von X auf Y gibt.*

Lemma 4 *Ist Y ein Retrakt des metrischen Raumes X und hat jede stetige Selbstabbildung von X einen Fixpunkt, so hat auch jede stetige Selbstabbildung von Y einen Fixpunkt.*

Beweis. Es sei $f : Y \to Y$ stetig und $r : X \to Y$ eine Retraktion. Dann ist $f \circ r : X \to Y$ stetig, besitzt also einen Fixpunkt $c \in X$. Insbesondere ist $c = f(r(c)) \in Y$, also $r(c) = c$ und $c = f(c)$. □

[3] Diese Unterteilungen lassen sich z. B. durch fortgesetzte baryzentrische Unterteilung gewinnen, vgl. Aufgabe 6. Zur Definition des Durchmessers fixieren wir ein Skalarprodukt.

Definition 8 *Ein Hilbertraum ist ein Vektorraum mit Skalarprodukt, der bezüglich der Norm* $\|x\| = \sqrt{x \cdot x}$ *vollständig ist.*

Wir brauchen den folgenden Satz nur für endlichdimensionale Hilberträume, aber er gilt auch für unendlichdimensionale.

Satz 3 *Es sei* K *eine nichtleere konvexe abgeschlossene Menge in einem Hilbertraum* H*. Dann gibt es genau eine Abbildung* $r : H \to K$ *mit*

$$\|x - r(x)\| = \inf\{\|x - y\| \mid y \in K\}.$$

Dabei gilt für alle x*,* $x' \in H$*, dass*

$$\|r(x) - r(x')\| \leq \|x - x'\|.$$

Für den Beweis brauchen wir die Parallelogramm-Identität

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2),$$

die man einfach durch Ausmultiplizieren nachprüft.

Beweis. Zum Nachweis, dass der Wert $d := \inf\{\|x - y\| \mid y \in K\}$ tatsächlich von einem Punkt $z \in K$ angenommen wird, genügt es, den Fall $x = 0$ zu betrachten, da man sonst alles um den Vektor $-x$ verschieben kann. Nach Definition des Infimums gibt es zu jedem n ein $y_n \in K$ mit $\|y_n\| < d + \frac{1}{n}$. Aus der Parallelogramm-Identität folgt

$$\left\|\frac{y_m + y_n}{2}\right\|^2 + \left\|\frac{y_m - y_n}{2}\right\|^2 = \frac{1}{2}(\|y_m\|^2 + \|y_n\|^2).$$

Da $\frac{1}{2}(y_m + y_n)$ in der konvexen Menge K liegt, ist der erste Term links mindestens d^2, so dass

$$\frac{1}{4}\|y_m - y_n\|^2 \leq \frac{1}{2}(\|y_m\|^2 + \|y_n\|^2) - d^2.$$

Nun sieht man ohne Weiteres, dass y_n eine Cauchyfolge ist, und wegen der Vollständigkeit von H konvergiert sie gegen einen Punkt z, der wegen der Abgeschlossenheit von K in K liegen muss. Dies gilt, wie gesagt, auch ohne die Annahme $x = 0$.

Für $y \in K$ und $0 \leq \lambda \leq 1$ ist auch $y_\lambda := \lambda y + (1 - \lambda)z \in K$, und

$$\begin{aligned}\|x - y_\lambda\|^2 &= \|(x - z) - \lambda(y - z)\|^2 \\ &= \|x - z\|^2 - 2\lambda(x - z) \cdot (y - z) + \lambda^2\|y - z\|^2.\end{aligned}$$

Diese quadratische Funktion von λ nimmt an der Stelle $\lambda = 0$ den Wert d an und ist auf dem Intervall $[0, 1]$ von unten durch d beschränkt. Es folgt, dass

$$(x - z) \cdot (y - z) \leq 0,$$

und hieraus wiederum, dass für $\lambda = 1$ und $y \neq z$ die strenge Ungleichung gilt. Das Minimum wird also nur in einem Punkt $z = r(x)$ angenommen, und für alle $x \in H$ und $y \in K$ gilt

$$(x - r(x)) \cdot (y - r(x)) \leq 0.$$

Hieraus erhalten wir für alle x, $x' \in H$

$$\begin{aligned}(r(x) - r(x')) \cdot (r(x) - r(x')) - (x - x') \cdot (r(x) - r(x')) \\ = (x - r(x)) \cdot (r(x') - r(x)) + (x' - r(x') \cdot (r(x) - r(x')) \leq 0.\end{aligned}$$

Mit Hilfe der Cauchy-Schwarz-Ungleichung folgt daraus

$$\|r(x) - r(x')\|^2 \leq \|x - x'\| \|r(x) - r(x')\|.$$

Ist $\|r(x) - r(x')\| \neq 0$, so können wir diesen Faktor herauskürzen, und andernfalls ist die letzte Behauptung des Satzes sowieso trivial. $\square$

Nun folgt der allgemeine Fall des Brouwerschen Fixpunktsatzes aus dem bewiesenen Spezialfall, indem wir die gegebene kompakte konvexe Teilmenge K in ein genügend großes Simplex S einbetten und die Retraktion aus Satz 3 benutzen.

Folgerung 1 *Es sei K die abgeschlossene Einheitskugel in $\mathbb{R}^n$ und S ihr Rand. Dann ist die Sphäre S kein Retrakt von K.*

Gäbe es nämlich eine Retraktion $r : K \to S$, so hätte nach dem Brouwerschen Fixpunktsatz auch die durch $f(x) = -r(x)$ definierte Abbildung $f : K \to K$ einen Fixpunkt c, der wegen $f(K) \subseteq S$ in S liegen müsste. Das steht aber im Widerspruch zu $f(c) = -c$.

Man kann umgekehrt den Brouwerschen Fixpunktsatz für die Kugel leicht aus der Folgerung herleiten. Hätte nämlich $f : K \to K$ keinen Fixpunkt, so könnten wir eine Retraktion $r : K \to S$ wie folgt definieren: Wegen $f(x) \neq x$ gibt es genau einen Strahl mit dem Anfangspunkt $f(x)$ durch den Punkt x, und es sei $r(x)$ sein Schnittpunkt mit S. Dann ist $r|_S = \mathrm{id}_S$, und aus der Stetigkeit von f kann man die von r folgern.

2 Der Satz von Hahn-Banach

Mit Hilfe dieses Satzes wollen wir beweisen, dass man zwei disjunkte konvexe Mengen durch eine Hyperebene trennen kann. Der Beweis wird durchsichtiger, wenn wir Funktionen mit Werten in der geordneten Menge $\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$ zulassen, deren Elemente man auf naheliegende Weise addiert (mit Ausnahme der Operation $\infty - \infty$) und mit reellen Zahlen $\lambda \neq 0$ multipliziert.

Definition 9 *Es sei V ein reeller Vektorraum. Ein Funktional*[4] $p : V \to [0, \infty]$ *heißt sublinear, wenn*

(i) $p(\lambda x) = \lambda p(x)$ *für alle* $\lambda > 0$ *und* $x \in V$,

(ii) $p(x + y) \leq p(x) + p(y)$ *für alle* x, $y \in V$.

Ist z. B. $l : V \to \mathbb{R}$ linear, so ist $p(x) = |l(x)|$ sublinear. Auch jede Norm ist ein sublineares Funktional. Ist K ein konvexer Kegel mit Spitze im Nullpunkt, so ist durch

$$p(x) = \begin{cases} 0 & \text{falls } x \in K, \\ \infty & \text{falls } x \notin K \end{cases}$$

ein sublineares Funktional gegeben.

Satz 4 (Hahn-Banach) *Es sei V ein reeller Vektorraum, U ein Unterraum, $p : V \to [0, \infty]$ sublinear und $l : U \to \mathbb{R}$ linear mit*

$$l(u) \leq p(u) \quad \textit{für alle } u \in U.$$

Dann existiert ein lineares Funktional $L : V \to \mathbb{R}$ mit $L|_U = l$ und

$$L(x) \leq p(x) \qquad \textit{für alle } x \in U.$$

Beweis. Im Fall $U = V$ ist nichts zu beweisen, also nehmen wir an, dass es ein $x_0 \in V$ mit $x_0 \notin U$ gibt. Dann ist für jedes $r \in \mathbb{R}$ durch

$$l'(u + \lambda x_0) = l(u) + \lambda r \qquad \text{für } u \in U \text{ und } \lambda \in \mathbb{R}$$

eine Fortsetzung l' von l auf $U' = U + \mathbb{R}x_0$ gegeben. Die Bedingung $l'(x) \leq p(x)$ bedeutet

$$l(u) + \lambda r \leq p(u + \lambda x_0) \qquad \text{für } u \in U \text{ und } \lambda \neq 0$$

[4] Funktionen auf unendlichdimensionalen Vektorräumen nennt man traditionell Funktionale.

bzw., wenn wir durch $|\lambda|$ teilen,

$$l(v) \pm r \leq p(v \pm x_0) \qquad \text{für alle } v \in U.$$

Wir müssen also r so wählen, dass

$$\sup_{v \in U}(l(v) - p(v - x_0)) \leq r \leq \inf_{v \in U}(p(v + x_0) - l(v)).$$

Nun gilt aber für alle u, $v \in U$

$$l(u) + l(v) = l(u + v) \leq p(u + v) \leq p(u - x_0) + p(v + x_0),$$

also

$$l(u) - p(u - x_0) \leq p(v + x_0) - l(v).$$

Das Supremem der linken Seite über $u \in U$ ist also nicht größer als das Infimum der rechten Seite über $v \in U$, und r kann in der Tat wie gefordert gewählt werden.

Ist V endlichdimensional, so folgt die Behauptung durch vollständige Induktion. Ist V unendlichdimensional, so benutzt man

Lemma 5 (Zorn) *Es sei $\mathcal{A}$ eine nichtleere Menge mit einer partiellen Ordnungsrelation $\prec$, so dass jede Kette*[5] *in $\mathcal{A}$ eine obere Schranke besitzt. Dann gibt es zu jedem Element a ein maximales Element m mit $a \preceq m$.*

Dieses Lemma kann mit Hilfe des Auswahlaxioms bewiesen werden, was wir hier nicht tun werden.

Es sei $\mathcal{A}$ die Menge aller Paare (U, l), wobei U ein Unterraum von V und l ein lineares Funktional auf U mit $l(u) \leq p(u)$ für alle $u \in U$ ist, und es sei $(U_1, l_1) \prec (U_2, l_2)$ genau dann, wenn $U_1 \subset U_2$ und $l_2|_{U_1} = l_1$. Nach Voraussetzung ist $\mathcal{A} \neq \emptyset$. Ist $\{(U_i, l_i) \mid i \in I\} \subset \mathcal{A}$ eine Kette, so setzen wir $U = \bigcup_{i \in I} U_i$ und definieren $l : U \to \mathbb{R}$ durch $l(u) = l_i(u)$, falls $u \in U_i$. Da zwei beliebige Elemente von U in einem der Räume U_i liegen, ist U ein linearer Unterraum und l ein lineares Funktional, das die geforderte Ungleichung erfüllt. Somit ist $(U, l) \in \mathcal{A}$ eine obere Schranke der Kette.

Nach dem Zornschen Lemma gibt es ein maximales Element $(U_{\max}, l_{\max}) \in \mathcal{A}$. Wäre $U_{\max} \neq V$, so gäbe es nach dem oben Bewiesenen ein $(U'_{\max}, l'_{\max}) \succ (U_{\max}, l_{\max})$ im Widerspruch zur Maximalität. $\square$

Ist $p : V \to [0, \infty]$ sublinear, so ist die Menge

$$K = \{x \in V \mid p(x) \leq 1\}$$

[5] also jede total geordnete Teilmenge

konvex, denn für x, $y \in K$ und $0 < \lambda < 1$ gilt

$$p(\lambda x + (1-\lambda)y) \leq \lambda p(x) + (1-\lambda)p(y) \leq 1.$$

In dem Fall, dass p eine Norm ist, nennt man K ihre Einheitskugel.

Lemma 6 *Ist K eine konvexe Teilmenge eines reellen Vektorraums V und $0 \in K$, so ist das Minkowskifunktional p_K, das durch*

$$p_K(x) = \inf\{\lambda > 0 \mid x \in \lambda K\}$$

definiert ist, sublinear. Für $y \in K$ gilt $p_K(y) \leq 1$, für $x \notin K$ gilt $p_K(x) \geq 1$.

Beweis. Für $\lambda > 0$ und $\mu > 0$ gilt $\mu x \in \lambda K$ genau dann, wenn $x \in \frac{\lambda}{\mu} K$, und durch Bildung des Infimums über λ folgt $p_K(\mu x) = \mu p_K(x)$. Zum Beweis von (ii) betrachten wir beliebige x, $y \in V$. Sind λ, $\mu > 0$ derart, dass $x \in \lambda K$ und $y \in \mu K$, so gilt wegen der Konvexität von K

$$\frac{x+y}{\lambda+\mu} = \frac{\lambda}{\lambda+\mu} \cdot \frac{x}{\lambda} + \frac{\mu}{\lambda+\mu} \cdot \frac{y}{\mu} \in K$$

und somit $p_K(x+y) \leq \lambda + \mu$. Bilden wir nun das Infimum über alle solchen λ und μ, so folgt die Sublinearität.

Ist $\lambda \leq 1$, so ist $\lambda K \subseteq K$, weil K konvex ist und 0 enthält. Daraus folgt $p_K(x) \geq 1$ für $x \notin K$. Die Behauptung über $y \in K$ ist offensichtlich. □

Satz 5 *Es sei K eine konvexe Teilmenge eines $\mathbb{R}$-Vektorraums V und $x_0 \in V$, $x_0 \notin K$. Dann existiert ein lineares Funktional $L \neq 0$ auf V, so dass für alle $y \in K$ gilt $L(y) \leq L(x_0)$.*

Beweis. Bei Anwendung einer Translation verändern sich beide Seiten der Ungleichung um dieselbe Konstante, also können wir annehmen, dass $0 \in K$ ist. Nach dem Lemma ist p_K dann sublinear. Auf $U = \mathbb{R}x_0$ definieren wir $l(\lambda x_0) = \lambda p_K(x_0)$. Dann ist $l(u) \leq p_K(u)$, und nach Satz 4 existiert ein lineares Funktional L mit

$$L(x) \leq p_K(x), \qquad L(x_0) = p_K(x_0).$$

Für alle $y \in K$ ist $p_K(y) \leq 1$, während $p_K(x_0) \geq 1$, und wir sind fertig.

Ist V ein Hilbertraum und x_0 nicht im Abschluss $\bar{K}$ von K, so können wir einen anderen Beweis geben. Nach dem Beweis von Satz 3 existiert ein $z \in \bar{K}$, so dass für alle $y \in K$ gilt

$$(x_0 - z) \cdot (y - z) \leq 0.$$

Setzen wir $L(x) = (x_0 - z) \cdot x$, so ist L nicht identisch gleich Null, und

$$L(y) - L(z) \leq 0, \qquad L(x_0) - L(z) = \|x_0 - z\|^2 \geq 0.$$

□

Die Hyperebene $\{x \in V \mid L(x) = L(x_0)\}$ heißt Stützhyperebene von K im Punkt x_0, sie existiert auch für $x_0 \in K$ mit $p_K(x_0) = 1$ und ist selbst dann nicht unbedingt eindeutig.

Ein lineares Funktional L auf einem normierten Vektorraum V ist genau dann stetig, wenn es ein $\mu \geq 0$ gibt, so dass für alle $x \in V$ gilt

$$|L(x)| \leq \mu \|x\|.$$

Auf einem endlichdimensionalen Vektorraum sind alle Normen äquivalent und ist jedes lineare Funktional stetig, aber auf unendlichdimensionalen normierten Vektorräumen gibt es unstetige lineare Funktionale.

Folgerung 2 *Ist in Satz 5 zusätzlich V ein normierter Vektorraum und K offen, so kann man L stetig wählen.*

Beweis. Nehmen wir wieder $0 \in K$ an und bezeichnen mit U die offene Einheitskugel, so folgt $\varepsilon U \subset K$ für genügend kleines ε, so dass $p_K \leq p_{\varepsilon U} = \frac{1}{\varepsilon} p_U$, also $L(x) \leq \frac{1}{\varepsilon}\|x\|$ für alle $x \in V$. Dasselbe gilt mit $-x$ statt x, also $|L(x)| \leq \frac{1}{\varepsilon}\|x\|$. □

Folgerung 3 *Es seien A und B disjunkte konvexe Teilmengen eines reellen Vektorraums V. Dann gibt es ein lineares Funktional $L \neq 0$ auf V und eine Zahl $\mu \in \mathbb{R}$, so dass*

$$L(a) \geq \mu \qquad \textit{und} \qquad L(b) \leq \mu$$

für alle $a \in A$ und $b \in B$.

Beweis. Die Menge

$$K = A - B = \{a - b \mid a \in A,\, b \in B\}$$

ist offensichtlich konvex, und da A und B disjunkt sind, ist $0 \notin A - B$. Nach dem vorigen Satz gibt es ein lineares Funktional L mit $L(a - b) \leq L(0) = 0$ für alle $a \in A$ und $b \in B$, und wir können μ so wählen, dass

$$\sup\{L(a) \mid a \in A\} \leq \mu \leq \inf\{L(b) \mid b \in B\}.$$

□

Folgerung 4 *Ist V ein normierter reeller Vektorraum, K eine konvexe abgeschlossene Teilmenge und $x_0 \in V$ mit $x \notin K$, so kann man ein stetiges lineares Funktional L so wählen, dass*

$$L(x_0) > \sup_{x \in K} L(x).$$

Beweis. Da das Komplement von K offen ist, gibt es ein $\varepsilon > 0$ mit $K \cap (x_0 + \varepsilon U) = \emptyset$. Nach der vorigen Folgerung gibt es ein lineares Funktional $L \neq 0$, so dass $\sup_{x \in K} L(x) \leq \inf_{x \in (x_0 + \varepsilon U)} L(x)$, und da $K - \varepsilon U$ offen ist, sieht man wie im Beweis der ersten Folgerung, dass L stetig ist. Wegen $L \neq 0$ gibt es ein $y \in \varepsilon U$ mit $y < 0$, und die strenge Ungleichung folgt. □

3 Der Satz von Krein-Milman

Wir wollen konvexe Mengen als konvexe Hülle von möglichst wenigen Punkten darstellen.

Definition 10 *Es sei V ein reeller Vektorraum und a, $b \in V$. Ein Punkt $x \in V$ heißt innerer Punkt der Strecke ab, wenn es ein $\lambda \in \mathbb{R}$ mit $0 < \lambda < 1$ gibt, so dass*

$$x = \lambda a + (1 - \lambda) b.$$

Nun sei M eine Menge in V. Ein Punkt $x \in M$ heißt Extremalpunkt von M, wenn für beliebige a, $b \in M$ gilt: Ist x ein innerer Punkt der Strecke ab, so ist $a = b = x$.

Eine Teilmenge N von M heißt extremale Teilmenge von M, wenn für beliebige a, $b \in M$ gilt: Enthält N einen inneren Punkt der Strecke ab, so gilt a, $b \in N$.

Ein Punkt $x \in M$ ist genau dann Extremalpunkt von M, wenn $\{x\}$ eine extremale Teilmenge von M ist. Ist P eine extremale Teilmenge von N und N eine extremale Teilmenge von M, so ist P eine extremale Teilmenge von M.

Lemma 7 *Es sei M eine Teilmenge des reellen Vektorraums V und L ein lineares Funktional auf V. Hat die Menge $\{L(x) \mid x \in M\}$ ein größtes Element μ, so ist die Menge $N = \{x \in M \mid L(x) = \mu\}$ eine extremale Teilmenge von M.*

Beweis. Angenommen, $x \in N$ ist ein innerer Punkt der Strecke ab mit a, $b \in M$, also $x = \lambda a + (1 - \lambda) b$, $0 < \lambda < 1$. Dann gilt $L(a) \leq \mu$ und $L(b) \leq \mu$. Wäre eine dieser Ungleichungen strikt, so wäre

$$\mu = L(x) = \lambda L(a) + (1 - \lambda) L(b) < \lambda \mu + (1 - \lambda) \mu = \mu$$

(Widerspruch), also $L(a) = L(b) = \mu$. $\square$

Satz 6 *Jede nichtleere kompakte Teilmenge K eines normierten Vektorraumes V besitzt Extremalpunkte.*

Beweis. Es sei $\mathcal{A}$ die Menge aller nichtleeren kompakten extremalen Teilmengen von K. Die Menge $\mathcal{A}$ ist partiell geordnet bezüglich Inklusion, und wegen $K \in \mathcal{A}$ ist $\mathcal{A}$ nichtleer.

Ist $\{N_i \mid i \in I\}$ eine Kette in $\mathcal{A}$, so ist

$$N = \bigcap_{i \in I} N_i$$

nichtleer, denn sonst würde sie von den offenen Mengen $V \setminus N_i$ überdeckt werden, aus denen man keine endliche Teilüberdeckung auswählen kann. Außerdem ist N als abgeschlossene Teilmenge von K kompakt. Ist $x \in N$ innerer Punkt einer Strecke ab mit a, $b \in K$, so ist a, $b \in N_i$ für ein beliebiges i wegen der Extremalität von N_i, also auch a, $b \in N$. Somit ist N eine extremale Teilmenge von K und somit $N \in \mathcal{A}$. Wir haben gezeigt, dass jede Kette in $\mathcal{A}$ eine untere Schranke hat, und nach dem Zornschen Lemma besitzt $\mathcal{A}$ ein minimales Element M.

Angenommen, M besitzt verschiedene Punkte $c \neq d$. Dann gibt es auf dem von $c-d$ aufgespannten eindimensionalen Unterraum ein nichtverschwindendes lineares Funktional l mit $l(c-d) \leq \|c-d\|$, das sich nach Satz 4 zu einem linearen Funktional L auf V mit

$$L(x) \leq \|x\| \qquad \text{für alle } x \in V$$

fortsetzen lässt. Die Ungleichung gilt auch für $-x$ statt x, und es folgt $|L(x)| \leq \|x\|$, d. h. L ist stetig. Somit ist $\{L(x) \mid x \in M\}$ kompakt, hat also ein größtes Element, und Lemma 7 liefert eine nichtleere extremale Teilmenge von M, die wegen $L(c) \neq L(d)$ eine echte Teilmenge ist. Dies widerspricht der Minimalität von M, also besteht M nur aus einem Punkt, und dieser ist ein Extremalpunkt von K. $\square$

Satz 7 (Krein-Milman) *Es sei K eine nichtleere kompakte konvexe Teilmenge eines reellen normierten Vektorraumes V und E die Menge der Extremalpunkte von K. Dann gilt*

$$K = \overline{\operatorname{conv} E}.$$

Beweis. Offensichtlich ist $\overline{\operatorname{conv} E} \subset K$. Angenommen, x_0 liegt in K, aber nicht in $K_0 = \overline{\operatorname{conv} E}$. Nach Satz 6 ist K_0 nicht leer, und als Abschluss einer konvexen Menge ist K_0 konvex. Nach Folgerung 4 aus Satz 5 gibt es ein stetiges lineares Funktional L auf V, so dass $L(x_0) > \sup\{L(y) \mid y \in K_0\}$. Somit schneidet $K_1 = \{x \in K \mid L(x) = \max_{y \in K} L(y)\}$ nicht die Menge K_0, also auch nicht E.

Da das stetige Funktional L auf der kompakten Menge K ein Maximum annimmt, ist K_1 nicht leer, und als Durchschnitt von K mit einer abgeschlossenen Hyperebene ist K_1 kompakt, besitzt also nach Satz 6 einen Extremalpunkt. Nach Lemma 7 ist K_1 eine extremale Teilmenge von K, also ist jeder Extremalpunkt von K_1 auch Extremalpunkt von K – Widerspruch. □

4 Der Schaudersche Fixpunktsatz

In einem unendlichdimensionalen normierten Vektorraum ist nicht jede beschränkte abgeschlossene Teilmenge kompakt, selbst wenn der Vektorraum vollständig ist. Es existieren auch fixpunktfreie Selbstabbildungen von beschränkten abgeschlossenen konvexen Teilmengen (siehe Aufgabe 10). Mit einer Kompaktheitsvoraussetzung erzwingt man dennoch die Existenz eines Fixpunktes.

Satz 8 (Schauder) *Es sei V ein normierter Vektorraum, K eine konvexe Teilmenge von V und C eine nichtleere kompakte Teilmenge von K. Dann besitzt jede stetige Abbildung $f : K \to C$ einen Fixpunkt.*

Zum Beweis benötigen wir zwei Lemmas.

Lemma 8 (Zerlegung der Eins) *Es sei V ein normierter Vektorraum und $U \subset V$ eine endliche Vereinigung offener Kugeln $U_1, \ldots, U_m$. Dann gibt es stetige reelle Funktionen $\psi_j \geq 0$ auf U, so dass $\psi_j = 0$ auf $U \setminus U_j$ und $\sum_{j=1}^m \psi_j = 1$.*

Beweis. Ist x_j der Mittelpunkt und r_j der Radius von U_j, so definieren wir $\varphi_j : V \to \mathbb{R}$ durch

$$\varphi_j(x) = \max\{0, r_j - \|x - x_j\|\}.$$

Dann ist $\varphi_j \geq 0$ stetig und $\varphi_j|_{U_j} > 0$. Die Funktion $\varphi = \sum_{j=1}^m \varphi_j$ ist auf U positiv, und wir können $\psi_j = \frac{\varphi_j}{\varphi}$ setzen. □

Ist V endlichdimensional, so kann man sogar unendlich oft differenzierbare Funktionen ψ_j wählen.

Lemma 9 *Es sei (X, d) ein metrischer Raum, C eine kompakte Teilmenge und f eine stetige Abbildung von X in C. Zu jedem $\varepsilon > 0$ gebe es einen Punkt $x \in X$ mit $d(f(x), x) < \varepsilon$. Dann besitzt f einen Fixpunkt.*

Beweis. Es gibt eine Folge x_n in X mit $d(f(x_n), x_n) < \frac{1}{n}$. Da C kompakt ist, gibt es eine Teilfolge x_{n_k}, so dass die Folge $f(x_{n_k})$ gegen einen Punkt $c \in C$ konvergiert. Wegen

$$d(x_{n_k}, c) \leq d(x_{n_k}, f(x_{n_k})) + d(f(x_{n_k}), c) \leq \frac{1}{n_k} + d(f(x_{n_k}), c)$$

konvergiert dann auch die Folge x_{n_k} gegen c. Da f stetig ist, folgt daraus, dass $\lim_{k\to\infty} f(x_{n_k}) = f(c)$, wegen $\lim_{k\to\infty} f(x_{n_k}) = c$ also $f(c) = c$. □

Beweis von Satz 4. Es sei $\varepsilon > 0$. Da C kompakt ist, existieren Punkte $x_1, \ldots, x_m \in C$ derart, dass C von den offenen Kugeln vom Radius ε um diese Punkte überdeckt wird. Nun wählen wir Funktionen $\psi_1, \ldots, \psi_m$ gemäß Lemma 8 und definieren eine stetige Abbildung $g : C \to V$ durch

$$g(x) = \sum_{j=1}^{m} \psi_j(x) x_j.$$

Nach Lemma 8 gilt für alle $x \in C$

$$g(x) - x = \sum_{j=1}^{m} \psi_j(x)(x_j - x).$$

Ist $\psi_j(x) \neq 0$, so gilt $\|x_j - x\| < \varepsilon$, also ist

$$\|g(x) - x\| < \varepsilon \sum_{j=1}^{m} \psi_j(x) = \varepsilon.$$

Das Bild von g liegt in der Menge

$$K_0 = \operatorname{conv}\{x_1, \ldots, x_m\} \subseteq K,$$

also schränkt sich $g \circ f : K \to K_0$ zu einer stetigen Abbildung von K_0 in sich ein. Da K_0 in dem von $x_1, \ldots, x_m$ aufgespannten endlichdimensionalen affinen Unterraum enthalten ist und als konvexe Hülle einer endlichen Menge abgeschlossen und beschränkt ist, besitzt $g \circ f$ nach Satz 1 einen Fixpunkt x. Wegen $f(x) \in C$ folgt

$$\|x - f(x)\| = \|g(f(x)) - f(x)\| < \varepsilon.$$

Nach Lemma 9 angewendet auf $X = K$ besitzt f daher einen Fixpunkt. $\square$

Eine Anwendungssituation wäre z. B. der Raum $V = C_{\mathrm{b}}(X)$ der beschränkten stetigen Funktionen auf einem metrischen Raum X, versehen mit der Supremumsnorm, wobei die Beschränktheit für kompaktes X automatisch folgt. Sind K, C und f wie in Satz 4, so ist $\overline{f(K)}$ als abgeschlossene Teilmenge von C kompakt; man könnte also C durch $\overline{f(K)}$ ersetzen. Wie kann man aber prüfen, ob der Abschluss einer Teilmenge von V kompakt ist?

Satz 9 (Arzelà-Ascoli) *Es sei X ein kompakter metrischer Raum und $M \subset C(X, \mathbb{R}^n)$ mit folgenden Eigenschaften:*

(i) M ist beschränkt,

(ii) M ist gleichgradig stetig, d. h. zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $\varphi \in M$ und x, $x' \in X$ gilt:

$$d(x, x') < \delta \quad \Longrightarrow \quad \|\varphi(x) - \varphi(x')\| < \varepsilon.$$

Dann ist der Abschluss $\bar{M}$ von M kompakt.

Wir beweisen die Folgenkompaktheit, die ja in Satz 4 benutzt wird. In der Tat ist für metrische Räume die Folgenkompaktheit äquivalent zur Kompaktheit.

Beweis. Es sei eine Folge von Funktionen φ_n in M gegeben. Da X kompakt ist, gibt es in X eine dichte abzälhbare Teilmenge $\{x_1, x_2, \dots\}$ (vgl. Aufgabe 17). Wegen (i) ist die Folge $\varphi_n(x_1)$ in $\mathbb{R}^n$ beschränkt, also gibt es eine Teilfolge $\varphi_{1,n}$ von φ_n, so dass $\varphi_{1,n}(x_1)$ konvergiert. Da auch $\varphi_{1,n}(x_2)$ beschränkt ist, finden wir eine Teilfolge $\varphi_{2,n}$ von $\varphi_{1,n}$, so dass $\varphi_{2,n}(x_2)$ konvergent ist. So fortfahrend, erhalten wir für jedes m eine Teilfolge $\varphi_{m,n}$, die an den Stellen $x_1, \dots, x_m$ konvergiert. Die Diagonalfolge $\psi_n = \varphi_{n,n}$ ist dann eine Teilfolge von φ_n und konvergiert an allen Punkten x_j.

Wir wollen zeigen, dass die Folge ψ_n eine Cauchyfolge bezüglich der Supremumsnorm ist. Sei $\varepsilon > 0$ und dazu $\delta > 0$ wie in (ii). Der kompakte Raum X ist Vereinigung der δ-Umgebungen von endlich vielen der Punkte x_j, sagen wir, von $x_1, \dots, x_p$. Wegen der Konvergenz in diesen Punkten können wir ein n_0 wählen, so dass

$$\|\psi_m(x_j) - \psi_n(x_j)\| < \varepsilon \qquad \text{für } m,\, n \geq n_0 \text{ und } 1 \leq j \leq p.$$

Ist nun $x \in X$ beliebig, so gibt es ein $j \in \{1, \dots, p\}$ mit $d(x, x_j) < \delta$, also nach (ii)

$$\|\psi_n(x) - \psi_n(x_j)\| < \varepsilon.$$

Nun folgt für m, $n \geq n_0$, dass

$$\|\psi_m(x)-\psi_n(x)\| \leq \|\psi_m(x)-\psi_m(x_j)\|+\|\psi_m(x_j)-\psi_n(x_j)\|+\|\psi_n(x)-\psi_n(x_j)\| < 3\varepsilon,$$

d. h. ψ_n ist eine Cauchyfolge.

Wegen der Vollständigkeit von $C(X, \mathbb{R}^n)$ ist die Funktionenfolge ψ_n konvergent. Somit ist $\bar{M}$ folgenkompakt (vgl. Aufgabe 16). □

Als Anwendung betrachten wir Systeme gewöhnlicher Differentialgleichungen erster Ordnung

$$\begin{cases} \varphi_1'(x) = f_1(x, \varphi_1(x), \ldots, \varphi_n(x)), \\ \qquad \vdots \\ \varphi_n'(x) = f_n(x, \varphi_1(x), \ldots, \varphi_n(x)). \end{cases}$$

Wir fassen die gesuchten Funktionen φ_1, …, φ_n wie auch die gegebenen Funktionen f_1, …, f_n zu vektorwertigen Funktionen φ bzw. f zusammen und schreiben das System in kompakter Form als

$$\varphi'(x) = f(x, \varphi(x)). \tag{1}$$

Satz 10 *Es sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^n$ stetig und $(x_0, y_0) \in U$. Dann gibt es ein offenes Intervall I, das x_0 enthält, und eine Lösung $\varphi : I \to \mathbb{R}^n$ der Differentialgleichung (1), die der Anfangsbedingung $\varphi(x_0) = y_0$ genügt. Ist f Lipschitz-stetig, so gibt es für genügend kleines I genau eine Lösung.*

Beweis. Ohne Beschränkung der Allgemeinheit ist $(x_0, y_0) = (0, 0)$. Da U offen ist, gibt es $c > 0$ und $r > 0$, so dass $J \times B \subseteq U$, wobei $J = [-c, c]$ und $B = \{y \in \mathbb{R}^n \mid \|y\| \leq r\}$.

Eine stetige Funktion $\varphi : J \to B$ ist genau dann Lösung von (1) und genügt der Bedingung $\varphi(0) = 0$, wenn für alle $x \in J$ gilt

$$\varphi(x) = \int_0^x f(t, \varphi(t))\, dt. \tag{2}$$

Dies folgt aus dem Hauptsatz der Differential- und Integralrechnung. Wir definieren $C(J, B) \xrightarrow{F} C(J, \mathbb{R}^n) \xrightarrow{A} C^1(J, \mathbb{R}^n)$ durch[6]

$$F\varphi(x) = f(x, \varphi(x)), \qquad A\psi(x) = \int_0^x \psi(t)\, dt.$$

Die Fixpunkte von AF sind gerade die Lösungen $\varphi : J \to B$ von (2).

[6] Wir schreiben $F\varphi$ statt $F(\varphi)$, um die doppelten Klammern zu vermeiden.

Die Abbildung A ist linear und Lipschitz-stetig, denn wenn wir den Integranden durch sein Supremum abschätzen, so erhalten wir bezüglich der Supremumsnormen

$$\|A\psi\| \leq c\|\psi\|.$$

Wegen der Kompaktheit von $J \times B$ existiert $r' = \max_{J\times B} \|f\|$, und F bildet $C(J, B)$ in $C(J, B')$ ab, wobei $B' = \{y \in \mathbb{R}^n \mid \|y\| \leq r'\}$. Wir können c durch eine kleinere positive Zahl ersetzen, so dass $cr' \leq r$ gilt; dann wird $C(J, B)$ von AF in sich selbst abgebildet.

Um die Stetigkeit von F zu beweisen, geben wir uns ein $\varepsilon > 0$ vor. Die stetige Funktion f ist auf der kompakten Menge $J \times B$ gleichmäßig stetig, also existiert ein $\delta > 0$, so dass für alle $x \in J$ und y_1, $y_2 \in B$ gilt

$$\|y_1 - y_2\| < \delta \quad \Longrightarrow \quad \|f(x, y_1) - f(x, y_2)\| < \varepsilon.$$

Es folgt, dass für φ_1, $\varphi_2 \in C(J, B)$ gilt

$$\|\varphi_1 - \varphi_2\| < \delta \quad \Longrightarrow \quad \|F\varphi_1 - F\varphi\| < \varepsilon.$$

Die Abbildung F ist also stetig.

Nach Aufgabe 19 ist das Bild von $C(J, B')$ unter A gleichgradig stetig. Nach dem Satz von Arzelà-Ascoli ist der Abschluss des Bildes kompakt. Die Teilmenge $C(J, B)$ des Banachraums $C(J, \mathbb{R}^n)$ ist konvex und nichtleer, und nach dem Satz von Schauder hat AF einen Fixpunkt.

Nun kommen wir zur Eindeutigkeit. Dazu genügt es, dass f Lipschitz-stetig bezüglich y ist, dass also ein $C > 0$ existiert, so dass für alle $x \in J$ und y_1, $y_2 \in B$ gilt

$$\|f(x, y_1) - f(x, y_2)\| \leq C\|y_1 - y_2\|.$$

Es folgt dann bezüglich der Supremumsnormen

$$\|F\varphi_1 - F\varphi_2\| \leq C\|\varphi_1 - \varphi_2\|.$$

Wir können c weiter verkleinern, so dass $cC < 1$ ist. Dann ist AF eine Kontraktion und hat nach dem Banachschen Fixpunktsatz genau einen Fixpunkt. □

Ist f lokal Lipschitz-stetig bezüglich y, so genügt der Fixpunktsatz von Banach[7] für den Beweis von Existenz und Eindeutigkeit. Der Beweis jenes Fixpunktsatzes läuft in diesem Fall auf die Picard-Iteration

$$\varphi_{k+1}(x) = y_0 + \int_{x_0}^{x} f(t, \varphi_k(t))\, dt$$

hinaus, und die betreffende Version von Satz 10 geht auf Picard und Lindelöf zurück. Auch der Existenzsatz für beliebige stetige f wurde schon von Peano bewiesen, bevor der Schaudersche Fixpunktsatz bekannt war.

[7] der allerdings schon lange vor Banach benutzt wurde

5 Fixpunktsatz von Kakutani

Der Fixpunktsatz von Kakutani verallgemeinert den Brouwerschen Fixpunktsatz auf Korrespondenzen.

Definition 11 *Es seien X und Y metrische Räume und $f : X \to Y$ eine Korrespondenz, d. h. eine Abbildung, die jedem Punkt $x \in X$ eine nichtleere Teilmenge $f(x)$ von Y zuordnet.*

1. *f heißt oberhalbstetig*[8] *an der Stelle x, falls es zu jeder Umgebung V von $f(x)$ eine Umgebung U von x gibt mit $f(u) \subseteq V$ für alle $u \in U$. Die Korrespondenz f heißt oberhalbstetig, wenn sie in jedem Punkt von X oberhalbstetig ist.*

2. *f heißt abgeschlossen, wenn ihr Graph*

$$G = \{(x, y) \in X \times Y \mid y \in f(x)\}$$

abgeschlossen ist.

Dabei versieht man $X \times Y$ mit der Metrik

$$d((x_1, y_1), (x_2, y_2)) = \max\{d(x_1, x_2), d(y_1, y_2)\}$$

oder einer äquivalenten Metrik.
Bemerkung. Ist f eine abgeschlossene Korrespondenz, so ist die Menge $f(x)$ für jedes x abgeschlossen. Ist nämlich y_n eine Folge in $f(x)$, die in Y konvergiert, so konvergiert die Folge (x, y_n) in $X \times Y$ gegen ein $(x, y) \in G$.

Lemma 10 *Es seien X und Y metrische Räume und $p : X \times Y \to X$ die Projektion $p(x, y) = x$. Ist Y kompakt, so ist für jede abgeschlossene Teilmenge $A \subseteq X \times Y$ die Menge $p(A)$ abgeschlossen.*

Beweis. Es sei W das Komplement von A in $X \times Y$, also eine offene Menge. Wir müssen zeigen, dass das Komplement von $p(A)$ in X offen ist. Sei also $x_0 \notin p(A)$. Für jedes $y \in Y$ ist $(x_0, y) \in W$, also gibt es Umgebungen U_y von x_0 und V_y von y, so dass $U_y \times V_y \subseteq W$. Der kompakte Raum Y wird von endlich vielen der Mengen V_y, sagen wir $V_{y_1}, \ldots, V_{y_n}$, überdeckt. Setzen wir $U = U_1 \cap \cdots \cap U_n$, so ist $U \times Y \subseteq W$, also $U \cap p(A) = \emptyset$. □

Der gegebene Beweis funktioniert auch für topologische Räume. Im Fall metrischer Räume sind abgeschlossene Mengen durch Folgenkonvergenz charakterisiert, und man kann auch einen Beweis mittels Folgenkompaktheit geben.

[8] Gemeint als Abkürzung des Ausdrucks „halbstetig von oben“.

Folgerung 5 *Ist $f : X \to Y$ eine abgeschlossene Korrespondenz und Y kompakt, so ist f oberhalbstetig.*

Beweis. Es sei V eine offene Teilmenge von Y. Dann ist $K = Y \setminus V$ abgeschlossen in Y, also kompakt. Außerdem ist $G \cap (X \times K) = G \setminus (X \times V)$ abgeschlossen in $X \times K$, also nach dem Lemma $p(G \setminus (X \times V))$ abgeschlossen. Das Komplement dieser Menge, also

$$U = \{x \in X \mid f(x) \subseteq V\},$$

ist somit offen. Nun folgt die obere Halbstetigkeit nach Definition. □

Satz 11 (Kakutani) *Es sei X eine nichtleere kompakte konvexe Teilmenge von $\mathbb{R}^n$ und $f : X \to X$ eine abgeschlossene konvexwertige Korrespondenz, d. h. für alle $x \in X$ sei $f(x)$ konvex. Dann besitzt f einen Fixpunkt, d. h. einen Punkt $y \in X$ mit $y \in f(y)$.*

Beweis. Es sei δ_k eine Nullfolge positiver reeller Zahlen. Da X kompakt ist, existieren Punkte $x_{k1}, \ldots, x_{km_k} \in X$ derart, dass X von den offenen Kugeln vom Radius δ_k um diese Punkte überdeckt wird. Zu dieser Überdeckung wählen wir eine Zerlegung der Eins $\psi_{k1}, \ldots, \psi_{km_k}$ gemäß Lemma 8. In jeder Menge $f(x_{kj})$ wählen wir einen Punkt z_{kj} und definieren eine Abbildung

$$f_k(x) = \sum_{j=1}^{m} \psi_{kj}(x) z_{kj}.$$

Da die Menge X konvex ist, wird sie von f_k in sich abgebildet, und nach dem Brouwerschen Fixpunktsatz besitzt f_k einen Fixpunkt y_k. Da X kompakt ist, dürfen wir annehmen, dass y_k gegen einen Punkt $y \in X$ konvergiert. Wir wollen zeigen, dass y ein Fixpunkt der Korrespondenz f ist.

Es sei $\varepsilon > 0$ und $U_\varepsilon \subset \mathbb{R}^n$ die offene Kugel vom Radius ε um den Nullpunkt. Die ε-Umgebung von $f(y)$ ist $V = f(y) + U_\varepsilon$; sie ist offen und wegen der Konvexität von $f(y)$ und U_ε ebenfalls konvex. Da f an der Stelle y oberhalbstetig ist, gibt es ein positives $\delta \leq \varepsilon$, so dass für alle $u \in X$ mit $\|u - y\| < \delta$ gilt $f(u) \subseteq V$.

Wir wählen k mit $\delta_k \leq \delta/2$ und $\|y_k - y\| < \delta/2$. In der Formel für $f_k(y_k)$ kommt es nur auf die Indizes $j \in \{1, \ldots, m_k\}$ mit $\psi_{kj}(y_k) \neq 0$ an, und für diese gilt

$$\|x_{kj} - y\| \leq \|x_{kj} - y_k\| + \|y_k - y\| < \delta_k + \frac{\delta}{2} \leq \delta.$$

Aus der obigen Eigenschaft von δ folgt $f(x_{kj}) \subseteq V$, insbesondere $z_{kj} \in V$. Wegen der Konvexität von V ist auch $y_k = f_k(y_k) \in V$. Daraus folgt, dass

der Abstand von y zur Menge $f(y)$ kleiner als $\frac{\delta}{2} + \varepsilon$, also kleiner als 2ε ist. Da $\varepsilon > 0$ beliebig gewählt war und $f(y)$ abgeschlossen ist, liegt somit y in $f(y)$. $\square$

Lemma 11 *Es sei X eine nichtleere kompakte konvexe Teilmenge von $\mathbb{R}^n$. Für $y \in \mathbb{R}^n$ definieren wir*

$$g(y) = \{z \in X \mid z \cdot y = \max_{x \in X} x \cdot y\}.$$

Dann ist $g : \mathbb{R}^n \to X$ eine abgeschlossene konvexwertige Korrespondenz.

Beweis. Die Menge $g(y)$ ist als Durchschnitt von X mit einer Hyperebene (oder ganz X) konvex. Um die Abgeschlossenheit des Graphen $G \subseteq \mathbb{R}^n \times X$ zu zeigen, müssen wir für einen beliebigen Punkt (y_0, x_0) außerhalb von G eine zu G disjunkte Umgebung finden. Wegen $y_0 \notin g(x_0)$ gibt es ein $z_0 \in g(y_0)$, so dass $z_0 \cdot y_0 > x_0 \cdot y_0$. Die Menge

$$\{(y, x) \in \mathbb{R}^n \times X \mid z_0 \cdot y > x \cdot y\}$$

ist offen in $\mathbb{R}^n \times X$, enthält (y_0, x_0) und ist disjunkt zu G. $\square$

Folgerung 6 (Gale, Nikaido, Debreu) *Es seien X und Y nichtleere kompakte konvexe Teilmengen von $\mathbb{R}^n$ und $f : X \to Y$ eine abgeschlossene konvexwertige Korrespondenz derart, dass*

$$x \cdot y \leq 0 \qquad \text{für alle } x \in X \text{ und } y \in f(x).$$

Dann gibt es Punkte $x_0 \in X$ und $y_0 \in f(x_0)$ derart, dass $x \cdot y_0 \leq 0$ für alle $x \in X$.

Beweis. Wir definieren Korrespondenzen $g : Y \to X$ wie im Lemma und $h : X \times Y \to X \times Y$ durch

$$h(x, y) = g(y) \times f(x).$$

Dann ist h abgeschlossen und konvexwertig (vgl. Aufgabe 23). Nach Satz 11 besitzt h einen Fixpunkt, d. h. es existieren $x_0 \in X$, $y_0 \in Y$, so dass

$$x_0 \in g(y_0), \qquad y_0 \in f(x_0).$$

Aus der ersten Inklusion folgt $x \cdot y_0 \leq x_0 \cdot y_0$ für alle $x \in X$, und nach Voraussetzung gilt $x_0 \cdot y_0 \leq 0$. $\square$

6 Existenz eines Wettbewerbsgleichgewichts

Für $x, y \in \mathbb{R}^n_+ = [0, \infty)^n$ schreiben wir $x \leq y$, wenn jede der n Komponenten von x höchstens so groß wie die entsprechende Komponente von y ist. Die zweistellige Relation $\leq$ ist reflexiv und transitiv, aber für $n > 1$ nicht total.

Definition 12 *Eine zweistellige Relation $\precsim$ auf $\mathbb{R}^n_+$ heißt Präferenzordnung, wenn sie folgende Eigenschaften hat:*

1. *Für alle $x \in \mathbb{R}^n_+$ gilt $x \precsim x$ (Reflexivität).*
2. *Sind x, y, $z \in \mathbb{R}^n_+$ mit $x \precsim y$ und $y \precsim z$, so ist $x \precsim z$ (Transitivität).*
3. *Für alle x, $y \in \mathbb{R}^n_+$ ist $x \precsim y$ oder $y \precsim x$ (Totalität).*

Gilt sowohl $x \precsim y$ als auch $x \succsim y$, so schreiben wir $x \sim y$. Gilt nicht $x \succsim y$, so schreiben wir $x \prec y$. Die Präferenzordnung $\precsim$ heißt

- *monoton, wenn aus $x \leq y$ und $x \neq y$ folgt, dass $x \prec y$.*
- *konvex, wenn für jedes $x \in \mathbb{R}^n_+$ die Menge $\{y \in \mathbb{R}^n_+ \mid x \precsim y\}$ konvex ist.*
- *stetig, wenn die Menge $\{(x, y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ \mid x \precsim y\}$ abgeschlossen ist.*

Ist $F : \mathbb{R}^n_+ \to \mathbb{R}$ eine stetige Funktion (Nutzenfunktion), so ist

$$x \precsim y \quad \Longleftrightarrow \quad F(x) \leq F(y)$$

ein Beispiel einer stetigen Präferenzordnung. Im Fall $F(x_1, \ldots, x_n) = p_1 x_1 + \cdots + p_n x_n$ für gegebenes $p \in \mathbb{R}^n_+$ ist sie auch konvex und monoton.

Wirtschaftstheoretische Interpretation: Gibt es n Güter und ist x_j die Menge des $j$ten Gutes, so stellt $(x_1, \ldots, x_n) \in \mathbb{R}^n_+$ ein Güterbündel dar. Ist p_j der Preis einer Einheit des $j$ten Gutes, so nennt man $(p_1, \ldots, p_n) \in \mathbb{R}^n_+$ den Preisvektor. Der Wert des Güterbündels x ist dann $p \cdot x = p_1 x_1 + \cdots + p_n x_n$. Hat ein Akteur zunächst das Güterbündel e (Erstausstattung), so sind für ihn am Markt alle Güterbündel in der sogenannten Budgetmenge[9]

$$B(p, e) = \{x \in \mathbb{R}^n_+ \mid p \cdot x \leq p \cdot e\}.$$

erschwinglich. Schätzt der Akteur das Güterbündel y mindestens so wie das Güterbündel x, so schreiben wir $x \precsim y$. Diese Relation sollte eine Präferenzordnung sein. Die Menge der dominanten Elemente in $B(p, e)$, also

$$D(p, e) = \{x \in B(p, e) \mid x \succsim y \text{ für alle } y \in B(p, e)\}$$

[9] Wir vernachlässigen die Tatsache, dass Mengen von Stückgütern ganze Zahlen sein müssten.

heißt auch Nachfragemenge, denn

$$D(p,e) = \{x \in B(p,e) \mid \text{Es gibt kein } y \in B(p,e) \text{ mit } y \succ x.\}$$

Ist die Präferenzordnung monoton, so gilt offenbar das Walrassche Gesetz

$$p \cdot x = p \cdot e \qquad \text{für jedes } x \in D(p,e),$$

denn würde hier die strenge Ungleichung gelten, so könnte man in $B(p,e)$ ein $y \neq x$ mit $y \geq x$ finden.

Lemma 12 *Ist $\precsim$ stetig und $p > 0$ (d. h. jede Komponente von p ist größer als Null), so ist $D(p,e) \neq \emptyset$ (und kompakt).*

Beweis. Die Menge $B(p,e)$ ist abgeschlossen und wegen $p > 0$ beschränkt, also kompakt. Für jedes $x \in B(p,e)$ sei

$$A_x = \{y \in B(p,e) \mid y \succsim x\}.$$

Wegen der Stetigkeit von $\precsim$ ist A_x abgeschlossen, also auch kompakt, und A_x ist nicht leer ($x \in A_x$). Sind $x_1, \ldots, x_k \in B(p,e)$, so dominiert eines von ihnen, sagen wir, x_k, die Übrigen. Also ist $A_{x_k} \subseteq A_{x_j}$ für $1 \leq j \leq k$, so dass $A_{x_1} \cap \cdots \cap A_{x_k} = A_{x_k}$. Die Familie der nichtleeren kompakten Mengen A_x für $x \in B(p,e)$ hat also die endliche Durchschnittseigenschaft, und somit ist

$$D(p,e) = \bigcap_{x \in B(p,e)} A_x$$

nichtleer und kompakt. □

Ist die Voraussetzung $p > 0$ nicht erfüllt (gibt es also ein kostenloses Gut), so ist bei monotoner Präferenzordnung übrigens $D(p,e) = \emptyset$, d. h. kein Güterbündel ist dominant. Nimmt man nämlich von dem kostenlosen Gut noch etwas hinzu, so erhält man ein Güterbündel $y \neq x$ mit $y \geq x$ (folglich $y \succ x$) und $p \cdot y = p \cdot x$.

Wir können also für jede Erstausstattung $e \in \mathbb{R}^n_+$ die Nachfragekorrespondenz $p \mapsto D(p,e)$ betrachten. Ihr Definitionsbereich ist nicht ganz $\mathbb{R}^n_+$, sondern nur $\{p \in \mathbb{R}^n_+ \mid p > 0\}$.

Lemma 13 *Ist $\precsim$ konvex, so ist die Nachfragekorrespondenz konvexwertig. Ist $\precsim$ monoton und stetig, so ist der Graph der Nachfragekorrespondenz abgeschlossen in $P_e \times \mathbb{R}^n_+$, wobei $P_e = \{p \in \mathbb{R}^n_+ \mid p \cdot e > 0\}$.*

(Für $e \neq 0$ ist dies stärker als die Abgeschlossenheit der Nachfragekorrespondenz.)

Beweis. Ist $\precsim$ konvex, so ist $D(p,e)$ der Durchschnitt der konvexen Mengen A_x für alle $x \in B(p,e)$, also auch konvex.

Zum Nachweis der Abgeschlossenheit müssen wir zu jedem Punkt im Komplement des Graphen eine Umgebung finden, die disjunkt zum Graphen ist. Es sei also $(p_0, x_0) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+$ mit $x_0 \notin D(p_0, e)$ und $p_0 \cdot e > 0$.

Erster Fall: $x_0 \notin B(p_0, e)$. Dann ist $p_0 \cdot x_0 > p_0 \cdot e$, und die Menge

$$\{(p,x) \in P_e \times \mathbb{R}^n_+ \mid p \cdot x > p \cdot e\}$$

ist offen, disjunkt zum Graphen der Korrespondenz, und enthält (p_0, x_0).

Zweiter Fall: $x_0 \in B(p_0, e)$. Dann gibt es ein $y_0 \in B(p_0, e)$ mit $x_0 \prec y_0$. Ist $\precsim$ stetig, so gibt es nach Aufgabe 27 Umgebungen U von x_0 und V von y_0 in $\mathbb{R}^n_+$, so dass $u \prec v$ für alle $u \in U$ und $v \in V$. Wegen $p_0 \cdot y_0 \leq p_0 \cdot e$ und $p_0 \cdot e > 0$ finden wir[10] $v_0 \in V$, so dass $p_0 \cdot v_0 < p_0 \cdot e$. Nun ist

$$\{p \in P_e \mid p \cdot v_0 < p \cdot e\} \times U$$

eine Umgebung von (p_0, x_0). Ist (p, u) in dieser Umgebung, so ist $u \prec v_0 \in B(p,e)$, also $u \notin D(p,e)$. □

Nehmen wir jetzt an, dass es m Akteure mit ihren individuellen Präferenzordnungen $\precsim_1, \ldots, \precsim_m$ und entsprechenden Nachfragemengen $D_i(p,e)$ gibt. Im allgemeinen haben sie verschiedene Erstausstattungen e_i. Wir fragen, ob es einen Preisvektor $p > 0$ gibt, so dass eine Umverteilung (Allokation) des Gesamtgüterbündeld $e = e_1 + \cdots + e_m$ existiert (z. B. als Ergebnis eines Tauschhandels, während dessen weder Güter produziert noch welche verbraucht werden), bei dem jeder Akteur ein aus seiner Sicht optimales Güterbündel a_i erhält.

Ein Wettbewerbsgleichgewicht oder Walrasgleichgewicht ist also gegeben durch einen Preisvektor $p > 0$ und eine Zuordnung $i \mapsto a_i \in D_i(p, e_i)$ für $1 \leq i \leq m$, so dass $a_1 + \cdots + a_m = e$.

Satz 12 *Sind die Präferenzordnungen $\precsim_1, \ldots, \precsim_m$ monoton, konvex und stetig, so existiert für beliebige $e_1, \ldots, e_m \in \mathbb{R}^n_+$ mit $e > 0$ ein Marktgleichgewicht.*

Beweis. Wir betrachten die für $p > 0$ definierten Korrespondenzen

$$f_i(p) = D_i(p, e_i) - e_i, \qquad f(p) = f_1(p) + \cdots + f_m(p) = \sum_{i=1}^{m} D_i(p, e_i) - e,$$

[10] z. B. $v_0 = \lambda y_0$ mit geeignetem $0 < \lambda < 1$

interpretierbar als Nachfrageüberschuss. Der Beweis ist beendet, wenn wir ein $p > 0$ mit $0 \in f(p)$ finden.

Da sich die Nachfragemengen bei Multiplikation des Presvektors mit einer positiven Zahl nicht ändern, genügt es, die kompakte konvexe Menge

$$P = \{p \in \mathbb{R}^n_+ \mid p \cdot (1, 1, \ldots, 1) = 1\}$$

(ein Simplex) zu betrachten. Die Nachfragekorrespondenzen brauchen aber nicht beschränkt zu sein, darum wählen wir eine Nullfolge von positiven Zahlen $\varepsilon_k \leq \frac{1}{n}$ und betrachten die kompakten konvexen Mengen

$$P_k = \{p \in P \mid p \geq (\varepsilon_k, \ldots, \varepsilon_k)\},$$

die alle mindestens den Punkt $p_0 = (\frac{1}{n}, \ldots, \frac{1}{n})$ enthalten. Die Einschränkung von f_i auf P_k hat Werte in der kompakten konvexen Menge

$$X_{ik} = \operatorname{conv} \bigcup_{p \in P_k} \big(B(p, e_i) - e_i\big) = \operatorname{conv} E_{ik};$$

dabei bezeichnet E_{ik} die Menge der Ecken aller der Simplexe $B(p, e_i) - e_i$, für die p eine Ecke von P_k ist.

Nach Lemma 13 ist $f_i|_{P_k}$ konvexwertig und abgeschlossen. Also ist auch der Graph G_k von $p \mapsto f_1(p) \times \cdots \times f_m(p)$ abgeschlossen in $P_k \times X_{1k} \times \cdots \times X_{mk}$ (vgl. Aufgabe 23). Projizieren wir die abgeschlossene Menge

$$\{(p, x_i, \ldots, x_m, y) \in G_k \times \mathbb{R}^n_+ \mid y = x_1 + \cdots + x_m\}$$

entlang der kompakten Menge $X_{1k} \times \cdots \times X_{mk}$ auf $P_k \times \mathbb{R}^n_+$, so erhalten wir den Graphen von $f|_{P_k}$. Nach Lemma 10 ist also die Korrespondenz $f|_{P_k}$ abgeschlossen, und natürlich ist sie konvexwertig mit Werten in der kompakten konvexen Menge $X_{1k} + \cdots + X_{mk}$.

Nach dem Walrasschen Gesetz gilt $p \cdot y = 0$ für alle $y \in f_i(p)$ und alle i, also auch für alle $y \in f(p)$. Folgerung 6 von Gale/Nikaido/Debreu liefert nun

$$p_k \in P_k, \quad y_k \in f(p_k) \quad \text{mit } q \cdot y_k \leq 0 \text{ für alle } q \in P_k.$$

Genauer ist

$$y_k = x_{1k} + \cdots + x_{mk} \quad \text{mit } x_{ik} \in f_i(p_k).$$

Wegen $y_k \geq -e$ sind die Komponenten von y_k von unten beschränkt. Wegen $p_0 \in P_k$ gilt $p_0 \cdot y_k \leq 0$, also sind diese Komponenten auch von oben beschränkt, und wegen $x_{ik} \geq -e_i$ gilt dasselbe für die Komponenten von x_{ik}. Schließlich liegt die Folge p_k in der kompakten Menge P. Wir können also nach dem Satz von Bolzano-Weierstraß annehmen, dass die Folgen p_k, y_k und

x_{ik} für $k \to \infty$ gegen ein $p \in P$, $y \in \mathbb{R}^n_+ - e$ bzw. $x_i \in \mathbb{R}^n_+ - e_i$ konvergieren, wobei $y = x_1 + \cdots + x_m$ gilt und (p, x_i) im Abschluss des Graphen von f_i liegt.

Wegen der Voraussetzung $e > 0$ ist $p \cdot e > 0$, also gibt es ein i mit $p \cdot e_i > 0$. Nach Lemma 13 liegt (p, x_i) im Graphen von f_i, also ist $p > 0$.

Für genügend großes k ist $p \in P_k$. Wegen der Abgeschlossenheit von $f|_{P_k}$ ergibt sich durch Grenzübergang

$$y \in f(p), \qquad q \cdot y \leq 0 \text{ für alle } q \in P_k.$$

Letztere Gleichung gilt für q in der Vereinigung aller P_k und nach Stetigkeit sogar für alle $q \in P$. Dies zeigt, dass $y \leq 0$ ist. Aus dem Walrasschen Gesetz hatten wir bereits $p \cdot y = 0$ gefolgert, und wegen $p > 0$ folgt $y = 0$. $\square$

Im Fall $m = 1$ ist $a_1 = e = e_1$, und ein Wettbewerbsgleichgewicht ist dann ein Preisvektor $p > 0$ mit $e \in D(p, e)$. Für einen einzelnen Akteur, der von jedem Gut etwas besitzt, besagt der Satz also, dass es einen solchen Preisvektor gibt, für den seine Erstaustattung dominant ist und kein Gut kostenlos ist.

7 Ringe, Algebren und Inhalte von Mengen

In der Maßtheorie geht es um die Beschreibung solcher Begriffe wie

- Flächen- und Rauminhalt,
- Masse und Ladung,
- Wahrscheinlichkeit.

Dabei wird jeweils einer Menge A eine Zahl $\mu(A)$ zugeordnet, wobei A Teilmenge einer vorgegebenen Gesamtmenge X ist. Zunächst geht es uns um den Definitionsbereich einer solchen Abbildung μ.

Wir bezeichnen die Potenzmenge einer Menge X, d. h. die Menge aller Teilmengen von X, mit $\mathcal{P}(X)$. Neben den üblichen Mengenoperationen benutzen wir auch die symmetrische Differenz, die definiert ist durch

$$A \triangle B = (A \cup B) \setminus (A \cap B).$$

Lemma 14 *Versieht man $\mathcal{P}(X)$ mit $\triangle$ als Addition und $\cap$ als Multiplikation, so ist $(\mathcal{P}(X), \triangle, \cap)$ ein kommutativer Ring mit dem Nullelement $\emptyset$ und dem Einselement X. Also*

$$\begin{aligned} &(A \triangle B) \triangle C = A \triangle (B \triangle C), \\ &A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C) \qquad \textit{usw.} \end{aligned}$$

Beweis. Die charakteristische Funktion[11] einer Menge A, nämlich

$$\chi_A(x) = \begin{cases} 1 & \text{für } x \in A, \\ 0 & \text{für } x \in X \setminus A \end{cases}$$

kann als Funktion mit Werten im Restklassenkörper $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ aufgefasst werden, und dann gilt (man beachte $1 + 1 = 0$)

$$\begin{aligned} &\chi_{A \triangle B} = \chi_A + \chi_B, \qquad &&\chi_{A \cap B} = \chi_A \cdot \chi_B, \\ &\chi_\emptyset(x) = 0, &&\chi_X(x) = 1 \end{aligned}$$

für alle $x \in X$. Die Abbildung $A \mapsto \chi_A$ ist eine Bijektion von $\mathcal{P}(X)$ auf den Ring aller $\mathbb{F}_2$-wertigen Funktionen auf X, denn das Urbild einer Funktion f ist $\{x \in X \mid f(x) = 1\}$. Damit überträgt sich die Ringstruktur auf $\mathcal{P}(X)$. $\square$

Da die Menge der $\mathbb{F}_2$-wertigen Funktionen auf X eine $\mathbb{F}_2$-Algebra ist, könnte man auch $\mathcal{P}(X)$ zu einer solchen machen, aber man benutzt den Begriff Algebra von Mengen in einem anderen Sinne.

[11] In der Wahrscheinlichkeitsrechnung sagt man „Indikatorfunktion“, da der Begriff „charakteristische Funktion“ dort eine andere Bedeutung hat.

Definition 13 *Eine Menge $\mathcal{R} \subseteq \mathcal{P}(X)$ heißt Ring, wenn $\mathcal{R}$ ein Unterring von $(\mathcal{P}(X), \triangle, \cap)$ ist. Ein Ring, der das Einselement X enthält, heißt Algebra über X.*

Ein Ring muss also die leere Menge enthalten und abgeschlossen unter den Operationen $\triangle$ und $\cap$ sein. Eine Algebra ist außerdem invariant unter der Bildung des Komplements $A^c = X \setminus A$ (wobei sich die Gesamtmenge X aus dem Zusammenhang ergeben muss). Es gelten die de Morganschen Regeln

$$(A \cap B)^c = A^c \cup B^c, \qquad (A \cup B)^c = A^c \cap B^c.$$

Satz 13 *Eine Menge $\mathcal{R} \subseteq \mathcal{P}(X)$ ist genau dann ein Ring, wenn sie die leere Menge enthält und abgeschlossen unter den Operationen Vereinigung und Differenz ist.*

Eine Menge $\mathcal{A} \subseteq \mathcal{P}(X)$ ist genau dann eine Algebra über X, wenn sie die die folgenden äquivalenten Eigenschaften besitzt:

(a) $\mathcal{A}$ enthält X und ist abgeschlossen unter den Operationen Vereinigung und Komplementbildung.

(b) $\mathcal{A}$ enthält X und ist abgeschlossen unter den Operationen Durchschnitt und Komplementbildung.

Beweis. Ein Ring hat die genannten Eigenschaften, denn

$$A \cup B = (A \triangle B) \triangle (A \cap B), \qquad A \setminus B = A \triangle (A \cap B).$$

Die Umkehrung folgt aus

$$A \triangle B = (A \setminus B) \cup (B \setminus A), \qquad A \cap B = A \setminus (A \setminus B)$$

(Abgeschlossenheit bezüglich $\triangle$ und $\cap$) und

$$A \triangle A = \emptyset$$

(Existenz des entgegengesetzten Elements zu A).

In jeder Algebra gelten offensichtlich (a) und (b). Wegen

$$A \cup B = (A^c \cap B^c)^c, \qquad A \cap B = (A^c \cup B^c)^c$$

sind (a) und (b) äquivalent. Gelten (a) und (b), so ist $\emptyset = X^c \in \mathcal{A}$, und aus A, $B \in \mathcal{A}$ folgt $A \setminus B = A \cap B^c \in \mathcal{A}$. □

Der Durchschnitt zweier Ringe (oder sogar einer beliebigen Familie von Ringen) ist ein Ring. Ist also $\mathcal{E} \subseteq \mathcal{P}(X)$, so gibt es einen kleinsten Ring, der $\mathcal{E}$ enthält. Man nennt ihn den von $\mathcal{E}$ erzeugten Ring. Analog definiert man die von $\mathcal{E}$ erzeugte Algebra über X.

Beispiele.

- Den Ring aller endlichen Teilmengen von X,
- den Ring aller Polytope in einem endlichdimensionalen reellen Vektorraum, erzeugt von den abgeschlossenen (oder den offenen) Halbräumen.

Wir wollen nun eine Verallgemeinerung der Begriffe Flächen- und Rauminhalt definieren.

Definition 14 *Es sei $\mathcal{M} \subseteq \mathcal{P}(X)$. Eine Abbildung $\mu : \mathcal{M} \to [0, \infty]$ heißt Inhalt, wenn sie nicht identisch gleich ∞ ist und für disjunkte $A_1, \ldots, A_m \in \mathcal{M}$, deren Vereinigung zu $\mathcal{M}$ gehört, gilt*

$$\mu(A_1 \cup \cdots \cup A_m) = \mu(A_1) + \cdots + \mu(A_m).$$

Der Inhalt μ heißt endlich, falls $\mu(A) < \infty$ für alle $A \in \mathcal{M}$ ist. Eine μ-Nullmenge ist eine Menge $A \in \mathcal{M}$ mit $\mu(A) = 0$.

Ist $\mathcal{M}$ ein Ring, so folgt obige Eigenschaft für beliebiges m aus dem Fall $m = 2$. Nehmen wir ein $A \in \mathcal{M}$ mit $\mu(A) \neq \infty$, so folgt aus $\mu(A) = \mu(A \cup \emptyset)$, dass $\mu(\emptyset) = 0$.

Beispiele.

- $\mathcal{M} = \mathcal{P}(X)$, $\mu(A)$ ist die Anzahl der Elemente von A. (Beim Würfeln ist $X = \{1, \ldots, 6\}$ der Ereignisraum und $\frac{1}{6}\mu$ die Wahrscheinlichkeit.)
- $a \in X$ fest, $\mu(A) = \chi_A(a)$. (Punktmasse der Größe 1 im Punkt a.)
- $\mathcal{M}$ ist die Menge aller Intervalle in $\mathbb{R}$, $\mu(I)$ ist die Länge von I.

Satz 14 *Es sei μ ein Inhalt auf einem Ring $\mathcal{R}$. Dann gilt für alle Mengen A, B, A_i aus $\mathcal{R}$:*

(i) *Aus $A \subseteq B$ folgt $\mu(A) \leq \mu(B)$.*

(ii) $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$.

(iii) $\mu(A_1 \cup \cdots \cup A_m) \leq \mu(A_1) + \cdots + \mu(A_m)$.

(iv) *Sind die A_i paarweise disjunkt und ist $\bigcup_{j=1}^{\infty} \in \mathcal{M}$, so gilt*

$$\sum_{i=1}^{\infty} \mu(A_i) \leq \mu\left(\bigcup_{i=1}^{\infty} A_i\right).$$

Beweis. (i) Wegen $\mu(B \setminus A) \geq 0$ ist $\mu(A) \leq \mu(A) + \mu(B \setminus A) = \mu(B)$.

(ii) Es ist $\mu(A \cup B) = \mu(A) + \mu(B \setminus A)$ und $\mu(B) = \mu(B \setminus A) + \mu(A \cap B)$.

(iii) Mit $B_0 = \emptyset$ und $B_n = A_n \setminus \bigcup_{i<n} A_i$ ist

$$\mu\left(\bigcup_{i=1}^{m} A_i\right) = \mu\left(\bigcup_{i=1}^{m} B_i\right) = \sum_{i=1}^{m} \mu(B_i) \leq \sum_{i=1}^{m} \mu(A_i).$$

(iv) Für alle $m \in \mathbb{N}$ ist

$$\sum_{i=1}^{m} \mu(A_i) = \mu\left(\bigcup_{i=1}^{m} A_i\right) \leq \mu\left(\bigcup_{i=1}^{\infty} A_i\right).$$

Die linke Seite konvergiert nach dem Monotoniekriterium und hat die rechte als obere Schranke. □

8 Der elementargeometrische Inhalt

Wir wollen einen Inhalt von Polytopen definieren. Dieses Problem führen wir durch Zerlegung auf den Fall von konvexen Polyedern zurück. Unter einem konvexen Polyeder in einem endlichdimensionalen reellen Vektorraum V verstehen wir eine beschränkte Menge K, die sich als Durchschnitt von abgeschlossenen Halbräumen darstellen lässt. Ersetzen wir einige (oder alle, oder keine) dieser Halbräume durch ihren Rand, so erhalten wir eine Seite von K; diese ist wieder ein konvexes Polyeder. Es gibt einen kleinsten affinen Unterraum, der K enthält, und in diesem ist K der Abschluss der Menge seiner inneren Punkte. Die Dimension dieses Unteraums nennen wir die Dimension von K.

Definition 15 *Ein endlicher Polyederkomplex in V ist eine Menge $\mathcal{K}$ von konvexen Polyedern in V mit folgenden Eigenschaften:*

1. *Jede Seite eines Polyeders $K \in \mathcal{K}$ gehört zu $\mathcal{K}$.*
2. *Der Durchschnitt zweier Polyeder K_1, $K_2 \in \mathcal{K}$ ist leer oder eine gemeinsame Seite von K_1 und K_2.*

Die Elemente von $\mathcal{K}$ nennen wir Zellen, die nulldimensionalen Zellen nennen wir Ecken. Der Träger eines Polyederkomplexes ist die Vereinigung seiner Zellen. Ein Polyederkomplex $\mathcal{U}$ heißt Unterteilung des Polyederkomplexes $\mathcal{K}$, wenn jede Zelle von $\mathcal{U}$ in einer Zelle von $\mathcal{K}$ enthalten ist und jede Zelle von $\mathcal{K}$ eine Vereinigung von Zellen von $\mathcal{U}$ ist.

Satz 15 *Jedes kompakte Polytop P ist Träger eines Polyederkomplexes $\mathcal{K}$. Mehr noch, sind kompakte Polytope $P_1, \ldots, P_k$ gegeben, so gibt es einen Polyederkomplex $\mathcal{K}$ mit Träger $P_1 \cup \cdots \cup P_k$, so dass jedes P_j Träger eines Teilkomplexes von $\mathcal{K}$ ist.*

Beweis. Nach Definition lassen sich die P_j durch Mengenoperationen aus endlich vielen Halbräumen gewinnen. Es sei $\mathcal{E}$ die Menge der Hyperebenen, die diese Halbräume beranden, sowie der offenen Halbräume zu beiden Seiten dieser Hyperebenen. Weiter sei $\mathcal{R}$ der von $\mathcal{E}$ erzeugte Ring und $\mathcal{M}_j$ die Menge der minimalen Elemente von $\mathcal{R}$, die in P_j enthalten sind. Dann sind die Elemente von $\mathcal{M}_j$ konvex, und ihre Vereinigung ist P_j. Die Menge der Abschlüsse der Elemente von $\mathcal{M}_1 \cup \cdots \cup \mathcal{M}_k$ ist der gesuchte Polyederkomplex $\mathcal{K}$. □

Folgerung 7 *Zwei Polyederkomplexe mit demselben Träger haben eine gemeinsame Unterteilung.*

Dies folgt, wenn wir den Satz auf die Menge der Zellen beider Komplexe anwenden.

Jeder Simplizialkomplex ist natürlich ein Polyederkomplex, aber nicht umgekehrt.

Satz 16 *Jeder Polyederkomplex $\mathcal{K}$ lässt sich in einen Simplizialkomplex $\mathcal{S}$ unterteilen.*

Beweis. Unter dem n-dimensionalen Skelett von $\mathcal{K}$ verstehen wir die Menge $\mathcal{K}_n$ aller Zellen der Dimension höchstens n. Dies ist ebenfalls ein Polyederkomplex, und für genügend große n ist $\mathcal{K}_n = \mathcal{K}$. Wir zeigen durch Induktion, dass sich $\mathcal{K}_n$ in einen Simplizialkomplex $\mathcal{S}_n$ unterteilen lässt. Natürlich können wir $\mathcal{S}_0 = \mathcal{K}_0$ setzen. Ist $n > 0$, so existiert nach Induktionsvoraussetzung eine simpliziale Unterteilung $\mathcal{S}_{n-1}$ von $\mathcal{K}_{n-1}$. Nun wählen wir im Inneren jeder n-dimensionalen Zelle K von $\mathcal{K}$ einen Punkt a_K. Wir fügen zu $\mathcal{S}_{n-1}$ die Simplexe $\operatorname{conv}(\{a_K\} \cup S)$ hinzu, wobei $S \in \mathcal{S}_{n-1}$ im Rand von K liegt. □

Ist $\mathcal{K}$ selbst ein Simplizialkomplex und wählen für a_K den Schwerpunkt von K, so erhalten wir die baryzentrische Unterteilung.

Man kann versuchen, mit weniger Simplexen auszukommen, indem man in einer Zelle K nicht einen inneren Punkt, sondern eine Ecke a wählt und K in die Pyramiden $\operatorname{conv}(\{a\} \cup L)$ zerlegt, wobei L die Seiten von K durchläuft, die a nicht enthalten. Dieselbe Methode kann man rekursiv auf diese Seiten L anwenden. Bei den Seiten, die a enthalten, gibt es aber Kompatibilitätsprobleme mit benachbarten Zellen.

Beispiel. Es sei K ein n-dimensionales Parallelepiped, also

$$K = \{a + \lambda_1 v_1 + \dots \lambda_n v_n \mid 0 \leq \lambda_i \leq 1 \text{ für alle } i\},$$

wobei die Vektoren $v_1, \dots, v_n$ linear unabhängig sind. Seine Ecken sind $a_I = a + \sum_{i \in I} v_i$ für alle $I \subseteq \{1, \dots, n\}$. Offensichtlich ist K die Vereinigung der Mengen

$$S_\pi = \{a + \lambda_1 v_1 + \dots + \lambda_n v_n \mid 0 \leq \lambda_{\pi(1)} \leq \dots \leq \lambda_{\pi(n)} \leq 1\}$$

über alle Permutationen π von $\{1, \dots, n\}$. Die Menge S_π ist ein Simplex mit den Ecken $a_{\pi(I_k)}$, wobei $I_k = \{i \mid k < i \leq n\}$ und $0 \leq k \leq n$. Die Menge $\mathcal{S}$ dieser Simplexe und ihrer Seiten ist ein Simplizialkomplex mit dem Träger K.

Definition 16 *Ist S ein n-Simplex in $\mathbb{R}^n$ mit den Ecken $a_0, \dots, a_n$ (geschrieben als Spaltenvektoren), so definieren wir*

$$\mu(S) = \frac{1}{n!} \left|\det A\right|, \qquad \textit{wobei} \qquad A = \begin{pmatrix} 1 & 1 & \dots & 1 \\ a_{01} & a_{11} & \dots & a_{n1} \\ \vdots & \vdots & & \vdots \\ a_{0n} & a_{1n} & \dots & a_{nn} \end{pmatrix}.$$

Ist $\mathcal{S}$ ein Simplizialkomplex in $\mathbb{R}^n$, so definieren wir $\mu(\mathcal{S})$ als Summe der Zahlen $\mu(S)$ über alle n-Simplexe S von $\mathcal{S}$.

Bemerkung 1. Es ist klar, dass $\mu(S)$ nicht von der Nummerierung der Ecken abhängt. Subtrahieren wir die erste Spalte von A von allen anderen und entwickeln dann $\det A$ nach der ersten Zeile, so erhalten wir

$$\mu(S) = \frac{1}{n!} \left|\det V\right|, \qquad \text{wobei} \qquad V = \begin{pmatrix} v_{11} & v_{21} & \dots & v_{n1} \\ \vdots & \vdots & & \vdots \\ v_{1n} & v_{2n} & \dots & v_{nn} \end{pmatrix},$$

wobei $v_i = a_i - a_0$.
Bemerkung 2. Ist K das von $v_1, \dots, v_n$ aufgespannte Parallelepiped wie im obigen Beispiel, so ist $a_I - a = \sum_{i \in I} v_i$, und durch elementare Spaltenoperationen folgt

$$\mu(S_\pi) = \frac{1}{n!} \left|\det V\right|$$

unabhängig von π. Ist also $\mathcal{S}$ der obige Simplizialkomplex mit dem Parallelepiped K als Träger, so ist

$$\mu(\mathcal{S}) = \left|\det V\right|.$$

Bemerkung 3. Ist ein Inhalt $\nu(T)$ für Simplexe $T \subseteq \mathbb{R}^{n-1}$ analog zu μ definiert und ist $a = (0, \ldots, 0, h)^\top$, so erhalten wir für den Pyramideninhalt

$$\mu(\operatorname{conv}(\{a\} \cup T)) = \frac{1}{n}\,|h|\nu(T),$$

indem wir A nach der Spalte a entwickeln.
Bemerkung 4. Ist $f : \mathbb{R}^n \to \mathbb{R}^n$ eine affine Abbildung, also $f(x) = Bx + v$ mit einer $n \times n$-Matrix B und einem Vektor $v \in \mathbb{R}^n$, so gilt

$$\mu(f(S)) = |\det B|\,\mu(S),$$

denn $f(S)$ wird von den Spaltenvektoren der Matrix BV aufgespannt.
Bemerkung 5. Wir können einen beliebigen affinen Raum E mit $\mathbb{R}^n$ identifizieren und einen Inhalt μ von Simplizialkomplexen in E definieren. Die Abbildung μ ist bis auf einen positiven skalaren Faktor bestimmt.

Satz 17 *Haben zwei Simplizialkomplexe $\mathcal{S}$ und $\mathcal{T}$ denselben Träger, so ist $\mu(\mathcal{S}) = \mu(\mathcal{T})$.*

Beweis. Nach Folgerung 7 und Satz 16 haben die beiden Simplexe eine gemeinsame simpliziale Unterteilung. Darum können wir annehmen, dass $\mathcal{T}$ eine Unterteilung von $\mathcal{S}$ ist. Weiter können wir annehmen, dass $\mathcal{S}$ aus einem einzigen n-dimensionalen Simplex S und seinen Seiten S_i besteht, wobei jeweils die Ecke a_i nicht zu S_i gehört. Es sei E_i die Hyperebene, die S_i enthält.

Ersetzen wir in der Matrix A die Koordinaten von a_i durch 0, so entsteht eine Matrix A_i. Entwickeln wir A nach der ersten Zeile und jeweils A_i nach der $i$ten Spalte, so erhalten wir

$$\det A = \sum_{i=0}^{n} \det A_i, \qquad \text{also} \qquad \mu(S) = \sum_{i=0}^{n} \varepsilon_i \mu(\operatorname{conv}(\{0\} \cup S_i)),$$

wobei $\varepsilon_i = \operatorname{sgn} \frac{\det A_i}{\det A}$ angibt, ob 0 und a_i auf derselben Seite von E_i liegen oder nicht. Eine analoge Formel gilt für die Simplexe von $\mathcal{T}$, und in der Summe $\mu(\mathcal{T})$ kürzen sich die Beiträge ihrer Seiten im Inneren von S weg. Es bleibt zu zeigen, dass für jedes i gilt

$$\mu(\operatorname{conv}(\{0\} \cup S_i)) = \sum_{T \subseteq S_i} \mu(\operatorname{conv}(\{0\} \cup T)),$$

wobei sich die Summe über alle $(n-1)$-dimensionalen Simplexe $T \in \mathcal{T}$ mit $T \subseteq S_i$ erstreckt. Diese bilden eine Unterteilung von S_i, und nach Induktionsvoraussetzung gilt für einen elementargeometrischen Inhalt ν_i auf E_i (vgl.

Bemerkung 5)

$$\nu_i(S_i) = \sum_{T \subseteq S_i} \nu_i(T).$$

Nun folgt die Behauptung aus der Formel für den Pyramideninhalt (Bemerkung 3). □

Definition 17 *Ist P ein beschränktes Polytop, so setzen wir $\mu(P) = \mu(\mathcal{S})$, wobei $\mathcal{S}$ ein Simplizialkomplex mit dem Träger $\bar{P}$ ist.*

Bemerkung 6. Da sich P als endliche Vereinigung konvexer Polytope schreiben lässt, gibt endlich viele Hyperebenen $H_1, \ldots, H_k$, so dass $\bar{P} \subseteq P \cup \bigcup_{j=1}^k H_j$. Jeder innere Punkt von $\bar{P}$ hat also eine Umgebung U, so dass $U \setminus \bigcup_{j-1}^k H_j$ eine (nichtleere) Teilmenge von P ist.

Satz 18 *Die Abbildung μ vom Ring der beschränkten Polytope in $\mathbb{R}^n$ nach $[0, \infty)$ ist ein endlicher Inhalt. Ist $f : \mathbb{R}^n \to \mathbb{R}^n$ eine affine Abbildung, also $f(x) = Bx + v$, so gilt*

$$\mu(f(P)) = |\det B|\, \mu(P).$$

Beweis. Zu gegebenen disjunkten beschränkten Polytopen P_1, P_2 gibt es nach Satz 15 und Satz 16 einen Simplizialkomplex $\mathcal{S}$ mit Träger $\bar{P}_1 \cup \bar{P}_2 = \overline{P_1 \cup P_2}$, so dass $\bar{P}_j$ der Träger eines Teilkomplexes $\mathcal{S}_j$ ist.

Gehört ein Simplex $S \in \mathcal{S}$ nicht zum Teilkomplex $\mathcal{S}_j$, so ist $S \cap \bar{P}_j$ eine Vereinigung von echten Seiten von S. Wegen $S = (S \cap \bar{P}_1) \cup (S \cap \bar{P}_2)$ folgt $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$.

Hätten $\bar{P}_1$ und $\bar{P}_2$ gemeinsamen innere Punkte, so hätten sie nach Bemerkung 6 gemeinsame Punkte, was der Voraussetzung widerspricht. Somit haben $\mathcal{S}_1$ und S_2 keine gemeinsamen n-Simplexe. Es folgt $\mu(\mathcal{S}_1) + \mu(\mathcal{S}_2) = \mu(\mathcal{S})$ und somit $\mu(P_1) + \mu(P_2) = \mu(P_1 \cup P_2)$.

Die behauptete Transformationsformel folgt aus dem bereits bewiesenen Spezialfall für Simplexe (Bemerkung 3). □

Lemma 15 *Für jedes beschränkte Polytop P in $\mathbb{R}^n$ und jedes $\varepsilon > 0$ gibt es eine kompaktes Polytop K, so dass $K \subseteq P$ und $\mu(P \setminus K) < \varepsilon$, sowie ein beschränktes offenes Polytop U, so dass $P \subseteq U$ und $\mu(U \setminus P) < \varepsilon$.*

Beweis. Es seien H_j wie in Bemerkung 6. Wir wählen $\delta > 0$, bezeichnen mit U_j die δ-Umgebung von H_j und setzen

$$K = \bar{P} \setminus \bigcup_{j=1}^k U_j.$$

Dies ist ein kompaktes Polytop, enthalten in P, und

$$P \setminus K \subseteq \bigcup_{j=1}^{k} U_j \cap P.$$

Da P beschränkt ist, ist $U_j \cap P$ in einem Quader enthalten, dessen Kanten parallel bzw. senkrecht zu H_j sind, wobei die zu H_j senkrechten Kanten die Länge δ haben. Den Inhalt des Quaders kann man nach Bemerkung 2 berechnen, und durch geeignete Wahl von δ kann man $\mu(U_j \cap P)$ so klein machen, dass $\mu(P \setminus K) < \varepsilon$.

Nun wählen wir einen offenen Quader V, der P enthält, und einen kompakten Quader Q, der V enthält. Die obige Konstruktion liefert ein kompaktes Polytop $L \subseteq Q \setminus P$ mit $\mu((Q \setminus P) \setminus L) < \varepsilon$, und wir setzen $U = V \setminus L$. Dann ist U ein offenes Polytop und enthält P. Wir können annehmen, dass L das kompakte Polytop $Q \setminus V$ enthält, so dass $Q \setminus L \subseteq V$ und $(Q \setminus P) \setminus L = (Q \setminus L) \setminus P = U \setminus P$. □

9 σ-Additivität

Definition 18 *Ein Inhalt $\mu : \mathcal{M} \to [0, \infty]$ heißt σ-additiv, wenn für beliebige disjunkte A_1, A_2, A_3, $\cdots \in \mathcal{M}$ mit der Eigenschaft $A = \bigcup_{j=1}^{\infty} \in \mathcal{M}$ gilt*

$$\sum_{i=1}^{\infty} \mu(A_i) = \mu(A).$$

Nach Satz 14 gilt immer

$$\sum_{j=1}^{\infty} \mu(A_j) \leq \mu(A).$$

Satz 19 *Es seien P, P_1, P_2, ... beschränkte Polytope in $\mathbb{R}^n$, so dass*

$$P \subseteq \bigcup_{j=1}^{\infty} P_j.$$

Dann gilt für den elementargeometrischen Inhalt μ

$$\mu(P) \leq \sum_{j=1}^{k} \mu(P_j).$$

Der elementargeometrische Inhalt ist also σ-additiv.

Beweis. Für ein beliebiges $\varepsilon > 0$ gibt es nach Lemma 15 ein kompaktes Polytop K und offene Polytope U_j, so dass $K \subseteq P$, $\mu(P \setminus K) < \varepsilon$, $P_j \subseteq U_j$ und $\mu(U_j \setminus P_j) < 2^{-j}\varepsilon$ für alle j. Nun bilden die U_j eine offene Überdeckung von K, und nach dem Satz von Heine-Borel gibt es ein k, so dass

$$K \subset \bigcup_{j=1}^{k} U_j.$$

Nach Satz 14 ist

$$\mu(P) - \varepsilon \leq \sum_{j=1}^{k} (\mu(P_j) + 2^{-j}\varepsilon).$$

Dies gilt auch, wenn wir k durch ∞ ersetzen. Berechnen wir die geometrische Reihe, so folgt

$$\mu(P) \leq \sum_{j=1}^{k} \mu(P_j) + 2\varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt die Behauptung. □

Satz 20 *Für einen Inhalt μ auf einem Ring $\mathcal{R}$ sind folgende Aussagen äquivalent:*

(a) μ ist σ-additiv.

(b) Für jede aufsteigende Folge $B_1 \subseteq B_2 \subseteq \ldots$ in $\mathcal{R}$ mit $\bigcup_{j=1}^{\infty} B_j = B \in \mathcal{R}$ gilt

$$\lim_{j\to\infty} \mu(B_j) = \mu(B).$$

Ist μ zudem ein endlicher Inhalt, so sind (a) und (b) äquivalent zu:

(c) Für jede absteigende Folge $C_1 \supseteq C_2 \supseteq \ldots$ in $\mathcal{R}$ mit $\bigcap_{j=1}^{\infty} C_j = C \in \mathcal{R}$ gilt

$$\lim_{j\to\infty} \mu(C_j) = \mu(C).$$

Beweis. (b)⇒(a): Sind die $A_j \in \mathcal{R}$ wie in der Definition der σ-Additivität und setzen wir $B_k = \bigcup_{j=1}^{\infty} A_j$, so bilden die B_k eine aufsteigende Folge mit $\bigcup_{k=1}^{\infty} B_K = A$. Aus Satz 14 folgt

$$\mu(B_k) = \sum_{j=1}^{k} \mu(A_j),$$

und gehen wir zum Grenzwert $k \to \infty$ über, so erhalten wir wegen (b) auf der linken Seite $\mu(A)$.

(a)$\Rightarrow$(b): Sind die B_j wie im Satz und setzen wir $B_0 = \emptyset$, $A_j = B_j \setminus B_{j-1}$, so sind die $A_j \in \mathcal{R}$ disjunkt und $\bigcup_{j=1}^k A_j = B_k$ für alle k, also nach Satz 14

$$\sum_{j=1}^k \mu(A_j) = \mu(B_k).$$

Gehen wir zum Grenzwert über, so erhalten wir

$$\sum_{j=1}^\infty \mu(A_j) = \lim_{k\to\infty} \mu(B_k),$$

und wegen der σ-Additivität ist die linke Seite gleich

$$\mu\left(\bigcup_{j=1}^\infty A_j\right) = \mu(B).$$

(b)$\Leftrightarrow$(c): Ist eine aufsteigende Folge B_j gegeben und $B = \bigcup_{j=1}^\infty B_j \in \mathcal{R}$, so ist $C_j = B \setminus B_j$ eine absteigende Folge. Ist eine absteigende Folge C_j gegeben und setzen wir $B = C_1$, so ist $B_j = B \setminus C_j$ eine aufsteigende Folge. In beiden Fällen gilt

$$\mu(B_j) + \mu(C_j) = \mu(B).$$

Ist $\mu(B) < \infty$, so folgt aus der Konvergenz eines Summanden auf der linken Seite die des anderen. $\square$

Beispiel. Es sei μ ein beliebiger endlicher Inhalt auf dem Ring der Polytope in $\mathbb{R}$. Wir halten $c \in \mathbb{R}$ fest und setzen

$$F(x) = \begin{cases} \mu((c,x]) & \text{falls } x \geq c, \\ -\mu((x,c]) & \text{falls } x < c. \end{cases}$$

Dann ist F monoton wachsend, und für $a \leq b$ gilt wegen der Additivität von μ

$$\mu((a,b]) = F(b) - F(a).$$

(Dies prüft man einzeln für die drei Fälle $c \leq a \leq b$, $a < c \leq b$ und $a \leq b < c$ nach.) Ist μ beschränkt, so kann man die Funktion F durch ihren Grenzwert für $c \to -\infty$ ersetzen. (Dies tut man in der Wahrscheinlichkeitstheorie und nennt F die Verteilungsfunktion von μ.) Ist μ σ-additiv, so gilt für jede monoton fallende Folge b_k mit Grenzwert b wegen Satz 20

$$\lim_{k\to\infty} F(b_k) - F(b) = \lim_{k\to\infty} \mu((b,b_k]) = \mu(\emptyset) = 0,$$

d. h. F ist rechtsseitig stetig, und für jede monoton wachsende Folge a_k mit Grenzwert a ist

$$F(a) - \lim_{k\to\infty} F(a_k) = \lim_{k\to\infty} \mu((a_k, a]) = \mu(\{a\}).$$

Man kann zeigen, dass umgekehrt zu jeder rechtsseitig stetigen monoton wachsenden Funktion F auf $\mathbb{R}$ durch

$$\mu((a,b]) = F(b) - F(a), \qquad \mu([a,b]) = F(b) - \lim_{\substack{x\to a\\ x<a}} F(x)$$

ein σ-additiver Inhalt auf der Menge der Intervalle definiert wird, der sich zu einem σ-additiven Inhalt auf dem Ring der beschränkten Polytope in $\mathbb{R}$ fortsetzt.

Definition 19 *Es sei $\mathcal{P}$ der Ring der beschränkten Polytope in $\mathbb{R}^n$. Für eine beschränkte Menge $A \subset \mathbb{R}^n$ sei*

$$\mu_*(A) = \sup_{\substack{P\in\mathcal{P}\\ P\subseteq A}} \mu(P), \qquad \mu^*(A) = \inf_{\substack{P\in\mathcal{P}\\ P\supseteq A}} \mu(P)$$

(innerer und äußerer Jordan-Inhalt). Im Falle $\mu_*(A) = \mu^*(A)$ nennen wir A Jordan-messbar und bezeichnen den gemeinsamen Wert mit $\mu(A)$.

Man kann zeigen (Aufgabe 38), dass die Menge der Jordan-messbaren Teilmengen von $\mathbb{R}^n$ ein Ring ist und μ ein Inhalt auf diesem Ring, genannt Jordan-Inhalt.

Satz 21 *Der Jordan-Inhalt ist σ-additiv.*

Beweis. Ist die Jordan-messbare Menge A die Vereinigung der Jordan-messbaren Mengen A_1, A_2, ... und ist $\varepsilon > 0$, so gibt es nach Definition beschränkte Polytope $P \subseteq A$, $P_1 \supseteq A_1$, $P_2 \supseteq A_2$, so dass

$$\mu(A \setminus P) < \varepsilon, \qquad \mu(P_j \setminus A_j) < 2^{-j}\varepsilon.$$

Aus Satz 19 folgt, dass

$$\mu(P) \le \sum_{j=1}^{\infty} \mu(P_j),$$

also

$$\mu(A) - \varepsilon \le \sum_{j=1}^{\infty} (\mu(A_j) + 2^{-j}\varepsilon),$$

und wir beenden den Beweis mit demselben Argument wie dort. □

10 σ-Algebren und Dynkin-Systeme

Ein Inhalt ist wesentlich einfacher zu handhaben, wenn man die Operation der abzählbaren Vereinigung, die in der Definition der σ-Additivität vorkommt, in seinem Definitionsbereich unbegrenzt ausführen kann.

Definition 20 *Ein Ring $\mathcal{R}$ heißt σ-Ring, wenn für jede Folge von Mengen A_n in $\mathcal{R}$ die Vereinigung $\bigcup_{n=1}^{\infty} A_n$ ebenfalls zu $\mathcal{R}$ gehört, und er heißt δ-Ring, wenn für jede Folge in $\mathcal{R}$ auch der Durchschnitt zu $\mathcal{R}$ gehört. Ein σ-Ring, der eine Algebra ist, heißt σ-Algebra.*

Jeder σ-Ring ist ein δ-Ring, denn ist $A := \bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$, so auch

$$\bigcap_{n=1}^{\infty} A_n = A \setminus \bigcup_{n=1}^{\infty} (A \setminus A_n) \in \mathcal{R}.$$

Lemma 16 *Eine Teilmenge $\mathcal{A} \subseteq \mathcal{P}(X)$ ist genau dann eine σ-Algebra, wenn sie folgenden äquivalenten Bedingungen genügt:*

(a) $X \in \mathcal{A}$, aus $A \in \mathcal{A}$ folgt $A^c \in \mathcal{A}$, und aus $A_1, A_2, \ldots \in \mathcal{A}$ folgt $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

(b) $X \in \mathcal{A}$, aus $A \in \mathcal{A}$ folgt $A^c \in \mathcal{A}$, und aus $A_1, A_2, \ldots \in \mathcal{A}$ folgt $\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$.

Beweis. Jede σ-Algebra genügt (a) und (b). Wegen

$$\left(\bigcup_{n=1}^{\infty} A_n \right)^c = \bigcap_{n=1}^{\infty} A_n^c$$

ist (a)$\Leftrightarrow$(b). Ist $A, B \in \mathcal{A}$ und setzt man $A_1 = A$, $A_n = B$ für $n > 1$, so folgt aus (a), dass $A \cup B \in \mathcal{A}$, also ist $\mathcal{A}$ eine Algebra. $\square$

Beispiele.

- Die Menge aller endlichen Teilmengen von X ist ein δ-Ring, aber für unendliches X kein σ-Ring.
- Die Menge aller abzählbaren[12] Teilmengen von X ist ein σ-Ring, aber für überabzählbares X keine σ-Algebra.

[12] Ich nenne auch endliche Mengen abzählbar.

- Der Durchschnitt einer beliebigen Familie von σ-Ringen ist ein σ-Ring, also gibt es für jede Teilmenge $\mathcal{E}$ von $\mathcal{P}(X)$ einen kleinsten σ-Ring, der $\mathcal{E}$ enthält, genannt der von $\mathcal{E}$ erzeugte σ-Ring. Analoges gilt für σ-Algebren.

- Der Ring der Jordan-messbaren Mengen in $\mathbb{R}^n$ ist kein σ-Ring. So ist z. B. die Menge $A = \mathbb{Q} \cap [0,1]$ eine abzählbare Vereinigung von einelementigen, also Jordan-messbaren Mengen, aber nicht Jordan-messbar: Ein Polytop in $\mathbb{R}$ ist eine endliche Vereinigung von disjunkten Intervallen (seinen Zusammenhangskomponenten). Die einzigen in A enthaltenen Intervalle sind Punkte, also ist $\mu_*(A) = 0$. Überdecken wir A durch ein Polytop P, so darf das Polytop $[0,1] \setminus P$ keine rationale Zahl enthalten, kann also nur aus endlich vielen Punkten bestehen, und somit ist $\mu^*(A) = 1$.

Definition 21 *Die von der Menge der offenen Teilmengen eines metrischen (oder topologischen) Raumes erzeute σ-Algebra nennt man die Algebra der Borel-Mengen von X.*

Dies ist natürlich gleichzeitig die von der Menge der abgeschlossenen Teilmengen erzeugte σ-Algebra. Im Fall von $X = \mathbb{R}^n$ ist es auch die vom Ring der (beschränkten) Polytope erzeugte σ-Algebra (Aufgabe 41).

Es ist mitunter schwierig festzustellen, ob eine Teilmenge von $\mathcal{P}(X)$ eine σ-Algebra ist. Manchmal gelingt zunächst der Nachweis eines Teils der geforderten Eigenschaften.

Definition 22 *Eine Teilmenge $\mathcal{D} \subseteq \mathcal{P}(X)$ heißt Dynkin-System (über X), wenn folgendes gilt:*

(i) $X \in \mathcal{D}$,

(ii) aus A, $B \in \mathcal{D}$ und $B \subseteq A$ folgt $A \setminus B \in \mathcal{D}$,

(iii) für jede Folge paarweise disjunkter Mengen $A_n \in \mathcal{D}$ ist $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.

Jede σ-Algebra ist ein Dynkin-System, aber nicht umgekehrt: Ist $\mathcal{D}$ die Menge aller Teilmengen von $\{1, 2, \ldots, 2n\}$ mit einer geraden Anzahl von Elementen, so ist $\mathcal{D}$ ein Dynkin-System, aber keine Algebra, denn $\mathcal{D}$ ist nicht durchschnittsstabil (vgl. Satz 13(b)).

Lemma 17 *Ein Dynkin-System ist genau dann eine σ-Algebra, wenn es durchschnittsstabil ist.*

Beweis. Jede σ-Algebra ist ein durchschnittsstabiles Dynkin-System. Umgekehrt sei $\mathcal{D}$ ein solches. Für beliebige $A, B \in \mathcal{D}$ gilt dann $A \setminus B = A \setminus (A \cap B) \in \mathcal{D}$. Ist eine Folge von beliebigen Teilmengen $A_n \in \mathcal{D}$ gegeben, so setzen wir $B_n = A_n \setminus A_{n-1} \setminus \cdots \setminus A_1$ (wobei die Operationen von links nach rechts ausgeführt werden). Dann sind die B_n disjunkt, und $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n \in \mathcal{D}$. Somit ist $\mathcal{D}$ eine σ-Algebra. □

Satz 22 *Es sei $\mathcal{E} \subseteq \mathcal{P}(X)$ durchschnittsstabil. Dann ist das von $\mathcal{E}$ erzeugte Dynkin-System $\mathcal{D}$ gleich der von $\mathcal{E}$ erzeugten σ-Algebra $\mathcal{A}$.*

Beweis. Da $\mathcal{A}$ ein Dynkin-System ist, gilt $\mathcal{D} \subseteq \mathcal{A}$. Für die umgekehrte Inklusion müssen wir zeigen, dass $\mathcal{D}$ eine σ-Algebra ist, wofür nach dem Lemma hinreicht, dass $\mathcal{D}$ durchschnittsstabil ist.

Für jedes $C \in \mathcal{D}$ sei[13]

$$\mathcal{D}_C = \{A \in \mathcal{D} \mid A \cap C \in \mathcal{D}\}.$$

Dann ist $\mathcal{D}_C$ ein Dynkin-System (Aufgabe 42). Für jedes $E \in \mathcal{E}$ gilt $\mathcal{E} \subseteq \mathcal{D}_E$, weil $\mathcal{E}$ durchschnittsstabil ist, und mit der Minimalität von $\mathcal{D}$ folgt $\mathcal{D} \subseteq \mathcal{D}_E$. Somit ist $C \cap E \in \mathcal{D}$ für alle $C \in \mathcal{D}$ und $E \in \mathcal{E}$. Mit anderen Worten, für alle $C \in \mathcal{D}$ gilt $\mathcal{E} \subseteq \mathcal{D}_C$, und wegen der Minimalität von $\mathcal{D}$ folgt $\mathcal{D} \subseteq \mathcal{D}_C$. Da $C \in \mathcal{D}$ beliebig war, ist $\mathcal{D}$ durchschnittsstabil. □

11 Maße

Definition 23 *Ein Maß auf einer Menge X ist ein σ-additiver Inhalt auf einer σ-Algebra $\mathcal{A}$ über X. (Man spricht auch von einem Maß auf $\mathcal{A}$.)*

Der Jordan-Inhalt ist kein Maß, weil der Ring der Jordan-messbaren Mengen keine σ-Algebra ist. Die Ursache liegt darin, dass der äußere Jordan-Inhalt nicht σ-subadditiv ist (Eigenschaft (iii) in der folgenden Definition).

Definition 24 *Ein äußeres Maß auf einer Menge X ist eine Abbildung $\eta : \mathcal{P}(X) \to [0, \infty]$ mit folgenden Eigenschaften:*

(a) $\eta(\emptyset) = 0$,

(b) aus $A \subseteq B \subseteq X$ folgt $\eta(A) \leq \eta(B)$,

[13] Solche Mengen A sind „gut für C“, daher nennt man diese Beweismethode das Prinzip der guten Mengen.

(c) für jede Folge von Teilmengen A_n von X gilt

$$\eta\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \eta(A_n).$$

Die Methode von Lebesgue zur Konstruktion eines äußeren Maßes (die eigentlich auf Borel zurückgeht) besteht darin, eine Menge nicht durch ein Polytop, sondern eine abzählbare Vereinigung von Polytopen zu approximieren.

Satz 23 *Es sei μ ein Inhalt auf einem Ring $\mathcal{R}$. Für $A \subseteq X$ setzen wir*

$$\eta(A) = \inf\left\{\sum_{n=1}^{\infty} \mu(B_n) \,\middle|\, B_1,\, B_2, \ldots \in \mathcal{R},\, A \subseteq \bigcup_{n=1}^{\infty} B_n\right\}.$$

(Beachte, dass $\inf \emptyset = \infty$.) Dann ist η ein äußeres Maß. Außerdem gilt genau dann $\eta|_{\mathcal{R}} = \mu$, wenn μ σ-additiv ist.

Bemerkung. Es genügt, das Infimum nur über Folgen von disjunkten Mengen $B_n \in \mathcal{R}$ zu nehmen, da man in einer Überdeckung von A durch beliebige Mengen $B_n \in \mathcal{R}$ diese durch die disjunkten Mengen $C_n = B_n \setminus \bigcup_{j<n} B_j$ ersetzen kann, wobei der Wert der Reihe nicht größer wird.

Beweis. Die Eigenschaften (a) und (b) sind offensichtlich erfüllt. Zum Beweis von (c) sei $\varepsilon > 0$. Dann gibt es für jedes $n \in \mathbb{N}$ Mengen $B_{nk} \in \mathcal{R}$ mit $A_n \subseteq \bigcup_{k=1}^{\infty} B_{nk}$ und

$$\sum_{k=1}^{\infty} \mu(B_{nk}) \leq \eta(A) + 2^{-n}\varepsilon.$$

Die Menge $A := \bigcup_{n=1}^{\infty} A_n$ wird überdeckt durch die Vereinigung der Mengen B_{nk} über alle n und k, also

$$\eta(A) \leq \sum_{n,k=1}^{\infty} \mu(B_{nk}) \leq \sum_{n=1}^{\infty} (\eta(A_n) + 2^{-n}\varepsilon) = \sum_{n=1}^{\infty} \eta(A_n) + \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt (c).

Nun zu $\eta|_{\mathcal{R}}$. Für $A \in \mathcal{R}$ können wir $B_1 = A$, $B_n = \emptyset$ für $n > 1$ nehmen, so dass $\eta(A) \leq \mu(A)$. Ist μ σ-additiv, so gilt für disjunkte $B_n \in \mathcal{R}$ mit $A \subseteq \bigcup_{n=1}^{\infty} B_n$, dass

$$\mu(A) = \sum_{n=1}^{\infty} \mu(A \cap B_n) \leq \sum_{n=1}^{\infty} \mu(B_n),$$

also $\mu(A) \leq \eta(A)$ wegen der Bemerkung vor dem Beweis, so dass $\eta|_{\mathcal{R}} = \mu$. Ohne die σ-Additivität von μ folgt für $B := \bigcup_{n=1}^{\infty} B_n \in \mathcal{R}$ zumindest

$$\eta(B) \leq \sum_{n=1}^{\infty} \mu(B_n) \leq \mu(B)$$

aus der Definition von η und Satz 14(iv). Ist $\eta|_{\mathcal{R}} = \mu$, so sind die äußeren Terme gleich, also ist auch die rechte Ungleichung eine Gleichung, und die σ-Additivität folgt. $\square$

Der Begriff des inneren Jordan-Inhalts war eigentlich überflüssig, denn man kann eine beschränkte Menge $A \subseteq \mathbb{R}^n$ in eine Jordan-messbare Menge Q mit $\mu(Q) < \infty$ einbetten (z. B. einen genügend großen n-dimensionalen Quader), und dann gilt

$$\mu_*(A) + \mu^*(Q \setminus A) = \mu(Q),$$

so dass sich μ_* durch μ^* ausdrücken lässt. Die Jordan-Messbarkeit von A ist dann gleichbedeutend mit

$$\mu^*(A) + \mu^*(Q \setminus A) = \mu^*(Q).$$

Ersetzt man hier μ^* durch ein äußeres Maß η, so erhält man Lebesgues Definition einer messbaren Teilmenge von Q. In einem weiteren Schritt muss man den Begriff der Messbarkeit auf unbeschränkte Mengen verallgemeinern.

Geht man aber von einem beliebigen äußeren Maß auf einer Menge X aus, so sind noch keine messbaren Mengen Q vorgegeben. Carathéodory hat einen Weg gefunden, die Bedingung, dass A in einer messbaren Menge Q von endlichem Maß enthalten sein muss, zu umgehen.

Satz 24 *Es sei η ein äußeres Maß auf X. Eine Teilmenge A von X nennen wir η-messbar, wenn für jede Teilmenge $Q \subseteq X$ gilt*

$$\eta(Q) \geq \eta(Q \cap A) + \eta(Q \cap A^c).$$

Dann gilt:

(i) Die Menge $\mathcal{A}$ der η-messbaren Mengen ist eine σ-Algebra, und $\eta|_{\mathcal{A}}$ ist ein Maß.

(ii) Kommt η wie in Satz 23 von einem Inhalt μ auf einem Ring $\mathcal{R}$, so ist $\mathcal{R} \subseteq \mathcal{A}$.

Bemerkung. Wegen Eigenschaft (c) von η gilt für η-messbares A sogar

$$\eta(Q) = \eta(Q \cap A) + \eta(Q \cap A^c).$$

Beweis. (i) Für A, $B \in \mathcal{A}$ und $Q \subseteq X$ gilt

$$\begin{aligned}
\eta(Q) &\geq \eta(Q \cap A) + \eta(Q \cap A^c) && \text{(weil } A \in \mathcal{A}\text{)} \\
&\geq \eta(Q \cap A) + \eta(Q \cap A^c \cap B) + \eta(Q \cap A^c \cap B^c) && \text{(weil } B \in \mathcal{A}\text{)} \\
&\geq \eta((Q \cap A) \cup (Q \cap A^c \cap B)) + \eta(Q \cap A^c \cap B^c) && \text{(wegen (c))} \\
&= \eta(Q \cap (A \cap B)) + \eta(Q \cap (A \cap B)^c),
\end{aligned}$$

so dass $A \cap B \in \mathcal{A}$. Natürlich ist auch $A^c \in \mathcal{A}$ und $X \in \mathcal{A}$, also ist $\mathcal{A}$ eine Algebra.

Nun sei eine Folge von paarweise disjunkten Mengen $A_n \in \mathcal{A}$ gegeben und $A = \bigcup_{n=1}^{\infty} A_n$. Die Messbarkeit von A_1 bedeutet, wenn wir Q durch $Q \cap (A_1 \cup A_2)$ ersetzen, dass

$$\eta(Q \cap (A_1 \cup A_2)) \geq \eta(Q \cap A_1) + \eta(Q \cap A_2).$$

Durch Induktion nach n folgt

$$\eta\left(Q \cap \bigcup_{j=1}^{n} A_j\right) \geq \sum_{j=1}^{n} \eta(Q \cap A_j).$$

Da $\mathcal{R}$ ein Ring ist, gehört $C_n := \bigcup_{j=1}^{n} A_j$ zu $\mathcal{A}$, und mit $A^c \subseteq C_n^c$ folgt

$$\eta(Q) \geq \eta(Q \cap C_n) + \eta(Q \cap C_n^c) \geq \sum_{j=1}^{n} \eta(Q \cap A_j) + \eta(Q \cap A^c).$$

Bilden wir das Supremum über n, so folgt mit Eigenschaft (c) von η

$$\eta(Q) \geq \sum_{j=1}^{\infty} \eta(Q \cap A_j) + \eta(Q \cap A^c) \geq \eta(Q \cap A) + \eta(Q \cap A^c).$$

Somit ist $A \in \mathcal{A}$, und wegen der Bemerkung vor dem Beweis sind die beiden äußeren Terme gleich. Insbesondere folgt für $Q = A$, dass

$$\sum_{j=1}^{\infty} \eta(A_j) = \eta(A).$$

d. h. $\eta|_{\mathcal{A}}$ ist σ-additiv.

(ii) Es sei $A \in \mathcal{R}$ und $Q \subseteq X$. Zur Berechnung von $\eta(Q)$ betrachten wir beliebige Mengen $B_n \in \mathcal{R}$ mit $Q \subseteq \bigcup_{n=1}^{\infty} B_n$. Dann ist $Q \cap A \subseteq \bigcup_{n=1}^{\infty}(B_n \cap A)$, und eine analoge Beziehung gilt für $Q \cap A^c$, so dass

$$\sum_{n=1}^{\infty} \mu(B_n) = \sum_{n=1}^{\infty} \big(\mu(B_n \cap A) + \mu(B_n \cap A^c)\big) \geq (Q \cap A) + \eta(Q \cap A^c).$$

Bilden wir das Infimum über alle solchen Folgen B_n, so erhalten wir auf der linken Seite $\eta(Q)$, und die η-Messbarkeit von A folgt. □

Die Eleganz von Carathéodorys Definition wird erkauft durch eine unendliche Zahl von Bedingungen für die Messbarkeit einer Menge. Man kann die Messbarkeit aber auch einfacher nachprüfen, wenn man (z. B. aus Satz 23) einen genügenden Vorrat an messbaren Mengen hat.

Lemma 18 *Es sei η ein äußeres Maß auf X und $A \subseteq X$. Gibt es zu jedem $\varepsilon > 0$ eine η-messbare Menge B mit $\eta(A \triangle B) < \varepsilon$, so ist A η-messbar.*

Beweis. Wegen $A \subseteq B \cup (A \triangle B)$ ist $\eta(A) \leq \eta(B) + \eta(A \triangle B)$, also gilt für jedes $Q \subseteq X$

$$\eta(Q \cap A) \leq \eta(Q \cap B) + \eta(A \triangle B),$$

und wegen $A^c \triangle B^c = A \triangle B$ gilt analog

$$\eta(Q \cap A^c) \leq \eta(Q \cap B^c) + \eta(A \triangle B).$$

In der Situation des Lemmas folgt durch Addition beider Ungleichungen

$$\eta(Q \cap A) + \eta(Q \cap A^c) \leq \eta(Q) + 2\varepsilon,$$

und da $\varepsilon > 0$ beleibig war, ist A η-messbar. □

Folgerung 8 (aus Satz 24) *Ist μ ein σ-additiver Inhalt auf einem Ring $\mathcal{R}$, η das zugehörige äußere Maß und $\mathcal{A}$ die σ-Algebra der η-messbaren Mengen, so ist $\lambda := \eta|_{\mathcal{A}}$ ein Maß, welches μ fortsetzt.*

Bemerkung. Ein Maß heißt vollständig, wenn jede Teilmenge einer Nullmenge zum Definitionsbereich des Maßes gehört. Das Maß λ ist nach Konstruktion vollständig, man nennt es die Lebesguesche Vervollständigung von μ. Man kann zeigen, dass jede vollständige Fortsetzung von μ eine Fortsetzung von λ ist, vorausgesetzt, μ ist σ-endlich.

Definition 25 *Ein Inhalt μ auf einem Ring über einer Menge X heißt σ-endlich, wenn es eine Folge von Teilmengen $X_n \subseteq X$ gibt, so dass $X = \bigcup_{n=1}^{\infty} X_n$ und $\mu(X_n) < \infty$ für alle n.*

Der elementargeometrische Inhalt ist z. B. σ-endlich, das Zählmaß auf einer überabzählbaren Menge ist es nicht.

Ohne die Bedingung der Vollständigkeit ist die Lebesguesche Vervollständigung nicht die einzige Fortsetzung auf die σ-Algebra $\mathcal{A}$. Der kleinstmögliche Definitionsbereich einer Fortsetzung von μ zu einem Maß ist die von $\mathcal{R}$ erzeugte σ-Algebra.

Satz 25 *Es sei $\mathcal{E} \subseteq \mathcal{P}(X)$ durchschnittsstabil, und β, γ seien Maße auf der von $\mathcal{E}$ erzeugten σ-Algebra $\mathcal{B}$, so dass $\beta|_{\mathcal{E}} = \gamma|_{\mathcal{E}}$. Gibt es eine Folge X_n in $\mathcal{E}$, so dass $X = \bigcup_{n=1}^{\infty} X_n$ und $\mu(X_n) < \infty$ für alle n, dann ist $\beta = \gamma$.*

Beweis. Wir betrachten zunächst den Fall $X \in \mathcal{E}$ und setzen

$$\mathcal{D} = \{A \in \mathcal{B} \mid \beta(A) = \gamma(A)\}.$$

Dann ist $X \in \mathcal{D}$, und für A, $B \in \mathcal{D}$ mit $B \subseteq A$ ist wegen $\beta(A) \leq \beta(X) < \infty$

$$\beta(A \setminus B) = \beta(A) - \beta(B) = \gamma(A) - \gamma(B) = \gamma(A \setminus B),$$

so dass $A \setminus B \in \mathcal{D}$. Schließlich gilt für jede Folge disjunkter Mengen $A_n \in \mathcal{D}$ wegen der σ-Additivität

$$\beta\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \beta(A_n) = \sum_{n=1}^{\infty} \gamma(A_n) = \gamma\left(\bigcup_{n=1}^{\infty} A_n\right),$$

also $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$. Somit ist $\mathcal{D}$ ein Dynkin-System.

Da $\mathcal{E}$ durchschnittsstabil ist, ist $\mathcal{B}$ gleich dem von $\mathcal{E}$ erzeugten Dynkin-System, und dieses ist in $\mathcal{D}$ enthalten. Also gilt $\beta(A) = \gamma(A)$ für alle $A \in \mathcal{B}$.

Im allgemeinen Fall ist $\mathcal{B}_n := \{A \in \mathcal{E} \mid A \subseteq X_n\}$ die von $\mathcal{E}_n := \{E \in \mathcal{E} \mid E \subseteq X_n\} \subseteq \mathcal{E}$ erzeugte σ-Algebra. Dei Einschränkungen β_n und γ_n von β und γ auf $\mathcal{B}_n$ sind Maße auf X_n und stimmen nach dem Bewiesenen überein. Für beliebiges $A \in \mathcal{B}$ folgt nun mit der σ-Additivität

$$\beta(A) = \sum_{n=1}^{\infty} \beta(A \cap X_n) = \sum_{n=1}^{\infty} \gamma(A \cap X_n) = \gamma(A).$$

□

Folgerung 9 *Ein σ-endlicher σ-additiver Inhalt μ auf einem Ring $\mathcal{R}$ hat genau eine Fortsetzung zu einem Maß β auf der von $\mathcal{R}$ erzeugten σ-Algebra $\mathcal{B}$.*

Zum Beweis der Existenz schränken wir einfach die Lebesguesche Vervollständigung auf $\mathcal{B}$ ein, die Eindeutigkeit folgt aus dem letzten Satz, da $\mathcal{R}$ durchschnittsstabil ist.

Beispiel. Ist μ der elementargeometrische Inhalt auf dem Ring $\mathcal{R}$ der beschränkten Polytope in $\mathbb{R}^n$, so nennt man die Lebesguesche Vervollständigung $\lambda : \mathcal{A} \to [0, \infty]$ von μ das Lebesgue-Maß auf $\mathbb{R}^n$. Offensichtlich gilt $\eta(A) \leq \mu^*(A)$, und man sieht leicht, dass die Einschränkung von λ auf den Ring $\mathcal{J}$ der Jordan-messbaren Mengen gleich dem Jordan-Inhalt ist. Die von $\mathcal{R}$ erzeugte σ-Algebra $\mathcal{B}$ ist nach Aufgabe 41 gleich der Algebra der Borel-Mengen. Man kann zeigen, dass für $n > 0$ sämtliche Inklusionen

$$\mathcal{R} \subset \mathcal{J} \subset \mathcal{A}, \qquad \mathcal{R} \subset \mathcal{B} \subset \mathcal{A} \subset \mathcal{P}(\mathbb{R}^n)$$

echt sind. Dies ist klar für die ersten drei Inklusion. Für die vierte zeigt man, dass $|\mathcal{B}| = |\mathbb{R}|$ und $|\mathcal{A}| = 2^{|\mathbb{R}|}$, für die letzte siehe unten.

Folgerung 10 *Ist $f : \mathbb{R}^n \to \mathbb{R}^n$ eine affine Abbildung, also $f(x) = Bx + v$, so gilt für jede Lebesgue-messbare Menge $A \subseteq \mathbb{R}^n$*

$$\lambda(f(A)) = |\det B|\, \lambda(A).$$

Insbesondere ist λ translationsinvariant.

Dies folgt aus Satz 18. Mit Satz 25 kann man leicht zeigen, dass $\beta = \lambda|_{\mathcal{B}}$ bis auf einen konstanten Faktor das einzige translationsinvariante Maß auf $\mathcal{B}$ ist.

Beispiel einer nicht Lebesgue-messbaren Menge $A \subset \mathbb{R}$. Nach dem Auswahlaxiom gibt es ein Vertretersystem A für den Quotientenraum $\mathbb{R}/\mathbb{Q}$, wobei wir annehmen können, dass $A \subset [0, 1)$. Dann ist $\mathbb{R}$ die disjunkte Vereinigung der Mengen $A + r$ mit $r \in \mathbb{Q}$. Wenn A messbar wäre, so müsste wegen σ-Additivität und Translationsinvarianz gelten

$$\infty = \lambda(\mathbb{R}) = \sum_{r \in \mathbb{Q}} \lambda(A + r) = \sum_{r \in \mathbb{Q}} \lambda(A),$$

also $\lambda(A) > 0$. Andererseits würde gelten

$$2 = \lambda([0, 2)) \geq \sum_{\substack{r \in \mathbb{Q} \\ 0 \leq r < 1}} \lambda(A + r) = \sum_{\substack{r \in \mathbb{Q} \\ 0 \leq r < 1}} \lambda(A)$$

(Widerspruch).

Satz 26 *Für jede Lebesgue-messbare Menge $A \subseteq \mathbb{R}^n$ und jedes $\varepsilon > 0$ gibt es eine offene Menge $U \supseteq A$ mit $\lambda(U \setminus A) < \varepsilon$ und eine abgeschlossene Menge $F \subseteq A$ mit $\lambda(A \setminus F) < \varepsilon$.*

Beweis. Zunächst sei $\lambda(A) < \infty$. Zu jedem vorgegebenen $\varepsilon > 0$ gibt es beschränkte Polytope P_j, so dass $A \subseteq \bigcup_{j=1}^{\infty} P_j$ und

$$\lambda(A) = \eta(A) > \sum_{j=1}^{\infty} \mu(P_j) - \varepsilon.$$

Nach Lemma 15 gibt es offene beschränkte Polytope $V_j \supseteq P_j$, so dass $\mu(V_j) \leq \mu(P_j) + 2^{-j}\varepsilon$. Nun ist $U := \bigcup_{j=1}^{\infty} V_j$ offen, enthält A, und

$$\lambda(U) \leq \sum_{j=1}^{\infty} \lambda(V_j) \leq \sum_{j=1}^{\infty} \mu(P_j) + \varepsilon < \lambda(A) + 2\varepsilon.$$

Ist $\lambda(A)$ beliebig, so betrachten wir $A_k = A \cap [-k, k]^n$. Nach dem Bewiesenen gibt es offene Mengen $U_k \supseteq A_k$ mit $\lambda(U_k \setminus A_k) < 2^{-k}\varepsilon$, und setzen wir $U = \bigcup_{k=1}^{\infty} U_k$, so folgt

$$\lambda(U \setminus A) \leq \sum_{k=1}^{\infty} \lambda(U_k \setminus A) \leq \sum_{k=1}^{\infty} 2^{-k}\varepsilon = \varepsilon.$$

Wenden wir dieses Ergebnis auf A^c an, so erhalten wir eine offene Menge $V \supseteq A^c$ mit $\lambda(V \setminus A^c) < \varepsilon$, und setzen wir $F = V^c$, so gilt $V \setminus A^c = A \setminus F$. $\square$