Konvexgeometrie

Vorlesung im WS 2004/05 (R. Schneider)

Konvexität ist ein Begriff, der in verschiedenen Zusammenhängen eine Rolle spielt und sich als nützlich erwiesen hat. Eine Punktmenge heißt *konvex*, wenn sie mit je zwei Punkten deren Verbindungsstrecke enthält. Dieser Begriff ist immer dann sinnvoll, wenn die Verbindungsstrecke definiert ist, also sicher für Teilmengen in beliebigen reellen Vektorräumen. Eine Funktion heißt *konvex*, wenn ihr Epigraph konvex ist; dies ist ein sinnvoller Begriff für reelle Funktionen, die auf Teilmengen reeller Vektorräume erklärt sind. Konvexe Mengen und konvexe Funktionen spielen in verschiedenen Teilgebieten der Funktionalanalysis eine wichtige Rolle. Konvexitätsbetrachtungen sind auch von Bedeutung in der Optimierungstheorie und in anderen anwendungsorientierten Gebieten, wie zum Beispiel in der Kontrolltheorie. Von diesen Dingen soll im folgenden nicht die Rede sein. Vielmehr geht es in dieser Vorlesung nur um die Geometrie konvexer Mengen im euklidischen Raum.

Konvexe Mengen sind ein geometrisch naheliegender und natürlicher Untersuchungsgegenstand. Wenn man unter „Geometrie“ zunächst einmal die Untersuchung der Gebilde im Anschauungsraum (und seinen höherdimensionalen Verallgemeinerungen) versteht, so ergibt sich die Frage, welche Gebilde überhaupt einer aussichtsreichen Untersuchung zugänglich sind. Das sind zunächst einmal die Gebilde der linearen Geometrie, wie Punkte, Geraden, Ebenen u.s.w. und die damit gebildeten Konfigurationen, ferner Gebilde zweiter Ordnung, wie Kegelschnitte und Quadriken. Geht man darüber hinaus, so muß man den Rahmen der zu betrachtenden Objekte einschränken, um inhaltsreiche Aussagen machen zu können. Verschiedene Möglichkeiten bieten sich an. In der algebraischen Geometrie beispielsweise beschränkt man sich auf Gebilde, die durch algebraische Gleichungen beschrieben werden können, in der Differentialgeometrie betrachtet man Kurven, Flächen und andere Untermannigfaltigkeiten, die eine Beschreibung durch genügend oft differenzierbare Funktionen zulassen. Eine weitere und vom Ansatz her besonders einfache Möglichkeit besteht nun darin, nur konvexe Mengen zu betrachten. Es ist erstaunlich, welcher Reichtum von geometrisch ansprechenden Ergebnissen allein unter Verwendung dieser einfachen Grundvoraussetzung erzielt werden kann.

In dieser Vorlesung werden wir uns auf einige Aspekte der Geometrie der konvexen Körper (der kompakten konvexen Mengen) beschränken. Nach den erforderlichen Grundlagen wollen wir vor allem Funktionale konvexer Körper wie Volumen und Oberfläche und deren Verallgemeinerungen betrachten und dafür Extremalprobleme behandeln.

Literaturhinweise

Die Vorlesung richtet sich überwiegend nach dem Buch

R. Schneider, *Convex Bodies: The Brunn-Minkowski Theory.* Cambridge University Press 1993.

Daraus werden zunächst Auszüge behandelt (zum teil wörtlich); später kommt auch anderer Stoff zur Sprache. An weiteren Einführungen in die Konvexgeometrie seien genannt:

R. Webster, *Convexity.* Oxford University Press 1994.

K. Leichtweiß, *Konvexe Mengen.* Springer, Berlin 1980.

Empfehlenswert wegen seiner reichhaltigen Einblicke in die Rolle der Konvexität in verschiedenen Gebieten ist das Buch

A. Barvinok, *A Course in Convexity.* Amer. Math. Soc., Providence, R.I. 2002.

Speziell über konvexe Polytope gibt es zwei Standardwerke:

B. Grünbaum, *Convex Polytopes.* Interscience, London 1967.

G.M. Ziegler, *Lectures on Polytopes.* Springer, New York 1995.

Inhalt (`vorläufig`)

Bezeichnungen

In diesem Vorspann sollen einige grundlegende Definitionen gegeben sowie Sprechweisen und Bezeichnungen festgelegt werden.

Wir arbeiten im n-dimensionalen reellen euklidischen Raum $\mathbb{E}^n$. Darunter verstehen wir den n-dimensionalen reellen Vektorraum $\mathbb{R}^n$, zusammen mit einem Skalarprodukt $\langle \cdot, \cdot \rangle$ und der induzierten euklidischen Norm $\| \cdot \|$. Wir machen keine formale Unterscheidung zwischen einem Vektorraum und dem zugehörigen affinen Raum. Die Elemente von $\mathbb{E}^n$, bezeichnet mit kleinen lateinischen Buchstaben, werden daher sowohl als Vektoren wie auch als Punkte bezeichnet. Welche dieser Bezeichnungen jeweils verwendet wird, ist aber nicht willkürlich; die Wahl der Bezeichnung soll andeuten, ob es sich um einen Sachverhalt der linearen Geometrie (Vektorraumgeometrie) oder der affinen Geometrie (Invarianz unter Translationen) handelt.

Der Vektor $x \in \mathbb{E}^n$ ist eine *lineare Kombination* der Vektoren $x_1, \dots, x_k \in \mathbb{E}^n$, wenn es Zahlen $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ gibt mit $x = \lambda_1 x_1 + \ldots + \lambda_k x_k$. Gibt es solche Zahlen mit $\lambda_1 + \ldots + \lambda_k = 1$, so ist der Punkt x eine *affine Kombination* der Punkte $x_1, \dots, x_k$. Für $A \subset \mathbb{E}^n$ ist $\operatorname{lin} A$ ($\operatorname{aff} A$) die *lineare Hülle* (*affine Hülle*) von A; das ist die Menge aller linearen (affinen) Kombinationen von Elementen von A und zugleich der kleinste lineare (affine) Unterraum von $\mathbb{E}^n$, der A enthält. Die Punkte $x_1, \dots, x_k \in \mathbb{E}^n$ sind *affin unabhängig* wenn keiner von ihnen eine affine Kombination der übrigen ist, also wenn aus

$$\sum_{i=1}^{k} \lambda_i x_i = 0 \quad \text{mit } \lambda_i \in \mathbb{R} \text{ und } \sum_{i=1}^{k} \lambda_i = 0$$

folgt, dass $\lambda_1 = \ldots = \lambda_k = 0$ ist. Dies ist äquivalent mit der linearen Unabhängigkeit der Vektoren $x_2 - x_1, \dots, x_k - x_1$. Erklären wir die Abbildung $\tau : \mathbb{E}^n \to \mathbb{E}^n \times \mathbb{R}$ durch $\tau(x) := (x, 1)$, so sind $x_1, \dots, x_k \in \mathbb{E}^n$ genau dann affin unabhängig, wenn $\tau(x_1), \dots, \tau(x_k)$ linear unabhängig sind.

Für $x, y \in \mathbb{E}^n$ ist

$$[x, y] := \{(1 - \lambda)x + \lambda y : 0 \le \lambda \le 1\}$$

die *abgeschlossene Strecke* mit Endpunkten x und y, und

$$\begin{aligned} [x, y) &:= \{(1 - \lambda)x + \lambda y : 0 \le \lambda < 1\}, \\ (x, y] &:= \{(1 - \lambda)x + \lambda y : 0 < \lambda \le 1\} \end{aligned}$$

sind halboffene Strecken.

Für Teilmengen $A, B \subset \mathbb{E}^n$ und für $\lambda \in \mathbb{R}$ definiert man

$$\begin{aligned} A + B &:= \{a + b : a \in A, b \in B\}, \\ \lambda A &:= \{\lambda a : a \in A\}. \end{aligned}$$

Die Menge $A + B$ heißt die *Vektorsumme* oder *Minkowski-Summe* der Mengen A und B. Zur Abkürzung schreibt man $(-1)A =: -A$, $A + (-B) =: A - B$ und $A + \{x\} =: A + x$. Sind die Mengen A und B in komplementären affinen Unterräumen von $\mathbb{E}$ enthalten, so wird die Vektorsumme $A+B$ auch in der Form $A \oplus B$ geschrieben und als die *direkte Summe* von A und B bezeichnet.

An topologischen Bezeichnungen verwenden wir $\operatorname{clos} A$ für die abgeschlossene Hülle, $\operatorname{int} A$ für das Innere, $\operatorname{bd} A$ für den Rand einer Teilmenge A von $\mathbb{E}^n$. Das Innere und der Rand einer Menge $A \subset \mathbb{E}^n$ bezüglich der affinen Hülle von A werden mit $\operatorname{relint} A$ bzw. $\operatorname{relbd} A$ bezeichnet.

Eine *Hyperebene* in $\mathbb{E}^n$ ist ein affiner Unterraum der Kodimension 1. Eine Hyperebene kann in der Form

$$H_{u,\alpha} = \{x \in \mathbb{E}^n : \langle x, u \rangle = \alpha\}$$

mit $u \in \mathbb{E}^n \setminus \{0\}$ und $\alpha \in \mathbb{R}$ dargestellt werden. Es gilt $H_{u,\alpha} = H_{v,\beta}$ genau dann, wenn $(v, \beta) = (\lambda u, \lambda\alpha)$ mit einem $\lambda \neq 0$ gilt. Jeden Vektor u mit $H = H_{u,\alpha}$ bezeichnen wir als einen *Normalenvektor* der Hyperebene H. Die Hyperebene $H_{u,\alpha}$ berandet die zwei abgeschlossenen Halbräume

$$\begin{aligned} H^-_{u,\alpha} &:= \{x \in \mathbb{E}^n : \langle x, u \rangle \leq \alpha\}, \\ H^+_{u,\alpha} &:= \{x \in \mathbb{E}^n : \langle x, u \rangle \geq \alpha\}. \end{aligned}$$

Einen affinen Unterraum des $\mathbb{E}^n$ bezeichnen wir auch als *Ebene*, und der Durchschnitt einer Ebene mit einem abgeschlossenen Halbraum, der die Ebene trifft, aber nicht ganz enthält, heißt *Halbebene*. Eine eindimensionale Ebene wird als *Gerade* bezeichnet und eine Halbgerade auch als *Strahl*.

Die folgenden Begriffsbildungen verwenden die euklidische Metrik. Für $x, y \in \mathbb{E}^n$ ist $\|x - y\|$ der *Abstand* von x und y, und für $\emptyset \neq A \subset \mathbb{E}^n$ und $x \in \mathbb{E}^n$ ist

$$d(A, x) := \inf \{\|x - a\| : a \in A\}$$

der *Abstand* des Punktes x von der Menge A. Für eine nichtleere beschränkte Menge $A \subset \mathbb{E}^n$ ist die Zahl

$$\operatorname{diam} A := \sup\{\|x - y\| : x, y \in A\}$$

der *Durchmesser* von A. Durch

$$B(z, \rho) := \{x \in \mathbb{E}^n : \|x - z\| \leq \rho\}$$

wird die *abgeschlossene Kugel* und durch

$$B_0(z, \rho) := \{x \in \mathbb{E}^n : \|x - z\| < \rho\}$$

die *offene Kugel* mit Mittelpunkt $z \in \mathbb{E}^n$ und Radius $\rho > 0$ definiert. Wir schreiben kurz $B(0,1) =: B^n$ und nennen dies die *Einheitskugel* des $\mathbb{E}^n$. Der Rand

$$S^{n-1} := \{x \in \mathbb{E}^n : \|x\| = 1\}$$

der Einheitskugel ist die *Einheitssphäre*.

Lineare Abbildungen, affine Abbildungen und Isometrien zwischen euklidischen Räumen sind aus der Linearen Algebra wohlbekannt. Insbesondere ist eine *Translation* $\varphi : \mathbb{E}^n \to \mathbb{E}^n$ eine Abbildung der Form $\varphi(x) = x + t$, $x \in \mathbb{E}^n$, mit einem festen Vektor $t \in \mathbb{E}^n$, der dann als *Translationsvektor* bezeichnet wird. Die Menge $A + t$ wird als ein *Translat* von A bezeichnet. Eine *Homothetie* $\varphi : \mathbb{E}^n \to \mathbb{E}^n$ ist eine Abbildung der Form $\varphi(x) = \lambda x + t$, $x \in \mathbb{E}^n$, mit festen $\lambda > 0$ und $t \in \mathbb{E}^n$. Die Menge $\lambda A + t$ heißt ein *Homothet* von A. Die Mengen $A, B \subset \mathbb{E}^n$ heißen *positiv homothetisch*, wenn $A = \lambda B + t$ mit $\lambda > 0$ und $t \in \mathbb{E}^n$ gilt (positive Homothetie ist offenbar eine Äquivalenzrelation), und sie heißen *homothetisch*, wenn sie entweder positiv homothetisch sind oder eine der beiden Mengen einpunktig ist. Eine *Bewegung* des $\mathbb{E}^n$ ist eine isometrische Abbildung von $\mathbb{E}^n$ auf sich, und eine *Drehung* ist eine Bewegung mit Fixpunkt 0. Die Drehungen sind genau die linearen Abbildungen des $\mathbb{E}^n$ auf sich, die das Skalarprodukt erhalten. Jede Bewegung ist die Komposition einer Drehung und einer Translation. Eine Bewegung heißt *eigentlich*, wenn sie die Orientierung des $\mathbb{E}^n$ erhält, und andernfalls *uneigentlich*. Eine *Ähnlichkeit* ist die Komposition einer Bewegung und einer *Streckung*, das heißt einer Abbildung der Form $x \mapsto \lambda x$ mit $\lambda > 0$.

Kapitel 1

Grundlagen der Konvexgeometrie

1.1 Konvexe Mengen und Kombinationen

Eine Teilmenge $A \subset \mathbb{E}^n$ heißt *konvex*, wenn sie mit je zwei Punkten x, y auch stets deren Verbindungsstrecke $[x, y]$ enthält. Die Menge A ist also genau dann konvex, wenn gilt:

$$\forall\, x, y \in A\; \forall\, \lambda \in [0,1] : (1-\lambda)x + \lambda y \in A.$$

Triviale Beispiele von konvexen Mengen im $\mathbb{E}^n$ sind $\emptyset, \mathbb{E}^n$, affine Unterräume, die abgeschlossenen Kugeln $B(z, \rho)$. Ein weniger triviales Beispiel ist die Vereinigung $B_0(z, \rho) \cup A$ einer offenen Kugel $B_0(z, \rho)$ und einer beliebigen Teilmenge A ihres Randes. Später werden wir aber überwiegend abgeschlossene konvexe Mengen betrachten. Unmittelbar aus der Definition leitet man her: Durchschnitte konvexer Mengen sind konvex, Bilder und Urbilder konvexer Mengen unter affinen Abbildungen sind konvex; wenn A und B konvex sind, so auch $A + B$ und λA für $\lambda \in \mathbb{R}$.

1.1.1 Bemerkung. Für $A \subset \mathbb{E}^n$ und $\lambda, \mu > 0$ gilt trivialerweise $\lambda A + \mu A \supset (\lambda + \mu)A$. Die Gleichheit $\lambda A + \mu A = (\lambda + \mu)A$ (für alle $\lambda, \mu > 0$) gilt genau dann, wenn A konvex ist. In der Tat, sei A konvex und $x \in \lambda A + \mu A$. Dann ist $x = \lambda a + \mu b$ mit $a, b \in A$, und es folgt

$$x = (\lambda + \mu)\left(\frac{\lambda}{\lambda + \mu} a + \frac{\mu}{\lambda + \mu} b\right) \in (\lambda + \mu)A,$$

also ist $\lambda A + \mu A = (\lambda + \mu)A$. Die umgekehrte Schlussrichtung ist trivial.

Eine Teilmenge $A \subset \mathbb{E}^n$ heißt *konvexer Kegel*, wenn A konvex und nichtleer ist und wenn mit $x \in A$, $\lambda \geq 0$ stets $\lambda x \in A$ gilt. Eine nichtleere Teilmenge $A \subset \mathbb{E}^n$

ist also genau dann ein konvexer Kegel, wenn sie abgeschlossen ist gegenüber Addition und Multiplikation mit nichtnegativen reellen Zahlen.

Indem wir die Definitionen von affinen und linearen Kombinationen auf nichtnegative Koeffizienten beschränken, erhalten wir die folgenden grundlegenden Begriffsbildungen. Der Punkt $x \in \mathbb{E}^n$ ist eine *konvexe Kombination* der Punkte $x_1, \dots, x_k \in \mathbb{E}^n$, wenn es Zahlen $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ gibt mit

$$x = \sum_{i=1}^{k} \lambda_i x_i, \quad \lambda_i \geq 0 \ \ (i = 1, \dots, k), \quad \sum_{i=1}^{k} \lambda_i = 1.$$

Analog ist der Vektor $x \in \mathbb{E}^n$ eine *positive Kombination* der Vektoren $x_1, \dots, x_k \in \mathbb{E}^n$, wenn es Zahlen $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ gibt mit

$$x = \sum_{i=1}^{k} \lambda_i x_i, \quad \lambda_i \geq 0 \ \ (i = 1, \dots, k).$$

Man beachte, dass man von einer positiven Kombination spricht, obwohl zugelassen ist, dass die Koeffizienten Null sind.

Für eine Teilmenge $A \subset \mathbb{E}^n$ ist die *konvexe Hülle* $\operatorname{conv} A$ (die *positive Hülle* $\operatorname{pos} A$) definiert als die Menge aller konvexen Kombinationen (aller positiven Kombinationen) von je endlich vielen Elementen von A.

1.1.2 Satz. *Für $A \subset \mathbb{E}^n$ ist $\operatorname{conv} A$ konvex. Ist $A \subset \mathbb{E}^n$ konvex, so gilt $\operatorname{conv} A = A$. Für eine beliebige Teilmenge $A \subset \mathbb{E}^n$ ist $\operatorname{conv} A$ der Durchschnitt aller konvexen Teilmengen von $\mathbb{E}^n$, die A enthalten. Für $A, B \subset \mathbb{E}^n$ gilt $\operatorname{conv}(A + B) = \operatorname{conv} A + \operatorname{conv} B$.*

Beweis. Dass $\operatorname{conv} A$ stets konvex ist, folgt unmittelbar aus den Definitionen. Sei $A \subset \mathbb{E}^n$ konvex. Trivialerweise gilt $A \subset \operatorname{conv} A$. Wir zeigen durch Induktion, dass A alle konvexen Kombinationen von je k Punkten von A enthält. Für $k = 2$ gilt das nach Definition der Konvexität. Sei $k > 2$, die Behauptung bewiesen für je $k - 1$ Punkte, und sei $x = \lambda_1 x_1 + \ldots + \lambda_k x_k$ mit $x_1, \dots, x_k \in A$, $\lambda_1, \dots, \lambda_k > 0$ (o.B.d.A.) und $\lambda_1 + \ldots + \lambda_k = 1$. Dann gilt

$$x = (1 - \lambda_k) \sum_{i=1}^{k-1} \frac{\lambda_i}{1 - \lambda_k} x_i + \lambda_k x_k \in A,$$

denn wegen

$$\frac{\lambda_i}{1 - \lambda_k} > 0, \quad \sum_{i=1}^{k-1} \frac{\lambda_i}{1 - \lambda_k} = 1$$

ist

$$\sum_{i=1}^{k-1} \frac{\lambda_i}{1 - \lambda_k} x_i \in A$$

nach Induktionsannahme. Damit ist $A = \operatorname{conv} A$ gezeigt.

Für beliebiges $A \subset \mathbb{E}^n$ bezeichne $D(A)$ den Durchschnitt der Familie aller konvexen Teilmengen $K \subset \mathbb{E}^n$ mit $A \subset K$. Da $A \subset \operatorname{conv} A$ gilt und $\operatorname{conv} A$ konvex ist, gilt $D(A) \subset \operatorname{conv} A$. Jede konvexe Menge K mit $A \subset K \subset \mathbb{E}^n$ erfüllt $\operatorname{conv} A \subset \operatorname{conv} K = K$, also ist $\operatorname{conv} A \subset D(A)$. Daraus folgt die Gleichheit $\operatorname{conv} A = D(A)$.

Sei $A, B \subset \mathbb{E}^n$. Sei $x \in \operatorname{conv}(A + B)$, also

$$x = \sum_{i=1}^{k} \lambda_i (a_i + b_i) \quad \text{mit } a_i \in A, b_i \in B, \quad \lambda_i \geq 0, \quad \sum_{i=1}^{k} \lambda_i = 1$$

und daher $x = \sum \lambda_i a_i + \sum \lambda_i b_i \in \operatorname{conv} A + \operatorname{conv} B$. Umgekehrt sei $x \in \operatorname{conv} A + \operatorname{conv} B$, also

$$x = \sum_i \lambda_i a_i + \sum_j \mu_j b_j$$

mit $a_i \in A$, $b_j \in B$, $\lambda_i, \mu_j \geq 0$ und $\sum \lambda_i = \sum \mu_j = 1$. Aus

$$x = \sum_{i,j} \lambda_i \mu_j (a_i + b_j)$$

folgt jetzt $x \in \operatorname{conv}(A + B)$. ■

1.1.3 Satz. *Für $\emptyset \neq A \subset \mathbb{E}^n$ ist* $\operatorname{pos} A$ *ein konvexer Kegel. Ist $A \subset \mathbb{E}^n$ ein konvexer Kegel, so gilt* $\operatorname{pos} A = A$. *Für eine beliebige nichtleere Teilmenge $A \subset \mathbb{E}^n$ ist* $\operatorname{pos} A$ *der Durchschnitt aller konvexen Kegel in $\mathbb{E}^n$, die A enthalten.*

Der Beweis ist völlig analog zu dem von Satz 1.1.2.

Wir kommen nun zu drei einfachen, aber grundlegenden Aussagen der kombinatorischen Konvexgeometrie, den Sätzen von Carathéodory, Radon und Helly.

1.1.4 Satz (von Carathéodory). *Ist $A \subset \mathbb{E}^n$ und $x \in \operatorname{conv} A$, so ist x konvexe Kombination von affin unabhängigen Punkten von A. Insbesondere ist x konvexe Kombination von $n + 1$ oder weniger Punkten von A.*

Beweis. Der Punkt $x \in \operatorname{conv} A$ hat eine Darstellung

$$x = \sum_{i=1}^{k} \lambda_i x_i \quad \text{mit } x_i \in A, \quad \lambda_i > 0, \quad \sum_{i=1}^{k} \lambda_i = 1$$

mit einem $k \in \mathbb{N}$, und wir können annehmen, dass hierbei k minimal gewählt ist. Angenommen, $x_1, \ldots, x_k$ wären affin abhängig. Dann gibt es Zahlen $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, die nicht alle Null sind, so dass

$$\sum_{i=1}^{k} \alpha_i x_i = 0 \quad \text{und} \quad \sum_{i=1}^{k} \alpha_i = 0$$

gilt. Wir können eine Zahl m wählen derart, dass λ_m/α_m positiv und dabei so klein wie möglich ist (denn alle λ_i sind positiv, und wenigstens ein α_i ist positiv). In der affinen Darstellung

$$x = \sum_{i=1}^{k} \left(\lambda_i - \frac{\lambda_m}{\alpha_m} \alpha_i \right) x_i$$

sind alle Koeffizienten nichtnegativ (trivialerweise, wenn $\alpha_i \leq 0$ ist, und andernfalls wegen der Wahl von m), und wenigstens ein Koeffizient ist Null. Dies ist ein Widerspruch zur Minimalität von k. Also sind $x_1, \ldots, x_k$ affin unabhängig, woraus insbesondere $k \leq n+1$ folgt. ■

Die konvexe Hülle von endlich vielen Punkten wird als *Polytop* bezeichnet. Ein *k-Simplex* ist die konvexe Hülle von $k+1$ affin unabhängigen Punkten, und diese Punkte heißen dann die *Ecken* des Simplex. Man kann den Satz von Carathéodory also auch so formulieren, dass $\operatorname{conv} A$ die Vereinigung aller Simplizes mit Ecken in A ist.

Ein weiterer einfacher, aber wichtiger Satz über konvexe Hüllen ist der Satz von Radon.

1.1.5 Satz (von Radon). *Jede nichtleere Menge von affin abhängigen Punkten (insbesondere also jede Menge von mindestens $n+2$ Punkten) im $\mathbb{E}^n$ kann zerlegt werden in zwei Mengen, deren konvexe Hüllen nichtleeren Durchschnitt haben.*

Beweis. Seien $x_1, \ldots, x_k \in \mathbb{E}^n$ affin abhängig. Dann gibt es Zahlen $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, die nicht alle Null sind, so dass

$$\sum_{i=1}^{k} \alpha_i x_i = 0 \quad \text{und} \quad \sum_{i=1}^{k} \alpha_i = 0$$

gilt. Nach eventueller Umnumerierung können wir annehmen, dass $\alpha_i > 0$ genau für $i = 1, \ldots, j$ gilt; dann ist $1 \leq j < k$ (denn wenigstens ein α_i ist > 0, aber nicht alle α_i sind > 0). Mit

$$\alpha := \alpha_1 + \ldots + \alpha_j = -(\alpha_{j+1} + \ldots + \alpha_k) > 0$$

gilt

$$x := \sum_{i=1}^{j} \frac{\alpha_i}{\alpha} x_i = \sum_{i=j+1}^{k} \left(-\frac{\alpha_i}{\alpha}\right) x_i$$

und daher $x \in \operatorname{conv}\{x_1, \ldots, x_j\} \cap \operatorname{conv}\{x_{j+1}, \ldots, x_k\}$. Daraus folgt die Behauptung. ■

Aus dem Satz von Radon kann man den nachfolgenden Satz von Helly herleiten, ein besonders hübsches Resultat der kombinatorischen Konvexgeometrie.

1.1.6 Satz (von Helly). *Seien $A_1, \ldots, A_k \subset \mathbb{E}^n$ konvexe Mengen. Haben je $n+1$ dieser Mengen nichtleeren Durchschnitt, so haben alle dieser Mengen nichtleeren Durchschnitt.*

Beweis. Sei $k \geq n+2$ (für $k < n+1$ ist nichts zu zeigen, und für $k = n+1$ ist die Behauptung trivial), und sei die Behauptung bewiesen für $k-1$ konvexe Mengen. Für jedes $i \in \{1, \ldots, k\}$ gibt es dann einen Punkt

$$x_i \in A_1 \cap \ldots \cap \check{A}_i \cap \ldots \cap A_k,$$

wobei $\check{A}_i$ bedeutet, dass die Menge A_i weggelassen wird. Die $k \geq n+2$ Punkte $x_1, \ldots, x_k$ sind affin abhängig. Nach dem Satz von Radon können wir daher nach geeigneter Umnumerierung annehmen, dass es einen Punkt

$$x \in \operatorname{conv}\{x_1, \ldots, x_j\} \cap \operatorname{conv}\{x_{j+1}, \ldots, x_k\}$$

gibt, wo $j \in \{1, \ldots, k-1\}$ geeignet gewählt ist. Wegen $x_1, \ldots, x_j \in A_{j+1}, \ldots, A_k$ gilt

$$x \in \operatorname{conv}\{x_1, \ldots, x_j\} \subset A_{j+1}, \ldots, A_k$$

und analog

$$x \in \operatorname{conv}\{x_{j+1}, \ldots, x_k\} \subset A_1, \ldots, A_j,$$

also $x \in A_1 \cap \ldots \cap A_k$. ■

Aus dem Satz von Helly kann man ähnliche Aussagen der kombinatorischen Geometrie herleiten; hierfür ein kleines Beispiel. Wir sagen, dass die Mengen $A_1, \ldots, A_m$ von K getroffen werden, wenn $A_i \cap K \neq \emptyset$ für $i = 1, \ldots, m$ gilt.

1.1.7 Satz. *Sei $\mathcal{M}$ eine endliche Familie konvexer Mengen im $\mathbb{E}^n$, sei $K \subset \mathbb{E}^n$ konvex. Werden je $n+1$ Elemente von $\mathcal{M}$ von einem Translat von K getroffen, so werden alle Elemente von $\mathcal{M}$ von einem Translat von K getroffen.*

Beweis. Sei $\mathcal{M} = \{A_1, \ldots, A_k\}$. Zu je $n+1$ der Zahlen $\{1, \ldots, k\}$, etwa zu $1, \ldots, n+1$, gibt es ein $t \in \mathbb{E}^n$ und Punkte $a_i \in A_i \cap (K+t)$. Es ist also $a_i = k_i + t$

mit $a_i \in A$ und $k_i \in K$. Es folgt $t = a_i - k_i \in A_i - K$ für $i = 1, \dots, n+1$. Also haben je $n+1$ Elemente der Familie $\{A_1 - K, \dots, A_k - K\}$ nichtleeren Durchschnitt. Da $A_i - K$ konvex ist $(i = 1, \dots, k)$, haben nach dem Satz von Helly alle Elemente dieser Familie nichtleeren Durchschnitt. Es gibt also ein $t \in \mathbb{E}^n$ mit $t \in A_i - K$, also $A_i \cap (K + t) \neq \emptyset$, für $i = 1, \dots, k$. ∎

Wir wenden uns nun den Beziehungen zwischen Konvexität und topologischen Eigenschaften zu und beginnen mit einer einfachen Beobachtung.

1.1.8 Hilfssatz. *Sei $A \subset \mathbb{E}^n$ konvex. Ist $x \in \operatorname{int} A$ und $y \in \operatorname{clos} A$, so gilt $[x, y) \subset \operatorname{int} A$.*

Beweis. Sei $z = (1 - \lambda)y + \lambda x$ mit $0 < \lambda < 1$. Wegen $x \in \operatorname{int} A$ gilt $B(x, \rho) \subset A$ für geeignetes $\rho > 0$. Setze $B(0, \rho) =: U$. Zunächst werde $y \in A$ angenommen. Sei $w \in \lambda U + z$, also $w = \lambda u + z$ mit $u \in U$. Wegen $x + u \in A$ gilt $w = (1 - \lambda)y + \lambda(x + u) \in A$. Es ist also $\lambda U + z \subset A$, daher ist $z \in \operatorname{int} A$.

Nun gelte lediglich $y \in \operatorname{clos} A$. Setze $V := [\lambda/(1-\lambda)]U + y$. Es gibt ein $a \in A \cap V$. Wegen $a = [\lambda/(1 - \lambda)]u + y$ mit $u \in U$ ist $z = (1 - \lambda)a + \lambda(x - u) \in A$. Damit ist $[x, y) \subset A$ gezeigt. Zusammen mit dem ersten Teil liefert das $[x, y) \subset \operatorname{int} A$. ∎

1.1.9 Satz. *Ist $A \subset \mathbb{E}^n$ konvex, so sind $\operatorname{int} A$ und $\operatorname{clos} A$ konvex. Ist $A \subset \mathbb{E}^n$ offen, so ist $\operatorname{conv} A$ offen.*

Beweis. Die Konvexität von $\operatorname{int} A$ folgt aus Hilfssatz 1.1.8. Die Konvexität von $\operatorname{clos} A$ für konvexes A und die Offenheit von $\operatorname{conv} A$ für offenes A sind einfache Übungsaufgaben. ∎

Man beachte, dass die konvexe Hülle einer abgeschlossenen Menge nicht abgeschlossen zu sein braucht. Dies zeigt das Beispiel einer Menge, die aus einer Geraden und einem nicht auf ihr liegenden Punkt besteht. Die konvexe Hülle einer kompakten (beschränkten und abgeschlossenen) Menge im $\mathbb{E}^n$ ist jedoch stets abgeschlossen. Dies kann man mit Hilfe des Satzes von Carathéodory zeigen:

1.1.10 Satz. *Für $A \subset \mathbb{E}^n$ gilt $\operatorname{conv} \operatorname{clos} A \subset \operatorname{clos} \operatorname{conv} A$. Ist A beschränkt, so gilt $\operatorname{conv} \operatorname{clos} A = \operatorname{clos} \operatorname{conv} A$. Insbesondere ist die konvexe Hülle einer kompakten Menge stets kompakt.*

Beweis. Sei $x \in \operatorname{conv} \operatorname{clos} A$. Dann gibt es eine Darstellung

$$x = \sum_{i=1}^{k} \lambda_i x_i \quad \text{mit } \lambda_i \geq 0, \quad x_i \in \operatorname{clos} A, \quad \sum_{i=1}^{k} \lambda_i = 1.$$

Sei $\epsilon > 0$. Zu $i \in \{1, \dots, k\}$ existiert $a_i \in A$ mit $\|x_i - a_i\| < \epsilon$. Setze $a := \sum_{i=1}^{k} \lambda_i a_i$. Dann gilt $a \in \operatorname{conv} A$ und $\|x - a\| \leq \sum \lambda_i \|x_i - a_i\| < \epsilon$. Da $\epsilon > 0$ beliebig war, ist $x \in \operatorname{clos} \operatorname{conv} A$.

Nun sei $A \subset \mathbb{E}^n$ beschränkt. Dann ist

$$\{(\lambda_1, \dots, \lambda_{n+1}, x_1, \dots, x_{n+1}) : \lambda_i \geq 0,\ x_i \in \operatorname{clos} A,\ \sum_{i=1}^{n+1} \lambda_i = 1\}$$

eine kompakte Teilmenge von $\mathbb{R}^{n+1} \times (\mathbb{E}^n)^{n+1}$; das Bild dieser Menge unter der stetigen Abbildung

$$(\lambda_1, \dots, \lambda_{n+1}, x_1, \dots, x_{n+1}) \mapsto \sum_{i=1}^{n+1} \lambda_i x_i \in \mathbb{E}^n$$

ist also kompakt. Nach dem Satz von Carathéodory stimmt dieses Bild überein mit $\operatorname{conv} \operatorname{clos} A$. Daher gilt $\operatorname{clos} \operatorname{conv} A \subset \operatorname{clos} \operatorname{conv} \operatorname{clos} A = \operatorname{conv} \operatorname{clos} A$. ■

Die Menge $\operatorname{clos} \operatorname{conv} A$, die nach 1.1.9 konvex ist, heißt die *abgeschlossene konvexe Hülle* von A. Dies ist zugleich der Durchschnitt aller abgeschlossenen konvexen Teilmengen von $\mathbb{E}^n$, die A enthalten.

Um Informationen über das relative Innere einer konvexen Menge zu erhalten, betrachten wir nun zunächst Simplizes.

1.1.11 Hilfssatz. *Seien $x_1, \dots, x_k \in \mathbb{E}^n$ affin unabhängig, sei $S := \operatorname{conv} \{x_1, \dots, x_k\}$ und $x \in \operatorname{aff} S$. Dann gilt $x \in \operatorname{relint} S$ genau dann, wenn in der eindeutigen affinen Darstellung*

$$x = \sum_{i=1}^{k} \lambda_i x_i \quad \textit{mit} \quad \sum_{i=1}^{k} \lambda_i = 1$$

alle Koeffizienten λ_i positiv sind.

Beweis. Wir können ohne Beschränkung der Allgemeinheit $k = n+1$ annehmen. Die Bedingung ist notwendig, denn andernfalls würde, weil die Darstellung eindeutig ist, jede Umgebung von x Punkte enthalten, die nicht zu S gehören. Um zu zeigen, dass die Bedingung hinreichend ist, setzen wir voraus, dass x wie oben dargestellt ist mit $\lambda_i > 0$ für $i = 1, \dots, n+1$. Da $x_1, \dots, x_{n+1}$ affin unabhängig sind, sind die Vektoren $(x_1, 1), \dots, (x_{n+1}, 1)$ in $\mathbb{E}^n \times \mathbb{R}$ linear unabhängig, bilden also eine Basis. Für $y \in \mathbb{E}^n$ sind die Koeffizienten $\mu_1, \dots, \mu_{n+1}$ in der affinen Darstellung

$$y = \sum_{i=1}^{n+1} \mu_i x_i \quad \text{mit} \quad \sum_{i=1}^{n+1} \mu_i = 1$$

gerade die Koordinaten von $(y, 1) \in \mathbb{E}^n \times \mathbb{R}$ bezüglich dieser Basis. Da Koordinatenfunktionen stetig sind, hängen die Koeffizienten $\mu_1, \dots, \mu_{n+1}$ stetig von y ab. Daher kann eine Zahl $\delta > 0$ so gewählt werden, dass für alle $y \in \mathbb{E}^n$ mit $\|y - x\| < \delta$ die Ungleichungen $\mu_i > 0$ $(i = 1, \dots, n+1)$ erfüllt sind und daher $y \in S$ gilt. Es ist also $x \in \operatorname{int} S$. ■

1.1.12 Satz. *Ist $A \subset \mathbb{E}^n$ konvex und nichtleer, so gilt* $\operatorname{relint} A \neq \emptyset$.

Beweis. Sei $\dim \operatorname{aff} A = k$. Dann gibt es $k+1$ affin unabhängige Punkte in A. Ihre konvexe Hülle S erfüllt $\operatorname{relint} S \neq \emptyset$ nach Hilfssatz 1.1.11, ferner gilt $S \subset A$ und $\operatorname{aff} S = \operatorname{aff} A$. Damit folgt die Behauptung. ■

In Anbetracht von 1.1.12 ist es sinnvoll, als *Dimension* einer konvexen Menge A, bezeichnet mit $\dim A$, die Dimension der affinen Hülle von A zu definieren.

Die Darstellung von $\operatorname{relint} \operatorname{conv} A$, die durch Hilfssatz 1.1.11 für affin unabhängige Mengen A gegeben wird, kann auf beliebige endliche Mengen ausgedehnt werden.

1.1.13 Satz. *Seien $x_1, \dots, x_k \in \mathbb{E}^n$, sei $P := \operatorname{conv} \{x_1, \dots, x_k\}$ und $x \in \mathbb{E}^n$. Dann gilt $x \in \operatorname{relint} P$ genau dann, wenn es eine Darstellung*

$$x = \sum_{i=1}^{k} \lambda_i x_i \quad \textit{mit } \lambda_i > 0 \ \ (i = 1, \dots, k), \quad \sum_{i=1}^{k} \lambda_i = 1$$

gibt.

Beweis. Ohne Beschränkung der Allgemeinheit können wir $\dim P = n$ annehmen. Sei $x \in \operatorname{int} P$. Setze

$$y := \sum_{i=1}^{k} \frac{1}{k} x_i.$$

Dann ist $y \in P$. Da $x \in \operatorname{int} P$ ist, können wir ein $z \in P$ wählen mit $x \in [y, z)$. Es gibt Darstellungen

$$\begin{aligned} z &= \sum_{i=1}^{k} \mu_i x_i \quad \text{mit } \mu_i \geq 0, \quad \sum_{i=1}^{k} \mu_i = 1, \\ x &= (1 - \lambda) y + \lambda z \quad \text{mit } 0 \leq \lambda < 1. \end{aligned}$$

Hieraus folgt

$$x = \sum_{i=1}^{k} \lambda_i x_i \quad \text{mit } \lambda_i = (1 - \lambda) \frac{1}{k} + \lambda \mu_i > 0, \quad \sum_{i=1}^{k} \lambda_i = 1;$$

das ist eine Darstellung der gewünschten Art.

Umgekehrt gelte

$$x = \sum_{i=1}^{k} \lambda_i x_i \quad \text{mit } \lambda_i > 0, \quad \sum_{i=1}^{k} \lambda_i = 1.$$

Wir können o.B.d.A. annehmen, dass $x_1, \ldots, x_{n+1}$ affin unabhängig sind. Setze $\lambda := \lambda_1 + \ldots + \lambda_{n+1}$ und

$$y := \sum_{i=1}^{n+1} \frac{\lambda_i}{\lambda} x_i.$$

Aus Hilfssatz 1.1.11 folgt

$$y \in \operatorname{int} \operatorname{conv} \{x_1, \ldots, x_{n+1}\} \subset \operatorname{int} P.$$

Ist $k = n + 1$, so ist $x = y \in \operatorname{int} P$. Andernfalls setze

$$z := \sum_{i=n+2}^{k} \frac{\lambda_i}{1 - \lambda} x_i.$$

Dann ist $z \in P$ und $x \in [y, z) \subset \operatorname{int} P$ nach Hilfssatz 1.1.8. ■

1.1.14 Satz. *Sei $A \subset \mathbb{E}^n$ konvex. Dann gilt*

(a) $\operatorname{relint} A = \operatorname{relint} \operatorname{clos} A$,
(b) $\operatorname{clos} A = \operatorname{clos} \operatorname{relint} A$,
(c) $\operatorname{relbd} A = \operatorname{relbd} \operatorname{clos} A = \operatorname{relbd} \operatorname{relint} A$.

Beweis. Wir können o.B.d.A. $\dim A = n$ annehmen. Teil (a): Trivialerweise gilt $\operatorname{int} A \subset \operatorname{int} \operatorname{clos} A$. Sei $x \in \operatorname{int} \operatorname{clos} A$. Wähle $y \in \operatorname{int} A$ (y existiert wegen $\dim A = n$ und 1.1.12). Es gibt ein $z \in \operatorname{clos} A$ mit $x \in [y, z)$, und nach Hilfssatz 1.1.8 folgt $x \in \operatorname{int} A$. Teil (b): Trivialerweise gilt $\operatorname{clos} A \supset \operatorname{clos} \operatorname{int} A$. Sei $x \in \operatorname{clos} A$. Wähle $y \in \operatorname{int} A$. Nach Hilfssatz 1.1.8 gilt $[y, x) \subset \operatorname{int} A$, also $x \in \operatorname{clos} \operatorname{int} A$. Teil (c): Nach (a) gilt $\operatorname{bd} \operatorname{clos} A = \operatorname{clos} (\operatorname{clos} A) \setminus \operatorname{int} (\operatorname{clos} A) = \operatorname{clos} A \setminus \operatorname{int} A = \operatorname{bd} A$. Mit (b) folgt $\operatorname{bd} \operatorname{int} A = \operatorname{clos} (\operatorname{int} A) \setminus \operatorname{int} (\operatorname{int} A) = \operatorname{clos} A \setminus \operatorname{int} A = \operatorname{bd} A$. ■

Jetzt definieren wir den zentralen Begriff dieser Vorlesung. Ein *konvexer Körper* im $\mathbb{E}^n$ ist eine nichtleere, kompakte, konvexe Teilmenge des $\mathbb{E}^n$. Die Menge aller konvexen Körper des $\mathbb{E}^n$ wird mit $\mathcal{K}^n$ bezeichnet, und $\mathcal{K}_0^n$ ist die Teilmenge der konvexen Körper mit inneren Punkten.

1.2 Die metrische Projektion

In diesem Abschnitt ist $A \subset \mathbb{E}^n$ eine feste nichtleere abgeschlossene konvexe Menge. Zu jedem Punkt $x \in \mathbb{E}^n$ gibt es einen eindeutig bestimmten Punkt $p(A, x) \in A$

mit
$$\|x - p(A,x)\| \le \|x - y\| \quad \text{für alle } y \in A.$$

Die Existenz folgt so: Für geeignetes $\rho > 0$ ist die kompakte Menge $B(x,\rho) \cap A$ nicht leer; die stetige Funktion $y \mapsto \|x - y\|$ nimmt also auf dieser Menge ein Minimum an, etwa in y_0. Dann gilt $\|x - y_0\| \le \|x - y\|$ für alle $y \in A$. Zum Beweis der Eindeutigkeit nehmen wir an, auch der Punkt $y_1 \in A$ erfülle $\|x - y_1\| \le \|x - y\|$ für alle $y \in A$. Für den Punkt $z := \frac{1}{2}(y_0 + y_1) \in A$ gilt dann $\|x - z\| < \|x - y_0\|$, falls nicht $y_0 = y_1$ ist. Also ist $y_0 =: p(A,x)$ eindeutig bestimmt.

Die damit definierte Abbildung $p(A,\cdot) : \mathbb{E}^n \to A$ heißt die *metrische Projektion* oder *Fußpunktabbildung* der Menge A. Sie ist in der Konvexgeometrie recht nützlich; deshalb sollen hier ihre wichtigsten Eigenschaften festgestellt werden.

Durch $\|x - p(A,x)\| = d(A,x)$ ist gerade der Abstand des Punktes x von der Menge A gegeben. Genau für die Punkte $x \in A$ gilt $d(A,x) = 0$. Für $x \in \mathbb{E}^n \setminus A$ setzen wir
$$u(A,x) := \frac{x - p(A,x)}{d(A,x)},$$
das ist also der vom Fußpunkt $p(A,x)$ zu x hindeutende Einheitsvektor. Die Menge
$$R(A,x) := \{p(A,x) + \lambda u(A,x) : \lambda \ge 0\}$$
ist der Strahl durch x mit Endpunkt $p(A,x)$.

1.2.1 Hilfssatz. *Für $x \in \mathbb{E}^n \setminus A$ und $y \in R(A,x)$ gilt $p(A,y) = p(A,x)$.*

Beweis. Angenommen, es wäre $p(A,y) \ne p(A,x)$. Gilt $y \in [x, p(A,x))$, so folgt
$$\begin{aligned} \|x - p(A,y)\| &\le \|x - y\| + \|y - p(A,y)\| \\ &< \|x - y\| + \|y - p(A,x)\| \\ &= \|x - p(A,x)\|, \end{aligned}$$
ein Widerspruch. Ist $x \in [y, p(A,x))$, so sei $q \in [p(A,x), p(A,y)]$ der Punkt mit der Eigenschaft, dass die Strecke $[x,q]$ parallel ist zur Strecke $[y, p(A,y)]$. Dann folgt
$$\frac{\|x - q\|}{\|x - p(A,x)\|} = \frac{\|y - p(A,y)\|}{\|y - p(A,x)\|} < 1,$$
was ebenfalls ein Widerspruch ist. ■

1.2.2 Satz. *Die metrische Projektion hat Lipschitz-Konstante* 1*, das heißt es gilt*
$$\|p(A,x) - p(A,y)\| \le \|x - y\| \quad \textit{für } x, y \in \mathbb{E}^n.$$

Beweis. Wir können $v := p(A,x) - p(A,y) \neq 0$ annehmen und behaupten, dass

$$\langle x - p(A,x), v\rangle \geq 0 \tag{1.1}$$

gilt. Angenommen, das wäre falsch. Dann ist $x \notin A$, und der Strahl $R(A,x)$ trifft die zu v senkrechte Hyperebene durch $p(A,y)$ in einem Punkt z. Aus Hilfssatz 1.2.1 folgt

$$\|z - p(A,y)\| < \|z - p(A,x)\| = \|z - p(A,z)\|,$$

ein Widerspruch. Also gilt (1.1). Analog ergibt sich $\langle y - p(A,y), v\rangle \leq 0$. Die Strecke $[x,y]$ trifft daher die beiden zu v orthogonalen Hyperebenen durch $p(A,x)$ bzw. $p(A,y)$. Daraus folgt die Behauptung. ■

Der folgende Hilfssatz wird für Anwendungen der metrischen Projektion im nächsten Abschnitt benötigt.

1.2.3 Hilfssatz. *Sei B eine Kugel mit $A \subset \operatorname{int} B$, sei $S := \operatorname{bd} B$. Dann gilt $p(A,S) = \operatorname{bd} A$.*

Beweis. Dass $p(A,S) \subset \operatorname{bd} A$ gelten muss, ist klar. Sei $x \in \operatorname{bd} A$. Für $i \in \mathbb{N}$ können wir $x_i \in \operatorname{int} B$ wählen, so dass $x_i \notin A$ und $\|x_i - x\| < 1/i$ gilt. Nach Satz 1.2.2 gilt

$$\|x - p(A,x_i)\| = \|p(A,x) - p(A,x_i)\| \leq \|x - x_i\| < 1/i.$$

Der Strahl $R(A,x_i)$ trifft S in einem Punkt y_i, und es gilt $p(A,y_i) = p(A,x_i)$, also $\|x - p(A,y_i)\| < 1/i$. Es gibt eine Teilfolge $(y_{i_j})_{j\in\mathbb{N}}$, die gegen einen Punkt $y \in S$ konvergiert. Aus $\lim p(A,y_i) = x$ und der Stetigkeit der metrischen Projektion folgt $x = p(A,y)$. Also gilt $\operatorname{bd} A \subset p(A,S)$. ■

1.3 Stützen und Trennen

In diesem Abschnitt beweisen wir grundlegende Stütz- und Trennungseigenschaften für konvexe Mengen. Sie sind anschaulich sehr plausibel und auch leicht zu beweisen, aber gerade wegen ihrer Einfachheit für verschiedene Anwendungen sehr nützlich.

Sei $A \subset \mathbb{E}^n$ eine zunächst beliebige Teilmenge. Sei H eine Hyperebene, und seien H^+, H^- die beiden abgeschlossenen Halbräume, die von H berandet werden. Wir sagen, dass H die Menge A *im Punkt x stützt*, wenn $x \in A \cap H$ und entweder $A \subset H^+$ oder $A \subset H^-$ gilt. Die Hyperebene H heißt eine *Stützebene* von A oder *stützt* A wenn H die Menge A in einem Punkt x stützt. Ist $H = H_{u,\alpha}$ eine Stützebene von A und gilt $A \subset H^-_{u,\alpha} = \{y \in \mathbb{E}^n : \langle y,u\rangle \leq \alpha\}$, so heißt $H^-_{u,\alpha}$

ein *Stützhalbraum* von A, und u heißt ein *äußerer Normalenvektor* sowohl von $H_{u,\alpha}$ als auch von $H^-_{u,\alpha}$. Wenn $H_{u,\alpha}$ außerdem die Menge A im Punkt x stützt, so heißt u ein *äußerer Normalenvektor von A in x*. Eine Ebene E *stützt A in x* wenn $x \in A \cap E$ gilt und E in einer Stützebene von A liegt.

1.3.1 Hilfssatz. *Sei $A \subset \mathbb{E}^n$ nichtleer, konvex und abgeschlossen, sei $x \in \mathbb{E}^n \setminus A$. Die zu $u(A,x)$ orthogonale Hyperebene H durch $p(A,x)$ stützt A.*

Beweis. Es gilt $H \cap A \neq \emptyset$. Sei H^- der von H berandete abgeschlossene Halbraum, der x nicht enthält. Angenommen, es gäbe ein $y \in A$ mit $y \notin H^-$. Sei z der zu x nächste Punkt der Strecke $[p(A,x), y]$. Dann gilt $\|x - z\| < \|x - p(A,x)\|$, was wegen $z \in A$ der Definition von $p(A,x)$ widerspricht. Es folgt $A \subset H^-$. ■

1.3.2 Satz. *Sei $A \subset \mathbb{E}^n$ konvex und abgeschlossen. Dann gibt es durch jeden Randpunkt von A eine Stützebene von A. Ist $A \neq \emptyset$ beschränkt, so gibt es zu jedem Vektor $u \in \mathbb{E}^n \setminus \{0\}$ eine Stützebene an A mit äußerem Normalenvektor u.*

Beweis. Sei $x \in \operatorname{bd} A$. Zunächst sei A beschränkt. Nach Hilfssatz 1.2.3 gibt es einen Punkt $y \in \mathbb{E}^n \setminus A$ mit $x = p(A,y)$. Nach Hilfssatz 1.3.1 ist die zu $y - x$ orthogonale Hyperebene durch $p(A,y) = x$ eine Stützebene von A in x.

Ist A unbeschränkt, so existiert durch x jedenfalls eine Stützebene H von $A \cap B(x,1)$. Sei H^- der zugehörige Stützhalbraum von $A \cap B(x,1)$. Gibt es einen Punkt $z \in A \setminus H^-$, so ist $[z,x] \subset A$, aber $[z,x) \cap B(x,1) \not\subset H^-$, ein Widerspruch. Also ist H Stützebene von A.

Sei A beschränkt und $u \in \mathbb{E}^n \setminus \{0\}$. Da A kompakt ist, gibt es einen Punkt $x \in A$ mit $\langle x,u \rangle = \max\{\langle y,u \rangle : y \in A\}$. Offenbar ist $\{y \in \mathbb{E}^n : \langle y,u \rangle = \langle x,u \rangle\}$ eine Stützebene von A mit äußerem Normalenvektor u. ■

Umgekehrt können die konvexen Mengen unter den abgeschlossenen Mengen mit nichtleerem Inneren dadurch charakterisiert werden, dass durch jeden Randpunkt eine Stützebene geht:

1.3.3 Satz. *Sei $A \subset \mathbb{E}^n$ eine abgeschlossene Menge mit $\operatorname{int} A \neq \emptyset$ und der Eigenschaft, dass durch jeden Randpunkt von A eine Stützebene von A geht. Dann ist A konvex.*

Beweis. Angenommen, A erfüllte die Voraussetzungen und wäre nicht konvex. Dann gibt es Punkte $x, y \in A$ und $z \in [x,y]$ mit $z \notin A$. Da $\operatorname{int} A \neq \emptyset$ (und o.B.d.A. $n \geq 2$) ist, können wir einen Punkt $a \in \operatorname{int} A$ wählen, so dass x, y, a affin unabhängig sind. Es gibt einen Punkt $b \in \operatorname{bd} A \cap [a,z)$. Nach Voraussetzung gibt es durch b eine Stützebene H an A, und wegen $a \in \operatorname{int} A$ gilt $a \notin H$. Die Hyperebene H schneidet daher die Ebene $\operatorname{aff}\{x,y,a\}$ in einer Geraden. Die

Punkte x, y, a müssen auf derselben Seite dieser Geraden liegen, was offenbar ein Widerspruch ist. ■

Nun behandeln wir Trennungsaussagen. Es seien $A, B \subset \mathbb{E}^n$ Teilmengen und $H_{u,\alpha} \subset \mathbb{E}^n$ eine Hyperebene. Die Hyperebene $H_{u,\alpha}$ *trennt* A und B, wenn $A \subset H^-_{u,\alpha}$ und $B \subset H^+_{u,\alpha}$ gilt oder umgekehrt. Diese Trennung wird als *eigentlich* bezeichnet, wenn A und B nicht beide in der Hyperebene $H_{u,\alpha}$ liegen. Die Mengen A und B werden von $H_{u,\alpha}$ *strikt getrennt*, wenn $A \subset \operatorname{int} H^-_{u,\alpha}$ und $B \subset \operatorname{int} H^+_{u,\alpha}$ gilt oder umgekehrt, und sie werden von $H_{u,\alpha}$ *stark getrennt*, wenn es ein $\epsilon > 0$ gibt, so dass $H_{u,\alpha-\epsilon}$ und $H_{u,\alpha+\epsilon}$ beide die Mengen A und B trennen. Dass eine Menge A und ein Punkt x getrennt werden, bedeutet, dass A und $\{x\}$ getrennt werden. Zunächst betrachten wir diesen speziellen Fall:

1.3.4 Satz. *Sei $A \subset \mathbb{E}^n$ konvex, und sei $x \in \mathbb{E}^n \setminus A$. Dann können A und x getrennt werden. Ist A abgeschlossen, so können A und x stark getrennt werden.*

Beweis. Zunächst sei A abgeschlossen. Die Hyperebene durch $p(A, x)$ senkrecht zu $u(A, x)$ stützt A und trennt daher A und x. Die parallele Hyperebene durch $\frac{1}{2}(x + p(A, x))$ trennt A und x stark. Jetzt sei A nicht notwendig abgeschlossen. Gilt $x \notin \operatorname{clos} A$, dann trennt eine Hyperebene, die $\operatorname{clos} A$ und x trennt, erst recht die Mengen A und x. Sei $x \in \operatorname{clos} A$, also $x \in \operatorname{bd} A$. Nach Satz 1.1.14 gilt dann $x \in \operatorname{bd} \operatorname{clos} A$, und nach Satz 1.3.2 gibt es eine Stützebene an $\operatorname{clos} A$ durch x; diese Stützebene trennt A und x. ■

1.3.5 Korollar. *Jede nichtleere abgeschlossene konvexe Teilmenge von $\mathbb{E}^n$ ist der Durchschnitt ihrer Stützhalbräume.*

Die Trennung von Paaren von Mengen kann zurückgeführt werden auf die Trennung einer Menge und eines Punktes:

1.3.6 Hilfssatz. *Seien $A, B \subset \mathbb{E}^n$ nichtleere Teilmengen. Die Mengen A und B können genau dann getrennt werden (stark getrennt werden), wenn $A - B$ und 0 getrennt (stark getrennt) werden können.*

Beweis. Wir beweisen das nur für starke Trennung; der andere Fall ergibt sich analog (indem man $\epsilon = 0$ setzt). Die Hyperebene $H_{u,\alpha}$ möge A und B stark trennen, gelte etwa $A \subset H^-_{u,\alpha-\epsilon}$ und $B \subset H^+_{u,\alpha+\epsilon}$ für ein $\epsilon > 0$. Sei $x \in A - B$, dann ist $x = a - b$ mit $a \in A$, $b \in B$. Wegen $\langle a, u\rangle \leq \alpha - \epsilon$ und $\langle b, u\rangle \geq \alpha + \epsilon$ gilt $\langle x, u\rangle \leq -2\epsilon$, also werden $A - B$ und 0 durch $H_{u,-\epsilon}$ stark getrennt.

Jetzt setzen wir voraus, dass $A - B$ und 0 stark getrennt werden können. Dann existieren $u \in \mathbb{E}^n \setminus \{0\}$ und $\epsilon > 0$, so dass $\langle x, u\rangle \leq -2\epsilon$ für alle $x \in A - B$ gilt. Setze

$$\alpha \;:=\; \sup\{\langle a, u\rangle : a \in A\},$$

$$\beta \quad := \quad \inf\{\langle b,u\rangle : b \in B\}.$$

Für $a \in A$, $b \in B$ gilt $\langle a,u\rangle - \langle b,u\rangle \leq -2\epsilon$, also $\beta - \alpha \geq 2\epsilon$. Durch die Hyperebene $H_{u,(\alpha+\beta)/2}$ werden daher A und B stark getrennt. ■

Sind $A, B \subset \mathbb{E}^n$ konvex, so ist $A - B$ konvex. Ist A kompakt und B abgeschlossen, so ist $A - B$ abgeschlossen, wie man leicht sieht. Die Bedingung $0 \notin A - B$ ist äquivalent mit $A \cap B = \emptyset$. Aus Hilfssatz 1.3.6 und Satz 1.3.4 folgt daher:

1.3.7 Satz. *Seien $A, B \subset \mathbb{E}^n$ nichtleere konvexe Mengen mit $A \cap B = \emptyset$. Dann können A und B getrennt werden. Ist außerdem A kompakt und B abgeschlossen, so können A und B stark getrennt werden.*

Die folgenden Beispiele zeigen, dass die Voraussetzungen nicht wesentlich abgeschwächt werden können. Es sei

$$\begin{aligned} A &:= \{(\xi,\eta) \in \mathbb{E}^2 : \xi > 0,\ \eta \geq 1/\xi\}, \\ B &:= \{(\xi,\eta) \in \mathbb{E}^2 : \xi > 0,\ \eta \leq -1/\xi\}, \\ G &:= \{(\xi,\eta) \in \mathbb{E}^2 : \eta = 0\}. \end{aligned}$$

Diese Mengen sind paarweise disjunkt, abgeschlossen und konvex. Die Mengen A und B können strikt getrennt werden (nämlich durch die Gerade G), aber nicht stark. Die Menge $A - B$ und der Punkt 0 können nicht strikt getrennt werden. Die Mengen A und G können getrennt werden, aber nicht strikt.

Andererseits können konvexe Mengen auch dann trennbar sein, wenn sie nicht disjunkt sind. Der folgende Satz gibt eine notwendige und hinreichende Bedingung für die eigentliche Trennbarkeit an.

1.3.8 Satz. *Seien $A, B \subset \mathbb{E}^n$ nichtleere konvexe Mengen. Sie können genau dann eigentlich getrennt werden, wenn*

$$\operatorname{relint} A \cap \operatorname{relint} B = \emptyset \tag{1.2}$$

ist.

Beweis. Wir nehmen an, dass (1.2) gilt, und setzen $C := \operatorname{relint} A - \operatorname{relint} B$. Dann gilt $0 \notin C$, und C ist konvex. Nach Satz 1.3.4 gibt es also eine Hyperebene $H_{u,0}$ mit $C \subset H^-_{u,0}$. Setze

$$\beta := \inf\{\langle b,u\rangle : b \in B\},$$

dann ist $B \subset H^+_{u,\beta}$. Angenommen, es gäbe einen Punkt $a \in A$ mit $\langle a,u\rangle > \beta$. Nach Satz 1.1.12 gibt es einen Punkt $z \in \operatorname{relint} A$, und nach Hilfssatz 1.1.8 ist $[z,a) \subset \operatorname{relint} A$. Daher gibt es einen Punkt $\overline{a} \in \operatorname{relint} A$ mit $\langle \overline{a},u\rangle > \beta$. Ferner gibt es einen Punkt $b \in B$ mit $\langle b, u\rangle < \langle \overline{a}, u\rangle$ und dann, nach einem analogen Argument wie zuvor, einen Punkt $\overline{b} \in \operatorname{relint} B$ mit $\langle \overline{b}, u\rangle < \langle \overline{a}, u\rangle$. Damit ist $\overline{a} - \overline{b} \in C$ und $\langle \overline{a} - \overline{b}, u\rangle > 0$, ein Widerspruch. Hieraus folgt $A \subset H^-_{u,\beta}$. Die Mengen A und B werden also von $H_{u,\beta}$ getrennt. Wenn $A \cup B$ in einer Hyperebene liegt, dann liefert dieselbe Argumentation eine Hyperebene bezüglich $\operatorname{aff}(A \cup B)$, die A und B trennt, und diese Ebene ist enthalten in einer Hyperebene des $\mathbb{E}^n$, die A und B eigentlich trennt.

Umgekehrt sei vorausgesetzt, dass A und B von einer Hyperebene H eigentlich getrennt werden; gelte etwa $A \subset H^-$ und $B \subset H^+$. Angenommen, es gäbe einen Punkt $x \in \operatorname{relint} A \cap \operatorname{relint} B$. Dann ist $x \in H$. Da $A \subset H^-$ und $x \in \operatorname{relint} A$ gilt, folgt $A \subset H$ und analog $B \subset H$, ein Widerspruch. Also gilt (1.2). ■

Gelegentlich werden wir auch Trennen und Stützen von konvexen Kegeln zu benutzen haben. Für diese gilt:

1.3.9 Satz. *Sei $C \subset \mathbb{E}^n$ ein abgeschlossener konvexer Kegel. Jede Stützebene von C geht durch 0. Ist $x \in \mathbb{E}^n \setminus C$, so existiert ein Vektor $u \in \mathbb{E}^n$ mit*

$$\langle c, u\rangle \geq 0 \quad \text{für alle } c \in C \text{ und } \langle x, u\rangle < 0.$$

Beweis. Sei H eine Stützebene an C. Es gibt einen Punkt $y \in H \cap C$. Daraus folgt $\lambda y \in C$ für alle $\lambda > 0$ und daher $0 \in H$. Der Rest der Behauptung ist klar nach Hilfssatz 1.3.1. ■

Wir beweisen nun zwei weitere Sätze vom kombinatorischen Typ der Sätze von Carathéodory und Helly. Sie werden erst in diesem Abschnitt behandelt, weil im Beweis des ersten Satzes Stützebenen verwendet werden und weil es im zweiten Satz um eine Trennungsaussage geht.

1.3.10 Satz (von Steinitz). *Sei $A \subset \mathbb{E}^n$ und $x \in \operatorname{int}\operatorname{conv} A$. Dann gilt $x \in \operatorname{int}\operatorname{conv} A'$ für eine geeignete Teilmenge $A' \subset A$ mit höchstens $2n$ Punkten.*

Beweis. Der Punkt x liegt im Inneren eines Simplex mit Ecken in $\operatorname{conv} A$, und jede dieser Ecken ist eine konvexe Kombination von endlich vielen Punkten aus A. Es gilt also jedenfalls $x \in \operatorname{int}\operatorname{conv} B$ für eine endliche Teilmenge B von A. Wir wählen eine Gerade G durch x, die nicht die affine Hülle von irgend $n-1$ Punkten aus B in einem Punkt $\neq x$ trifft. Seien x_1, x_2 die Endpunkte der Strecke $G \cap \operatorname{conv} B$. Nach Satz 1.3.2 gibt es durch x_j eine Stützebene H_j an $\operatorname{conv} B$ ($j = 1, 2$). Es ist klar, dass $x_j \in \operatorname{conv}(B \cap H_j)$ gilt. Nach dem Satz von Carathéodory gibt es daher eine Darstellung

$$x_j = \sum_{i=1}^{n} \lambda_{ji} y_{ji} \quad \text{mit } y_{ji} \in B \cap H_j, \quad \lambda_{ji} \geq 0, \quad \sum_{i=1}^{n} \lambda_{ji} = 1$$

$(j = 1, 2)$. Wegen der Wahl von G gilt hier notwendigerweise $\lambda_{ji} > 0$. Mit geeigneten $\lambda \in (0, 1)$ gilt

$$\begin{aligned} x &= (1-\lambda)x_1 + \lambda x_2 = \sum_{i=1}^{n} [(1-\lambda)\lambda_{1i} y_{1i} + \lambda\lambda_{2i} y_{2i}] \\ &\in \operatorname{relint} \operatorname{conv} \{y_{11}, \ldots, y_{1n}, y_{21}, \ldots, y_{2n}\} \end{aligned}$$

nach Satz 1.1.13. Hier kann relint ersetzt werden durch int, denn wegen der Wahl von G sind die Punkte $y_{11}, \ldots, y_{1n}$ affin unabhängig, und für wenigstens ein k sind auch $y_{11}, \ldots, y_{1n}, y_{2k}$ affin unabhängig. ■

1.3.11 Satz (von Kirchberger). *Seien $A, B \subset \mathbb{E}^n$ kompakte Mengen. Wenn für jede Teilmenge $M \subset A \cup B$ mit höchstens $n+2$ Punkten die Mengen $M \cap A$ und $M \cap B$ stark getrennt werden können, so können auch die Mengen A und B stark getrennt werden.*

Beweis. Zunächst nehmen wir an, dass A und B endliche Mengen sind. Mit der durch $\tau(x) := (x, 1)$ definierten Abbildung $\tau : \mathbb{E}^n \to \mathbb{E}^n \times \mathbb{R}$ setzen wir für $x \in \mathbb{E}^n$

$$H_x^{\pm} := \{v \in \mathbb{E}^n \times \mathbb{R} : \pm\langle v, \tau(x)\rangle > 0\},$$

wo $\langle \cdot, \cdot \rangle$ das induzierte Skalarprodukt in $\mathbb{E}^n \times \mathbb{R}$ bezeichnet (also $\langle (w, a), (x, 1)\rangle = \langle w, x\rangle + a$).

Sei $M \subset A \cup B$ und $\operatorname{card} M \leq n + 2$. Nach Voraussetzung können $M \cap A$ und $M \cap B$ stark getrennt werden, d.h. es gibt einen Vektor $u \in \mathbb{E}^n$ und eine Zahl $\alpha \in \mathbb{R}$ mit

$$\begin{aligned} \langle u, a\rangle > \alpha \quad &\text{für} \quad a \in M \cap A, \\ \langle u, b\rangle < \alpha \quad &\text{für} \quad b \in M \cap B. \end{aligned}$$

Setzen wir $v := (u, -\alpha)$, so ist also $\langle v, \tau(a)\rangle = \langle u, a\rangle - \alpha > 0$ und daher $v \in H_a^+$ für $a \in M \cap A$. Analog gilt $v \in H_b^-$ für $b \in M \cap B$. Also ist

$$\{H_a^+ : a \in A\} \cup \{H_b^- : b \in B\}$$

eine Familie von endlich vielen konvexen Mengen in $\mathbb{E}^n \times \mathbb{R}$, von denen je $n + 2$ Mengen nichtleeren Durchschnitt haben. Nach dem Satz von Helly haben alle Mengen dieser Familie nichtleeren Durchschnitt. Da dieser Durchschnitt offen ist, enthält er ein Element der Form $v = (u, -\alpha)$ mit $u \neq 0$. Für $a \in A$ gilt $v \in H_a^+$, also $\langle u, a\rangle > \alpha$, und für $b \in B$ gilt analog $\langle u, b\rangle < \alpha$. Die endlichen Mengen A und B werden also durch $H_{u,\alpha}$ stark getrennt.

Nun seien A und B kompakte Mengen, die die Voraussetzung des Satzes erfüllen. Angenommen, A und B können nicht stark getrennt werden. Dann können auch

$\operatorname{conv} A$ und $\operatorname{conv} B$ nicht stark getrennt werden. Nach Satz 1.1.10 sind diese Mengen kompakt, nach Satz 1.3.7 sind sie also nicht disjunkt. Sei $x \in \operatorname{conv} A \cap \operatorname{conv} B$. Dann gilt $x \in \operatorname{conv} A' \cap \operatorname{conv} B'$ mit geeigneten endlichen Teilmengen $A' \subset A$ und $B' \subset B$. Die Mengen A' und B' können nicht stark getrennt werden, ein Widerspruch. ■

1.4 Extremaldarstellungen

In diesem Abschnitt geht es um die Darstellung einer abgeschlossenen konvexen Menge als konvexe Hülle einer möglichst kleinen Menge. Ein erster Kandidat für eine solche möglichst kleine Menge ist der relative Rand. Dabei müssen aber einige triviale Fälle ausgeschlossen werden. Es sei daran erinnert, dass wir unter einer *Ebene* des $\mathbb{E}^n$ einen affinen Unterraum von $\mathbb{E}^n$ verstehen (das kann auch der $\mathbb{E}^n$ selbst sein) und unter einer *Halbebene* den Durchschnitt einer Ebene mit einem abgeschlossenen Halbraum, der die Ebene nicht ganz enthält.

1.4.1 Hilfssatz. *Ist $A \subset \mathbb{E}^n$ eine abgeschlossene konvexe Menge mit $A \neq \operatorname{conv} \operatorname{relbd} A$, so ist A entweder eine Ebene oder eine Halbebene.*

Beweis. O.B.d.A. können wir $\dim A = n$ annehmen. Es gibt einen Punkt $x \in \operatorname{int} A$ mit $x \notin \operatorname{conv} \operatorname{bd} A$ (denn andernfalls wäre $A = \operatorname{int} A \cup \operatorname{bd} A = \operatorname{conv} \operatorname{bd} A$). Nach dem Trennungssatz 1.3.4 gibt es einen abgeschlossenen Halbraum H^- mit $x \in H^-$ und $\operatorname{conv} \operatorname{bd} A \subset H^+$. Jeder Punkt $y \in \operatorname{int} H^-$ erfüllt $[x, y] \cap \operatorname{bd} A = \emptyset$ und daher $y \in \operatorname{int} A$, also gilt $H^- \subset A$. Aus der Konvexität und Abgeschlossenheit von A folgert man, dass jedes Translat von H^-, das einen Punkt von A im Rand enthält, in A enthalten sein muss. Es folgt, dass A entweder ganz $\mathbb{E}^n$ oder ein abgeschlossener Halbraum ist. ■

Um die in Hilfssatz 1.4.1 vorkommenden Ausnahmefälle auszuschließen, fordern wir, dass A *geradenfrei* ist, d.h. keine Gerade enthält. Wegen des unten folgenden Hilfssatzes 1.4.2 ist das keine starke Einschränkung. Zunächst sei $A \subset \mathbb{E}^n$ eine abgeschlossene konvexe Menge. Die Menge A möge einen Strahl

$$G_{x,u} := \{x + \lambda u : \lambda \geq 0\}$$

mit $x \in \mathbb{E}^n$ und $u \in \mathbb{E}^n \setminus \{0\}$ enthalten. Sei $y \in A$, $z \in G_{y,u}$ und $w \in [x, z)$. Der Strahl durch w mit Endpunkt y trifft $G_{x,u}$, daher ist $w \in A$. Es gilt also $[x, z) \subset A$ und daher $z \in A$. Damit ist $G_{y,u} \subset A$ gezeigt. Aus diesem Grunde ist es sinnvoll, die Menge

$$\operatorname{rec} A := \{u \in \mathbb{E}^n \setminus \{0\} : G_{x,u} \subset A\} \cup \{0\}$$

mit einem beliebigen $x \in A$ zu definieren; nach der vorstehenden Bemerkung ist diese Menge unabhängig von der Wahl von x. Die Menge $\operatorname{rec} A$ ist offensichtlich ein konvexer Kegel und wird als der *Rezessionskegel* von A bezeichnet. Eine äquivalente Definition ist

$$\operatorname{rec} A = \{u \in \mathbb{E}^n : A + u \subset A\}.$$

1.4.2 Hilfssatz. *Jede abgeschlossene konvexe Menge $A \subset \mathbb{E}^n$ kann dargestellt werden in der Form $A = \overline{A} \oplus V$, wobei V ein linearer Unterraum von $\mathbb{E}^n$ ist und $\overline{A}$ eine geradenfreie abgeschlossene konvexe Menge in einem zu V komplementären Unterraum von $\mathbb{E}^n$.*

Beweis. Wenn A geradenfrei ist, kann $V = \{0\}$ gewählt werden; sei also A nicht geradenfrei. Dann ist

$$V := \operatorname{rec} A \cap (-\operatorname{rec} A) \neq \{0\}$$

ein linearer Unterraum; er besteht aus allen Vektoren, die parallel sind zu einer in A enthaltenen Geraden. Sei U ein zu V komplementärer linearer Unterraum. Setzen wir $\overline{A} := A \cap U$, so gilt $\overline{A} + V \subset A$. Sei $x \in A$. Durch x gibt es eine Gerade $G \subset A$; da sie parallel zu V ist, trifft sie U in einem Punkt y. Es ist also $x = y + (x - y)$ mit $y \in \overline{A}$ und $x - y \in V$, daher ist $x \in \overline{A} + V$. Damit ist $A = \overline{A} \oplus V$ gezeigt. Es ist klar, dass $\overline{A}$ abgeschlossen, konvex und geradenfrei ist. ■

Die Darstellung durch konvexe Hüllen möglichst kleiner Mengen erfordert einige Definitionen. Sei $A \subset \mathbb{E}^n$ eine konvexe Menge. Eine *Seite* von A ist eine konvexe Teilmenge $F \subset A$ derart, dass jede Strecke $[x, y] \subset A$ mit $F \cap \operatorname{relint}[x, y] \neq \emptyset$ ganz in F enthalten ist. Äquivalent damit ist, dass aus $x, y \in A$ und $\frac{1}{2}(x+y) \in F$ stets $x, y \in F$ folgt. Wenn $\{z\}$ eine Seite von A ist, heißt z ein *Extrempunkt* von A. Mit anderen Worten, z ist genau dann ein Extrempunkt von A, wenn es keine Darstellung $z = (1 - \lambda)x + \lambda y$ mit $x, y \in A \setminus \{z\}$ und $\lambda \in (0, 1)$ gibt. Die Menge aller Extrempunkte von A wird mit $\operatorname{ext} A$ bezeichnet. Ein *Extremstrahl* von A ist ein Strahl (eine abgeschlossene Halbgerade), der eine Seite von A ist. Die Vereinigung aller Extremstrahlen von A wird mit $\operatorname{extr} A$ bezeichnet.

1.4.3 Satz. *Jede geradenfreie abgeschlossene konvexe Menge $A \subset \mathbb{E}^n$ ist die konvexe Hülle ihrer Extrempunkte und Extremstrahlen, d.h. es gilt*

$$A = \operatorname{conv}(\operatorname{ext} A \cup \operatorname{extr} A).$$

Beweis. Die Behauptung ist klar für $n \leq 1$. Sei $n \geq 2$, $\dim A = n$ (o.B.d.A.), und sei die Behauptung bewiesen für konvexe Mengen kleinerer Dimension. Nach

Hilfssatz 1.4.1 gilt $A = \operatorname{conv} \operatorname{bd} A$. Sei $x \in \operatorname{bd} A$. Nach dem Stützsatz 1.3.2 liegt x in einer Stützebene H an A. Nach Induktionsannahme liegt x in der konvexen Hülle der Extrempunkte und Extremstrahlen von $H \cap A$, und dies sind auch Extrempunkte bzw. Extremstrahlen von A, wie sofort aus der Definition der Seiten folgt. Damit ergibt sich die Behauptung. ■

1.4.4 Korollar. *Ist $A \subset \mathbb{E}^n$ eine geradenfreie abgeschlossene konvexe Menge, so gilt*

$$A = \operatorname{conv} \operatorname{ext} A + \operatorname{rec} A.$$

Beweis. Nach Satz 1.4.3 kann ein gegebener Punkt $x \in A$ geschrieben werden in der Form

$$x = \sum_{i=1}^{k} \lambda_i x_i + \sum_{i=k+1}^{m} \lambda_i v_i \quad \text{mit } \lambda_i \geq 0, \quad \sum_{i=1}^{m} \lambda_i = 1,$$

wobei $x_i \in \operatorname{ext} A$ und $v_i \in \operatorname{extr} A$ gilt. Für $i \in \{k+1, \dots, m\}$ liegt v_i in einem Extremstrahl von A, also ist $v_i = x_i + u_i$, wobei x_i der Endpunkt des Extremstrahls und daher ein Extrempunkt von A ist und wo $u_i \in \operatorname{rec} A$ gilt. Daraus folgt die Behauptung. ■

Das folgende Korollar ist ein grundlegend wichtiger Satz:

1.4.5 Satz (von Minkowski). *Jeder konvexe Körper $K \in \mathcal{K}^n$ ist die konvexe Hülle seiner Extrempunkte.*

Hierbei kann die Menge der Extrempunkte nicht durch eine kleinere Menge ersetzt werden, da ein Punkt $x \in K$ genau dann Extrempunkt von K ist, wenn $K \setminus \{x\}$ konvex ist.

Extrempunkte können auch in etwas anderer Weise charakterisiert werden. Sei $A \subset \mathbb{E}^n$ abgeschlossen und konvex, und sei $x \in A$. Unter einer *Kappe von A um x* verstehen wir eine Menge der Form $A \cap H^+$, wobei H^+ ein abgeschlossener Halbraum mit $x \in \operatorname{int} H^+$ ist. Es genügt im folgenden, konvexe Körper zu betrachten.

1.4.6 Hilfssatz. *Sei $K \in \mathcal{K}^n$, sei $x \in K$. Der Punkt x ist genau dann ein Extrempunkt von K, wenn jede Umgebung von x eine Kappe von K um x enthält.*

Beweis. Sei $x \in \operatorname{ext} K$. Eine gegebene Umgebung von x enthält eine offene Kugel B_0 mit Mittelpunkt x. Setze $\overline{K} := \operatorname{conv}(K \setminus B_0)$. Dann gilt $x \notin \overline{K}$, also gibt es eine Hyperebene H, die $\overline{K}$ und x stark trennt. Ist H^+ der abgeschlossene Halbraum, der von H berandet wird und x enthält, dann ist $K \cap H^+$ die gewünschte Kappe.

Gelte $x \notin \operatorname{ext} K$. Dann gilt $x \in \operatorname{relint}[y, z]$ für geeignete $y, z \in K$ mit $y \neq z$. Jede Kappe von K um x enthält y oder z, daher kann eine hinreichend kleine Umgebung von y keine Kappe von K um x enthalten. ■

Der Begriff des Extrempunktes einer konvexen Menge verwendet konvexe Kombinationen von Punkten und gehört damit zu einer „internen" Beschreibung von A. Betrachtet man eine konvexe Menge von außen her, so wird man zu einer verwandten, aber unterschiedlichen Klasse von speziellen Randpunkten geführt. Der Punkt $x \in A$ heißt *exponierter Punkt* von A, wenn es eine Stützebene H von A mit $H \cap A = \{x\}$ gibt. Die Menge der exponierten Punkte von A wird mit $\exp A$ bezeichnet. Trivialerweise gilt $\exp A \subset \operatorname{ext} A$, aber im allgemeinen ist $\exp A \neq \operatorname{ext} A$. Dies zeigt das Beispiel eines Rechteckes, bei dem an einer Seite ein Halbkreis angesetzt ist. Andererseits ist jeder Extrempunkt einer abgeschlossenen konvexen Menge Limes von exponierten Punkten. Es genügt wieder, dieses Resultat für konvexe Körper zu beweisen:

1.4.7 Satz (von Straszewicz). *Für $K \in \mathcal{K}^n$ gilt*

$$\operatorname{ext} K \subset \operatorname{clos} \exp K,$$

also

$$K = \operatorname{clos} \operatorname{conv} \exp K.$$

Beweis. Sei $x \in \operatorname{ext} K$, und sei U eine Umgebung von x. Nach Hilfssatz 1.4.6 enthält U eine Kappe $K \cap H^+$ von K um x. Sei G der zu $H = \operatorname{bd} H^+$ orthogonale Strahl mit Endpunkt x, der H trifft. Zu jedem Punkt $z \in G$ gibt es einen Punkt $y_z \in K$ mit maximalem Abstand von z. Offenbar ist $y_z \in \exp K$. Ist $\|z - x\|$ genügend groß, so gilt $y_z \in H^+$ (Elementargeometrie). Es gilt dann also $y_z \in K \cap H^+ \subset U$. Somit ist $\operatorname{ext} K \subset \operatorname{clos} \exp K$.

Mit dem Satz von Minkowski (und 1.1.10) ergibt sich

$$\begin{aligned} K &= \operatorname{conv} \operatorname{ext} K \subset \operatorname{conv} \operatorname{clos} \exp K \\ &= \operatorname{clos} \operatorname{conv} \exp K \subset \operatorname{clos} \operatorname{conv} \operatorname{ext} K = K, \end{aligned}$$

also $K = \operatorname{clos} \operatorname{conv} \exp K$. ■

1.5 Konvexe Funktionen

Die Untersuchung konvexer Mengen hängt eng zusammen mit der Untersuchung konvexer Funktionen. Wir behandeln hier konvexe Funktionen in etwas größerer

Allgemeinheit als unbedingt erforderlich, weil sie auch in anderen Gebieten (z.B. Optimierungstheorie) eine wichtige Rolle spielen.

Es ist bei konvexen Funktionen zweckmäßig, als Wertebereich das erweiterte System $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ der reellen Zahlen mit den üblichen Regeln zuzulassen. Diese Regeln sind die folgenden Vereinbarungen. Für $\lambda \in \mathbb{R}$ gilt $-\infty < \lambda < \infty$, $\infty + \infty = \lambda + \infty = \infty + \lambda = \infty$, $-\infty - \infty = -\infty + (-\infty) = \lambda - \infty = -\infty + \lambda = -\infty$, schließlich $\lambda\infty = \infty$ bzw. 0 bzw. $-\infty$, je nachdem ob $\lambda > 0$, $\lambda = 0$ oder $\lambda < 0$ ist (man beachte $0 \cdot \infty := 0$). Für eine gegebene Funktion $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ und für $\alpha \in \overline{\mathbb{R}}$ benutzen wir die Abkürzung

$$\{f = \alpha\} := \{x \in \mathbb{E}^n : f(x) = \alpha\},$$

und $\{f \leq \alpha\}$, $\{f < \alpha\}$ u.s.w. werden entsprechend definiert. Eine Funktion $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ heißt *eigentlich*, wenn $\{f = -\infty\} = \emptyset$ und $\{f = \infty\} \neq \mathbb{E}^n$ ist.

Eine Funktion $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ heißt *konvex*, wenn sie eigentlich ist und wenn

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y)$$

für alle $x, y \in \mathbb{E}^n$ und für $0 \leq \lambda \leq 1$ gilt. Eine Funktion $f : D \to \overline{\mathbb{R}}$ mit einem Definitionsbereich $D \subset \mathbb{E}^n$ heißt *konvex*, wenn ihre durch

$$\tilde{f}(x) := \begin{cases} f(x) & \text{für } x \in D \\ \infty & \text{für } x \in \mathbb{E}^n \setminus D \end{cases}$$

definierte Fortsetzung $\tilde{f}$ konvex ist. Eine Funktion f heißt *konkav*, wenn $-f$ konvex ist.

Triviale Beispiele für konvexe Funktionen sind die affinen Funktionen, also die Funktionen der Form $f(x) = \langle u, x\rangle + \alpha$ mit $u \in \mathbb{E}^n$ und $\alpha \in \mathbb{R}$. Eine $\overline{\mathbb{R}}$-wertige Funktion auf $\mathbb{E}^n$ ist genau dann affin, wenn sie zugleich konvex und konkav ist.

Die folgenden Behauptungen ergeben sich unmittelbar aus der Definition. Das Supremum einer beliebigen Familie konvexer Funktionen ist konvex, falls es eigentlich ist. Sind f und g konvexe Funktionen, so auch $f + g$ und αf für $\alpha \geq 0$, falls sie eigentlich sind.

Bemerkung. Ist f konvex, so gilt

$$f(\lambda_1 x_1 + \ldots + \lambda_k x_k) \leq \lambda_1 f(x_1) + \ldots + \lambda_k f(x_k)$$

für alle $k \in \mathbb{N}$, $x_1, \ldots, x_k \in \mathbb{E}^n$ und $\lambda_1, \ldots, \lambda_k \in [0, 1]$ mit $\lambda_1 + \ldots + \lambda_k = 1$. Diese Ungleichung, die sich durch Induktion ergibt, wird auch als *Jensensche Ungleichung* bezeichnet.

Konvexe Funktionen haben zum Beispiel die wichtige Eigenschaft, dass jedes lokale Minimum ein globales Minimum ist. In der Tat, sei $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ konvex, und seien $x_0 \in \mathbb{E}^n$ und $\rho > 0$ derart, dass $f(x_0) < \infty$ und $f(x_0) \leq f(x)$ für $\|x - x_0\| \leq \rho$ gilt. Für $x \in \mathbb{E}^n$ mit $\|x - x_0\| > \rho$ setzen wir

$$y := \frac{\rho}{\|x - x_0\|} x + \left(1 - \frac{\rho}{\|x - x_0\|}\right) x_0;$$

dann gilt $\|y - x_0\| = \rho$ und daher

$$f(x_0) \leq f(y) \leq \frac{\rho}{\|x - x_0\|} f(x) + \left(1 - \frac{\rho}{\|x - x_0\|}\right) f(x_0),$$

woraus $f(x_0) \leq f(x)$ folgt.

Eine konvexe Funktion bestimmt in natürlicher Weise einige konvexe Mengen. Sei $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ konvex. Dann sind

$$\operatorname{dom} f := \{f < \infty\},$$

der *effektive Bereich* von f, und für $\alpha \in \mathbb{R}$ die *Subniveaumengen* $\{f < \alpha\}$ und $\{f \leq \alpha\}$ konvexe Mengen. Der *Epigraph* von f, definiert durch

$$\operatorname{epi} f := \{(x, \zeta) \in \mathbb{E}^n \times \mathbb{R} : f(x) \leq \zeta\},$$

ist eine konvexe Teilmenge von $\mathbb{E}^n \times \mathbb{R}$. Die behauptete Konvexität ist jeweils leicht zu sehen. Umgekehrt bestimmt eine nichtleere konvexe Menge $A \subset \mathbb{E}^n$ eine konvexe Funktion durch

$$I_A(x) := \begin{cases} 0 & \text{für } x \in A, \\ \infty & \text{für } x \in \mathbb{E}^n \setminus A. \end{cases}$$

Sie heißt die *Indikatorfunktion* von A. Man beachte, dass dies eine andere Definition der Indikatorfunktion ist als in der Maß- und Integrationstheorie üblich.

Wir wollen nun die allgemeinen analytischen Eigenschaften konvexer Funktionen untersuchen und beginnen dazu mit der Stetigkeit.

1.5.1 Satz. *Eine konvexe Funktion $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ ist stetig auf* $\operatorname{int} \operatorname{dom} f$ *und Lipschitz-stetig auf jeder kompakten Teilmenge von* $\operatorname{int} \operatorname{dom} f$.

Beweis. Zunächst zeigen wir die Stetigkeit. Sei $x_0 \in \operatorname{int} \operatorname{dom} f$. Wir können ein Simplex S mit $x_0 \in \operatorname{int} S \subset S \subset \operatorname{int} \operatorname{dom} f$ und eine Zahl $\rho > 0$ mit $B(x_0, \rho) \subset S$ wählen. Für $x \in S$ gibt es eine Darstellung

$$x = \sum_{i=1}^{n+1} \lambda_i x_i \quad \text{mit} \quad \lambda_i \geq 0, \quad \sum_{i=1}^{n+1} \lambda_i = 1,$$

wo $x_1, \ldots, x_{n+1}$ die Ecken von S sind. Es folgt

$$f(x) \leq \sum_{i=1}^{n+1} \lambda_i f(x_i) \leq c := \max\{f(x_1), \ldots, f(x_{n+1})\}.$$

Nun sei $y = x_0 + \alpha u$ mit $\alpha \in [0,1]$ und $\|u\| = \rho$. Aus $y = (1-\alpha)x_0 + \alpha(x_0 + u)$ folgt

$$f(y) \leq (1-\alpha)f(x_0) + \alpha f(x_0 + u),$$

also

$$f(y) - f(x_0) \leq \alpha(c - f(x_0)),$$

wegen $x_0 + u \in S$. Andererseits ist

$$x_0 = \frac{1}{1+\alpha} y + \frac{\alpha}{1+\alpha}(x_0 - u)$$

und daher

$$f(x_0) \leq \frac{1}{1+\alpha} f(y) + \frac{\alpha}{1+\alpha} f(x_0 - u),$$

also

$$f(x_0) - f(y) \leq \alpha(c - f(x_0)).$$

Somit gilt

$$|f(y) - f(x_0)| \leq \frac{1}{\rho}[c - f(x_0)]\|y - x_0\|$$

für $y \in B(x_0, \rho)$. Dies zeigt in verschärfter Form die Stetigkeit von f in x_0. Also ist f stetig auf $\operatorname{int}\operatorname{dom} f$.

Nun sei $C \subset \operatorname{int}\operatorname{dom} f$ eine kompakte Teilmenge. Wegen der Kompaktheit existiert eine Zahl $\rho > 0$ mit $C_\rho := C + B(0, \rho) \subset \operatorname{int}\operatorname{dom} f$. Auf der kompakten Menge C_ρ nimmt die stetige Funktion $|f|$ ein Maximum a an. Seien $x, y \in C$ gegeben. Dann ist

$$z := y + \frac{\rho}{\|y - x\|}(y - x) \in C_\rho$$

und

$$y = (1-\lambda)x + \lambda z \quad \text{ mit } \quad \lambda = \frac{\|y - x\|}{\rho + \|y - x\|};$$

aus $f(y) \leq (1-\lambda)f(x) + \lambda f(z)$ folgt also

$$f(y) - f(x) \leq \lambda[f(z) - f(x)] \leq \frac{2a}{\rho}\|y - x\|.$$

Vertauschung von x und y liefert die Ungleichung $|f(y) - f(x)| \leq b\|y - x\|$ mit einer von x und y unabhängigen Konstanten b. Die Funktion f hat also auf C die Lipschitz-Eigenschaft. ■

Die Differenzierbarkeitseigenschaften konvexer Funktionen untersuchen wir zunächst im Fall $n = 1$.

1.5.2 Satz. *Sei $f : \mathbb{R} \to \overline{\mathbb{R}}$ konvex. Auf* $\operatorname{int}\operatorname{dom} f$ *gelten die folgenden Aussagen. Die rechtsseitige Ableitung f'_r und die linksseitige Ableitung f'_ℓ existieren und sind (schwach) monoton wachsende Funktionen. Es gilt $f'_\ell \le f'_r$, und bis auf höchstens abzählbar viele Punkte gilt $f'_\ell = f'_r$, d.h. f ist dort differenzierbar. Die Funktion f'_r ist rechtsseitig stetig, und f'_ℓ ist linksseitig stetig. Ist also f differenzierbar auf* $\operatorname{int}\operatorname{dom} f$*, so ist f dort stetig differenzierbar.*

Beweis. Alle im folgenden auftretenden Argumente sollen zu $\operatorname{int}\operatorname{dom} f$ gehören. Sei $0 < \lambda < \mu$. Dann ist

$$f(x+\lambda) = f\left(\frac{\mu-\lambda}{\mu}x + \frac{\lambda}{\mu}(x+\mu)\right) \le \frac{\mu-\lambda}{\mu}f(x) + \frac{\lambda}{\mu}f(x+\mu),$$

also

$$\frac{f(x+\lambda)-f(x)}{\lambda} \le \frac{f(x+\mu)-f(x)}{\mu}.$$

Analog ist

$$f(x-\lambda) = f\left(\frac{\mu-\lambda}{\mu}x + \frac{\lambda}{\mu}(x-\mu)\right) \le \frac{\mu-\lambda}{\mu}f(x) + \frac{\lambda}{\mu}f(x-\mu),$$

also

$$\frac{f(x)-f(x-\mu)}{\mu} \le \frac{f(x)-f(x-\lambda)}{\lambda}.$$

Für beliebige $\lambda, \mu > 0$ ist

$$f(x) = f\left(\frac{\lambda}{\lambda+\mu}(x-\mu) + \frac{\mu}{\lambda+\mu}(x+\lambda)\right) \le \frac{\lambda}{\lambda+\mu}f(x-\mu) + \frac{\mu}{\lambda+\mu}f(x+\lambda),$$

also

$$\frac{f(x)-f(x-\mu)}{\mu} \le \frac{f(x+\lambda)-f(x)}{\lambda}.$$

Aus den damit erhaltenen Monotonie-Eigenschaften und Ungleichungen folgt sowohl die Existenz der Grenzwerte

$$\begin{aligned} f'_r(x) &= \lim_{\lambda\downarrow 0}\frac{f(x+\lambda)-f(x)}{\lambda}, \\ f'_\ell(x) &= \lim_{\mu\downarrow 0}\frac{f(x-\mu)-f(x)}{-\mu} \end{aligned}$$

als auch die Ungleichung $f'_\ell(x) \le f'_r(x)$.

Für $x < y$ gilt also

$$f'_\ell(x) \le f'_r(x) \le \frac{f(y) - f(x)}{y - x} \le f'_\ell(y) \le f'_r(y).$$

Hieran liest man ab, dass f'_ℓ und f'_r (schwach) monoton wachsend sind; sie haben daher höchstens abzählbar viele Unstetigkeitsstellen. An jeder Stetigkeitsstelle von f'_ℓ gilt nach den obigen Ungleichungen $f'_\ell = f'_r$, d.h. dort existiert die Ableitung f'.

Sei $x < y$. Es gilt

$$\frac{f(y) - f(x)}{y - x} = \lim_{z \downarrow x} \frac{f(y) - f(z)}{y - z} \ge \lim_{z \downarrow x} f'_r(z),$$

also

$$\lim_{z \downarrow x} f'_r(z) \le f'_r(x) \le \lim_{z \downarrow x} f'_r(z)$$

wegen der Monotonie von f'_r. Also ist f'_r rechtsseitig stetig. Analog ist f'_ℓ linksseitig stetig. ∎

1.5.3 Bemerkung. Sei $f : \mathbb{R} \to \mathbb{R}$ konvex, sei $x_0 \in \mathbb{R}$ und m eine Zahl mit $f'_\ell(x_0) \le m \le f'_r(x_0)$. Wie im obigen Beweis festgestellt, gilt

$$\frac{f(x) - f(x_0)}{x - x_0} \begin{cases} \ge f'_r(x_0) \ge m, & \text{falls } x > x_0, \\ \le f'_\ell(x_0) \le m, & \text{falls } x < x_0. \end{cases}$$

Zusammen ergibt das

$$f(x) \ge f(x_0) + m(x - x_0) \quad \text{für alle } x \in \mathbb{R}.$$

Dies zeigt, dass die Gerade

$$\{(x, y) \in \mathbb{R} \times \mathbb{R} : y = f(x_0) + m(x - x_0)\}$$

eine Stützgerade an den Epigraphen von f im Punkt $(x_0, f(x_0))$ ist.

Nun sei $n \ge 2$ und $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ eine konvexe Funktion. Satz 1.5.2 impliziert insbesondere für jeden Punkt $x \in \operatorname{int} \operatorname{dom} f$ die Existenz aller einseitigen Richtungsableitungen

$$f'(x; u) := \lim_{\lambda \downarrow 0} \frac{f(x + \lambda u) - f(x)}{\lambda},$$

wo $u \in \mathbb{E}^n \setminus \{0\}$ ist. Der Untersuchung dieser Richtungsableitungen schicken wir einige Definitionen voraus.

Eine Funktion $f : \mathbb{E}^n \to \mathbb{R}$ heißt *positiv homogen*, wenn

$$f(\lambda x) = \lambda f(x) \quad \text{für alle } \lambda \ge 0 \text{ und alle } x \in \mathbb{E}^n$$

gilt; sie heißt *subadditiv*, wenn

$$f(x+y) \leq f(x) + f(y) \qquad \text{für alle } x, y \in \mathbb{E}^n$$

gilt. Eine Funktion, die positiv homogen und subadditiv ist, wird als *sublinear* bezeichnet. Sublineare Funktionen sind also spezielle konvexe Funktionen.

1.5.4 Satz. *Sei* $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ *konvex, und sei* $x \in \operatorname{int} \operatorname{dom} f$. *Dann ist die Richtungsableitung*

$$f'(x;\cdot) : \mathbb{E}^n \to \mathbb{R}$$

(*mit* $f'(x;0) := 0$) *eine sublineare Funktion.*

Beweis. Sei $u \in \mathbb{E}^n \setminus \{0\}$. Für $\lambda, \tau > 0$ ist

$$\frac{f(x+\tau\lambda u) - f(x)}{\tau} = \lambda \frac{f(x+\tau\lambda u) - f(x)}{\tau\lambda},$$

und mit $\tau \downarrow 0$ folgt $f'(x;\lambda u) = \lambda f'(x;u)$. Für $u, v \in \mathbb{E}^n$ gilt wegen der Konvexität von f

$$\begin{aligned} f(x+\tau(u+v)) &= f\left(\frac{1}{2}(x+2\tau u) + \frac{1}{2}(x+2\tau v)\right) \\ &\leq \frac{1}{2} f(x+2\tau u) + \frac{1}{2} f(x+2\tau v), \end{aligned}$$

also

$$\frac{f(x+\tau(u+v)) - f(x)}{\tau} \leq \frac{f(x+2\tau u) - f(x)}{2\tau} + \frac{f(x+2\tau v) - f(x)}{2\tau}$$

für $\tau > 0$. Mit $\tau \downarrow 0$ folgt jetzt $f'(x;u+v) \leq f'(x;u) + f'(x;v)$. ■

Mehr als die Existenz der einseitigen Richtungsableitungen lässt sich bei einer konvexen Funktion im allgemeinen nicht aussagen; insbesondere brauchen die partiellen Ableitungen nicht zu existieren. Wenn sie aber an einer Stelle existieren, dann ist die konvexe Funktion dort sogar (total) differenzierbar.

1.5.5 Satz. *Sei* $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ *konvex, sei* $x \in \operatorname{int} \operatorname{dom} f$. *Hat* f *bei* x *partielle Ableitungen* (*erster Ordnung*), *so ist* f *bei* x *differenzierbar.*

Beweis. Die i-te partielle Ableitung von f bei x, bezeichnet mit $\partial_i f(x)$, ist die zweiseitige Richtungsableitung $f'(x;e_i) = -f'(x;-e_i)$, wo $(e_1, \ldots, e_n)$ die orthonormierte Standardbasis von $\mathbb{E}^n$ bezeichnet. Wenn die partiellen Ableitungen von f bei x existieren, können wir den Gradienten von f bei x erklären durch

$$\operatorname{grad} f(x) := \sum_{i=1}^{n} \partial_i f(x) e_i.$$

Wir müssen für

$$g(h) := f(x+h) - f(x) - \langle \operatorname{grad} f(x), h \rangle, \quad h \in \mathbb{E}^n,$$

die Limesbeziehung

$$\lim_{\|h\|\to 0} \frac{g(h)}{\|h\|} = 0 \tag{1.3}$$

zeigen. Die Funktion g ist konvex und endlich in einer Umgebung von 0. Für $h = \sum_{i=1}^n \eta_i e_i$ gilt

$$g(h) = g\left(\frac{1}{n}\sum_{i=1}^n n\eta_i e_i\right) \leq \frac{1}{n}\sum_{i=1}^n g(n\eta_i e_i).$$

Unter Verwendung von

$$\langle h, v \rangle \leq \|h\|\|v\| \leq \|h\| \sum_{i=1}^n |\nu_i| \quad \text{für } v = \sum_{i=1}^n \nu_i e_i$$

erhalten wir

$$g(h) \leq \sum_{i=1}^n \eta_i \frac{g(n\eta_i e_i)}{n\eta_i} \leq \|h\| \sum_{i=1}^n \left|\frac{g(n\eta_i e_i)}{n\eta_i}\right|$$

(wobei die Summanden mit $\eta_i = 0$ durch 0 zu ersetzen sind) sowie eine entsprechende Ungleichung, in der h durch $-h$ ersetzt ist. Da $g(h) + g(-h) \geq g(0) = 0$ wegen der Konvexität von g gilt, erhalten wir für $h \neq 0$

$$-\sum_{i=1}^n \left|\frac{g(-n\eta_i e_i)}{n\eta_i}\right| \leq \frac{-g(-h)}{\|h\|} \leq \frac{g(h)}{\|h\|} \leq \sum_{i=1}^n \left|\frac{g(n\eta_i e_i)}{n\eta_i}\right|.$$

Weil die partiellen Ableitungen von g bei 0 existieren und Null sind, konvergieren die linke und die rechte Seite dieser Ungleichungskette für $h \to 0$ gegen 0, also folgt (1.3). ∎

Für differenzierbare Funktionen gibt es einige nützliche Konvexitätskriterien, die wir jetzt zusammenstellen.

1.5.6 Satz. *Sei $D \subset \mathbb{R}$ ein offenes Intervall, sei $f : D \to \mathbb{R}$ eine differenzierbare Funktion. Dann ist f genau dann konvex, wenn die Ableitung f' (schwach) monoton wachsend ist.*

Beweis. Ist f konvex, so ist f' nach Satz 1.5.2 wachsend. Sei f' wachsend, und seien $x, y \in D$ Punkte mit $x < y$, sei $\lambda \in (0,1)$. Nach dem Mittelwertsatz gibt es

Punkte $\theta_1 \in [x, (1-\lambda)x + \lambda y]$ und $\theta_2 \in [(1-\lambda)x + \lambda y, y]$ mit

$$\begin{aligned} f'(\theta_1) &= \frac{f((1-\lambda)x + \lambda y) - f(x)}{\lambda(y-x)}, \\ f'(\theta_2) &= \frac{f(y) - f((1-\lambda)x + \lambda y)}{(1-\lambda)(y-x)}. \end{aligned}$$

Aus $f'(\theta_1) \le f'(\theta_2)$ folgt

$$f((1-\lambda)x + \lambda y) \le (1-\lambda)f(x) + \lambda f(y).$$

Also ist f konvex. ■

1.5.7 Korollar. *Sei $D \subset \mathbb{R}$ ein offenes Intervall und $f : D \to \mathbb{R}$ eine zweimal differenzierbare Funktion. Dann ist f genau dann konvex, wenn $f'' \ge 0$ ist.*

Diese Kriterien lassen sich leicht auf differenzierbare Funktionen auf dem $\mathbb{E}^n$ ausdehnen. Dabei wird die (unmittelbar aus der Definition folgende) Tatsache benutzt, dass eine Funktion auf dem $\mathbb{E}^n$ genau dann konvex ist, wenn ihre Einschränkung auf jede Gerade konvex ist.

1.5.8 Satz. *Sei $D \subset \mathbb{E}^n$ konvex und offen, sei $f : D \to \mathbb{R}$ eine differenzierbare Funktion. Dann sind die folgenden Aussagen äquivalent:*

(a) *f ist konvex,*

(b) *$f(y) - f(x) \ge \langle \operatorname{grad} f(x), y - x\rangle$ für alle $x, y \in D$,*

(c) *$\langle \operatorname{grad} f(y) - \operatorname{grad} f(x), y - x\rangle \ge 0$ für alle $x, y \in D$.*

Beweis. Es gelte (a). Sei $x, y \in D$ und $0 < \lambda < 1$. Aus

$$f(x + \lambda(y-x)) = f((1-\lambda)x + \lambda y) \le (1-\lambda)f(x) + \lambda f(y)$$

erhalten wir

$$\frac{f(x + \lambda(y-x)) - f(x)}{\lambda} \le f(y) - f(x).$$

Für $\lambda \to 0$ konvergiert die linke Seite gegen $f'(x; y-x) = \langle \operatorname{grad} f(x), y - x\rangle$, also folgt (b).

Es gelte (b). Vertauschen wir x und y und addieren beide Ungleichungen, so erhalten wir (c).

Nun gelte (c). Sei $x, y \in D$. Setze

$$g(\lambda) := f(x + \lambda(y-x)) \qquad \text{für } 0 \le \lambda \le 1.$$

Für $0 \leq \lambda_0 < \lambda_1 \leq 1$ ergibt sich unter Verwendung der Kettenregel

$$\begin{aligned}
&g'(\lambda_1) - g'(\lambda_0) \\
&= \langle \operatorname{grad} f(x + \lambda_1(y-x)), y-x\rangle - \langle \operatorname{grad} f(x+\lambda_0(y-x)), y-x\rangle \\
&\geq 0.
\end{aligned}$$

Also ist g' eine wachsende Funktion und daher g konvex. Da $x, y \in D$ beliebig waren, ist f konvex. ■

1.5.9 Satz. *Sei $D \subset \mathbb{E}^n$ konvex und offen, und sei $f : D \to \mathbb{R}$ zweimal differenzierbar. Dann ist f genau dann konvex, wenn die Hessesche Matrix*

$$\operatorname{Hess} f(x) := (\partial_i \partial_j f(x))_{i,j=1}^n$$

für alle $x \in D$ positiv semidefinit ist.

Beweis. Für $x, y \in D$ definieren wir

$$g(\lambda) := f(x + \lambda(y-x)) \qquad \text{für } x + \lambda(y-x) \in D.$$

Die Funktion f ist genau dann konvex, wenn alle so definierten Funktionen g konvex sind, und letzteres ist äquivalent mit $g'' \geq 0$. Es genügt, die Bedingung $g''(0) \geq 0$ nachzuprüfen. In der Tat, sei $\lambda_1 \in (0,1)$. Setzen wir $x_1 := x + \lambda_1(y-x)$ und $h(\tau) := f(x_1 + \tau(y-x_1))$, so ist $g(\lambda_1+\lambda) = h(\lambda/(1-\lambda_1))$, also gilt $g''(\lambda_1) \geq 0$ genau dann, wenn $h''(0) \geq 0$ ist.

Nun gilt für $x = \sum_{i=1}^n \xi_i e_i$ und $y = \sum_{i=1}^n \eta_i e_i$

$$\begin{aligned}
g''(0) &= \lim_{\lambda\to 0} \frac{g'(\lambda) - g'(0)}{\lambda} \\
&= \lim_{\lambda \to 0} \sum_{j=1}^n \frac{1}{\lambda}[\partial_j f(x+\lambda(y-x)) - \partial_j f(x)](\eta_j - \xi_j) \\
&= \sum_{i=1}^n \sum_{j=1}^n \partial_i \partial_j f(x)(\eta_i - \xi_i)(\eta_j - \xi_j).
\end{aligned}$$

Hieran liest man die Behauptung ab. ■

Für konvexe Funktionen, die nicht notwendig differenzierbar sind, lassen sich Gradient und Differential in natürlicher Weise verallgemeinern. Diese Verallgemeinerungen sind insbesondere für die Optimierungstheorie von Bedeutung. Wir betrachten sie hier vor allem wegen ihrer einfachen geometrischen Interpretation.

Ist zunächst $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ konvex und differenzierbar im Punkt x, so gilt nach Satz 1.5.8

$$f(y) \geq f(x) + \langle \operatorname{grad} f(x), y - x \rangle \quad \text{für alle } y \in \mathbb{E}^n.$$

Ist f nicht differenzierbar bei x, so betrachtet man statt des Gradienten die Menge aller Vektoren, die in diesem Sinne dasselbe leisten wie der Gradient. Für eine konvexe Funktion $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ und für einen Punkt $x \in \operatorname{dom} f$ definiert man

$$\partial f(x) := \{v \in \mathbb{E}^n : f(y) \geq f(x) + \langle v, y - x \rangle \text{ für alle } y \in \mathbb{E}^n\}.$$

Man nennt die Menge $\partial f(x)$ das *Subdifferential* von f bei x, und jedes Element von $\partial f(x)$ heißt ein *Subgradient* von f bei x. Es ist klar, dass das Subdifferential von f bei x eine abgeschlossene konvexe Teilmenge von $\mathbb{E}^n$ ist. Diese Menge parametrisiert die Stützebenen an den Epigraphen von f über x, im folgenden präzisen Sinn.

1.5.10 Satz. *Sei $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ konvex, und sei $x \in \operatorname{dom} f$. Der Vektor $v \in \mathbb{E}^n$ ist genau dann ein Subgradient von f bei x, wenn der Vektor $(v, -1) \in \mathbb{E}^n \times \mathbb{R}$ ein äußerer Normalenvektor an $\operatorname{epi} f$ im Punkt $(x, f(x))$ ist.*

Beweis. Der Vektor $(v, -1)$ ist genau dann ein äußerer Normalenvektor an $\operatorname{epi} f$ im Punkt $(x, f(x))$, wenn

$$\langle (y, \eta) - (x, f(x)), (v, -1) \rangle \leq 0$$

für alle $(y, \eta) \in \operatorname{epi} f$ gilt. Dies ist äquivalent mit

$$\langle (y, f(y)) - (x, f(x)), (v, -1) \rangle \leq 0$$

für alle $y \in \mathbb{E}^n$, also mit

$$f(y) \geq f(x) + \langle v, y - x \rangle$$

für alle $y \in \mathbb{E}^n$. ■

Differenzierbarkeit einer konvexen Funktion ist äquivalent mit der Eindeutigkeit der Subgradienten:

1.5.11 Satz. *Sei $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ konvex, und sei $x \in \operatorname{int} \operatorname{dom} f$. Dann ist f genau dann bei x differenzierbar, wenn f bei x nur einen Subgradienten hat.*

Beweis. Sei f differenzierbar bei x. Sei v ein Subgradient von f bei x. Dann gilt

$$\frac{f(x + \lambda u) - f(x)}{\lambda} \geq \langle v, u \rangle$$

für alle $\lambda > 0$, also $f'(x;u) \geq \langle v, u\rangle$ für alle $u \in \mathbb{E}^n$. Wegen der Differenzierbarkeit von f bei x gilt $f'(x;u) = -f'(x;-u) \leq -\langle v, -u\rangle = \langle v, u\rangle$, also $f'(x;u) = \langle v, u\rangle$. Da diese Gleichung für alle $u \in \mathbb{E}^n$ gilt, ist v eindeutig bestimmt.

Umgekehrt sei vorausgesetzt, dass f bei x höchstens einen Subgradienten hat. Sei $u \in \mathbb{E}^n \setminus \{0\}$. Setze

$$g(\lambda) := f(x + \lambda u) \qquad \text{für } \lambda \in \mathbb{R}.$$

Sei m eine Zahl mit $g'_\ell(0) \leq m \leq g'_r(0)$. Nach Bemerkung 1.5.3 gilt

$$g(\lambda) \geq g(0) + m\lambda,$$

also $f(x + \lambda u) \geq f(x) + m\lambda$. Die Gerade

$$G := \{(x + \lambda u, f(x) + m\lambda) : \lambda \in \mathbb{R}\} \subset \mathbb{E}^n \times \mathbb{R}$$

erfüllt daher $G \cap \operatorname{int}\operatorname{epi} f = \emptyset$. Nach Satz 1.3.7 gibt es in $\mathbb{E}^n \times \mathbb{R}$ eine Stützebene $H_{(v,-1),\alpha}$, die G und $\operatorname{int}\operatorname{epi} f$ trennt (der Normalenvektor der Stützebene kann o.B.d.A. von der Form $(v, -1)$ angenommen werden, da er wegen $x \in \operatorname{int}\operatorname{dom} f$ nicht von der Form $(v, 0)$ sein kann). Da $H_{(v,-1),\alpha}$ den Epigraphen $\operatorname{epi} f$ im Punkt $(x, f(x))$ stützt, ist v ein Subgradient von f bei x und daher eindeutig bestimmt. Wegen $G \subset H_{(v,-1),\alpha}$ gilt

$$\langle (v, -1), (x + \lambda u, f(x) + m\lambda)\rangle = \alpha$$

für alle $\lambda \in \mathbb{R}$ und daher $m = \langle v, u\rangle$. Da v eindeutig ist, ist auch die Zahl m eindeutig. Also ist g bei 0 differenzierbar. Da $u \in \mathbb{E}^n \setminus \{0\}$ beliebig war, hat f bei x zweiseitige Richtungsableitungen und ist daher nach Satz 1.5.5 bei x differenzierbar. ■

1.6 Dualität

Konvexen Mengen, Kegeln und Funktionen kann man unter schwachen Zusatzvoraussetzungen duale Objekte derselben Art zuordnen. Diese Dualität ist oft von Nutzen und soll hier untersucht werden.

Zunächst sei $K \in \mathcal{K}^n$ ein konvexer Körper mit $0 \in \operatorname{int} K$. Wir definieren

$$K^* := \{x \in \mathbb{E}^n : \langle x, y\rangle \leq 1 \text{ für alle } y \in K\}.$$

K^* heißt der *Polarkörper* oder *duale Körper* von K. Der nachfolgende Satz zeigt, dass dies wieder ein konvexer Körper ist und dass in der Tat eine Dualität vorliegt.

1.6.1 Satz. *Sei $K \in \mathcal{K}^n$ und $0 \in \operatorname{int} K$. Dann gilt $K^* \in \mathcal{K}^n$, $0 \in \operatorname{int} K^*$ und $K^{**} = K$.*

Beweis. Für $x_1, x_2 \in K^*$ und $\lambda \in [0,1]$ gilt

$$\langle (1-\lambda)x_1 + \lambda x_2, y\rangle \leq 1 \qquad \text{für alle } y \in K,$$

also $(1-\lambda)x_1+\lambda x_2 \in K^*$. Also ist K^* konvex. Trivialerweise ist K^* abgeschlossen. Für die Kugel $B(0,\epsilon)$ mit Radius $\epsilon > 0$ gilt $B(0,\epsilon)^* = B(0,1/\epsilon)$, wie unmittelbar aus der Definition folgt. Ebenfalls aus der Definition ergibt sich, dass aus $K_1 \subset K_2$ stets $K_1^* \supset K_2^*$ folgt. Wir können $\epsilon, \rho > 0$ wählen mit $B(0,\epsilon) \subset K \subset B(0,\rho)$; dann ist $B(0,1/\rho) \subset K^* \subset B(0,1/\epsilon)$, also gilt $0 \in \operatorname{int} K^*$, und K^* ist beschränkt und daher kompakt.

Sei $y \in K$. Für beliebiges $x \in K^*$ gilt $\langle x,y\rangle \leq 1$, also ist $y \in K^{**}$. Somit gilt $K \subset K^{**}$ (übrigens haben wir bis jetzt weder die Konvexität noch die Abgeschlossenheit von K benutzt).

Sei $z \in \mathbb{E}^n \setminus K$. Da K konvex und abgeschlossen ist, gibt es nach Satz 1.3.4 eine Hyperebene $H_{u,\alpha}$ mit $K \subset H^-_{u,\alpha}$ und $\langle z,u\rangle > \alpha$; wegen $0 \in \operatorname{int} K$ ist hier $\alpha > 0$. Für alle $y \in K$ gilt $\langle y, u/\alpha\rangle \leq 1$, also ist $u/\alpha \in K^*$. Wegen $\langle z, u/\alpha\rangle > 1$ folgt jetzt $z \notin K^{**}$. Damit ist der Beweis von $K^{**} = K$ erbracht. ■

Der folgende Satz beschreibt, inwiefern beim Übergang zu Polarkörpern Durchschnitte und Vereinigungen vertauscht werden.

1.6.2 Satz. *Gilt* $K_1, K_2 \in \mathcal{K}^n$ *und* $0 \in \operatorname{int} K_i$ $(i = 1,2)$, *so ist*

$$\begin{aligned} (K_1 \cap K_2)^* &= \operatorname{conv}(K_1^* \cup K_2^*), \\ [\operatorname{conv}(K_1 \cup K_2)]^* &= K_1^* \cap K_2^*. \end{aligned}$$

Ist außerdem $K_1 \cup K_2$ *konvex, so ist* $K_1^* \cup K_2^*$ *konvex, also gilt*

$$\begin{aligned} (K_1 \cap K_2)^* &= K_1^* \cup K_2^*, \\ (K_1 \cup K_2)^* &= K_1^* \cap K_2^*. \end{aligned}$$

Beweis. Aus $K_1 \cap K_2 \subset K_i$ folgt $K_i^* \subset (K_1 \cap K_2)^*$, also $\operatorname{conv}(K_1^* \cup K_2^*) \subset (K_1 \cap K_2)^*$. Aus $K_i \subset \operatorname{conv}(K_1 \cup K_2)$ folgt $[\operatorname{conv}(K_1 \cup K_2)]^* \subset K_i^*$, also $[\operatorname{conv}(K_1 \cup K_2)]^* \subset K_1^* \cap K_2^*$. Wendet man beide Inklusionen auf K_i^* statt K_i an und benutzt $K^{**} = K$, so erhält man die ersten beiden Gleichungen des Satzes.

Sei jetzt $K_1 \cup K_2$ konvex. Sei $x \in \mathbb{E}^n \setminus (K_1^* \cup K_2^*)$. Nach Satz 1.3.4 gibt es $u_i \in \mathbb{E}^n \setminus \{0\}$ und $\alpha_i \in \mathbb{R}$ mit $\langle x, u_i\rangle > \alpha_i$ und $\langle y, u_i\rangle \leq \alpha_i$ für alle $y \in K_i^*$ (also ist $\alpha_i > 0$), woraus $u_i/\alpha_i \in K_i^{**} = K_i$ folgt $(i = 1,2)$. Da $K_1 \cup K_2$ konvex ist, gibt es einen Punkt $z \in [u_1/\alpha_1, u_2/\alpha_2] \cap K_1 \cap K_2$. Da z eine konvexe Kombination von u_1/α_1 und u_2/α_2 ist, gilt $\langle x, z\rangle > 1$ und somit $x \notin (K_1 \cap K_2)^*$. Damit ist

$(K_1 \cap K_2)^* \subset K_1^* \cup K_2^*$ gezeigt. Da die entgegengesetzte Inklusion trivial ist, folgt $(K_1 \cap K_2)^* = K_1^* \cup K_2^*$ und damit die Konvexität von $K_1^* \cup K_2^*$. ■

Eine analoge Dualität gibt es für konvexe Kegel. Sei $C \subset \mathbb{E}^n$ ein abgeschlossener konvexer Kegel. Durch

$$C^* := \{x \in \mathbb{E}^n : \langle x, y \rangle \leq 0 \text{ für alle } y \in C\}$$

wird der *duale Kegel* von C definiert. Dies ist wieder ein abgeschlossener konvexer Kegel, und es gilt $C^{**} = C$. Der Beweis ist ganz analog zu demjenigen für duale konvexe Körper, abgesehen davon, dass statt des Trennungssatzes 1.3.4 der Satz 1.3.9 zu benutzen ist. Der Beweis des folgenden Satzes ist analog zu dem von Satz 1.6.2 und braucht daher nicht ausgeführt zu werden.

1.6.3 Satz. *Seien $C_1, C_2 \subset \mathbb{E}^n$ abgeschlossene konvexe Kegel. Dann gilt*

$$\begin{aligned}(C_1 \cap C_2)^* &= C_1^* + C_2^*,\\ (C_1 + C_2)^* &= C_1^* \cap C_2^*.\end{aligned}$$

Ist $C_1 \cup C_2$ konvex, so ist $C_1^* \cup C_2^*$ konvex, also gilt

$$\begin{aligned}(C_1 \cap C_2)^* &= C_1^* \cup C_2^*,\\ (C_1 \cup C_2)^* &= C_1^* \cap C_2^*.\end{aligned}$$

Der Behandlung der Dualität für konvexe Funktionen schicken wir einige Vorbemerkungen voraus. Beim Nachweis von $K^{**} = K$ und $C^{**} = C$ wurde benutzt, dass eine abgeschlossene konvexe Menge der Durchschnitt ihrer Stützhalbräume ist. Im Fall konvexer Funktionen müssen wir die entsprechende Eigenschaft für den Epigraphen ausnutzen. Aus diesem Grunde betrachtet man nur solche Funktionen, deren Epigraph abgeschlossen ist; sie werden selbst als *abgeschlossen* bezeichnet. Die Stützebenen an den Epigraphen einer konvexen Funktion sind von zwei Arten. Allgemein ist ein abgeschlossener Halbraum in $\mathbb{E}^n \times \mathbb{R}$, der den Epigraphen einer eigentlichen Funktion enthält, von der Form

$$\{(x, \zeta) \in \mathbb{E}^n \times \mathbb{R} : \langle (x, \zeta), (u, \eta) \rangle \leq \alpha\}$$

mit $\eta \leq 0$. Ist $\eta = 0$, so wird dieser Halbraum als *vertikal* bezeichnet. Ist er nicht vertikal, so können wir o.B.d.A. $\eta = -1$ annehmen, dann ist der Halbraum also gegeben durch

$$\begin{aligned}H^-_{(u,-1),\alpha} &= \{(x, \zeta) \in \mathbb{E}^n \times \mathbb{R} : \langle x, u \rangle - \alpha \leq \zeta\}\\ &= \operatorname{epi} h\end{aligned}$$

mit $h(x) := \langle x, u \rangle - \alpha$. Da $\operatorname{epi} f \subset \operatorname{epi} h$ für beliebige Funktion h, f äquivalent ist mit $h \le f$, wird man so auf den folgenden Hilfssatz geführt.

1.6.4 Hilfssatz. *Sei $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ konvex und abgeschlossen. Dann gilt*

$$f = \sup\{h : h \text{ affine Funktion auf } \mathbb{E}^n,\ \ h \le f\}.$$

Beweis. Die abgeschlossene konvexe Menge $\operatorname{epi} f$ ist der Durchschnitt ihrer Stützhalbräume. Die Behauptung des Hilfssatzes ist nach der vorstehenden Bemerkung äquivalent damit, dass sie bereits der Durchschnitt ihrer nichtvertikalen Stützhalbräume ist. Sei

$$H_1^- := \{(x, \zeta) \in \mathbb{E}^n \times \mathbb{R} : h_1(x) := \langle x, u_1 \rangle - \alpha_1 \le 0\}$$

ein vertikaler Stützhalbraum von $\operatorname{epi} f$, und sei $(x_0, \zeta_0) \in \mathbb{E}^n \times \mathbb{R}$ ein Punkt mit $(x_0, \zeta_0) \notin H_1^-$. Es gibt jedenfalls einen nichtvertikalen Stützhalbraum an $\operatorname{epi} f$ (zum Beispiel in einem Punkt $(x, f(x))$ mit $x \in \operatorname{relint} \operatorname{dom} f$), etwa

$$H_2^- := \{(x, \zeta) \in \mathbb{E}^n \times \mathbb{R} : h_2(x) := \langle x, u_2 \rangle - \alpha_2 \le \zeta\}.$$

Wir setzen $h := \lambda h_1 + h_2$ mit einer Zahl $\lambda > 0$, die unten festgelegt wird. Für $x \in \operatorname{dom} f$ gilt $h_1(x) \le 0$ und $h_2(x) \le f(x)$, also $h(x) \le f(x)$. Ist $x \in \mathbb{E}^n \setminus \operatorname{dom} f$, so ist $f(x) = \infty$ und daher ebenfalls $h(x) \le f(x)$. Also gilt

$$\operatorname{epi} f \subset H^- := \{(x, \zeta) \in \mathbb{E}^n \times \mathbb{R} : h(x) \le \zeta\}.$$

Da $(x_0, \zeta_0) \notin H_1^-$ gilt, ist $h_1(x_0) > 0$, daher kann $\lambda > 0$ so gewählt werden, dass $h(x_0) > \zeta_0$ und damit $(x_0, \zeta_0) \notin H^-$ gilt. Es gibt also einen nichtvertikalen Stützhalbraum (ein Translat von H^-) an $\operatorname{epi} f$, der (x_0, ζ_0) nicht enthält. Somit ist $\operatorname{epi} f$ der Durchschnitt seiner nichtvertikalen Stützhalbräume; das ist äquivalent damit, dass f das Supremum aller affinen Funktionen $\le f$ ist. ■

Damit können wir nun Dualität für konvexe Funktionen behandeln. Für eine abgeschlossene konvexe Funktion $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ definieren wir die *konjugierte Funktion* durch

$$f^*(u) := \sup\{\langle u, x \rangle - f(x) : x \in \mathbb{E}^n\} \qquad \text{für } u \in \mathbb{E}^n.$$

Die folgende Umformulierung verdeutlicht die anschauliche Bedeutung der konjugierten Funktion. Für gegebenes $u \in \mathbb{E}^n$ bestimmt jede reelle Zahl α einen nichtvertikalen Halbraum $H^-_{(u,-1),\alpha}$ in $\mathbb{E}^n \times \mathbb{R}$, und es gilt

$$(x, f(x)) \in H^-_{(u,-1),\alpha} \Leftrightarrow \langle u, x \rangle - f(x) \le \alpha,$$

also

$$\mathrm{epi}\, f \subset H^-_{(u,-1),\alpha} \Leftrightarrow \forall\, x \in \mathbb{E}^n : \langle u, x\rangle - f(x) \le \alpha \Leftrightarrow f^*(u) \le \alpha.$$

Es gilt also auch

$$f^*(u) = \inf\{\alpha \in \mathbb{R} : \mathrm{epi}\, f \subset H^-_{(u,-1),\alpha}\}.$$

Dies kann man so interpretieren, dass die konjugierte Funktion die Stützhalbräume an den Epigraphen von f parametrisiert. Nun zeigen wir, dass wieder eine Dualität vorliegt.

1.6.5 Satz. *Sei $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ eine abgeschlossene konvexe Funktion. Dann ist f^* eine abgeschlossene konvexe Funktion, und es gilt $f^{**} = f$.*

Beweis. Wegen $\{f = \infty\} \ne \mathbb{E}^n$ ist $\{f^* = -\infty\} = \emptyset$. Der Epigraph $\mathrm{epi}\, f$ besitzt einen nichtvertikalen Stützhalbraum $\{(x,\zeta) : \langle u, x\rangle - \alpha \le \zeta\}$. Es gilt also $\langle u, x\rangle - f(x) \le \alpha$ für $x \in \mathbb{E}^n$ und daher $f^*(u) \le \alpha$, somit ist $\{f^* = \infty\} \ne \mathbb{E}^n$. Also ist f^* eine eigentliche Funktion. Als Supremum einer Familie affiner Funktionen ist f^* konvex. Der Epigraph von f^* ist Durchschnitt einer Familie von Epigraphen affiner Funktionen, also abgeschlossen. Somit ist f^* abgeschlossen.

Die Definition von f^* kann umgeschrieben werden in die Form

$$f^*(u) = \sup\{\langle u, x\rangle - \zeta : (x, \zeta) \in \mathrm{epi}\, f\}.$$

Hiernach ist also

$$f^{**}(x) = \sup\{\langle x, u\rangle - \alpha : (u, \alpha) \in \mathrm{epi}\, f^*\}.$$

Nach Definition von f^* ist $(u, \alpha) \in \mathrm{epi}\, f^*$ äquivalent mit $\langle \cdot, u\rangle - \alpha \le f$. Also gilt für $x \in \mathbb{E}^n$

$$\begin{aligned} f^{**}(x) &= \sup\{\langle x, u\rangle - \alpha : u \in \mathbb{E}^n,\ \ \langle \cdot, u\rangle - \alpha \le f\} \\ &= \sup\{h(x) : h \text{ affine Funktion auf } \mathbb{E}^n,\ \ h \le f\} \\ &= f(x) \end{aligned}$$

nach Hilfssatz 1.6.4. ■

1.7 Die Stützfunktion

Da eine abgeschlossene konvexe Menge der Durchschnitt ihrer Stützhalbräume ist, kann eine solche Menge bequem beschrieben werden, indem man die Lage ihrer Stützhalbräume in Abhängigkeit vom äußeren Normalenvektor angibt. Diese Beschreibung wird geliefert von der Stützfunktion, die in der Theorie der konvexen

Körper ein unentbehrliches Hilfsmittel ist. Die Stützfunktion kann für nichtleere abgeschlossene konvexe Mengen definiert werden; wir beschränken uns hier aber auf konvexe Körper, da wir die Stützfunktion nur dafür verwenden werden.

Für einen konvexen Körper $K \in \mathcal{K}^n$ wird die *Stützfunktion* $h(K, \cdot) = h_K$ definiert durch

$$h(K,u) := \max\{\langle x,u\rangle : x \in K\} \qquad \text{für } u \in \mathbb{E}^n.$$

Das Maximum existiert, weil stetige Funktionen auf kompakten Mengen ihr Maximum annehmen.

Für $u \in \mathbb{E}^n \setminus \{0\}$ setzen wir

$$\begin{aligned} H(K,u) &:= \{x \in \mathbb{E}^n : \langle x,u\rangle = h(K,u)\}, \\ H^-(K,u) &:= \{x \in \mathbb{E}^n : \langle x,u\rangle \leq h(K,u)\}, \\ F(K,u) &:= H(K,u) \cap K. \end{aligned}$$

Die Mengen $H(K,u)$, $H^-(K,u)$, $F(K,u)$ werden der Reihe nach als *Stützebene*, *Stützhalbraum* und *Stützmenge* von K bezeichnet, jeweils mit äußerem Normalenvektor u.

Die Stützfunktion hat eine einfache anschauliche Bedeutung: Für einen Einheitsvektor $u \in S^{n-1} \cap \mathbb{E}^n$ ist $h(K,u)$ der mit Vorzeichen versehene Abstand der Stützebene $H(K,u)$ mit äußerem Normalenvektor u vom Ursprung 0; der Abstand ist dabei genau dann negativ, wenn u in den offenen (von $H(K,u)$ berandeten) Halbraum weist, der 0 enthält.

Aus der Definition der Stützfunktion ergeben sich unmittelbar die folgenden Eigenschaften: Es gilt

$$h(K,\cdot) = \langle z,\cdot\rangle \Leftrightarrow K = \{z\},$$
$$h(K+t,u) = h(K,u) + \langle t,u\rangle \qquad \text{für } t \in \mathbb{E}^n,$$

ferner

$$h(K,\lambda u) = \lambda h(K,u) \qquad \text{für } \lambda \geq 0,$$
$$h(K,u+v) \leq h(K,u) + h(K,v) \qquad \text{für } u,v \in \mathbb{E}^n.$$

Also ist $h(K,\cdot)$ eine sublineare Funktion. Ebenfalls unmittelbar aus der Definition folgen die Eigenschaften

$$\begin{aligned} h_K \leq h_L &\Leftrightarrow K \subset L, \\ h(\lambda K,\cdot) &= \lambda h(K,\cdot) \qquad \text{für } \lambda \geq 0, \\ h(-K,u) &= h(K,-u). \end{aligned}$$

Ist $E \subset \mathbb{E}^n$ ein linearer Unterraum und bezeichnet $(\cdot)|E$ die Orthogonalprojektion auf E, so gilt

$$h(K|E,u) = h(K,u) \qquad \text{für } u \in E.$$

Der folgende wichtige Satz zeigt, dass die Zuordnung von sublinearen Funktionen zu konvexen Körpern bijektiv ist.

1.7.1 Satz. *Ist $f : \mathbb{E}^n \to \mathbb{R}$ eine sublineare Funktion, so gibt es einen eindeutig bestimmten konvexen Körper $K \in \mathcal{K}^n$ mit Stützfunktion f.*

Beweis. Die Eindeutigkeit ist klar. Nach Satz 1.5.1 ist die Funktion f stetig und damit abgeschlossen. Für jedes $\lambda > 0$ gilt für die konjugierte Funktion f^*

$$\begin{aligned} f^*(x) &= \sup\{\langle x, u\rangle - f(u) : u \in \mathbb{E}^n\} \\ &= \sup\{\langle x, \lambda u\rangle - f(\lambda u) : u \in \mathbb{E}^n\} \\ &= \lambda f^*(x). \end{aligned}$$

Dies ist nur möglich, wenn $f^*(x) \in \{0, \infty\}$ gilt ($x \in \mathbb{E}^n$). Nach Satz 1.6.5 ist f^* abgeschlossen und konvex, also ist f^* die Indikatorfunktion I_K einer abgeschlossenen konvexen Menge K. Wegen $\{f^* = \infty\} \neq \mathbb{E}^n$ ist $K \neq \emptyset$. Aus Satz 1.6.5 folgt weiter für $u \in \mathbb{E}^n$

$$\begin{aligned} f(u) = f^{**}(u) &= \sup\{\langle x, u\rangle - f^*(x) : x \in \mathbb{E}^n\} \\ &= \sup\{\langle x, u\rangle : x \in K\} \\ &= h(K, u), \end{aligned}$$

also $f = h(K, \cdot)$. Es ist klar, dass K beschränkt ist. ■

Der Beweis hat gezeigt, dass die Stützfunktion eines konvexen Körpers und seine Indikatorfunktion zueinander konjugiert sind.

Wir geben für Satz 1.7.1 noch einen zweiten Beweis, der nicht die Dualitätstheorie verwendet (aber dafür auch nicht die vorstehende Information liefert).

Zweiter Beweis. Wir setzen

$$K := \{x \in \mathbb{E}^n : \langle x, v\rangle \leq f(v) \text{ für alle } v \in \mathbb{E}^n\}.$$

Als Durchschnitt von abgeschlossenen Halbräumen (einer für jede äußere Normalenrichtung) ist K konvex und kompakt. Gilt $K \neq \emptyset$, so ist offenbar $h(K, u) \leq f(u)$ für alle $u \in \mathbb{E}^n$. Also muss nur gezeigt werden, dass $K \neq \emptyset$ ist und $h(K, u) \geq f(u)$ für $u \in \mathbb{E}^n$ gilt.

Da f sublinear und stetig ist, ist der Epigraph $\operatorname{epi} f$ ein abgeschlossener konvexer Kegel in $\mathbb{E}^n \times \mathbb{R}$. Sei $u \in \mathbb{E}^n \setminus \{0\}$. Da $(u, f(u)) \in \operatorname{bd} \operatorname{epi} f$ ist, gibt es nach Satz 1.3.2 eine Stützebene $H_{(y,\eta),\alpha}$ an $\operatorname{epi} f$ durch $(u, f(u))$, so dass $\operatorname{epi} f \subset H^-_{(y,\eta),\alpha}$ gilt. Nach Satz 1.3.9 ist $\alpha = 0$. Wegen $\operatorname{dom} f = \mathbb{E}^n$ ist $\eta \neq 0$, also kann $\eta = -1$

angenommen werden. Aus $\operatorname{epi} f \subset H^-_{(y,-1),0}$ folgt $\langle y, v\rangle \le f(v)$ für alle $v \in \mathbb{E}^n$, also $y \in K$ und damit $K \ne \emptyset$. Aus $(u, f(u)) \in H_{(y,-1),0}$ folgt $\langle y, u\rangle = f(u)$, also $h(K, u) \ge f(u)$. ■

Da die Stützfunktion eines konvexen Körpers K eine konvexe Funktion ist, existieren ihre Richtungsableitungen. Die Richtungsableitung $h'_K(u;\cdot)$ ist nach Satz 1.5.4 eine sublineare Funktion, also nach Satz 1.7.1 selbst die Stützfunktion eines konvexen Körpers. Der folgende Satz gibt an, um welchen Körper es sich dabei handelt.

1.7.2 Satz. *Sei $K \in \mathcal{K}^n$ und $u \in \mathbb{E}^n \setminus \{0\}$. Dann gilt*

$$h'_K(u;x) = h(F(K,u),x) \qquad \text{für } x \in \mathbb{E}^n.$$

Beweis. Wie gerade erwähnt, gibt es einen konvexen Körper $K_u \in \mathcal{K}^n$ mit

$$h'_K(u;\cdot) = h(K_u,\cdot).$$

Wegen $h_K(u+\lambda v) \le h_K(u) + \lambda h_K(v)$ ist

$$\frac{h_K(u+\lambda v) - h_K(u)}{\lambda} \le h_K(v)$$

für $\lambda > 0$, und mit $\lambda \downarrow 0$ folgt $h'_K(u;v) \le h_K(v)$ für $v \in \mathbb{E}^n$. Also gilt $K_u \subset K$.

Sei $y \in K_u$. Dann ist

$$\begin{aligned} \langle y,u\rangle &\le h(K_u,u) \le h(K,u), \\ \langle y,-u\rangle &\le h(K_u,-u) = h'_K(u;-u) = -h(K,u) \end{aligned}$$

(letzteres nach Definition der Richtungsableitung und wegen $h_K(\lambda u) = \lambda h_K(u)$ für $\lambda > 0$). Also ist $\langle y,u\rangle = h(K,u)$ und daher $y \in H(K,u)$. Somit gilt $K_u \subset K \cap H(K,u) = F(K,u)$.

Zum Nachweis der entgegengesetzten Inklusion sei $y \in F(K,u)$. Dann gilt

$$\begin{aligned} \langle y,u\rangle &= h(K,u), \\ \langle y,v\rangle &\le h(K,v) \qquad \text{für alle } v \in \mathbb{E}^n. \end{aligned}$$

Sei $x \in \mathbb{E}^n$. Mit $v := u + \lambda x$, $\lambda > 0$, folgt

$$\langle y,x\rangle \le \frac{h(K,u+\lambda x) - h(K,u)}{\lambda},$$

also $\langle y,x\rangle \le h'_K(u;x) = h(K_u,x)$. Da dies für alle $x \in \mathbb{E}^n$ gilt, ist $y \in K_u$. Somit gilt $F(K,u) \subset K_u$, und es ist $K_u = F(K,u)$ gezeigt. ■

1.7.3 Korollar. *Sei $K \in \mathcal{K}^n$ und $u \in \mathbb{E}^n \setminus \{0\}$. Genau dann ist die Stützfunktion h_K differenzierbar bei u, wenn $F(K,u) = \{x\}$ für ein $x \in \mathbb{E}^n$ gilt, und in diesem Fall ist $x = \operatorname{grad} h_K(u)$.*

Beweis. Sei $(e_1, \dots, e_n)$ die Standardbasis von $\mathbb{E}^n$. Mit Satz 1.5.5 und Satz 1.7.2 ergibt sich

$$\begin{aligned}
& h_K \text{ ist differenzierbar bei } u \\
\Leftrightarrow \quad & h_K \text{ ist partiell differenzierbar bei } u \\
\Leftrightarrow \quad & h_K'(u; e_i) = -h_K'(u; -e_i) \qquad \text{für } i = 1, \dots, n \\
\Leftrightarrow \quad & h(F(K,u), e_i) + h(F(K,u), -e_i) = 0 \qquad \text{für } i = 1, \dots, n \\
\Leftrightarrow \quad & F(K,u) \subset \bigcap_{i=1}^n H(F(K,u), e_i) \\
\Leftrightarrow \quad & F(K,u) = \{x\} \qquad \text{für ein } x \in \mathbb{E}^n \\
\Rightarrow \quad & \partial_i h_K(u) = h_K'(u; e_i) = h(F(K,u), e_i) = \langle x, e_i \rangle \qquad \text{für } i = 1, \dots, n \\
\Rightarrow \quad & x = \operatorname{grad} h_K(u).
\end{aligned}$$

■

Schließlich haben die Subdifferentiale der Stützfunktion eine einfache Bedeutung:

1.7.4 Satz. *Für $K \in \mathcal{K}^n$ und $u \in \mathbb{E}^n \setminus \{0\}$ gilt*

$$\partial h_K(0) = K \quad \textit{und} \quad \partial h_K(u) = F(K,u).$$

Beweis. Nach Definition des Subdifferentials gilt für $x \in \mathbb{E}^n$

$$x \in \partial h_K(0) \Leftrightarrow h_K(v) \geq \langle x, v \rangle \text{ für alle } v \in \mathbb{E}^n \Leftrightarrow x \in K,$$

also ist $\partial h_K(0) = K$. Ferner gilt für $u \neq 0$

$$x \in \partial h_K(u) \Leftrightarrow h_K(v) \geq h_K(u) + \langle x, v - u \rangle \text{ für alle } v \in \mathbb{E}^n.$$

Sei $x \in F(K,u)$. Dann gilt $\langle x, u \rangle = h_K(u)$ und $\langle x, v \rangle \leq h_K(v)$ für alle $v \in \mathbb{E}^n$. Es folgt $h_K(v) \geq h_K(u) + \langle x, v - u \rangle$, also $x \in \partial h_K(u)$.

Sei $x \in \partial h_K(u)$, also $h_K(v) \geq h_K(u) + \langle x, v - u \rangle$ für alle $v \in \mathbb{E}^n$. Dann ist $h_{K-x}(v) \geq h_{K-x}(u)$. Mit $v = 0$ ergibt sich $h_{K-x}(u) \leq 0$, und mit $v = 2u$ ergibt sich $h_{K-x}(u) \geq 0$, zusammen also $h_{K-x}(u) = 0$ und somit $h_K(u) = \langle x, u\rangle$. Wegen $h_{K-x}(v) \geq 0$ für alle $v \in \mathbb{E}^n$ folgt $0 \in K-x$, also $x \in K$, und wegen $h_K(u) = \langle x, u\rangle$ ist $x \in F(K,u)$. Damit ist $\partial h_K(u) = F(K,u)$ gezeigt. ■

Eine besonders wichtige Eigenschaft der Stützfunktion ist ihre Additivität im ersten Argument, im folgenden Sinn. Sind K und L konvexe Körper, so ist auch $K+L$ ein konvexer Körper. Die Konvexität ist klar, und die Kompaktheit folgt aus der Kompaktheit von $K \times L$ und der Stetigkeit der Addition als Abbildung von $\mathbb{E}^n \times \mathbb{E}^n$ in $\mathbb{E}^n$. Die Addition konvexer Körper wird auch als *Minkowski-Addition* bezeichnet. Eine Funktion f auf $\mathcal{K}^n$ mit Werten in einer abelschen Halbgruppe wird als *Minkowski-additiv* bezeichnet, wenn

$$f(K+L) = f(K) + f(L) \qquad \text{für alle } K, L \in \mathcal{K}^n$$

gilt. Beispiele hierfür liefert schon der folgende Satz.

1.7.5 Satz. *Für $K, L \in \mathcal{K}^n$ gilt*

(a) $h(K+L,\cdot) = h(K,\cdot) + h(L,\cdot)$,

(b) $H(K+L,\cdot) = H(K,\cdot) + H(L,\cdot)$,

(c) $F(K+L,\cdot) = F(K,\cdot) + F(L,\cdot)$.

Beweis. Sei $u \in \mathbb{E}^n \setminus \{0\}$. Es gibt Punkte $x \in K$ und $y \in L$ mit $h(K,u) = \langle x, u\rangle$ und $h(L,u) = \langle y, u\rangle$. Daraus folgt $h(K,u) + h(L,u) = \langle x+y, u\rangle \leq h(K+L,u)$. Jeder Punkt $z \in K+L$ hat eine Darstellung $z = x+y$ mit $x \in K$, $y \in L$, und es folgt $\langle z, u\rangle = \langle x, u\rangle + \langle y, u\rangle \leq h(K,u) + h(L,u)$. Da $z \in K+L$ beliebig war, folgt $h(K+L,u) \leq h(K,u) + h(L,u)$. Damit ist (a) gezeigt, und (b) ist eine triviale Folgerung hieraus. Aus (a) und Satz 1.7.2 folgt (c). ■

Eine erste Folgerung aus Satz 1.7.5 (a) ist, dass aus $K+M = L+M$ für konvexe Körper $K, L, M \in \mathcal{K}^n$ stets $K = L$ folgt. Es ist also $(\mathcal{K}^n, +)$ eine kommutative Halbgruppe mit Kürzungsregel.

Neben der Stützfunktion können auch andere Funktionen zur analytischen Beschreibung konvexer Körper dienen. Für einen konvexen Körper $K \in \mathcal{K}^n$ mit $0 \in \operatorname{int} K$ definiert man durch

$$g(K,x) := \min\{\lambda \geq 0 : x \in \lambda K\}, \qquad x \in \mathbb{E}^n \setminus \{0\},$$

die *Distanzfunktion* von K, und durch

$$\begin{aligned}\rho(K,x) &:= \max\{\lambda \geq 0 : \lambda x \in K\}, \qquad x \in \mathbb{E}^n \setminus \{0\},\\ &= \frac{1}{g(K,x)}\end{aligned}$$

wird die *Radiusfunktion* $\rho(K,\cdot)$ von K erklärt. Unmittelbar aus den Definitionen folgt

$$K = \{x \in \mathbb{E}^n : g(K,x) \leq 1\},$$
$$g(K,x) = \frac{\|x\|}{\|\lambda x\|}, \text{ falls } x \in \mathbb{E}^n \setminus \{0\}, \lambda > 0, \lambda x \in \operatorname{bd} K,$$
$$\rho(K,x)x \in \operatorname{bd} K \qquad \text{für } x \in \mathbb{E}^n \setminus \{0\}.$$

Die Distanzfunktion $g(K,\cdot)$ ist sublinear. Das ergibt sich leicht aus der Definition, aber auch aus dem folgenden Satz.

1.7.6 Satz. *Für* $K \in \mathcal{K}^n$ *mit* $0 \in \operatorname{int} K$ *gilt*

$$g(K,\cdot) = h(K^*,\cdot).$$

Beweis. Setzen wir $\overline{K} := \{x \in \mathbb{E}^n : h(K^*,x) \leq 1\}$, so gilt $g(\overline{K},\cdot) = h(K^*,\cdot)$ nach Definition der Distanzfunktion. Sei $x \in \overline{K}$. Für $u \in K^*$ gilt $\langle u,x\rangle \leq h(K^*,x) \leq 1$, also ist $x \in K^{**} = K$. Somit gilt $\overline{K} \subset K$. Sei $x \in K$. Es gibt ein $v \in K^*$ mit $h(K^*,x) = \langle v,x\rangle \leq 1$, also ist $x \in \overline{K}$. Dies zeigt $K = \overline{K}$. ■

1.7.7 Bemerkung. Nach Satz 1.7.6 gilt

$$\rho(K^*,u) = \frac{1}{h(K,u)} \qquad \text{für } u \in S^{n-1}.$$

Dies zeigt, wie man aus den Stützebenen von K die Randpunkte von K^* findet.

1.7.8 Bemerkung. Sei $0 \in \operatorname{int} K$ und $x \in \operatorname{bd} K$, sei u ein äußerer Normalenvektor von K in x, sei $y = \lambda u \in \operatorname{bd} K^*$ mit $\lambda > 0$. Wegen $1 = g(K^*,y) = g(K^*,\lambda u) = \lambda g(K^*,u) = \lambda h(K,u)$ ist

$$\langle y,x\rangle = \frac{\langle u,x\rangle}{h(K,u)} = 1 = g(K,x) = h(K^*,x),$$

also ist x ein äußerer Normalenvektor an K^* in y.

1.7.9 Bemerkung. Sei $K \in \mathcal{K}^n$ mit $0 \in \operatorname{int} K$ und $K = -K$, d.h. K ist symmetrisch zu 0. Dann wird durch

$$\|x\|_K := g(K,x) \qquad \text{für } x \in \mathbb{E}^n$$

eine Norm $\|\cdot\|_K$ auf $\mathbb{E}^n$ definiert, für die K die Normkugel ist. Wir identifizieren den Dualraum von $\mathbb{E}^n$ vermöge des Skalarprodukts $\langle\cdot,\cdot\rangle$ mit $\mathbb{E}^n$. Dann ist der duale Banachraum des Banachraums $(\mathbb{E}^n, \|\cdot\|_K)$ gegeben durch $(\mathbb{E}^n, \|\cdot\|_K^*)$ mit

$$\begin{aligned}
\|y\|_K^* &:= \sup\{\langle y,x\rangle : x \in \mathbb{E}^n, \|x\|_K \leq 1\} \\
&= \sup\{\langle y,x\rangle : x \in K\} \\
&= h(K,y) = g(K^*,y) = \|y\|_{K^*}.
\end{aligned}$$

Es ist also $\|\cdot\|_K^* = \|\cdot\|_{K^*}$, und K^* ist die Normkugel der dualen Norm $\|\cdot\|_K^*$.

1.8 Die Hausdorff-Metrik

Auf der Menge $\mathcal{K}^n$ der konvexen Körper wird nun eine Metrik δ eingeführt. Damit wird $(\mathcal{K}^n, \delta)$ ein vollständiger metrischer Raum, in dem beschränkte, abgeschlossene Mengen kompakt sind. Diese Tatsache ist von grundlegender Bedeutung, zum Beispiel für den Nachweis der Existenz konvexer Körper, die bestimmte Eigenschaften haben, etwa Extremalprobleme zu lösen. Die Konvergenz im Sinne der Metrik wird ferner oft verwendet, um Eigenschaften allgemeiner konvexer Körper mittels Approximation durch spezielle Körper nachzuweisen.

Wir betrachten zunächst allgemeiner die Menge $\mathcal{C}^n$ der nichtleeren kompakten Teilmengen von $\mathbb{E}^n$ und führen hierauf eine Metrik ein. Das erfordert keinen zusätzlichen Aufwand und ist ebenfalls oft von Nutzen.

Für $K, L \in \mathcal{C}^n$ wird der *Hausdorff-Abstand* definiert durch

$$\delta(K, L) := \max\{\sup_{x \in K} \inf_{y \in L} \|x - y\|, \sup_{x \in L} \inf_{y \in K} \|x - y\|\}. \tag{1.4}$$

Äquivalent hiermit ist die Definition

$$\delta(K, L) = \min\{\lambda \geq 0 : K \subset L + \lambda B^n, L \subset K + \lambda B^n\}. \tag{1.5}$$

In (1.4) können dabei, weil wir nur kompakte Mengen betrachten, sup und inf durch max bzw. min ersetzt werden. Von den beiden angegebenen Definitionen hat die erste den Vorteil, dass sie auch in allgemeineren metrischen Räumen verwendet werden kann, wenn $\|x - y\|$ durch den dort gegebenen Abstand ersetzt wird. Die Äquivalenz von (1.4) und (1.5) in $\mathbb{E}^n$ wird folgendermaßen eingesehen. Wir bezeichnen die rechte Seite von (1.5) mit α. Für $x \in K$ gilt dann $x \in L + \alpha B^n$, also $x = y + \alpha b$ mit geeigneten $y \in L$ und $b \in B^n$; es folgt $\|x - y\| \leq \alpha$ und damit $\inf_{y \in L} \|x - y\| \leq \alpha$. Da dies für alle $x \in K$ gilt, folgt $\sup_{x \in K} \inf_{y \in L} \|x - y\| \leq \alpha$. Vertauschung von K und L ergibt $\delta(K, L) \leq \alpha$. Nun sei o.B.d.A. $\alpha > 0$ und $0 < \lambda < \alpha$, etwa $K \not\subset L + \lambda B^n$. Dann gilt $x \notin L + \lambda B^n$ für geeignetes $x \in K$, also $\|x - y\| > \lambda$ für alle $y \in L$. Daher ist $\delta(K, L) \geq \lambda$. Da $\lambda < \alpha$ beliebig war, folgt $\delta(K, L) \geq \alpha$ und damit $\delta(K, L) = \alpha$, was zu zeigen war.

Dass δ eine Metrik auf $\mathcal{C}^n$ ist, ist leicht zu sehen. Zum Nachweis der Dreiecksungleichung (die anderen Eigenschaften einer Metrik sind trivialerweise erfüllt) seien $K, L, M \in \mathcal{C}^n$. Wir setzen $\delta(K, L) = \alpha$ und $\delta(L, M) = \beta$. Dann gilt $K \subset L + \alpha B^n$ und $L \subset M + \beta B^n$, also $K \subset M + \beta B^n + \alpha B^n = M + (\alpha + \beta) B^n$. Analog gilt $M \subset K + (\alpha + \beta) B^n$ und daher $\delta(K, M) \leq \alpha + \beta$.

Im folgenden beziehen sich alle metrischen und topologischen Begriffe, die im Zusammenhang mit $\mathcal{C}^n$ und $\mathcal{K}^n$ verwendet werden, auf die Hausdorff-Metrik und die davon induzierte Topologie, auch wenn dies nicht ausdrücklich gesagt wird.

Wir wollen zeigen, dass in dem metrischen Raum $(\mathcal{C}^n, \delta)$ jede abgeschlossene, beschränkte Menge kompakt ist. Insbesondere ist dieser Raum also lokalkompakt.

1.8.1 Hilfssatz. *Ist $(K_i)_{i\in\mathbb{N}}$ eine abnehmende Folge in $\mathcal{C}^n$, d.h. gilt $K_{i+1} \subset K_i$ für $i \in \mathbb{N}$, so ist*

$$\lim_{i\to\infty} K_i = \bigcap_{j=1}^{\infty} K_j.$$

Beweis. Die Menge $K := \bigcap_{j=1}^{\infty} K_j$ ist kompakt und nicht leer. Wäre die Behauptung falsch, so gäbe es ein $\epsilon > 0$ und unendlich viele $i \in \mathbb{N}$ mit $\delta(K_i, K) > \epsilon$. Wegen $K \subset K_i$ gilt also $K_i \not\subset K + \epsilon B^n$ für diese i, und wegen $K_{i+1} \subset K_i$ gilt sogar $K_i \not\subset K + \epsilon B^n$ für alle i. Setze

$$A_i := K_i \setminus \operatorname{int}(K + \epsilon B^n),$$

dann ist $(A_i)_{i\in\mathbb{N}}$ eine abnehmende Folge von nichtleeren kompakten Mengen, hat also nichtleeren Durchschnitt A. Es gilt $A \cap K = \emptyset$, aber aus $A_i \subset K_i$ folgt $A \subset K$, ein Widerspruch. ∎

1.8.2 Satz. *Der metrische Raum $(\mathcal{C}^n, \delta)$ ist vollständig.*

Beweis. Sei $(K_i)_{i\in\mathbb{N}}$ eine Cauchy-Folge in $\mathcal{C}^n$. Setze $A_m := \operatorname{clos} \bigcup_{i=m}^{\infty} K_i$. Dann ist $(A_m)_{m\in\mathbb{N}}$ eine abnehmende Folge von nichtleeren kompakten Mengen (aus der Cauchy-Eigenschaft ergibt sich sofort die Beschränktheit von $\bigcup K_i$). Aus Hilfssatz 1.8.1 folgt also $A_m \to A := \bigcap_{j\in\mathbb{N}} A_j$ für $m \to \infty$. Zu gegebenem $\epsilon > 0$ existiert also ein $n_0 \in \mathbb{N}$ mit $A_m \subset A + \epsilon B^n$ für $m \geq n_0$, also mit $K_i \subset A + \epsilon B^n$ für $i \geq n_0$. Da $(K_i)_{i\in\mathbb{N}}$ eine Cauchy-Folge ist, gibt es ein $n_1 \geq n_0$ mit $K_j \subset K_i + \epsilon B^n$ für $i, j \geq n_1$. Für $i, m \geq n_1$ gilt also $\bigcup_{j=m}^{\infty} K_j \subset K_i + \epsilon B^n$ und daher $A_m \subset K_i + \epsilon B^n$, woraus $A \subset K_i + \epsilon B^n$ folgt. Damit ist $\delta(K_i, A) \leq \epsilon$ für $i \geq n_1$ gezeigt, woraus sich die Behauptung ergibt. ∎

1.8.3 Satz. *In $\mathcal{C}^n$ besitzt jede beschränkte Folge eine konvergente Teilfolge.*

Beweis. Sei $(K_i^0)_{i\in\mathbb{N}}$ eine beschränkte Folge in $\mathcal{C}^n$. Dann gibt es eine beschränkte Menge $C \subset \mathbb{R}^n$ mit $K_i^0 \subset C$ für alle $i \in \mathbb{N}$. Zu gegebenem $\epsilon > 0$ gibt es eine endliche Menge $E \subset C$ mit $C \subset E + \epsilon B^n$. Für $K \in \mathcal{C}^n$ mit $K \subset C$ sei $E(K) := (K + \epsilon B^n) \cap E$. Dann gilt

$$\delta(K, E(K)) \leq \epsilon. \tag{1.6}$$

(Zu $x \in K$ existiert nämlich wegen $K \subset E + \epsilon B^n$ ein $e \in E$ mit $\|x - e\| \leq \epsilon$, es ist $e \in K + \epsilon B^n$, also $e \in E(K)$. Somit ist $K \subset E(K) + \epsilon B^n$, und nach Definition gilt $E(K) \subset K + \epsilon B^n$.)

Wir wählen nun zu $\epsilon = 1/m$ ($m \in \mathbb{N}$) eine derartige Menge $E = E_m$ und setzen $E_m(K) := \left(K + \frac{1}{m} B^n\right) \cap E_m$.

Da E_1 endlich ist, gibt es unendlich viele Indizes $i \in \mathbb{N}$, für die $E_1\left(K_i^0\right)$ dieselbe Teilmenge von E_1 ist. Die Folge $(K_i^0)_{i\in\mathbb{N}}$ hat also eine Teilfolge $(K_i^1)_{i\in\mathbb{N}}$ derart, dass

$$E_1\left(K_i^1\right) =: T_1$$

unabhängig von i ist. Analog hat $(K_i^1)_{i\in\mathbb{N}}$ eine Teilfolge $(K_i^2)_{i\in\mathbb{N}}$ derart, dass

$$E_2\left(K_i^2\right) =: T_2$$

unabhängig von i ist. Auf diese Weise kann man rekursiv eine Folge $(T_m)_{m\in\mathbb{N}}$ von endlichen Mengen $T_m \subset E_m$ definieren und zu jedem m eine Folge $(K_i^m)_{i\in\mathbb{N}}$ mit

$$E_m\left(K_i^m\right) = T_m \qquad \text{für } i \in \mathbb{N} \tag{1.7}$$

und derart, dass gilt:

$$\text{Für } k > m \text{ ist } \left(K_i^k\right)_{i\in\mathbb{N}} \text{ Teilfolge von } \left(K_i^m\right)_{i\in\mathbb{N}}. \tag{1.8}$$

Nach (1.6), (1.7), (1.8) gilt

$$\delta\left(K_i^k, T_m\right) \leq \frac{1}{m} \qquad \text{für } k \geq m \text{ und } i \in \mathbb{N},$$

also

$$\delta\left(K_i^m, K_j^k\right) \leq \frac{2}{m} \qquad \text{für } k \geq m \text{ und } i, j \in \mathbb{N}.$$

Für $K_k := K_k^k$ folgt

$$\delta\left(K_m, K_k\right) \leq \frac{2}{m} \qquad \text{für } k \geq m.$$

Die Folge $(K_m)_{m\in\mathbb{N}}$ ist also eine Cauchy-Folge und daher nach Satz 1.8.2 konvergent. Außerdem ist sie eine Teilfolge von $(K_i^0)_{i\in\mathbb{N}}$. ■

Nach einem bekannten Kompaktheitskriterium in metrischen Räumen folgt aus Satz 1.8.3:

1.8.4 Satz. *In $\mathcal{C}^n$ ist jede abgeschlossene beschränkte Menge kompakt.*

Jetzt wollen wir uns wieder auf die Menge $\mathcal{K}^n \subset \mathcal{C}^n$ der konvexen Körper beschränken.

1.8.5 Satz. *$\mathcal{K}^n$ ist eine abgeschlossene Teilmenge von $\mathcal{C}^n$.*

Beweis. Sei $K \in \mathcal{C}^n \setminus \mathcal{K}^n$. Dann gibt es Punkte $x, y \in K$ und Zahlen $\lambda \in (0,1)$, $\epsilon > 0$ mit $B(z,\epsilon) \cap K = \emptyset$ für $z := (1-\lambda)x + \lambda y$. Sei $K' \in \mathcal{C}^n$ eine Menge mit $\delta(K,K') < \epsilon/2$. Es gibt Punkte $x', y' \in K'$ mit $\|x' - x\| < \epsilon/2$ und $\|y' - y\| < \epsilon/2$, für den Punkt $z' := (1-\lambda)x' + \lambda y'$ gilt also $\|z' - z\| < \epsilon/2$. Wäre $z' \in K'$, so gäbe es einen Punkt $w \in K$ mit $\|w - z'\| < \epsilon/2$, also mit $\|w - z\| < \epsilon$, ein Widerspruch. Also ist K' nicht konvex. Damit ist gezeigt, dass $\mathcal{C}^n \setminus \mathcal{K}^n$ in $\mathcal{C}^n$ offen ist, also ist $\mathcal{K}^n$ abgeschlossen. ■

Die Sätze 1.8.3 und 1.8.5 zusammen ergeben den *Auswahlsatz von Blaschke*:

1.8.6 Satz. *Aus jeder beschränkten Folge konvexer Körper kann man eine Teilfolge auswählen, die gegen einen konvexen Körper konvergiert.*

Dieser Satz wird oft verwendet, um die Existenz von konvexen Körpern mit bestimmten Eigenschaften nachzuweisen.

Es ist manchmal nützlich, eine Beschreibung der Konvergenz von Folgen konvexer Körper zu haben, die nur die Konvergenz von Punktfolgen verwendet.

1.8.7 Satz. *Die Konvergenz $\lim_{i\to\infty} K_i = K$ in $\mathcal{K}^n$ gilt genau dann, wenn die folgenden beiden Bedingungen* (a) *und* (b) *gelten*:

(a) *Jeder Punkt in K ist Limes einer Folge $(x_i)_{i\in\mathbb{N}}$ mit $x_i \in K_i$ für $i \in \mathbb{N}$.*

(b) *Der Limes jeder konvergenten Folge $(x_{i_j})_{j\in\mathbb{N}}$ mit $x_{i_j} \in K_{i_j}$ für $j \in \mathbb{N}$ gehört zu K.*

Beweis. Es gelte $K_i \to K$ für $i \to \infty$. Sei $x \in K$. Setze $x_i := p(K_i, x)$. Dann gilt $x_i \in K_i$ und $\|x - x_i\| \leq \delta(K, K_i) \to 0$ für $i \to \infty$. Also gilt (a). Sei $(i_j)_{j\in\mathbb{N}}$ eine Folge in $\mathbb{N}$, und gelte $x_{i_j} \in K_{i_j}$ $(j \in \mathbb{N})$ sowie $x_{i_j} \to x$ für $j \to \infty$. Angenommen, es wäre $x \notin K$. Dann gilt $B(x,\rho) \cap (K + \rho B^n) = \emptyset$ für geeignetes $\rho > 0$. Für genügend großes j gilt aber $\|x_{i_j} - x\| < \rho$ und $x_{i_j} \in K_{i_j} \subset K + \rho B^n$, ein Widerspruch. Also gilt (b).

Umgekehrt seien jetzt (a) und (b) erfüllt. Sei $\epsilon > 0$. Wir müssen zeigen, dass

$$K \subset K_i + \epsilon B^n \qquad \text{für alle genügend großen } i, \tag{1.9}$$

$$K_i \subset K + \epsilon B^n \qquad \text{für alle genügend großen } i \tag{1.10}$$

gilt. Ist (1.9) falsch, so gibt es eine Folge $(y_{i_j})_{j\in\mathbb{N}}$ in K, die gegen ein $y \in K$ konvergiert und $d(K_{i_j}, y_{i_j}) \geq \epsilon$ $(j \in \mathbb{N})$ erfüllt. Nach (a) gibt es Punkte $x_i \in K_i$ mit $x_i \to y$ für $i \to \infty$. Daraus folgt $\|x_{i_j} - y_{i_j}\| \to 0$ für $j \to \infty$, ein Widerspruch. Also gilt (1.9). Angenommen, (1.10) wäre falsch. Dann gibt es eine Folge $(y_{i_j})_{j\in\mathbb{N}}$

mit $y_{i_j} \in K_{i_j}$ und $y_{i_j} \notin K + \epsilon B^n$ für $j \in \mathbb{N}$. Nach (a) gibt es Punkte $x_{i_j} \in K_{i_j}$ mit $x_{i_j} \in K + \epsilon B^n$ für alle genügend großen j. Wegen der Konvexität von K_{i_j} gibt es Punkte $z_{i_j} \in K_{i_j} \cap \mathrm{bd}\,(K + \epsilon B^n)$. Die Folge $(z_{i_j})_{j \in \mathbb{N}}$ hat eine konvergente Teilfolge; ihr Limes z muss wegen (b) zu K gehören, ein Widerspruch zu $z \in \mathrm{bd}\,(K + \epsilon B^n)$. Also gilt (1.10). ■

Nachdem wir $\mathcal{C}^n$ und $\mathcal{K}^n$ mit einer Metrik und der induzierten Topologie versehen haben, ist die Frage nach der Stetigkeit einiger früher betrachteten Abbildungen sinnvoll. Einige dieser Abbildungen sind sogar Lipschitz-stetig, aus trivialen Gründen. Dies gilt zum Beispiel für die Bildung der konvexen Hülle als Abbildung von $\mathcal{C}^n$ nach $\mathcal{K}^n$, denn es ist

$$\delta(\mathrm{conv}\, K, \mathrm{conv}\, L) \leq \delta(K, L) \qquad \text{für } K, L \in \mathcal{C}^n.$$

Es gilt auch für die Addition und die Vereinigungsbildung als Abbildungen von $\mathcal{C}^n \times \mathcal{C}^n$ nach $\mathcal{C}^n$, denn es ist

$$\begin{aligned} \delta(K + K', L + L') &\leq \delta(K, L) + \delta(K', L'), \\ \delta(K \cup K', L \cup L') &\leq \max\{\delta(K, L), \delta(K', L')\}. \end{aligned}$$

Die Durchschnittsbildung ist allerdings nicht stetig. Für konvexe Körper lässt sich jedoch folgendes zeigen.

1.8.8 Satz. *Seien $K, L \in \mathcal{K}^n$ konvexe Körper, die nicht durch eine Hyperebene getrennt werden können. Sind $K_i, L_i \in \mathcal{K}^n$ $(i \in \mathbb{N})$ konvexe Körper mit $K_i \to K$ und $L_i \to L$ für $i \to \infty$, so gilt $K_i \cap L_i \neq \emptyset$ für fast alle i und $K_i \cap L_i \to K \cap L$ für $i \to \infty$.*

Beweis. Wir verwenden das Kriterium 1.8.7. Sei $x \in K \cap L$ (es gilt $K \cap L \neq \emptyset$, denn andernfalls könnten K und L getrennt werden). Setze $x_i := p(K_i \cap L_i, x)$ für die i mit $K_i \cap L_i \neq \emptyset$. Wir behaupten, dass letzteres für fast alle i gilt und dass $x_i \to x$ für $i \to \infty$ gilt. Angenommen, dies wäre falsch. Dann gibt es eine Kugel B mit Mittelpunkt x, so dass $K_i \cap L_i \cap B = \emptyset$ für unendlich viele i gilt. Für genügend große i ist $K_i \cap B \neq \emptyset$ wegen $x \in K$ und $K_i \to K$; analog ist $L_i \cap B \neq \emptyset$ für große i. Daher gibt es eine Folge $(H_{i_j})_{j \in \mathbb{N}}$ von Hyperebenen H_{i_j}, die $K_{i_j} \cap B$ und $L_{i_j} \cap B$ trennen. Eine Teilfolge konvergiert gegen eine Hyperebene H, und H trennt $K \cap B$ und $L \cap B$ (wie man leicht mit Verwendung von Satz 1.8.7 zeigen kann). Wegen $x \in K \cap L \cap B$ gilt $x \in H$. Aus der Konvexität von K und L folgt nun, dass H die Körper K und L trennt, ein Widerspruch. Es folgt $x_i \to x$ für $i \to \infty$.

Ist andererseits $(i_j)_{j \in \mathbb{N}}$ eine Folge in $\mathbb{N}$ und gilt $x_{i_j} \in K_{i_j} \cap L_{i_j}$ $(j \in \mathbb{N})$ und $x_{i_j} \to y$ für $j \to \infty$, so gilt $y \in K$ nach Satz 1.8.7 und ebenso $y \in L$, also $y \in K \cap L$. Jetzt folgt aus Satz 1.8.7, dass $K_i \cap L_i \to K \cap L$ für $i \to \infty$ gilt. ■

Das folgende Beispiel zeigt, dass die Voraussetzung der Nicht-Trennbarkeit in Satz 1.8.8 nicht entbehrlich ist. Sei $K = L = L_i = [x, y]$ mit $x \neq y$, und sei $K_i = [x, y_i]$ eine Strecke mit $[x, y_i] \cap [x, y] = \{x\}$ und $y_i \to y$ für $i \to \infty$. Dann gilt $K_i \to K$ und $L_i \to L$, aber nicht $K_i \cap L_i \to K \cap L$.

Nun zeigen wir die Stetigkeit der Abbildungen

$$\begin{array}{rccc} p: & \mathcal{K}^n \times \mathbb{E}^n & \to & \mathbb{E}^n, \\ & (K, x) & \mapsto & p(K, x) \end{array} \qquad \begin{array}{rccc} h: & \mathcal{K}^n \times \mathbb{E}^n & \to & \mathbb{E}^n, \\ & (K, x) & \mapsto & h(K, x) \end{array}$$

also der metrischen Projektion und der Stützfunktion, jeweils als Funktion von zwei Veränderlichen. Die folgenden beiden Sätze zeigen die Stetigkeit in verschärfter Form.

1.8.9 Satz. *Sei $K, L \in \mathcal{K}^n$ und $x, y \in \mathbb{E}^n$, setze $D := \operatorname{diam}(K \cup L \cup \{x, y\})$. Dann gilt*

$$\|p(K, x) - p(L, y)\| \leq \|x - y\| + \sqrt{5D\delta(K, L)}.$$

Beweis. Wegen Satz 1.2.2 und der Dreiecksungleichung brauchen wir nur die Ungleichung

$$\|p(K, x) - p(L, x)\| \leq \sqrt{5D\delta(K, L)}$$

nachzuweisen. Wenn $x \in K \cap L$ gilt, ist das trivial. Sei also etwa $x \notin K$. Setze $d(K, x) =: d$ und $\delta(K, L) =: \delta$. Die Kugel $B(p(K, x), \delta)$ enthält einen Punkt aus L, also gilt $d(L, x) \leq d + \delta$ und daher $p(L, x) \in B(x, d + \delta)$. Sei H^- der Stützhalbraum von K mit äußerem Normalenvektor $u(K, x)$. Dann ist L enthalten in dem Translat $H^- + \delta u(K, x)$. Daher gilt

$$p(L, x) \in B(x, d + \delta) \cap [H^- + \delta u(K, x)].$$

Es folgt

$$\|p(K, x) - p(L, x)\| \leq \|p(K, x) - z\|,$$

wenn

$$z \in \operatorname{bd} B(x, d + \delta) \cap \operatorname{bd} [H^- + \delta u(K, x)]$$

gewählt wird. Bezeichnet c den Abstand des Punktes z von der Verbindungsstrecke $[x, p(K, x)]$, so ergibt zweimalige Anwendung des Satzes von Pythagoras

$$\begin{aligned} \|p(K, x) - z\|^2 &= \delta^2 + c^2 = \delta^2 + (d + \delta)^2 - (d - \delta)^2 \\ &= \delta^2 + 4d\delta \leq 5D\delta \end{aligned}$$

wegen $d \leq D$ und $\delta \leq D$. ∎

1.8.10 Satz. *Für $K, L \in \mathcal{K}^n$ mit $K, L \subset RB^n$ $(R > 0)$ und für $u, v \in \mathbb{E}^n$ gilt*

$$|h(K,u) - h(L,v)| \le R\|u - v\| + \max\{\|u\|, \|v\|\}\, \delta(K,L).$$

Beweis. Wir setzen $\delta(K,L) =: \delta$ und wählen einen Punkt $x \in K$ mit $h(K,u) = \langle x, u\rangle$. Aus $x \in K \subset L + \delta B^n$ ergibt sich $\langle x, v\rangle \le h(L + \delta B^n, v) = h(L,v) + \delta\|v\|$, also $h(K,u) - h(L,v) \le \langle x, u - v\rangle + \delta\|v\| \le \|x\|\,\|u - v\| + \delta\|v\| \le R\|u - v\| + \delta\|v\|$. Vertauschung von (K,u) und (L,v) ergibt jetzt die Behauptung. ■

Wenn man die Stützfunktion verwendet, lässt sich die Konvergenz konvexer Körper sehr bequem behandeln, aufgrund der folgenden einfachen Beobachtung. Hierin bezeichnet $\|\cdot\|$ für Funktionen auf der Sphäre S^{n-1} die Maximumsnorm. Für eine Funktion $h : \mathbb{E}^n \to \mathbb{R}$ bezeichnet $\overline{h} := h|S^{n-1}$ die Einschränkung auf S^{n-1}.

1.8.11 Satz. *Für $K, L \in \mathcal{K}^n$ gilt*

$$\delta(K,L) = \sup_{u \in S^{n-1}} |h(K,u) - h(L,u)| =: \|\overline{h}_K - \overline{h}_L\|.$$

Beweis. Sei $\delta(K,L) \le \alpha$. Dann gilt $K \subset L + \alpha B^n$, also $h(K,u) \le h(L + \alpha B^n, u) = h(L,u) + \alpha$ für $u \in S^{n-1}$. Vertauschung von K und L ergibt $|h(K,u) - h(L,u)| \le \alpha$ für $u \in S^{n-1}$ und daher $\|\overline{h}_K - \overline{h}_L\| \le \alpha$. Das Argument kann umgekehrt werden. ■

Satz 1.8.11 zeigt, dass die Konvergenz konvexer Körper äquivalent ist mit der auf S^{n-1} gleichmäßigen Konvergenz der entsprechenden Stützfunktionen. Wir zeigen nun, dass diese gleichmäßige Konvergenz schon aus der punktweisen Konvergenz gefolgert werden kann.

1.8.12 Satz. *Ist eine Folge von Stützfunktionen auf S^{n-1} punktweise konvergent, so konvergiert sie auf S^{n-1} gleichmäßig gegen eine Stützfunktion.*

Beweis. Sei $(h(K_i, \cdot))_{i \in \mathbb{N}}$ eine Folge von Stützfunktionen, die auf S^{n-1} und daher wegen der Homogenität auf ganz $\mathbb{E}^n$ punktweise konvergiert. Die Grenzfunktion h ist offenbar sublinear und daher nach Satz 1.7.1 eine Stützfunktion. Für jedes $u \in S^{n-1}$ ist $\alpha(u) := \sup_i h(K_i, u)$ endlich. Wir können eine Zahl $R > 0$ wählen mit

$$\bigcap_{u \in S^{n-1}} H^-_{u,\alpha(u)} \subset RB^n,$$

dann gilt $K_i \subset RB^n$ für $i \in \mathbb{N}$. Sei h_i die Einschränkung von $h(K_i, \cdot)$ auf S^{n-1}; dann gilt $|h_i(u) - h_i(v)| \le R\|u - v\|$ nach Satz 1.8.10. Sei $\epsilon > 0$ gegeben. Wir wählen eine endliche Menge $S \subset S^{n-1}$ derart, dass zu jedem $u \in S^{n-1}$ ein $v \in S$ existiert mit $\|u - v\| < \epsilon/3R$. Da S endlich ist, gibt es ein $n_0 \in \mathbb{N}$ mit $|h_i(v) - h_j(v)| < \epsilon/3$ für $i, j \geq n_0$ und alle $v \in S$. Nun sei $u \in S^{n-1}$. Wir können ein $v \in S$ wählen mit $\|u - v\| < \epsilon/3R$. Dann folgt

$$\begin{aligned} |h_i(u) - h_j(u)| &\leq |h_i(u) - h_i(v)| + |h_i(v) - h_j(v)| + |h_j(v) - h_j(u)| \\ &\leq R\|u - v\| + \epsilon/3 + R\|u - v\| < \epsilon \end{aligned}$$

für $i, j \geq n_0$, also gilt $|h_i(u) - h(u)| \leq \epsilon$ für $i \geq n_0$. Da $u \in S^{n-1}$ beliebig war, ist damit die gleichmäßige Konvergenz auf S^{n-1} der Folge $(h_i)_{i \in \mathbb{N}}$ gegen h gezeigt. ■

Da wir auf $\mathcal{K}^n$ eine Metrik zur Verfügung haben, können wir jetzt die Approximation konvexer Körper behandeln. Die Approximation allgemeiner konvexer Körper durch spezielle, wie Polytope, ist ein oft benutztes Beweishilfsmittel.

1.8.13 Satz. *Sei $K \in \mathcal{K}^n$ und $\epsilon > 0$. Dann gibt es ein Polytop $P \in \mathcal{K}^n$ mit $P \subset K \subset P + \epsilon B^n$, insbesondere also mit $\delta(K, P) \leq \epsilon$.*

Beweis. Als kompakte Menge kann K überdeckt werden durch endlich viele Kugeln mit Radius ϵ und Mittelpunkten in K. Die konvexe Hülle der Mittelpunkte dieser Kugeln leistet das Gewünschte. ■

Bemerkung. Im vorstehenden Beweis kann man zusätzlich fordern, dass alle verwendeten Kugelmittelpunkte rationale Koordinaten haben. Es folgt, dass der metrische Raum $(\mathcal{K}^n, \delta)$ separabel ist, d.h. es gibt in ihm eine abzählbare dichte Teilmenge.

1.8.14 Hilfssatz. *Seien $K_1, K_2 \in \mathcal{K}^n$ konvexe Körper mit $K_2 \subset \operatorname{int} K_1$. Dann gibt es eine Zahl $\eta > 0$ derart, dass für jeden konvexen Körper $K \in \mathcal{K}^n$ mit $\delta(K_1, K) < \eta$ auch noch $K_2 \subset K$ gilt.*

Beweis. Da $K_2 \subset \operatorname{int} K_1$ gilt, ist die Funktion $h(K_1, \cdot) - h(K_2, \cdot)$ auf $\mathbb{E}^n \setminus \{0\}$ positiv, sie nimmt also auf S^{n-1} ein positives Minimum η an. Sei $K \in \mathcal{K}^n$ ein konvexer Körper mit $\delta(K_1, K) < \eta$. Dann gilt $|h(K_1, u) - h(K, u)| < \eta$ und daher $h(K_2, u) \leq h(K_1, u) - \eta < h(K, u)$ für $u \in S^{n-1}$; es ist also $K_2 \subset K$. ■

1.8.15 Satz. *Sei $K \in \mathcal{K}^n$ und $0 \in \operatorname{int} K$. Zu jedem $\lambda > 1$ existiert ein Polytop $P \in \mathcal{K}^n$ mit $P \subset K \subset \lambda P$.*

Beweis. Wir wählen eine Zahl $\rho > 0$ mit $B(0, \rho) \subset \operatorname{int} K$ und eine Zahl ϵ mit $0 < \epsilon \leq (\lambda - 1)\rho$. Nach Satz 1.8.13 gibt es ein Polytop $P \subset K$ mit $\delta(K, P) \leq \epsilon$. Nach Hilfssatz 1.8.14 gilt, wenn ϵ genügend klein gewählt wird, auch noch $B(0, \rho) \subset P$. Dann ist also

$$B(0, \rho) \subset P \subset K \subset P + \epsilon B^n.$$

Für $u \in S^{n-1}$ folgt dann

$$\begin{aligned} h(\lambda P, u) &= h(P,u) + (\lambda - 1) h(P,u) \\ &\geq h(P,u) + \epsilon = h(P + \epsilon B^n, u) \\ &\geq h(K,u), \end{aligned}$$

also $K \subset \lambda P$. ■

Nach diesen Vorbereitungen können wir die Stetigkeit des Volumens auf $\mathcal{K}^n$ zeigen. Das *Volumen-Funktional* (oder kurz *Volumen*) auf $\mathcal{K}^n$ ist dadurch erklärt, dass $V_n(K)$ das Lebesgue-Maß von $K \in \mathcal{K}^n$ ist.

1.8.16 Satz. *Das Volumen-Funktional V_n ist stetig auf $\mathcal{K}^n$.*

Beweis. Sei $K \in \mathcal{K}^n$ gegeben. Zunächst sei $V_n(K) = 0$; dann liegt K in einer Hyperebene. Ist $\overline{K} \in \mathcal{K}^n$ ein konvexer Körper mit $\delta(K, \overline{K}) = \alpha \leq 1$, so gilt $\overline{K} \subset K + \alpha B^n$, also $V_n(\overline{K}) \leq V_n(K + \alpha B^n) \leq c(K)\alpha$ mit einer von α unabhängigen Zahl $c(K)$, wie der Satz von Fubini und eine einfache Abschätzung ergeben. Daraus folgt die Stetigkeit von V_n bei K.

Sei jetzt $V_n(K) > 0$. Wir können $0 \in \operatorname{int} K$ annehmen. Sei $\epsilon > 0$ gegeben. Wir wählen $\lambda > 1$ mit $(\lambda^n - 1)\lambda^n V_n(K) < \epsilon$ und ein $\rho > 0$ mit $\rho B^n \subset \operatorname{int} K$. Nach Hilfssatz 1.8.14 gibt es eine Zahl $\alpha > 0$ mit $\alpha \leq (\lambda - 1)\rho$ und derart, dass für alle $\overline{K} \in \mathcal{K}^n$ mit $\delta(K, \overline{K}) < \alpha$ noch $\rho B^n \subset \overline{K}$ gilt. Wenn das erfüllt ist, folgt

$$K \subset \overline{K} + \alpha B^n \subset \overline{K} + (\lambda - 1)\rho B^n \subset \overline{K} + (\lambda - 1)\overline{K} = \lambda \overline{K}$$

und analog $\overline{K} \subset \lambda K$. Es folgt

$$V_n(K) \leq V_n(\lambda \overline{K}) = \lambda^n V_n(\overline{K}),$$

also

$$\begin{aligned} V_n(K) - V_n(\overline{K}) &\leq (\lambda^n - 1) V_n(\overline{K}) \leq (\lambda^n - 1)\lambda^n V_n(K), \\ V_n(\overline{K}) - V_n(K) &\leq (\lambda^n - 1) V_n(K) \leq (\lambda^n - 1)\lambda^n V_n(K) \end{aligned}$$

und daher

$$|V_n(K) - V_n(\overline{K})| \leq (\lambda^n - 1)\lambda^n V_n(K) < \epsilon.$$

Daraus ergibt sich die Stetigkeit. ■

Man beachte, dass das (analog erklärte) Volumen-Funktional auf $\mathcal{C}^n$ keineswegs stetig ist: Jede kompakte Menge ist Limes einer Folge endlicher Mengen, und diese haben Volumen Null.

Kapitel 2

Randstruktur und Polytope

2.1 Seitenstruktur

In Abschnitt 1.4 wurden bereits die Begriffe Seite, Extrempunkt und exponierter Punkt einer konvexen Menge definiert. Sie gehören zu einer Reihe von ähnlichen Begriffen, die dazu dienen sollen, die Randstruktur einer konvexen Menge näher zu beschreiben. Eine derartige Beschreibung soll in diesem Kapitel erfolgen. In diesem Abschnitt ist $K \subset \mathbb{E}^n$ eine nichtleere abgeschlossene konvexe Menge.

Zunächst soll an die Definition einer Seite erinnert werden. Eine *Seite* von K ist eine konvexe Teilmenge $F \subset K$ mit der Eigenschaft, dass aus $[x, y] \subset K$ und $F \cap \operatorname{relint}[x, y] \neq \emptyset$ stets $[x, y] \subset F$ folgt. Äquivalent damit ist, dass aus $x, y \in K$ und $\frac{1}{2}(x + y) \in F$ stets $x, y \in F$ folgt. Eine Seite hat als konvexe Menge eine wohlbestimmte Definition. Eine i-dimensionale Seite wird kurz als *i-Seite* bezeichnet. $\mathcal{F}_i(K)$ sei die Menge aller i-Seiten von K und $\mathcal{F}(K)$ die Menge aller Seiten von K. Auch die leere Menge $\emptyset$ und K sind Seiten von K; die übrigen Seiten heißen *eigentliche Seiten*. Es folgt aus der Definition einer Seite und aus Hilfssatz 1.1.8, dass Seiten stets abgeschlossen sind. Ist $F \neq K$ eine Seite von K, so gilt $F \cap \operatorname{relint} K = \emptyset$. (Wäre nämlich $z \in F \cap \operatorname{relint} K$, so wählen wir ein $y \in K \setminus F$. Es gibt einen Punkt $x \in K$ mit $z \in \operatorname{relint}[x, y]$. Dann folgt $[x, y] \subset F$, ein Widerspruch.) Insbesondere gilt dann also $F \subset \operatorname{relbd} K$ und $\dim F < n$.

2.1.1 Satz. *Ist F_i eine Seite von K für $i \in I$ ($I \neq \emptyset$ eine Indexmenge), so ist $\bigcap_{i \in I} F_i$ eine Seite von K. Ist F Seite von K und G Seite von F, so ist G Seite von K.*

Beweis. Das folgt unmittelbar aus der Definition einer Seite.

2.1.2 Satz. *Sind $F_1, F_2 \in \mathcal{F}(K)$ verschiedene Seiten, so gilt* $\operatorname{relint} F_1 \cap \operatorname{relint} F_2 = \emptyset$. *Zu jeder nichtleeren relativ offenen konvexen Teilmenge A von K gibt es eine* *eindeutig bestimmte Seite $F \in \mathcal{F}(K)$ mit $A \subset \operatorname{relint} F$. Das System*

$$\{\operatorname{relint} F : F \in \mathcal{F}(K)\}$$

ist eine disjunkte Zerlegung von K.

Beweis. Angenommen, es wäre $F_1, F_2 \in \mathcal{F}(K)$, $F_1 \neq F_2$, aber $z \in \operatorname{relint} F_1 \cap \operatorname{relint} F_2$. Es gibt etwa einen Punkt $x \in F_1 \setminus F_2$. Wir können ein $y \in F_1$ wählen mit $z \in \operatorname{relint}[x, y]$. Aus $[x, y] \subset F_1 \subset K$, $z \in F_2$ und $F_2 \in \mathcal{F}(K)$ folgt dann $[x, y] \subset F_2$, ein Widerspruch. Also ist $\operatorname{relint} F_1 \cap \operatorname{relint} F_2 = \emptyset$.

Sei $A \neq \emptyset$ eine relativ offene konvexe Teilmenge von K. Sei F der Durchschnitt aller die Menge A enthaltenden Seiten von K; dann ist F eine Seite von K. Angenommen, es gäbe einen Punkt $x \in A \setminus \operatorname{relint} F$. Dann gibt es eine Stützebene H an F mit $x \in H$ und $H \cap F \neq F$. Da $x \in A \subset F$ gilt und A relativ offen ist, folgt $A \subset H$. Nun ist $H \cap F$ offenbar eine Seite von F und daher eine Seite von K. Aber dann folgt $F \subset H \cap F$ wegen der Definition von F, ein Widerspruch. Somit gilt $A \subset \operatorname{relint} F$. Aus dem ersten Teil folgt, dass F die einzige Seite mit dieser Eigenschaft ist.

Die Menge A kann insbesondere einpunktig gewählt werden. Also gibt es zu jedem Punkt $x \in K$ eine eindeutig bestimmte Seite $F \in \mathcal{F}(K)$ mit $x \in \operatorname{relint} F$. ■

Im vorstehenden Beweis haben wir die offensichtliche Tatsache benutzt, dass jede Stützmenge $H \cap K$, wo H eine Stützebene von K ist, auch eine Seite von K ist. Die Stützmengen werden auch als *exponierte Seiten* bezeichnet. Die einpunktige Menge $\{x\}$ ist genau dann eine exponierte Seite, wenn x ein exponierter Punkt ist. Wir bezeichnen mit $\mathcal{S}_i(K)$ die Menge aller i-dimensionalen exponierten Seiten von K und mit $\mathcal{S}(K)$ die Menge aller exponierten Seiten von K. Da nicht jeder Extrempunkt ein exponierter Punkt ist, sieht man durch Bildung direkter Summen, dass auch höherdimensionale Seiten keine exponierten Seiten zu sein brauchen. Andererseits ist jede eigentliche Seite enthalten in einer exponierten Seite. Sei nämlich F eine eigentliche Seite von K. Wegen $F \cap \operatorname{relint} K = \emptyset$ gibt es eine Hyperebene H, die F und K trennt. Wegen $F \subset K$ ist $F \subset H$, also ist F enthalten in der exponierten Seite $H \cap K$.

2.1.3 Satz. *Ist F_i eine exponierte Seite von K für $i \in I$ ($I \neq \emptyset$ eine Indexmenge), so ist $\bigcap_{i \in I} F_i$ entweder leer oder eine exponierte Seite von K.*

Beweis. Sei $F := \bigcap_{i \in I} F_i \neq \emptyset$. Nach einer Translation können wir $0 \in F$ annehmen. Zu jedem $i \in I$ existiert ein Vektor $u_i \in \mathbb{E}^n \setminus \{0\}$ mit $F_i = K \cap H_{u_i,0}$ und $K \subset H^-_{u_i,0}$. Wir können annehmen, dass die Vektoren $u_1, \ldots, u_r$ linear unabhängig sind und dass jeder andere Vektor u_i von diesen Vektoren linear abhängt. Dann gilt $F = \bigcap_{i=1}^r F_i$. Setze $u := \sum_{i=1}^r u_i$. Dann gilt $K \subset H^-_{u,0}$ und $0 \in H_{u,0}$, also ist $H_{u,0}$ eine Stützebene von K. Für $x \in F$ gilt $\langle x, u_i \rangle = 0$ für $i \in I$ und daher $\langle x, u \rangle = 0$; also gilt $F \subset H_{u,0} \cap K$. Für $y \in K \setminus F$ ist $\langle y, u_j \rangle < 0$ für ein $j \in \{1, \ldots, r\}$, also $\langle y, u \rangle < 0$ und damit $y \notin H_{u,0}$. Damit ist $H_{u,0} \cap K = F$ gezeigt, also ist $F \in \mathcal{S}(K)$. ∎

Man beachte, dass es zum zweiten Teil von Satz 2.1.1 kein Gegenstück für exponierte Seiten gibt: Eine exponierte Seite einer exponierten Seite von K braucht nicht selbst exponierte Seite von K zu sein (ist aber natürlich eine Seite von K).

2.2 Singularitäten

Durch jeden Randpunkt eines konvexen Körpers gibt es eine Stützebene, aber im allgemeinen nicht nur eine. Randpunkte, in denen mehr als eine Stützebene existiert, heißen singulär. Die singulären Punkte sollen in diesem Abschnitt näher untersucht werden.

Zunächst definieren wir konvexe Kegel, die zur Beschreibung singulärer Randpunkte dienen. Sei $K \in \mathcal{K}^n$ ein konvexer Körper, und sei $x \in K$. Wir definieren

$$P(K, x) := \{\lambda(y - x) : y \in K, \lambda > 0\} = \bigcup_{\lambda > 0} \lambda(K - x).$$

Dann ist $P(K, x)$ ein konvexer Kegel, und es gilt $P(K, x) = \operatorname{aff} K - x$ genau dann, wenn $x \in \operatorname{relint} K$ ist. Der abgeschlossene konvexe Kegel $S(K, x) := \operatorname{clos} P(K, x)$ kann offenbar dargestellt werden in der Form

$$S(K, x) = \bigcap_{x \in H(K,u)} H^-(K, u) - x,$$

wo der Durchschnitt erstreckt wird über alle $u \in \mathbb{E}^n \setminus \{0\}$ (oder $u \in S^{n-1}$), für die x in der Stützebene $H(K, u)$ liegt. (Die Darstellung gilt auch für $x \in \operatorname{int} K$, wenn wir wie üblich vereinbaren, dass der Durchschnitt einer leeren Familie von Teilmengen des $\mathbb{E}^n$ der ganze Raum $\mathbb{E}^n$ sein soll.) Der Kegel $S(K, x)$ heißt der *Stützkegel* von K bei x. Ferner definieren wir den *Normalenkegel* von K bei x durch

$$N(K, x) := \{u \in \mathbb{E}^n \setminus \{0\} : x \in H(K, u)\} \cup \{0\}.$$

Ist $x \in K \cap H(K, u)$, so heißt u ein *äußerer Normalenvektor* von K bei x. Somit besteht $N(K, x)$ aus allen äußeren Normalenvektoren von K bei x und dem Nullvektor. Es ist klar, dass $N(K, x)$ ein abgeschlossener konvexer Kegel ist und dass

$$N(K, x) = S(K, x)^* \tag{2.1}$$

gilt, d.h. $N(K,x)$ und $S(K,x)$ sind ein Paar von dualen konvexen Kegeln. Der Normalenkegel hängt auch eng mit der metrischen Projektion zusammen. In der Tat folgt aus Hilfssatz 1.3.1 die Gleichung

$$N(K,x) = p_K^{-1}(x) - x$$

für $x \in K$.

Der folgende Satz beschreibt das Verhalten von Stützkegeln und Normalenkegeln unter Minkowski-Addition und Durchschnittsbildung von konvexen Körpern.

2.2.1 Satz. *Sei $K, L \in \mathcal{K}^n$.*

(a) *Ist $x \in K$ und $y \in L$, so gilt*

$$\begin{aligned} S(K+L, x+y) &= S(K,x) + S(L,y), \\ N(K+L, x+y) &= N(K,x) \cap N(L,y). \end{aligned}$$

(b) *Ist $x \in K \cap L$ und* $\operatorname{relint} K \cap \operatorname{relint} L \neq \emptyset$, *so gilt*

$$\begin{aligned} S(K \cap L, x) &= S(K,x) \cap S(L,x), \\ N(K \cap L, x) &= N(K,x) + N(L,x). \end{aligned}$$

Beweis. (a) Nach Ausübung von Translationen können wir o.B.d.A. annehmen, dass $x = y = 0$ gilt. Sei $u \in N(K+L, 0) \setminus \{0\}$. Dann gilt $K + L \subset H^-_{u,0}$. Aus $K = K + \{0\} \subset H^-_{u,0}$ und $0 \in K$ folgt $u \in N(K,0)$. Analog gilt $u \in N(L,0)$.

Sei $u \in N(K,0) \cap N(L,0) \setminus \{0\}$. Aus $u \in N(K,0)$ folgt $K \subset H^-_{u,0}$; analog ist $L \subset H^-_{u,0}$ und daher $K+L \subset H^-_{u,0}$. Wegen $0 \in K+L$ erhalten wir $u \in N(K+L,0)$. Damit ist $N(K+L,0) = N(K,0) \cap N(L,0)$ gezeigt.

Aus (2.1) und Satz 1.6.3 folgt jetzt $S(K+L,0) = S(K,0) + S(L,0)$.

(b) Sei $x \in K \cap L$ und $y \in \operatorname{relint} K \cap \operatorname{relint} L$. Die erste Gleichung von (b) ist klar im Fall $x = y$; wir nehmen daher $x \neq y$ an. Sei $z \in S(K,x) \cap S(L,x)$. Wegen $y - x \in \operatorname{relint} K - x \subset \operatorname{relint} P(K,x)$ und $z \in \operatorname{clos} P(K,x)$ folgt aus Hilfssatz 1.1.8, dass $[y-x, z) \subset \operatorname{relint} P(K,x)$ gilt. Daher existiert zu $w \in [y-x,z)$ eine Zahl $\lambda > 0$ mit $[x, x+\lambda w] \subset K$. Analog gibt es eine Zahl $\mu > 0$ mit $[x, x+\mu w] \subset L$. Es folgt $w \in P(K \cap L, x)$. Da $w \in [y-x,z)$ beliebig war, folgt $z \in \operatorname{clos} P(K \cap L, x) = S(K \cap L, x)$. Also gilt $S(K,x) \cap S(L,x) \subset S(K \cap L, x)$. Die entgegengesetzte Inklusion ist trivial. Dies beweist die erste Gleichung von (b). Die zweite Gleichung ergibt sich dann wieder aus (2.1) und Satz 1.6.3. ■

Sei $K \in \mathcal{K}^n$ und $x \in \operatorname{bd} K$. Der Punkt x heißt *m-Kantenpunkt* von K, wenn

$$\dim N(K,x) = n - m$$

ist. Ein 0-Kantenpunkt heißt *Eckpunkt*, ein m-Kantenpunkt mit $m \leq n-2$ heißt *singulär*, ein $(n-1)$-Kantenpunkt heißt *regulär*. Ein konvexer Körper, der nur reguläre Randpunkte hat, heißt *glatt*. Die regulären Randpunkte sind also genau diejenigen Punkte von K, durch die genau eine Stützebene geht.

Bemerkung. Es gibt konvexe Körper, auf deren Rand die singulären Randpunkte dicht liegen.

Wir wollen nun zeigen, dass ein konvexer Körper nicht „zu viele" singuläre Randpunkte haben kann, in einem Sinn, der sich präzisieren lässt. Diese Präzisierung erfolgt unter Verwendung des Hausdorff-Maßes. Darunter ist folgendes zu verstehen.

Sei $p \geq 0$, sei $M \subset \mathbb{E}^n$ eine beliebige Teilmenge. Für $\delta > 0$ definiert man

$$\mathcal{H}^p_\delta(M) := \inf \left\{ \sum_{i=1}^\infty D(G_i)^p : \begin{array}{l} (G_i)_{i\in\mathbb{N}} \text{ Folge offener Mengen} \\ \text{mit } D(G_i) \leq \delta \text{ und } M \subset \bigcup_{i=1}^\infty G_i \end{array} \right\},$$

wo D den Durchmesser bezeichnet. Wird δ verkleinert, so werden zur Überdeckung weniger Mengenfolgen zugelassen, das Infimum wird also höchstens größer. Daher existiert in $\overline{\mathbb{R}}$ der Limes

$$\mathcal{H}^p(M) := \sup_{\delta>0} \mathcal{H}^p_\delta(M) = \lim_{\delta\to 0+} \mathcal{H}^p_\delta(M).$$

Man nennt $\mathcal{H}^p(M)$ das p-dimensionale Hausdorff-Maß von M.

2.2.2 Satz. *Die Menge der m-Kantenpunkte des konvexen Körpers K kann überdeckt werden durch abzählbar viele kompakte Mengen von endlichem m-dimensionalen Hausdorff-Maß* $(m = 0, \dots, n-1)$.

Beweis. Sei B eine Kugel mit $K \subset \operatorname{int} B$. Sei x ein m-Kantenpunkt von K. Dann ist $N(K, x) + x$ eine $(n-m)$-dimensionale konvexe Menge. Wir können eine m-dimensionale Ebene E wählen, die $N(K, x) + x$ in einem inneren Punkt von $B \setminus K$ trifft. Diese Ebene kann so gewählt werden, dass sie von $m+1$ Punkten mit rationalen Koordinaten aufgespannt wird. Für einen Punkt $y \in E \cap (N(K, x)+x)$ gilt $p(K, y) = x$, also $x \in p(K, E \cap B)$. Die Menge $E \cap B$ ist kompakt und hat endliches m-dimensionales Hausdorff-Maß. Da die metrische Projektion nach Satz 1.2.2 kontrahierend ist, ist (wie man leicht aus der Definition des Hausdorff-Maßes folgert) auch $p(K, E \cap B)$ von endlichem m-dimensionalen Hausdorff-Maß, außerdem kompakt. Da es nur abzählbar viele m-dimensionale Ebenen gibt, die von rationalen Punkten aufgespannt werden, folgt die Behauptung. ■

Satz 2.2.2 hat einige wichtige Konsequenzen. Im Fall $m = 0$ ergibt sich, da offenbar $\mathcal{H}^0(M) = \operatorname{card} M$ ist: Jeder konvexe Körper hat höchstens abzählbar viele Eckpunkte. Da ferner aus $\mathcal{H}^m(M) < \infty$ für ein $m < n-1$ stets $\mathcal{H}^{n-1}(M) = 0$ folgt, ergibt sich: Die Menge der singulären Randpunkte eines konvexen Körpers ist von $(n-1)$-dimensionalem Hausdorff-Maß Null. Daraus lässt sich insbesondere folgern:

2.2.3 Satz. *Sei $K \in \mathcal{K}^n$ ein n-dimensionaler konvexer Körper. Die regulären Randpunkte von K liegen auf $\operatorname{bd} K$ dicht. K ist der Durchschnitt aller Stützhalbräume von K, deren Rand einen regulären Randpunkt von K enthält.*

Beweis. Es ist leicht zu sehen, dass eine nichtleere, relativ offene Teilmenge von $\operatorname{bd} K$ positives $(n-1)$-dimensionales Hausdorff-Maß hat. Daraus folgt die erste Behauptung. Zum Beweis der zweiten Behauptung sei $B \subset \operatorname{int} K$ eine Kugel und $x \in \mathbb{R}^n \setminus K$. Dann ist $(B \cup \{x\}) \cap \operatorname{bd} K$ eine nichtleere, relativ offene Teilmenge von $\operatorname{bd} K$, enthält also einen regulären Randpunkt y. Eine Stützebene an K durch y trennt x und K. Hieraus folgt die zweite Behauptung. ■

Eine weitere wichtige Anwendung betrifft konvexe Funktionen:

2.2.4 Satz. *Jede konvexe Funktion $f : \mathbb{E}^n \to \overline{\mathbb{R}}$ ist fast überall auf $\operatorname{int} \operatorname{dom} f$ differenzierbar.*

Beweis. Sei $x \in \operatorname{int} \operatorname{dom} f$. Nach Satz 1.5.11 ist f in x genau dann differenzierbar, wenn f in x nur einen Subgradienten hat. Nach Satz 1.5.10 ist dies äquivalent damit, dass der Epigraph $\operatorname{epi} f$ in $(x, f(x))$ nur einen äußeren Normaleneinheitsvektor hat, also dass $(x, f(x))$ regulärer Randpunkt von $\operatorname{epi} f$ ist. Da die Menge der singulären Randpunkte von $\operatorname{epi} f$ von n-dimensionalem Hausdorff-Maß 0 ist (das folgt sofort aus 2.2.2, auch obwohl $\operatorname{epi} f$ nicht kompakt ist), folgt leicht, dass die Menge der $x \in \operatorname{int} \operatorname{dom} f$, für die $(x, f(x))$ singulärer Punkt von $\operatorname{epi} f$ ist, von n-dimensionalem Hausdorff-Maß 0 und damit auch von n-dimensionalem Lebesgue-Maß 0 ist. ■

2.3 Polytope

Konvexe Körper mit besonders einfacher Randstruktur sind die Polytope, denn sie haben nur endlich viele Seiten. In diesem Abschnitt werden grundlegende Eigenschaften von Polytopen behandelt.

Nach Definition ist ein Polytop im $\mathbb{E}^n$ die konvexe Hülle einer nichtleeren endlichen Punktmenge. Ist $P = \operatorname{conv}\{x_1, \dots, x_k\}$, so müssen die Extrempunkte von P unter den Punkten $x_1, \dots, x_k$ vorkommen; nach dem Satz von Minkowski sind Polytope also genau die konvexen Körper mit endlich vielen Extrempunkten. Ist

H eine Stützebene von P, so gilt

$$H \cap P = \operatorname{conv}(H \cap \{x_1, \ldots, x_k\}),$$

also ist jede Stützmenge eines Polytops ebenfalls ein Polytop. Sind P_1, P_2 Polytope, so ist jeder Extrempunkt der Minkowskischen Summe $P_1 + P_2$ Summe eines Extrempunktes von P_1 und eines Extrempunktes von P_2, also ist auch $P_1 + P_2$ ein Polytop.

Zuerst zeigen wir nun, dass bei einem Polytop die Seiten identisch sind mit den exponierten Seiten (Stützmengen).

2.3.1 Satz. *Sei $P \subset \mathbb{E}^n$ ein Polytop, F_1 eine Stützmenge von P und F eine Stützmenge von F_1. Dann ist F eine Stützmenge von P.*

Beweis. Wir können $0 \in F$ annehmen. Es gibt eine Stützebene $H_{u,0}$ an P mit $H_{u,0} \cap P = F_1$ und $P \subset H^-_{u,0}$. Ferner gibt es eine Stützebene $H_{v,0}$ an F_1 mit $H_{v,0} \cap F_1 = F$ und $F_1 \subset H^-_{v,0}$. Setze

$$\eta_0 := \max\left\{-\frac{\langle x, v\rangle}{\langle x, u\rangle} : x \in \operatorname{ext} P \setminus \operatorname{ext} F_1\right\}$$

und $H(\eta) := H_{\eta u+v,0}$ mit $\eta > \eta_0$. Es gilt $\langle x, u\rangle < 0$ für $x \in \operatorname{ext} P \setminus \operatorname{ext} F_1$ und daher

$$\langle x, \eta u + v\rangle < \eta_0\langle x, u\rangle + \langle x, v\rangle \leq 0$$

nach Definition von η_0. Für $x \in \operatorname{ext} F_1 \setminus \operatorname{ext} F$ gilt

$$\langle x, \eta u + v\rangle = \langle x, v\rangle < 0,$$

und für $x \in \operatorname{ext} F$ gilt $\langle x, \eta u+v\rangle = 0$. Damit ist $\operatorname{ext} F \subset H(\eta)$ und $\operatorname{ext} P \setminus \operatorname{ext} F \subset \operatorname{int} H^-_{\eta u+v,0}$ gezeigt. Also ist $H(\eta)$ eine Stützebene an P mit $H(\eta) \cap P = F$, das heißt F ist eine Stützmenge von P. ■

2.3.2 Korollar. *Jede eigentliche Seite eines Polytops ist eine Stützmenge.*

Beweis. Wir nehmen an, P sei ein Polytop der Dimension $n \geq 1$ und die Behauptung sei bereits bewiesen für Polytope kleinerer Dimension (für nulldimensionale Polytope ist nichts zu beweisen). Sei F eine eigentliche Seite von P. Dann ist F enthalten in einer Stützmenge F_1 von P und ist eine Seite von F_1. Nun gilt entweder $F = F_1$, dann ist F eine Stützmenge von P, oder F ist eine eigentliche Seite von F_1 und daher nach Induktionsannahme eine Stützmenge von F_1. Nach Satz 2.3.1 ist F eine Stützmenge von P. ■

Insbesondere ist damit jeder Extrempunkt von P exponierter Punkt von P. Die Extrempunkte eines Polytops nennt man seine *Ecken*; die Menge der Ecken von

P wird oft mit $\operatorname{vert} P$ bezeichnet. Die eindimensionalen Seiten eines Polytops P sind seine *Kanten*, und die $(\dim P - 1)$-dimensionalen Seiten heißen *Facetten.* Sind $x_1, \dots, x_k$ die Ecken des Polytops P und ist F eine eigentliche Seite von P, so gibt es eine Hyperebene H mit

$$F = H \cap P = \operatorname{conv}(H \cap \{x_1, \dots, x_k\}).$$

Jede Seite von P ist also die konvexe Hülle einer Teilmenge der Eckenmenge, daher ist $\mathcal{F}(P)$ endlich.

Ein Polytop war zunächst definiert worden als die konvexe Hülle von endlich vielen Punkten. Es kann auch dargestellt werden als Durchschnitt von endlich vielen abgeschlossenen Halbräumen.

2.3.3 Satz. *Jedes Polytop ist der Durchschnitt von endlich vielen abgeschlossenen Halbräumen.*

Beweis. Sei $P \subset \mathbb{E}^n$ ein Polytop. Wir nehmen $\dim P = n$ an; daraus folgt der allgemeine Fall, weil Ebenen und Halbebenen als Durchschnitte von endlich vielen abgeschlossenen Halbräumen dargestellt werden können. Seien $F_1, \dots, F_k$ die Facetten von P; dann gilt $F_i = H_i \cap P$ mit einer geeigneten Stützebene H_i von P. Sei H_i^- der von H_i berandete abgeschlossene Halbraum, der P enthält $(i = 1, \dots, k)$. Wir behaupten, dass

$$P = H_1^- \cap \dots \cap H_k^- \tag{2.2}$$

gilt. Die Inklusion $P \subset H_1^- \cap \dots \cap H_k^-$ ist trivial. Sei $x \in \mathbb{E}^n \setminus P$. Sei A die Vereinigung der affinen Hüllen von x und irgend $n-1$ Ecken von P. Wir wählen einen Punkt $y \in (\operatorname{int} P) \setminus A$. Es gibt einen Punkt $z \in \operatorname{bd} P \cap [x, y]$; er liegt in einer Stützebene an P und daher in einer Seite F von P. Angenommen, es wäre $\dim F =: j \leq n-2$. Nach dem Satz von Carathéodory liegt z in der konvexen Hülle von $j + 1 \leq n - 1$ Ecken von P und daher in A. Aber daraus folgt $y \in A$, ein Widerspruch. Damit ist gezeigt, dass F eine Facette ist, also gilt $F = F_i$ für ein $i \in \{1, \dots, k\}$. Aus $y \in \operatorname{int} P \subset \operatorname{int} H_i^-$ folgt $x \notin H_i^-$. Damit ist (2.2) gezeigt. ■

Sei $P \in \mathcal{K}_0^n$ ein Polytop und F eine Facette von P. Dann gilt $F = P \cap H_{u,\alpha}$ und $P \subset H_{u,\alpha}^-$ für geeignetes u und α. Wir können $u \in S^{n-1}$ annehmen; dann heißt u *äußerer Normaleneinheitsvektor* der Facette F. Die äußeren Normaleneinheitsvektoren der Facetten von P werden kurz als die *Normalenvektoren von* P bezeichnet.

2.3.4 Korollar. *Ist $P \in \mathcal{K}_0^n$ ein Polytop und sind $u_1, \dots, u_k$ seine Normalenvektoren, so gilt*

$$P = \bigcap_{i=1}^{k} H_{u_i, h(P, u_i)}^-.$$

Insbesondere ist P durch die Zahlen $h(P, u_1), \dots, h(P, u_k)$ eindeutig bestimmt.

Beweis. Ist F_i die Facette von P mit Normalenvektor u_i, so gilt $F_i = P \cap H_{u_i, h(P, u_i)}$. Die Behauptung ergibt sich jetzt aus dem Beweis von 2.3.3. ■

Im weiteren Verlauf ist es zweckmäßig, Dualität konvexer Körper zu verwenden. Der folgende Satz zeigt insbesondere, dass der Polarkörper eines Polytops (mit 0 im Innern) wieder ein Polytop ist.

2.3.5 Satz. *Sei $u_1, \dots, u_k \in \mathbb{E}^n \setminus \{0\}$. Ist*

$$P := \bigcap_{i=1}^{k} H_{u_i,1}$$

beschränkt, so ist

$$P^* = \operatorname{conv}\{u_1, \dots, u_k\}.$$

Beweis. Zunächst bemerken wir, dass $0 \in \operatorname{int} P$ gilt, also kann P^* gebildet werden.

Setze $Q := \operatorname{conv}\{u_1, \dots, u_k\}$. Für $x \in P$ gilt $\langle x, u_i \rangle \leq 1$, also $u_i \in P^*$ für $i = 1, \dots, k$. Also ist $Q \subset P^*$.

Sei $v \in \mathbb{E}^n \setminus Q$. Dann gibt es eine Hyperebene $H_{z,\alpha}$, die Q und v stark trennt. Wir können annehmen, dass $\langle z, u_i \rangle < \alpha$ für $i = 1, \dots, k$ und $\langle z, v \rangle > \alpha$ gilt. Da P beschränkt ist, gilt $\operatorname{pos}\{u_1, \dots, u_k\} = \mathbb{E}^n$ (andernfalls hätte der abgeschlossene konvexe Kegel $\operatorname{pos}\{u_1, \dots, u_k\}$ einen Randpunkt und damit eine Stützebene durch 0, jeder ihrer äußeren Normalenvektoren gehört zu P). Es gibt also eine Darstellung

$$0 = \sum_{i=1}^{k} \lambda_i u_i \quad \text{mit} \quad \lambda_i \geq 0, \quad \sum_{i=1}^{k} \lambda_i > 0.$$

Das ergibt $0 < \sum \lambda_i \alpha$ und daher $\alpha > 0$, o.B.d.A. kann $\alpha = 1$ angenommen werden. Damit gilt $z \in P$ und folglich $v \notin P^*$. Also ist $Q = P^*$. ■

2.3.6 Satz. *Jeder nichtleere, beschränkte Durchschnitt von endlich vielen abgeschlossenen Halbräumen ist ein Polytop.*

Beweis. Sei $P = \bigcap_{i=1}^{k} H_i^-$ nichtleer und beschränkt, wobei $H_1^-, \dots, H_k^- \subset \mathbb{E}^n$ abgeschlossene Halbräume sind. Wir können $\dim P = n$ annehmen (andernfalls ersetzen wir $\mathbb{E}^n$ durch $\operatorname{aff} P$) und nach einer Translation dann $0 \in \operatorname{int} P$. Damit gilt $0 \in \operatorname{int} H_i^-$, und wir können $H_i^- = H_{u_i,1}^-$ annehmen. Nach Satz 2.3.5 is P^* ein Polytop, also nach Satz 2.3.3 ein Durchschnitt von endlich vielen abgeschlossenen Halbräumen (und es ist $0 \in \operatorname{int} P^*$). Nach derselben Schlussweise ist $P^{**} = P$ ein Polytop. ■

Aus Satz 2.3.6 folgt insbesondere, dass der Durchschnitt von endlich vielen Polytopen ein Polytop ist und dass der Durchschnitt eines Polytops mit einer Ebene ein Polytop ist.

Wir können nun leicht weitere Informationen über die Seiten von Polytopen erhalten.

2.3.7 Satz. *Jede eigentliche Seite eines Polytops P ist enthalten in einer Facette von P.*

Beweis. Wir können $\dim P = n$ annehmen. Im Beweis von Satz 2.3.3 wurde die Existenz einer Darstellung

$$P = \bigcap_{i=1}^{k} H_i^-$$

gezeigt, wo H_i^- ein abgeschlossener Halbraum und $F_i := P \cap H_i$ eine Facette von P ist $(i = 1, \ldots, k)$.

Wegen $\operatorname{bd} P \subset \bigcup_{i=1}^{k} (P \cap H_i)$ ist jeder Randpunkt von P enthalten in einer Facette F_j.

Nun sei F eine eigentliche Seite von P. Wir wählen einen Punkt $x \in \operatorname{relint} F$. Dann gilt $x \in F_j = P \cap H_j$ für ein $j \in \{1, \ldots, k\}$. Da $x \in \operatorname{relint} F$ gilt und H_j eine Stützebene an P ist, folgt $F \subset H_j$ und damit $F \subset F_j$. ■

2.3.8 Korollar. *Sei P ein Polytop, seien $F^j \in \mathcal{F}_j(P)$ und $F^k \in \mathcal{F}_k(P)$ Seiten von P mit $F^j \subset F^k$. Dann gibt es Seiten $F^i \in \mathcal{F}_i(P)$ $(i = j+1, \ldots, k-1)$ mit $F^j \subset F^{j+1} \subset \ldots \subset F^{k-1} \subset F^k$.*

Beweis. Da F^j Stützmenge von P und damit von F^k ist, ist F^j Seite von F^k. Sei o.B.d.A. F^j eine eigentliche Seite von F^k. Dann ist F^j enthalten in einer Facette F^{k-1} von F^k, und nach Satz 2.1.1 ist F^{k-1} Seite von P. Wiederholung dieses Arguments ergibt die Behauptung. ■

Für ein n-dimensionales Polytop P bezeichne $f_i(P)$ die Anzahl seiner i-dimensionalen Seiten. Dabei wird auch P selbst als Seite mitgezählt, es ist also $f_n(P) = 1$. Zwischen diesen Anzahlen besteht eine berühmte Relation, die für $n = 3$ auf Euler zurückgeht und in höheren Dimensionen auch mit den Namen Schläfli und Poincaré verknüpft wird.

2.3.9 Satz (Eulersche Relation). *Für jedes n-dimensionale Polytop P gilt*

$$\sum_{i=0}^{n} (-1)^i f_i(P) = 1.$$

Beweis. Wir bezeichnen als *ro-Polytop* das relative Innere eines Polytops und als *ro-Polyeder* jede disjunkte Vereinigung von endlich vielen ro-Polytopen. Nach Satz 2.1.2 ist für jedes Polytop P das System $\{\operatorname{relint} F : F \in \mathcal{F}(P)\}$ eine disjunkte Zerlegung von P, also ist auch P ein ro-Polyeder. Wir zählen auch die leere Menge $\emptyset$ zu den ro-Polyedern.

Wir zeigen zunächst, dass es eine eindeutig bestimmte reelle Funktion χ auf der Menge der ro-Polyeder gibt, die folgende Eigenschaften hat:

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

(b) Ist P ein Polytop, so gilt $\chi(P) = 1$,

(c) Ist Q ein ro-Polytop, so gilt $\chi(Q) = (-1)^{\dim Q}$,

(d) Sind $Q_1, \dots, Q_k$ paarweise disjunkte ro-Polytope, so gilt $\chi(Q_1 \cup \ldots \cup Q_k) = \chi(Q_1) + \ldots + \chi(Q_k)$.

Die Eindeutigkeit einer solchen Funktion ist klar nach (d) und (c). Die Existenz beweisen wir durch Induktion nach der Dimension, wobei wir mit dem nulldimensionalen Raum $\{0\}$ anfangen. In diesem Raum ist $\{0\}$ das einzige Polytop und das einzige ro-Polytop; die durch $\chi(\emptyset) = 0$ und $\chi(\{0\}) = 1$ definierte Funktion erfüllt also (a) – (d). Nun sei $n \geq 1$ und die Existenz von χ bewiesen im reellen affinen Raum der Dimension $n-1$. Wir wählen $u \in \mathbb{E}^n \setminus \{0\}$ und betrachten die Hyperebenen $H_\lambda := H_{u,\lambda}$ für $\lambda \in \mathbb{R}$. Ist Q ein Polytop (ro-Polytop), so ist $Q \cap H_\lambda$ entweder leer oder ein Polytop (ro-Polytop). Für ein ro-Polyeder Q im $\mathbb{E}^n$ definieren wir nun

$$\chi(Q) := \sum_{\lambda \in \mathbb{R}} [\chi(Q \cap H_\lambda) - \lim_{\mu \downarrow \lambda} \chi(Q \cap H_\mu)], \tag{2.3}$$

wo χ auf der rechten Seite die Funktion ist, deren Existenz nach Induktionsannahme in $(n-1)$-dimensionalen Räumen schon gezeigt ist. Für $Q \neq \emptyset$ gibt es eine Darstellung $Q = Q_1 \cup \ldots \cup Q_k$ mit paarweise disjunkten ro-Polytopen $Q_1, \dots, Q_k$. Es gibt daher Zahlen $\lambda_1 < \lambda_2 < \ldots < \lambda_m$ derart, dass jede Funktion $\lambda \mapsto \chi(Q_i \cap H_\lambda)$ in jedem offenen Intervall $(\lambda_r, \lambda_{r+1})$ konstant ist. Zusammen mit der Induktionsannahme zeigt dies, dass die Grenzwerte in (2.3) existieren und dass die Summe endlich ist. Die Funktion χ ist also wohldefiniert. Unmittelbar aus der Induktionsannahme ergibt sich, dass χ die Eigenschaften (a) und (d) hat. Ist Q ein Polytop, so ergibt die rechte Seite von (2.3) den Wert $(1-1)+(1-0) = 1$, falls Q nicht in einer Hyperebene H_λ liegt, und den Wert $1 - 0 = 1$ im anderen Fall. Ist Q ein ro-Polytop, so ergibt die rechte Seite von (2.3) entweder $0 - (-1)^{(\dim Q)-1} = (-1)^{\dim Q}$ oder $(-1)^{\dim Q} - 0 = (-1)^{\dim Q}$, also gilt (c). Damit ist die Induktion beendet und die Existenz von χ gezeigt.

Ist nun P ein n-dimensionales Polytop in $\mathbb{E}^n$, so gilt $\chi(P) = 1$ nach (b). Andererseits ist $P = \bigcup_{F \in \mathcal{F}(P)} \operatorname{relint} F$ eine disjunkte Vereinigung, und (d) ergibt

$$\chi(P) = \sum_{F \in \mathcal{F}(P)} \chi(\operatorname{relint} F).$$

Wegen (c) folgt nun die Behauptung. ■

Nach den Seiten von Polytopen untersuchen wir nun ihre Normalenkegel.

2.3.10 Satz. *Sei $P \subset \mathbb{E}^n$ ein n-dimensionales Polytop, seien $F_1, \dots, F_k$ die Facetten von P, und sei u_i der äußere Normaleneinheitsvektor in F_i $(i = 1, \dots, k)$. Für $x \in \operatorname{bd} P$ gilt dann*

$$N(P, x) = \operatorname{pos} \{u_i : x \in F_i\}.$$

Der Punkt x ist genau dann ein m-Kantenpunkt, wenn $x \in \operatorname{relint} F$ für eine m-dimensionale Seite F von P gilt.

Beweis. Sei $x \in \operatorname{bd} P$. Wir können o.B.d.A. annehmen, dass $x = 0$ ist und dass $F_1, \dots, F_r$ genau die Facetten von P sind, die x enthalten. Dann gilt $\operatorname{aff} F_i = H_{u_i,0}$ und $P \subset H^-_{u_i,0}$ $(i = 1, \dots, r)$.

Sei nun $u \in \operatorname{pos} \{u_1, \dots, u_r\}$ und o.B.d.A. $u \neq 0$, also

$$u = \sum_{i=1}^{r} \lambda_i u_i \qquad \text{mit } \lambda_i \geq 0.$$

Wegen $\langle x, u_i \rangle = 0$ für $i = 1, \dots, r$ ist $\langle x, u \rangle = 0$, und für $y \in P$ ist $\langle y, u_i \rangle \leq 0$ für $i = 1, \dots, r$ und daher $\langle y, u \rangle \leq 0$. Somit ist $H_{u,0}$ Stützebene an P, $P \subset H^-_{u,0}$ und $x \in H_{u,0}$, folglich $u \in N(P, x)$.

Umgekehrt sei nun $u \in N(P, x)$. Wir können annehmen, dass $u \notin \operatorname{pos}\{u_i\}$, $i = 1, \dots, r$, gilt. Die Menge $H_{u,0} \cap P =: F$ ist eine Seite von P. Wegen $x \in F$ und $u \notin \operatorname{pos}\{u_i\}$ ist F keine Facette von P, also ist $\dim F \leq n - 2$. Da also der zu F orthogonale lineare Unterraum mindestens zweidimensional ist, gibt es linear unabhängige Vektoren v, w senkrecht zu F mit $u = v + w$. Wegen $\operatorname{ext} P \setminus \operatorname{ext} F \subset \operatorname{int} H^-_{u,0}$ können v, w so gewählt werden, dass auch

$$\operatorname{ext} P \setminus \operatorname{ext} F \subset (\operatorname{int} H^-_{v,0}) \cap (\operatorname{int} H^-_{w,0})$$

gilt. Bei dieser Wahl gilt also $v, w \in N(P, x)$. Der Vektor u ist somit positive Kombination von zwei linear unabhängigen Vektoren des Kegels $N(P, x)$; daher ist $\operatorname{pos}\{u\}$ kein Extremstrahl von $N(P, x)$. Höchstens $\operatorname{pos}\{u_1\}, \dots, \operatorname{pos}\{u_r\}$ sind also Extremalstrahlen von $N(P, x)$. Da der abgeschlossene konvexe Kegel $N(P, x)$ wegen $\dim P = n$ geradenfrei ist, wird er nach Satz 1.4.3 von seinen Extremstrahlen aufgespannt, also ist $u \in \operatorname{pos} \{u_1, \dots, u_r\}$.

Zum Nachweis der letzten Behauptung setzen wir $G := \bigcap_{i=1}^r F_i$; dann ist G eine Seite von P. Wegen

$$G = \bigcap_{i=r+1}^{k} \left(H^-_{u_i,0} \cap \operatorname{aff} G \right)$$

und $x \notin H_{u_i,0}$ für $i = r+1,\dots,k$ gilt $x \in \operatorname{relint} G$. Somit ist G die eindeutig bestimmte Seite von P, die x im relativen Inneren enthält. Ist $n-m$ die Maximalzahl linear unabhängiger Vektoren unter $u_1,\dots,u_r$, so ist einerseits $\dim N(P,x) = n-m$, also x ein m-Kantenpunkt von P, und andererseits $\dim G = m$. ■

Ist P ein n-Polytop, F eine eigentliche Seite von P und $x \in \operatorname{relint} F$, so gilt nach dem Vorstehenden und mit denselben Bezeichnungen

$$N(P,x) = \operatorname{pos}\{u_i : F \subset F_i\}.$$

Dieser Normalenkegel hängt also nicht von der Wahl des Punktes $x \in \operatorname{relint} F$ ab, und wir können

$$N(P,F) := N(P,x)$$

definieren. Wir nennen $N(P,F)$ den *Normalenkegel von P bei F*. Für eigentliche Seiten F,G von P gilt offenbar

$$F \subset G \Leftrightarrow N(P,F) \supset N(P,G). \tag{2.4}$$

Auch die nachstehenden Aussagen über Seiten F von P ergeben sich ohne Schwierigkeiten: Es ist

$$\dim N(P,F) = n - \dim F. \tag{2.5}$$

Für $v \in \mathbb{E}^n \setminus \{0\}$ gilt

$$F = F(P,v) \Leftrightarrow v \in \operatorname{relint} N(P,F). \tag{2.6}$$

Es ist

$$N(P,F) = \bigcap_{x \in \operatorname{ext} F} N(P,x).$$

Nun gehen wir auf die Seiten dualer Polytope ein. Sei $P \subset \mathbb{E}^n$ ein Polytop mit $0 \in \operatorname{int} P$. Wir wissen bereits, dass der durch

$$P^* = \{u \in \mathbb{E}^n : \langle u,y\rangle \le 1 \text{ für alle } y \in P\}$$

definierte duale Körper ebenfalls ein Polytop ist. Für eine eigentliche Seite F von P definieren wir

$$\hat{F} := \{u \in P^* : \langle u,y\rangle = 1 \text{ für alle } y \in F\}.$$

Man beachte, dass $\hat{F}$ nicht nur von F, sondern auch von P abhängt, obwohl das in der Bezeichnung nicht zum Ausdruck kommt.

2.3.11 Satz. *Sei P ein Polytop mit $0 \in \operatorname{int} P$, seien F und G eigentliche Seiten von P. Dann gilt:*

(a) $\hat{F}$ *ist eigentliche Seite von* P^*,

(b) $(\hat{F})\hat{} = F$,

(c) $F \subset G \Rightarrow \hat{G} \subset \hat{F}$,

(d) $\dim \hat{F} = n - 1 - \dim F$,

(e) $N(P, F) = \operatorname{pos} \hat{F}$.

Beweis. (a) Für $y \in \operatorname{bd} P$ ist $H_{y,1}$ Stützebene von P^* (z.B. nach Satz 1.7.6), also ist $P^* \cap H_{y,1}$ eine Seite von P^*. Wegen

$$\hat{F} = \bigcap_{y \in F} (P^* \cap H_{y,1})$$

ist also $\hat{F}$ Seite von P^*.

(b) Für $y \in F$ gilt $y \in P = P^{**}$ und $\langle u, y \rangle = 1$ für alle $u \in \hat{F}$, also $y \in (\hat{F})\hat{}$. Somit gilt $F \subset (\hat{F})\hat{}$.

Sei $F = P \cap H_{u,1}$ (o.B.d.A.), dann ist $u \in \hat{F}$.

Für $z \in (\hat{F})\hat{}$ gilt $z \in P^{**} = P$ und $\langle z, v \rangle = 1$ für alle $v \in \hat{F}$, insbesondere $\langle z, u \rangle = 1$, also $z \in F$. Somit gilt $(\hat{F})\hat{} \subset F$.

(c) Gelte $F \subset G$. Sei $u \in \hat{G}$, also $u \in P^*$ und $\langle u, y \rangle = 1$ für alle $y \in G$. Dann ist auch $\langle u, y \rangle = 1$ für alle $y \in F$, also $u \in \hat{F}$. Somit gilt $\hat{G} \subset \hat{F}$.

(d) Gilt $F^0 \subset F^1 \subset \ldots \subset F^{n-1}$ mit $F^j \in \mathcal{F}_j(P)$, so folgt mit (c)

$$\hat{F}^{n-1} \subset \ldots \subset \hat{F}^1 \subset \hat{F}^0.$$

Hier sind (z.B. wegen (b)) alle Inklusionen strikt; daraus folgt die Behauptung.

(e) Sei $u \in N(P, F) \setminus \{0\}$. Wähle $x \in \operatorname{relint} F$. Die Hyperebene $H_{v,1}$ mit $v := h(P, u)^{-1} u$ ist Stützebene an P mit $x \in P \cap H_{v,1}$. Wegen $x \in \operatorname{relint} F$ folgt $F \subset H_{v,1}$. Wegen $\langle y, v \rangle \leq 1$ für alle $y \in P$ und $\langle y, v \rangle = 1$ für alle $y \in F$ ist $v \in \hat{F}$, also $u \in \operatorname{pos} \hat{F}$.

Die Schlüsse lassen sich umkehren. ■

Mit Satz 2.3.11 lässt sich leicht eine Verallgemeinerung der Eulerschen Relation zeigen. Sei P ein Polytop und G eine eigentliche Seite von P. Wir können o.B.d.A. $\dim P = n$ und $0 \in \operatorname{int} P$ annehmen. Dann ist

$$\begin{aligned} \sum_{G \subset F \subset P} (-1)^{\dim F} &= (-1)^n + \sum_{G \subset F \neq P} (-1)^{\dim F} = (-1)^n + \sum_{\hat{F} \subset \hat{G}} (-1)^{n-1-\dim \hat{F}} \\ &= (-1)^{n-1} \left[-1 + \sum_{\hat{F} \subset \hat{G}} (-1)^{\dim \hat{F}} \right]. \end{aligned}$$

Mit Anwendung der Eulerschen Relation auf $\hat{G}$ folgt also

$$\sum_{G \subset F \subset P} (-1)^{\dim F} = 0.$$

Wir wissen bereits aus Abschnitt 1.8, dass ein konvexer Körper durch Polytope beliebig genau approximiert werden kann. Wir wollen diese fast triviale Aussage nun für spätere Anwendungen wesentlich verschärfen, indem wir zeigen, dass mehrere konvexe Körper simultan durch Polytope approximiert werden können, die in einem zu präzisierenden Sinn von ähnlicher Struktur sind. Das erfordert einige Vorbereitungen.

Zwei Polytope $P_1, P_2 \subset \mathbb{E}^n$ heißen *stark isomorph*, wenn

$$\dim F(P_1, u) = \dim F(P_2, u) \quad \text{für alle } u \in S^{n-1}$$

gilt. Offensichtlich ist das eine Äquivalenzrelation.

2.3.12 Hilfssatz. *Sind P_1, P_2 stark isomorphe Polytope, so sind für jedes $u \in S^{n-1}$ die Stützmengen $F(P_1, u)$ und $F(P_2, u)$ stark isomorph.*

Beweis. Sei $u \in S^{n-1}$ gegeben, sei $v \in S^{n-1}$. Setze $F(F(P_i, u), v) =: F_i'$ $(i = 1, 2)$. Wir können $0 \in F_1' \cap F_2'$ annehmen. Im Beweis von Satz 2.3.1 wurde gezeigt, dass für genügend großes η die Hyperebene $H_{\eta u+v,0}$ das Polytop P_i stützt und dass $F_i' = F(P_i, \eta u + v)$ ist. Da P_1 und P_2 stark isomorph sind, erhalten wir $\dim F_1' = \dim F_2'$ und damit die Behauptung. ■

2.3.13 Hilfssatz. *Die Polytope $P_1, P_2 \in \mathcal{K}^n$ sind genau dann stark isomorph, wenn*

$$\{N(P_1, x) : x \in \operatorname{ext} P_1\} = \{N(P_2, x) : x \in \operatorname{ext} P_2\} \tag{2.7}$$

gilt.

Beweis. Zunächst seien P_1 und P_2 stark isomorph. Sei $x \in \operatorname{ext} P_1$. Wir wählen $u \in \operatorname{int} N(P_1, x)$ und $y \in F(P_2, u)$, dann gilt $\dim F(P_2, u) = \dim F(P_1, u) = \dim \{x\} = 0$, also $y \in \operatorname{ext} P_2$. Sei $v \in N(P_1, x)$ und $\lambda \in [0, 1)$; dann gilt

$$v_\lambda := (1 - \lambda)u + \lambda v \in \operatorname{int} N(P_1, x)$$

und daher wie oben $\dim F(P_2, v_\lambda) = 0$, also $F(P_2, v_\lambda) = \{y_\lambda\}$ für ein $y_\lambda \in \operatorname{ext} P_2$. Dann muss $v_\lambda \in \operatorname{int} N(P_2, y_\lambda)$ sein. Da dies für alle $\lambda \in [0, 1)$ gilt, folgt $y_\lambda = y$. Also ist $v \in N(P_2, y)$, womit $N(P_1, x) \subset N(P_2, y)$ gezeigt ist. Da $x \in \operatorname{ext} P_1$ beliebig war und die Rollen von P_1 und P_2 vertauscht werden können, folgt $N(P_1, x) = N(P_2, y)$. Daraus folgt (2.7).

Umgekehrt sei jetzt (2.7) erfüllt. Sei $u \in S^{n-1}$. Wir wählen eine Ecke $x_1 \in F(P_1, u)$. Es gibt eine Ecke x_2 von P_2 mit $N(P_2, x_2) = N(P_1, x_1)$. Der Vektor u liegt im relativen Inneren einer Seite von $N(P_i, x_i)$ der Dimension $n - \dim F(P_i, u)$ (wie aus (2.5) und (2.6) folgt), also ist $\dim F(P_1, u) = \dim F(P_2, u)$. Da u beliebig war, sind P_1 und P_2 stark isomorph. ■

Wir wollen den Begriff „stark isomorph" erläutern. Seien P_1 und P_2 stark isomorphe Polytope. Sei F eine eigentliche Seite von P_1. Der Normalenkegel $N(P_1, F)$ ist Seite eines Normalenkegels $N(P_1, x)$ für eine geeignete Ecke x von F. Nach Hilfssatz 2.3.13 ist $N(P_1, F)$ Seite von $N(P_2, y)$ für eine geeignete Ecke y von P_2 und daher der Normalenkegel einer eindeutig bestimmten Seite von P_2, die wir mit $\varphi(F)$ bezeichnen. Wenn wir φ noch auf uneigentliche Seiten fortsetzen, ist damit eine bijektive Abbildung $\varphi : \mathcal{F}(P_1) \to \mathcal{F}(P_2)$ erklärt, und es folgt aus (2.4), dass diese Abbildung Inklusionen erhält. Eine derartige Abbildung zwischen den Seitenmengen von zwei Polytopen nennt man einen *kombinatorischen Isomorphismus*. Zwei stark isomorphe Polytope sind also kombinatorisch isomorph, aber in sehr spezieller Weise, insofern als entsprechende Seiten dieselben Normalenkegel haben.

Ein n-dimensionales Polytop (kurz n-Polytop) P heißt *einfach*, wenn jede Ecke von P in genau n Facetten von P enthalten ist.

2.3.14 Hilfssatz. *Sei P ein einfaches n-Polytop mit Normalenvektoren $u_1, \dots, u_k$. Dann gibt es eine Zahl $\beta > 0$ derart, dass jedes Polytop der Form*

$$P' := \bigcap_{i=1}^{k} H^-_{u_i, h(P,u_i)+\alpha_i} \tag{2.8}$$

mit $|\alpha_i| \le \beta$ einfach und stark isomorph zu P ist.

Beweis. Wir können ein $\rho > 0$ wählen, so dass die Kugeln $B(x, \rho)$, $x \in \operatorname{ext} P$, paarweise disjunkt sind und jede Kugel $B(x, \rho)$ nur diejenigen Facetten von P trifft, die x enthalten. Sodann können wir eine Zahl $\beta > 0$ mit folgender Eigenschaft wählen: Ist x eine Ecke von P und sind etwa $u_1, \dots, u_n$ die äußeren Normalenvektoren der Facetten von P, die x enthalten, so liegt für $|\alpha_i| \le \beta$ der Schnittpunkt der Hyperebenen

$$H_{u_i, h(P,u_i)+\alpha_i}, \qquad i = 1, \dots, n,$$

in $B(x, \rho)$, und es gilt

$$H_{u_j, h(P,u_j)+\alpha_j} \cap B(x, \rho) = \emptyset \text{ für } j \notin \{1, \dots, n\}.$$

Wird dann P' durch (2.8) definiert mit $|\alpha_i| \le \beta$, so hat P' in jeder der Kugeln $B(x, \rho)$, $x \in \operatorname{ext} P$, eine Ecke x' mit $N(P', x') = N(P, x)$. Daraus folgt, dass P'

keine weiteren Ecken haben kann, einfach ist und wegen Hilfssatz 2.3.13 stark isomorph zu P ist. ■

2.3.15 Hilfssatz. *Sei P ein n-Polytop mit Normalenvektoren $u_1, \dots, u_k$. Dann gibt es zu jeder Zahl $\beta > 0$ Zahlen $\alpha_1, \dots, \alpha_k$ mit $|\alpha_i| \leq \beta$ und folgender Eigenschaft: Das Polytop*

$$P' := \bigcap_{i=1}^{k} H^-_{u_i, h(P,u_i)+\alpha_i}$$

ist ein einfaches n-Polytop mit Normalenvektoren $u_1, \dots, u_k$ und derart, dass jeder Normalenkegel einer Ecke von P' enthalten ist in einem Normalenkegel von P.

Beweis. Wir wählen Kugeln $B(x, \rho)$, $x \in \operatorname{ext} P$, wie im Beweis von Hilfssatz 2.3.14. Sind $|\alpha_1|, \dots, |\alpha_k|$ hinreichend klein, so ist P' ein n-Polytop mit Normalenvektoren $u_1, \dots, u_k$ und derart, dass jede Ecke x' von P' enthalten ist in einer Kugel $B(x, \rho)$. Nach Satz 2.3.10 ist $N(P', x')$ die positive Hülle von Normalenvektoren von Facetten von P, die x enthalten, daher gilt $N(P', x') \subset N(P, x)$. Indem man die α_i rekursiv geeignet wählt, kann man offenbar erreichen, dass das Polytop P' einfach wird. ■

Nun können wir das angekündigte Resultat über die simultane Approximation konvexer Körper durch stark isomorphe Polytope beweisen.

2.3.16 Satz. *Seien $K_1, \dots, K_m \in \mathcal{K}^n$ konvexe Körper. Zu jedem $\epsilon > 0$ gibt es einfache, stark isomorphe Polytope $P_1, \dots, P_m \in \mathcal{K}^n$ mit $\delta(K_i, P_i) < \epsilon$ für $i = 1, \dots, m$.*

Beweis. Sei $\epsilon > 0$ gegeben. Nach Satz 1.8.13 gibt es Polytope $Q_1, \dots, Q_m \in \mathcal{K}^n$ mit $\delta(K_i, Q_i) < \epsilon/2$ für $i = 1, \dots, m$. Setze $P := Q_1 + \ldots + Q_m$. Dann ist P ein Polytop; seien $u_1, \dots, u_k$ seine Normalenvektoren. Ausgehend von P, konstruieren wir P' wie in Hilfssatz 2.3.15 und setzen $P_i := Q_i + \alpha P'$, wobei $\alpha > 0$ so klein sei, dass $\delta(P_i, Q_i) < \epsilon/2$ und daher $\delta(K_i, P_i) < \epsilon$ für $i = 1, \dots, m$ gilt. Sei x eine Ecke eines Polytops P_i. Dann gilt $x = q + \alpha p$ mit geeigneten $q \in \operatorname{ext} Q$ und $p \in \operatorname{ext} P'$. Nach Satz 2.2.1 gilt

$$N(P_i, x) = N(Q_i, q) \cap N(P', p).$$

Nach der Konstruktion von P' ist der Normalenkegel $N(P', p)$ enthalten in einem Normalenkegel von P, und nach Satz 2.2.1 ist der letztere enthalten in einem Normalenkegel von Q_i; dies kann nur $N(Q_i, q)$ sein, weil $N(Q_i, q) \cap N(P', p)$ von der Dimension n ist. Also gilt $N(P_i, x) = N(P', p)$. Da x eine beliebige Ecke von P_i war, schließen wir aus Hilfssatz 2.3.13, dass P_i stark isomorph zu P' ist. Hieraus folgt auch, dass P_i einfach ist. ■

Kapitel 3

Funktionale und Extremalprobleme

Nachdem die wesentlichen Grundlagen der Konvexgeometrie nun gelegt sind, wollen wir Funktionale konvexer Körper behandeln, wie Volumen und Oberfläche, und die Ungleichungen, die zwischen ihnen bestehen. Insbesondere soll so die sogenannte isoperimetrische Ungleichung, derzufolge unter allen konvexen Körpern gegebenen Volumens genau die Kugeln kleinste Oberfläche haben, in eine allgemeine Theorie eingebettet werden, die viele andere Anwendungen gestattet.

Ein einfaches Grundprinzip, aus dem sich eine interessante Theorie ergibt, ist die Verknüpfung von zwei elementaren Begriffen, dem Volumen und der Minkowskischen Addition. Das Verhalten des Volumens bei der Addition oder Linearkombination konvexer Körper zeigt einige wichtige Gesetzmäßigkeiten. Sie sollen in den folgenden Abschnitten studiert werden.

3.1 Der Satz von Brunn-Minkowski

Um das Verhalten des Volumens bei der Addition von zwei konvexen Körpern zu untersuchen, betrachten wir zunächst einen sehr speziellen Fall. Zwei konvexe Körper K_0, K_1 heißen *homothetisch*, wenn einer von ihnen aus dem anderen durch Streckung mit einem nichtnegativen Faktor und Translation hervorgeht. Seien K_0, K_1 homothetisch, etwa $K_0 = \lambda K$, $K_1 = \mu K + t$ mit einem konvexen Körper K und $\lambda, \mu \geq 0$. Dann ist $K_0 + K_1 = (\lambda + \mu)K + t$, für das Volumen V_n gilt also

$$V_n(K_0 + K_1) = V_n((\lambda + \mu)K + t) = (\lambda + \mu)^n V_n(K)$$

und daher

$$V_n(K_0 + K_1)^{1/n} = V_n(K_0)^{1/n} + V_n(K_1)^{1/n}. \tag{3.1}$$

Wegen der Homogenität des Volumens vom Grad n, also $V_n(\lambda K) = \lambda^n V_n(K)$, ist es sinnvoll, die n-te Wurzel des Volumens zu betrachten, und sie verhält sich nach (3.1) bei homothetischen Körpern additiv. Der Satz von Brunn-Minkowski beschreibt, was anstelle von (3.1) bei beliebigen konvexen Körpern K_0, K_1 gilt. Dies ist die Ungleichung

$$V_n(K_0 + K_1)^{1/n} \geq V_n(K_0)^{1/n} + V_n(K_1)^{1/n}. \tag{3.2}$$

Diese *Brunn-Minkowskische Ungleichung* wird oft für die verbindende Körperschar $\{(1-\lambda)K_0 + \lambda K_1 : 0 \leq \lambda \leq 1\}$ formuliert. Mit $\lambda \in [0,1]$ folgt aus (3.2)

$$V_n((1-\lambda)K_0 + \lambda K_1)^{1/n} \geq (1-\lambda)V_n(K_0)^{1/n} + \lambda V_n(K_1)^{1/n}.$$

In dieser Form wollen wir die Ungleichung beweisen.

3.1.1 Satz (Brunn-Minkowski). *Für konvexe Körper $K_0, K_1 \in \mathcal{K}^n$ und für $0 \leq \lambda \leq 1$ gilt*

$$V_n((1-\lambda)K_0 + \lambda K_1)^{1/n} \geq (1-\lambda)V_n(K_0)^{1/n} + \lambda V_n(K_1)^{1/n}. \tag{3.3}$$

Gleichheit für ein $\lambda \in (0,1)$ gilt genau dann, wenn entweder K_0 und K_1 in parallelen Hyperebenen liegen oder homothetisch sind.

Beweis. Wir setzen $K_\lambda := (1-\lambda)K_0 + \lambda K_1$ für $0 \leq \lambda \leq 1$. Liegen K_0 und K_1 in parallelen Hyperebenen, so liegt auch K_λ in einer Hyperebene, also gilt $V_n(K_\lambda) = 0$ für $0 \leq \lambda \leq 1$. Es gilt also Gleichheit in (3.3). Trivialerweise gilt auch Gleichheit, wenn K_0 und K_1 homothetisch sind.

Zunächst nehmen wir nun $\dim K_0 < n$ und $\dim K_1 < n$ an. Dann gilt (3.3) trivialerweise, da die rechte Seite verschwindet. Gilt Gleichheit für ein $\lambda \in (0,1)$, so ist $V_n(K_\lambda) = 0$, also liegt K_λ in einer Hyperebene. Dann liegen K_0 und K_1 in parallelen Hyperebenen.

Gelte nun etwa $\dim K_0 < n$ und $\dim K_1 = n$. Dann gilt $K_\lambda \supset (1-\lambda)x + \lambda K_1$ für beliebiges $x \in K_0$, also

$$V_n(K_\lambda) \geq V_n((1-\lambda)x + \lambda K_1) = \lambda^n V_n(K_1),$$

mit Gleichheit genau dann, wenn $K_0 = \{x\}$ ist. In diesem Fall sind K_0 und K_1 nach Definition homothetisch.

Von jetzt an können wir also $\dim K_0 = n$ und $\dim K_1 = n$ voraussetzen. Ferner können wir $V_n(K_0) = V_n(K_1) = 1$ annehmen. Ist nämlich (3.3) für diesen Fall bewiesen und sind dann K_0 und K_1 beliebige n-dimensionale konvexe Körper, so setzen wir

$$\begin{aligned} \bar{K}_i &:= V_n(K_i)^{-1/n} K_i \qquad \text{für } i = 0,1, \\ \bar{\lambda} &:= \frac{\lambda V_n(K_1)^{1/n}}{(1-\lambda)V_n(K_0)^{1/n} + \lambda V_n(K_1)^{1/n}}. \end{aligned}$$

Aus $V_n((1-\bar\lambda)\bar K_0+\bar\lambda\bar K_1)^{1/n}\geq 1$ folgt dann (3.3); die Aussagen über den Gleichheitsfall übertragen sich ebenso.

Satz 3.1.1 wird nun durch Induktion nach der Dimension bewiesen. Der Fall $n=1$ ist trivial; wir nehmen daher an, dass $n\geq 2$ ist und der Satz in der Dimension $n-1$ schon bewiesen ist. Wir wählen einen Vektor $u\in S^{n-1}$ und schreiben

$$H_{u,\alpha}=:H(\alpha),\qquad H^-_{u,\alpha}=:H^-(\alpha),$$

$$\alpha_\lambda:=-h(K_\lambda,-u),\qquad \beta_\lambda:=h(K_\lambda,u).$$

Dann gilt

$$\alpha_\lambda=(1-\lambda)\alpha_0+\lambda\alpha_1,\qquad \beta_\lambda=(1-\lambda)\beta_0+\lambda\beta_1.$$

Für $\zeta\in\mathbb{R}$ und $i=0,1$ definieren wir

$$\begin{aligned} v_i(\zeta) &:= V_{n-1}(K_i\cap H(\zeta)),\\ w_i(\zeta) &:= V_n(K_i\cap H^-(\zeta)), \end{aligned}$$

so dass also nach dem Satz von Fubini

$$w_i(\zeta)=\int_{\alpha_i}^{\zeta} v_i(t)\,dt$$

gilt. Auf (α_i,β_i) ist die Funktion v_i stetig, denn nach Satz 1.8.8 gilt $K_i\cap H(\zeta_j)\to K_i\cap H(\zeta)$ für $\zeta_j\to\zeta\in(\alpha_i,\beta_i)$; man kann dann in geeigneter Form Satz 1.8.16 anwenden. Nach dem Hauptsatz der Differential- und Integralrechnung folgt, dass die Funktion w_i differenzierbar ist und dass

$$w_i'(\zeta)=v_i(\zeta)>0\qquad \text{für } \alpha_i<\zeta<\beta_i$$

gilt. Sei z_i die Umkehrfunktion von w_i. Es ist also $z_i(\tau)$ dadurch charakterisiert, dass der Abschnitt $K_i\cap H^-(z_i(\tau))$ das Volumen τ hat. Es gilt

$$z_i'(\tau)=\frac{1}{v_i(z_i(\tau))}\qquad \text{für } 0<\tau<1.$$

Schreiben wir

$$\begin{aligned} k_i(\tau) &:= K_i\cap H(z_i(\tau)),\qquad i=0,1,\\ z_\lambda(\tau) &:= (1-\lambda)z_0(\tau)+\lambda z_1(\tau), \end{aligned}$$

so gilt

$$K_\lambda\cap H(z_\lambda(\tau))\supset(1-\lambda)k_0(\tau)+\lambda k_1(\tau).$$

Es folgt

$$\begin{aligned} V_n(K_\lambda) &= \int\limits_{\alpha_\lambda}^{\beta_\lambda} V_{n-1}(K_\lambda \cap H(\zeta))\,d\zeta \\ &= \int\limits_0^1 V_{n-1}(K_\lambda \cap H(z_\lambda(\tau)))z_\lambda'(\tau)\,d\tau \\ &\geq \int\limits_0^1 V_{n-1}((1-\lambda)k_0(\tau)+\lambda k_1(\tau))\left[\frac{1-\lambda}{v_0(z_0(\tau))}+\frac{\lambda}{v_1(z_1(\tau))}\right]d\tau. \end{aligned}$$

Setzen wir zur Abkürzung $v_i(z_i(\tau)) = v_i$ $(i = 0, 1)$, so ist nach Induktionsannahme

$$V_{n-1}((1-\lambda)k_0(\tau)+\lambda k_1(\tau))^{1/(n-1)} \geq (1-\lambda)v_0^{1/(n-1)} + \lambda v_1^{1/(n-1)}.$$

Nun gilt allgemein für $0 < \lambda < 1$ und $\mu, \nu, p > 0$ die Ungleichung

$$((1-\lambda)\mu^p + \lambda\nu^p)^{1/p}\left(\frac{1-\lambda}{\mu}+\frac{\lambda}{\nu}\right) \geq 1. \tag{3.4}$$

Sie ergibt sich aus der Konkavität der log-Funktion:

$$\begin{aligned} &\log\left[((1-\lambda)\mu^p + \lambda\nu^p)^{1/p}\left(\frac{1-\lambda}{\mu}+\frac{\lambda}{\nu}\right)\right] \\ &= \frac{1}{p}\log\left((1-\lambda)\mu^p+\lambda\nu^p\right) + \log\left(\frac{1-\lambda}{\mu}+\frac{\lambda}{\nu}\right) \\ &\geq \frac{1}{p}\left((1-\lambda)\log\mu^p + \lambda\log\nu^p\right) + (1-\lambda)\log\frac{1}{\mu} + \lambda\log\frac{1}{\nu} \\ &= 0. \end{aligned}$$

Daraus folgt (3.4). Das Gleichheitszeichen gilt offenbar genau für $\mu = \nu$.

Damit erhalten wir nun

$$\begin{aligned} V_n(K_\lambda) &\geq \int\limits_0^1 \left[(1-\lambda)v_0^{1/(n-1)} + \lambda v_1^{1/(n-1)}\right]^{n-1}\left[\frac{1-\lambda}{v_0}+\frac{\lambda}{v_1}\right]d\tau \\ &\geq \int\limits_0^1 1\,d\tau = 1. \end{aligned}$$

Die behauptete Ungleichung gilt also auch in der Dimension n und ist damit allgemein bewiesen.

Nun nehmen wir an, für ein $\lambda \in (0,1)$ gelte in (3.3) das Gleichheitszeichen. Nach der obigen Herleitung muss dann $v_0 = v_1$ sein, wegen $z_i'(\tau) = 1/v_i(z_i(\tau))$ also

$$z_1(\tau) - z_0(\tau) = \text{const.} \tag{3.5}$$

Wir verschieben K_0 und K_1 so, dass beide Körper ihren Schwerpunkt im Nullpunkt haben. Dann ist ($dy = d\lambda_n(y)$ bezeichnet Integration nach dem Lebesgue-Maß)

$$0 = \int_{K_i} \langle y, u\rangle dy = \int_{\alpha_i}^{\beta_i} V_{n-1}(K_i \cap H(\zeta))\zeta \, d\zeta = \int_0^1 z_i(\tau)\, d\tau.$$

Die Konstante in (3.5) ist daher Null. Aus $z_0(1) = z_1(1)$ folgt $\beta_0 = \beta_1$, also $h(K_0, u) = h(K_1, u)$. Da u hier beliebig gewählt werden kann, folgt $K_0 = K_1$. ■

Man kann den Satz von Brunn-Minkowski auch folgendermaßen formulieren:

Folgerung. *Die durch*

$$f(\lambda) := V_n((1-\lambda)K_0 + \lambda K_1)^{1/n}, \qquad \lambda \in [0,1],$$

definierte Funktion f ist konkav. Sie ist genau dann linear, wenn K_0 und K_1 in parallelen Hyperebenen liegen oder homothetisch sind.

Dies ergibt sich aus (3.3) in der folgenden Weise. Wir setzen wieder

$$K_\lambda := (1-\lambda)K_0 + \lambda K_1 \qquad \text{für } 0 \le \lambda \le 1.$$

Für $\sigma, \tau \in [0,1]$ und $\lambda \in [0,1]$ gilt dann

$$(1-\lambda)K_\sigma + \lambda K_\tau = (1-\alpha)K_0 + \alpha K_1$$

mit $\alpha = (1-\lambda)\sigma + \lambda\tau$. Aus Satz 3.1.1, angewandt auf die Körper K_σ und K_τ, folgt also

$$\begin{aligned}
&f((1-\lambda)\sigma + \lambda\tau) = f(\alpha) = V_n((1-\alpha)K_0 + \alpha K_1)^{1/n} \\
&= V_n((1-\lambda)K_\sigma + \lambda K_\tau)^{1/n} \\
&\ge (1-\lambda)V_n(K_\sigma)^{1/n} + \lambda V_n(K_\tau)^{1/n} \\
&= (1-\lambda)f(\sigma) + \lambda f(\tau).
\end{aligned}$$

Die Brunn-Minkowskische Ungleichung kann in den folgenden äquivalenten Formulierungen ausgedrückt werden.

(a) $\forall\, K_0, K_1 \in \mathcal{K}^n: \quad V_n(K_0 + K_1)^{1/n} \geq V_n(K_0)^{1/n} + V_n(K_1)^{1/n}.$

(b) $\forall\, K_0, K_1 \in \mathcal{K}^n \;\; \forall\, s, t \geq 0: \quad V_n(sK_0 + tK_1)^{1/n} \geq sV_n(K_0)^{1/n} + tV_n(K_1)^{1/n}.$

(c) $\forall\, K_0, K_1 \in \mathcal{K}^n \;\; \forall\, \lambda \in [0,1]: \quad V_n(K_0) = V_n(K_1) = 1 \Rightarrow V_n((1-\lambda)K_0 + \lambda K_1) \geq 1.$

(d) $\forall\, K_0, K_1 \in \mathcal{K}^n \;\; \forall\, \lambda \in [0,1]: \quad V_n((1-\lambda)K_0 + \lambda K_1) \geq \min\{V_n(K_0), V_n(K_1)\}.$

Aus (a) folgt (b) wegen der Homogenität des Volumens, und aus (b) folgen (c) und (d) trivialerweise. Dass aus (c) wieder (a) folgt, wurde für $V_n(K_i) > 0$ im Beweis von Satz 3.1.1 gezeigt; daraus folgt es allgemein. Aus (d) folgt sofort (c).

Wir wollen nun bereits auf einige Folgerungen aus der Brunn-Minkowskischen Ungleichung eingehen, obwohl das später in systematischerer Weise geschieht. Zu den verblüffendsten geometrischen Anwendungen der Brunn-Minkowskischen Ungleichung gehört die isoperimetrische Ungleichung. Sie besagt (bei Beschränkung auf konvexe Körper), dass unter allen konvexen Körpern gegebenen Volumens die Kugeln die kleinste Oberfläche haben. Die *Oberfläche* eines konvexen Körpers $K \in \mathcal{K}^n$ kann definiert werden durch

$$S(K) := \lim_{\epsilon \downarrow 0} \frac{V_n(K + \epsilon B^n) - V_n(K)}{\epsilon}. \tag{3.6}$$

Dass dieser Grenzwert existiert, wird im nächsten Abschnitt in allgemeinerer Form gezeigt. Im Moment wollen wir nur bemerken, dass die Definition (3.6) anschaulich sehr plausibel ist und dass die so definierte Oberfläche für konvexe Körper zusammenfällt mit anderen geläufigen Oberflächenbegriffen (differentialgeometrischer Flächeninhalt von ∂K, falls ∂K eine glatte Hyperfläche ist, $(n-1)$-dimensionales Hausdorff-Maß von ∂K, bei geeigneter Normierung).

Die isoperimetrische Ungleichung kann jetzt in der Form

$$\left(\frac{S(K)}{S(B^n)}\right)^{\frac{1}{n-1}} \geq \left(\frac{V_n(K)}{V_n(B^n)}\right)^{1/n} \tag{3.7}$$

geschrieben werden. Sie ergibt sich sofort aus der Ungleichung von Brunn-Minkowski:

$$\begin{aligned} S(K) &= \lim_{\epsilon \downarrow 0} \frac{V_n(K + \epsilon B^n) - V_n(K)}{\epsilon} \\ &\geq \lim_{\epsilon \downarrow 0} \frac{\left(V_n(K)^{1/n} + \epsilon V(B^n)^{1/n}\right)^n - V_n(K)}{\epsilon} \\ &= nV_n(K)^{\frac{n-1}{n}} V_n(B^n)^{1/n}. \end{aligned}$$

Wegen $S(B^n) = nV_n(B^n)$ ergibt das genau die Ungleichung (3.7).

Bei dieser sehr einfachen Herleitung erhält man allerdings keinen Aufschluss darüber, dass das Gleichheitszeichen in (3.7) nur dann gilt, wenn K eine Kugel ist. Die folgende Variante des Beweises liefert auch diese Information.

In (3.6) setzen wir $\epsilon = t/(1-t)$ und erhalten

$$\begin{aligned} S(K) &= \lim_{t\downarrow 0} \frac{V_n((1-t)K + tB^n) - (1-t)^n V_n(K)}{t(1-t)^{n-1}} \\ &= \lim_{t\downarrow 0} \frac{V_n((1-t)K + tB^n) - V_n(K)}{t} + \lim_{t\downarrow 0} \frac{(1-(1-t)^n)V_n(K)}{t} \\ &= \lim_{t\downarrow 0} \frac{V_n((1-t)K + tB^n) - V_n(K)}{t} + nV_n(K). \end{aligned}$$

Für die durch

$$f(t) := V_n((1-t)K + tB^n)^{1/n} \qquad \text{für } 0 \le t \le 1$$

definierte Funktion erhält man damit

$$f'(0) = \frac{1/nS(K) - V_n(K)}{V_n(K)^{(n-1)/n}}.$$

Nach dem Satz von Brunn-Minkowski ist f eine konkave Funktion, daher ist

$$f'(0) \ge f(1) - f(0).$$

Das ergibt wieder die Ungleichung

$$S(K) \ge nV_n(K)^{\frac{n-1}{n}} V_n(B^n)^{1/n}.$$

Gilt hier das Gleichheitszeichen, so ist $f'(0) = f(1) - f(0)$. Da f konkav ist, muss dann

$$\frac{f(t) - f(0)}{t} = f(1) - f(0)$$

für $0 < t \le 1$ gelten. Dies bedeutet gerade Gleichheit in der Brunn-Minkowskischen Ungleichung

$$V_n((1-t)K + tB^n)^{1/n} \ge (1-t)V_n(K)^{1/n} + tV_n(B^n)^{1/n}.$$

Dann sind aber K und B^n homothetisch, d.h. K ist eine Kugel.

Als weiteres Anwendungsbeispiel behandeln wir die Frage, welche konvexen Körper gegebenen Durchmessers das größte Volumen haben. Allgemein bezeichnet man Probleme dieser Art, bei denen ein Funktional konstant gehalten und ein anderes zum Extremum gemacht werden soll, als *isoperimetrische Probleme*.

Es ist nicht trivial, dass ein solches isoperimetrisches Problem überhaupt eine Lösung besitzt. Wir wollen daher für unser Problem zuerst die Existenz von Lösungen zeigen. Bei solchen Existenzaussagen ist der Auswahlsatz von Blaschke ein grundlegendes Hilfmittel, und die im folgenden benutzte Schlußweise ist typisch.

Zunächst überlegen wir uns, dass wir nur eine geeignete beschränkte Menge konvexer Körper zur Konkurrenz zulassen müssen. Da Durchmesser D und Volumen V_n homogen sind (vom Grad 1 bzw. n, d.h. für $\lambda \geq 0$ gilt $D(\lambda K) = \lambda D(K)$ und $V_n(\lambda K) = \lambda^n V_n(K)$), brauchen wir nur konvexe Körper vom Durchmesser der Einheitskugel, also 2, zu betrachten. Da Durchmesser und Volumen translationsinvariant sind, brauchen wir nur Körper K mit $0 \in K$ zu betrachten. Mit $\mathcal{K}^*$ bezeichnen wir die Menge der Körper $K \in \mathcal{K}^n$ mit $0 \in K$ und $D(K) = 2$. Es genügt also der Nachweis, daß in $\mathcal{K}^*$ ein Körper maximalen Volumens existiert. Dazu setzen wir

$$\alpha := \sup\{V_n(K) : K \in \mathcal{K}^*\}.$$

Da alle Körper aus $\mathcal{K}^*$ in einer festen Kugel liegen, ist $\alpha < \infty$. Es gibt eine Folge $(K_i)_{i\in\mathbb{N}}$ in $\mathcal{K}^*$ mit $V_n(K_i) \to \alpha$ für $i \to \infty$. Nach dem Auswahlsatz von Blaschke (Satz 1.8.6) besitzt diese Folge eine konvergente Teilfolge, o.B.d.A. sei dies die Folge $(K_i)_{i\in\mathbb{N}}$ selbst, und K_0 sei ihr Grenzkörper. Der Durchmesser ist stetig. Es ist also $D(K_0) = 2$, und es ist klar, daß $0 \in K_0$ gilt. Also ist $K_0 \in \mathcal{K}^*$. Wegen der Stetigkeit des Volumens (Satz 1.8.16) ist $V_n(K_0) = \lim_{i\to\infty} V_n(K_i) = \alpha$. Damit ist die Existenz von konvexen Körpern gezeigt, die Durchmesser 2 und dabei maximales Volumen haben.

Nun zeigen wir, daß K_0 eine Kugel ist. Angenommen, das wäre falsch. Dann gibt es eine Hyperebene H derart, daß K_0 keine zu H parallele Symmetrieebene hat. (Es ist nämlich leicht zu zeigen, daß ein konvexer Körper, der zu jeder Hyperebene eine dazu parallele Symmetrieebene besitzt, eine Kugel sein muß.) Sei K_1 das Spiegelbild von K_0 an der Hyperebene H. Der Durchmesser des Körpers $\frac{1}{2}K_0 + \frac{1}{2}K_1$ ist ≤ 2. Nach dem Satz von Brunn-Minkowski gilt für sein Volumen

$$V_n\left(\frac{1}{2}K_0 + \frac{1}{2}K_1\right)^{1/n} \geq \frac{1}{2}V_n(K_0)^{1/n} + \frac{1}{2}V_n(K_1)^{1/n},$$

also $V_n(\frac{1}{2}K_0 + \frac{1}{2}K_1) \geq \alpha$. Nach Definition von α muß hier das Gleichheitszeichen stehen. Nach dem Satz von Brunn-Minkowski sind dann K_0 und K_1 homothetisch (da sicher $\alpha > 0$ ist und die Körper daher nicht in parallelen Hyperebenen liegen), wegen $V_n(K_0) = V_n(K_1)$ also Translate voneinander. Das bedeutet aber, daß K_0 eine zu H parallele Symmetrieebene besitzt, ein Widerspruch. Somit ist K_0 eine Kugel. Unter allen konvexen Körpern gegebenen Durchmessers haben also genau die Kugeln das größte Volumen. Diese Aussage läßt sich in Form der Ungleichung

$$D(K)^n \geq \frac{2^n}{\kappa_n} V_n(K),$$

mit Gleichheit genau für Kugeln, ausdrücken (isodiametrische Ungleichung).

Wir kehren zurück zur Brunn-Minkowskischen Ungleichung und schließen einige Bemerkungen über ihre Verallgemeinerungsfähigkeit an. Für die Gültigkeit der Brunn-Minkowskischen Ungleichung spielt die Konvexität der beteiligten Mengen keine Rolle. In der Tat gilt die Ungleichung

$$V_n(A+B)^{1/n} \geq V_n(A)^{1/n} + V_n(B)^{1/n}, \tag{3.8}$$

wenn V_n das Lebesgue-Maß bezeichnet und wenn A, B und $A+B$ beschränkte, Lebesgue-messbare Mengen sind (sogar das lässt sich noch verallgemeinern). Der oben für konvexe Körper gegebene Beweis hat den Vorteil, dass er für diesen Fall eine vollständige Diskussion des Gleichheitsfalles gestattet. Die Gleichheitsdiskussion ist für allgemeinere Mengen schwieriger; die Brunn-Minkowskische Ungleichung selbst hingegen lässt sich für geeignete Klassen nichtkonvexer Mengen recht einfach beweisen. Wir wollen hier vorübergehend den Bereich der Konvexität verlassen und einen derartigen Beweis geben, weil die Brunn-Minkowskische Ungleichung in Geometrie und Analysis eine gewisse Bedeutung hat.

Wir zeigen die Ungleichung (3.8) für Elementarmengen. Unter einer *Elementarmenge* verstehen wir eine Vereinigung von endlich vielen (achsenparallelen) Quadern, die paarweise keine inneren Punkte gemeinsam haben. Dabei ist ein *Quader* im $\mathbb{E}^n$ (mit einer gegebenen orthonormierten Basis) eine Menge der Form

$$\{(\xi_1, \dots, \xi_n) \in \mathbb{E}^n : a_i \leq \xi_i \leq b_i \quad \text{für } i = 1, \dots, n\}$$

mit $a_i < b_i$ $(i = 1, \dots, n)$.

Zuerst seien A, B zwei Quader mit Kantenlängen a_i bzw. b_i $(i = 1, \dots, n)$. Dann gilt

$$V_n(A) = \prod_{i=1}^n a_i, \quad V_n(B) = \prod_{i=1}^n b_i, \quad V_n(A+B) = \prod_{i=1}^n (a_i + b_i).$$

Nun gilt

$$\left(\prod_{i=1}^n \frac{a_i}{a_i+b_i}\right)^{1/n} + \left(\prod_{i=1}^n \frac{b_i}{a_i+b_i}\right)^{1/n} \leq \frac{1}{n}\sum_{i=1}^n \frac{a_i}{a_i+b_i} + \frac{1}{n}\sum_{i=1}^n \frac{b_i}{a_i+b_i} = 1,$$

wobei die Ungleichung für das arithmetische und geometrische Mittel benutzt wurde (die sich aus der Konkavität der Logarithmus-Funktion ergibt). Das liefert (3.8) für Quader.

Jetzt beweist man die Ungleichung für Elementarmengen A, B durch Induktion nach der Anzahl k der Quader, die zur Darstellung von A und B insgesamt benötigt werden. Sei $k \geq 3$, und sei (3.8) schon bewiesen für Elementarmengen, die zur Darstellung insgesamt weniger als k Quader benötigen. Nach passender

Verschiebung von A gibt es eine Koordinatenhyperebene H durch 0, die zwei Quader von A trennt. Seien H^+, H^- die von H berandeten abgeschlossenen Halbräume. Sei A_+ (bzw. A_-) die Vereinigung der Durchschnitte der Quader von A mit H^+ (bzw. H^-). Analog seien B_+ und B_- definiert. Nach Verschiebung der Menge B kann angenommen werden, dass

$$\frac{V_n(A_+)}{V_n(A)} = \frac{V_n(B_+)}{V_n(B)}, \quad \text{also auch} \quad \frac{V_n(A_-)}{V_n(A)} = \frac{V_n(B_-)}{V_n(B)} \tag{3.9}$$

ist. Es gilt

$$A_+ + B_+ \subset H^+, \quad A_- + B_- \subset H^-.$$

Die zur Darstellung von A_+ und B_+ (A_- und B_-) insgesamt benötigte Anzahl von Quadern ist kleiner als k. Daher kann man die Induktionsannahme verwenden. Zusammen mit (3.9) ergibt das

$$\begin{aligned}
V_n(A+B) &\geq V_n(A_+ + B_+) + V_n(A_- + B_-) \\
&\geq \left(V_n(A_+)^{\frac{1}{n}} + V_n(B_+)^{1/n}\right)^n + \left(V_n(A_-)^{1/n} + V_n(B_-)^{1/n}\right)^n \\
&= V_n(A_+)\left(1 + \frac{V_n(B)^{1/n}}{V_n(A)^{1/n}}\right)^n + V_n(A_-)\left(1 + \frac{V_n(B)^{1/n}}{V_n(A)^{1/n}}\right)^n \\
&= V_n(A)\left(1 + \frac{V_n(B)^{1/n}}{V_n(A)^{1/n}}\right)^n \\
&= \left(V_n(A)^{1/n} + V_n(B)^{1/n}\right)^n.
\end{aligned}$$

Dies beendet den Induktionsbeweis.

Das Resultat für Elementarmengen genügt, um mit Kenntnissen aus der Maßtheorie die Ungleichung (3.8) für beschränkte, Lebesgue-messbare Mengen $A, B, A+B$ zu beweisen. Die Diskussion des Gleichheitszeichens ist allerdings schwieriger.

Die Brunn-Minkowskische Ungleichung für beschränkte messbare Mengen kann auch aus einer allgemeineren analytischen Ungleichung hergeleitet werden. Auch das soll hier erläutert werden, da dieser Zusammenhang und Verallgemeinerungen der analytischen Ungleichung in jüngster Zeit besondere Bedeutung erlangt haben. Diese analytische Ungleichung sieht zunächst etwas merkwürdig aus:

3.1.2 Satz (Prékopa-Leindler-Ungleichung). *Sei $0 < \lambda < 1$, seien f, g, h nichtnegative integrierbare Funktionen auf $\mathbb{E}^n$ mit*

$$h((1-\lambda)x + \lambda y) \geq f(x)^{1-\lambda} g(y)^{\lambda} \quad \text{für } x, y \in \mathbb{E}^n.$$

Dann gilt

$$\int_{\mathbb{E}^n} h(x)\,dx \geq \left(\int_{\mathbb{E}^n} f(x)\,dx \right)^{1-\lambda} \left(\int_{\mathbb{E}^n} g(x)\,dx \right)^{\lambda}. \tag{3.10}$$

Die Prékopa-Leindler-Ungleichung erinnert an die Höldersche Ungleichung. Diese besagt in einem Spezialfall, dass für die durch

$$h(x) := f(x)^{1-\lambda} g(x)^{\lambda}$$

definierte Funktion h die Ungleichung

$$\int_{\mathbb{E}^n} h(x)\,dx \leq \left(\int_{\mathbb{E}^n} f(x)\,dx \right)^{1-\lambda} \left(\int_{\mathbb{E}^n} g(x)\,dx \right)^{\lambda}$$

gilt. Die Prékopa-Leindler-Ungleichung geht also in die andere Richtung, was natürlich nur unter starken Voraussetzungen möglich sein kann.

Vor dem Beweis der Prékopa-Leindler-Ungleichung soll gezeigt werden, wie hieraus die Brunn-Minkowskische Ungleichung folgt. Sei $0 < \lambda < 1$, seien $A, B \subset \mathbb{E}^n$ beschränkte messbare Mengen derart, dass auch $(1-\lambda)A + \lambda B$ messbar ist. Setze $f = \mathbf{1}_A$, $g = \mathbf{1}_B$ und $h = \mathbf{1}_{(1-\lambda)A+\lambda B}$ (wo $\mathbf{1}$ die charakteristische Funktion bezeichnet). Für $x, y \in \mathbb{E}^n$ gilt

$$\begin{aligned} & f(x)^{1-\lambda} g(y)^{\lambda} > 0 \Leftrightarrow x \in A \wedge y \in B \\ \Rightarrow \quad & (1-\lambda)x + \lambda y \in (1-\lambda)A + \lambda B \Leftrightarrow h((1-\lambda)x + \lambda y) = 1. \end{aligned}$$

Es folgt

$$h((1-\lambda)x + \lambda y) \geq f(x)^{1-\lambda} g(y)^{\lambda}$$

für $x, y \in \mathbb{E}^n$, nach der Prékopa-Leindler-Ungleichung also

$$V_n((1-\lambda)A + \lambda B) \geq V_n(A)^{1-\lambda} V_n(B)^{\lambda}$$

und damit

$$V_n((1-\lambda)A + \lambda B) \geq \min\{V_n(A), V_n(B)\}.$$

Aus der Gültigkeit dieser Ungleichung für alle $\lambda \in (0,1)$ folgt, wie früher bemerkt, die Brunn-Minkowskische Ungleichung; dabei wurde die Konvexität nicht benutzt.

Beweis von Satz 3.1.2. Es genügt, die Ungleichung für stetige, positive, integrierbare Funktionen zu beweisen; mit Kenntnissen aus der Integrationstheorie folgt sie daraus allgemein.

Wir beweisen die Ungleichung (3.10) durch Induktion nach der Dimension. Zunächst sei $n = 1$. Setze

$$\int_{\mathbb{R}} f(x)\,dx =: F, \qquad \int_{\mathbb{R}} g(x)\,dx =: G,$$

dann ist $F, G > 0$. Für $t \in (0,1)$ seien $u(t)$, $v(t)$ die kleinsten Zahlen mit

$$\frac{1}{F}\int_{-\infty}^{u(t)} f(x)\,dx = \frac{1}{G}\int_{-\infty}^{v(t)} g(x)\,dx = t.$$

Dadurch sind Funktionen $u, v : (0,1) \to \mathbb{R}$ definiert, die monoton wachsend und differenzierbar sind. Differentiation ergibt

$$\frac{f(u(t))u'(t)}{F} = \frac{g(v(t))v'(t)}{G} = 1.$$

Setze

$$w(t) := (1-\lambda)u(t) + \lambda v(t).$$

Für alle $t \in (0,1)$ gilt (mit Verwendung der Ungleichung für das arithmetische und geometrische Mittel)

$$\begin{aligned} w'(t) &= (1-\lambda)u'(t) + \lambda v'(t) \geq u'(t)^{1-\lambda} v'(t)^{\lambda} \\ &= \left(\frac{F}{f(u(t))}\right)^{1-\lambda} \left(\frac{G}{g(v(t))}\right)^{\lambda} \end{aligned}$$

und daher

$$\begin{aligned} \int_{\mathbb{R}} h(x)\,dx &\geq \int_0^1 h(w(t))w'(t)\,dt \\ &\geq \int_0^1 f(u(t))^{1-\lambda} g(v(t))^{\lambda} \left(\frac{F}{f(u(t))}\right)^{1-\lambda} \left(\frac{G}{g(v(t))}\right)^{\lambda} dt \\ &= F^{1-\lambda} G^{\lambda}. \end{aligned}$$

Damit ist (3.10) für $n = 1$ bewiesen.

Jetzt sei $n > 1$ und die Behauptung schon bewiesen in Räumen kleinerer Dimension. Wir identifizieren $\mathbb{E}^n$ mit $\mathbb{R}^{n-1} \times \mathbb{R}$. Für $s \in \mathbb{R}$ definieren wir dann

$h_s(z) := h(z, s)$ für $z \in \mathbb{R}^{n-1}$, und f_s und g_s seien analog erklärt. Sei $x, y \in \mathbb{R}^{n-1}$, $a, b \in \mathbb{R}$ und $c := (1-\lambda)a + \lambda b$. Dann gilt

$$\begin{aligned} &h_c((1-\lambda)x + \lambda y) = h((1-\lambda)x + \lambda y, (1-\lambda)a + \lambda b) \\ &= h((1-\lambda)(x, a) + \lambda(y, b)) \\ &\geq f(x, a)^{1-\lambda} g(y, b)^{\lambda} = f_a(x)^{1-\lambda} f_b(y)^{\lambda}. \end{aligned}$$

Nach Induktionsannahme ist also

$$\int_{\mathbb{R}^{n-1}} h_c(x)\, dx \geq \left(\int_{\mathbb{R}^{n-1}} f_a(x)\, dx \right)^{1-\lambda} \left(\int_{\mathbb{R}^{n-1}} g_b(x)\, dx \right)^{\lambda}.$$

Setze

$$H(c) := \int_{\mathbb{R}^{n-1}} h_c(x)\, dx, \quad F(a) := \int_{\mathbb{R}^{n-1}} f_a(x)\, dx, \quad G(b) := \int_{\mathbb{R}^{n-1}} g_b(x)\, dx.$$

Dann schreibt sich die vorhergehende Ungleichung in der Form

$$H((1-\lambda)a + \lambda b) \geq F(a)^{1-\lambda} G(b)^{\lambda}.$$

Mit dem Satz von Fubini und dem bereits bewiesenen Fall $n = 1$ ergibt sich also

$$\begin{aligned} \int_{\mathbb{E}^n} h(x)\, dx &= \int_{\mathbb{R}} \int_{\mathbb{R}^{n-1}} h_c(z)\, dz\, dc = \int_{\mathbb{R}} H(c)\, dc \\ &\geq \left(\int_{\mathbb{R}} F(a)\, da \right)^{1-\lambda} \left(\int_{\mathbb{R}} G(b)\, db \right)^{\lambda} \\ &= \left(\int_{\mathbb{E}^n} f(x)\, dx \right)^{1-\lambda} \left(\int_{\mathbb{E}^n} g(x)\, dx \right)^{\lambda}. \end{aligned}$$

■

3.2 Steiner-Formel und innere Volumina

Im vorigen Abschnitt haben wir die Oberfläche eines konvexen Körpers $K \in \mathcal{K}^n$ definiert durch

$$S(K) := \lim_{\epsilon \downarrow 0} \frac{V_n(K + \epsilon B^n) - V_n(K)}{\epsilon},$$

aber es bleibt noch die Existenz dieses Grenzwertes zu zeigen. Diese wird sich aus der viel weiter gehenden Aussage ergeben, dass die Funktion $\epsilon \mapsto V_n(K+\epsilon B^n)$ ein Polynom vom Grad n in ϵ ist. Durch die Koeffizienten werden wichtige Maßzahlen für konvexe Körper definiert. Die Verknüpfung von Volumen und Minkowskischer Addition führt noch wesentlich weiter, wie im nächsten Abschnitt klar werden wird. Hier wollen wir aber zunächst nur den Spezialfall der Addition von Kugeln behandeln, weil er zu bewegungsinvarianten und damit geometrisch besonders natürlichen und wichtigen Funktionalen führt.

Wir schreiben im folgenden für $K \in \mathcal{K}^n$ und $\rho \geq 0$

$$K_\rho := K + \rho B^n = \{x \in \mathbb{E}^n : d(K, x) \leq \rho\};$$

dies ist der *Parallelkörper* von K im Abstand ρ.

Zunächst sei $P \subset \mathbb{E}^n$ ein Polytop. Wir benutzen die metrische Projektion $p(P, \cdot)$. Für $x \in \mathbb{R}^n \setminus P$ liegt der nächste Punkt $p(P, x)$ im Rand von P und daher im relativen Inneren einer eindeutig bestimmten Seite. Daher ist durch

$$P_\rho = \bigcup_{F \in \mathcal{F}(P)} \left[P_\rho \cap p(P, \cdot)^{-1}(\operatorname{relint} F)\right] \tag{3.11}$$

eine disjunkte Zerlegung des Paralellkörpers P_ρ gegeben (man beachte, dass auch $P \in \mathcal{F}(P)$ ist). Sei nun $F \in \mathcal{F}_m(P)$ mit $m \in \{0, \ldots, n-1\}$. Aus bekannten Eigenschaften der metrischen Projektion folgt

$$P_\rho \cap p(P, \cdot)^{-1}(\operatorname{relint} F) = \operatorname{relint} F \oplus (N(P, F) \cap \rho B^n), \tag{3.12}$$

wo $\oplus$ eine direkte orthogonale Summe bezeichnet. Hier ist $N(P, F)$ der Normalenkegel von P bei F. Im folgenden bezeichnen wir mit λ_k das k-dimensionale Lebesgue-Maß in k-dimensionalen affinen Unterräumen. Aus dem Satz von Fubini ergibt sich für das Volumen der Menge (3.12)

$$V_n(P_\rho \cap p(P, \cdot)^{-1}(\operatorname{relint} F)) = \lambda_m(F)\lambda_{n-m}(N(P, F) \cap \rho B^n).$$

Da $N(P, F)$ ein Kegel ist, folgt aus bekannten Eigenschaften des Lebesgue-Maßes die Gleichung

$$\lambda_{n-m}(N(P, F) \cap \rho B^n) = \rho^{n-m}\lambda_{n-m}(N(P, F) \cap B^n).$$

Definition. Durch

$$\gamma(F, P) := \frac{\lambda_{n-m}(N(P, F) \cap B^n)}{\lambda_{n-m}(B^{n-m})}$$

ist der *äußere Winkel* von P bei F definiert.

Der Nenner kann auch in der Form

$$\lambda_{n-m}(B^{n-m}) = \lambda_{n-m}(\operatorname{lin} N(P,F) \cap B^n)$$

geschrieben werden. Er dient zur Normierung, mit dem Effekt, dass $\gamma(F,F) = 1$ ist. Man definiert daher noch $\gamma(P,P) = 1$.

Mit dieser Bezeichnung ist jetzt also

$$V_n(P_\rho \cap p(P,\cdot)^{-1}(\operatorname{relint} F)) = \rho^{n-m}\kappa_{n-m}\lambda_m(F)\gamma(F,P),$$

wobei $\kappa_k := \lambda_k(B^k)$ das Volumen der k-dimensionalen Einheitskugel bezeichnet. Es sei daran erinnert, dass

$$\kappa_n = \frac{\pi^{n/2}}{\Gamma\left(1+\frac{1}{n}\right)}$$

ist.

Insgesamt ergibt sich mit (3.11)

$$V_n(P_\rho) = \sum_{m=0}^{n} \rho^{n-m}\kappa_{n-m} \sum_{F\in\mathcal{F}_m(P)} \lambda_m(F)\gamma(F,P). \tag{3.13}$$

Definition. Für ein Polytop $P \in \mathcal{K}^n$ und für $m \in \{0,\dots,n\}$ ist

$$V_m(P) := \sum_{F\in\mathcal{F}_m(P)} \lambda_m(F)\gamma(F,P) \tag{3.14}$$

das m-te *innere Volumen* von P.

Offenbar ist $V_n(P)$ das Volumen von P; die gewählte Bezeichnung ist also konsistent. Ist $\dim P = m$, so ist $V_m(P) = \lambda_m(P)$, also das m-dimensionale Volumen von P, und es ist $V_k(P) = 0$ für $k > m$. Ist $\dim P = n$ und $F \in \mathcal{F}_{n-1}(P)$, so ist $\gamma(F,P) = \frac{1}{2}$. Daher ist

$$2V_{n-1}(P) = \sum_{F\in\mathcal{F}_{n-1}(P)} \lambda_{n-1}(F).$$

Diese Zahl wird als die *Oberfläche* von P bezeichnet.

Wir dehnen dies nun auf allgemeine konvexe Körper aus.

3.2.1 Satz und Definition. *Es gibt stetige Funktionale $V_m : \mathcal{K}^n \to \mathbb{R}$ ($m = 0,\dots,n$), so dass für $K \in \mathcal{K}^n$ und $\rho \geq 0$ die Gleichung*

$$V_n(K_\rho) = \sum_{m=0}^{n} \rho^{n-m}\kappa_{n-m}V_m(K) \tag{3.15}$$

(*Steiner-Formel*) *gilt.* V_m *heißt das m-te innere Volumen.*

Beweis. Zunächst sei $P \in \mathcal{K}^n$ ein Polytop. Nach (3.13) und (3.14) gilt für $\rho \geq 0$

$$V_n(P_\rho) = \sum_{m=0}^{n} \rho^{n-m} \kappa_{n-m} V_m(P).$$

Hier setzen wir der Reihe nach die Werte $\rho = 1, \ldots, n+1$ ein. Die so erhaltenen $n+1$ Gleichungen fassen wir als inhomogenes lineares Gleichungssystem zur Bestimmung der Größen $\kappa_n V_0(P), \kappa_{n-1} V_1(P), \ldots, \kappa_0 V_n(P)$ auf. Die Determinante dieses Systems ist eine Vandermondesche und daher nicht Null. Das System ist also eindeutig lösbar, und wir erhalten Darstellungen der Form

$$V_m(P) = \sum_{\nu=1}^{n+1} a_{m\nu} V_n(P_\nu), \qquad m = 0, \ldots, n,$$

mit Koeffizienten $a_{m\nu}$, die nicht von P abhängen. Mit diesen Konstanten definieren wir

$$V_m(K) := \sum_{\nu=1}^{n+1} a_{m\nu} V_n(K_\nu) \tag{3.16}$$

für beliebige konvexe Körper $K \in \mathcal{K}^n$. Die Abbildung $K \mapsto V_n(K_\nu) = V_n(K + \nu B^n)$ ist stetig, wie aus der Stetigkeit der Minkowskischen Addition und des Volumens folgt. Aus (3.16) folgt also die Stetigkeit der Funktion V_m auf $\mathcal{K}^n$. Da die Gleichung (3.15) für Polytope gilt, ergibt sie sich für beliebige konvexe Körper K, indem man K durch eine Folge von Polytopen approximiert und die Stetigkeit benutzt. ■

Aus der Darstellung (3.16) ergeben sich sogleich weitere Eigenschaften des Funktionals V_m (ohne dass wir die Koeffizienten $a_{m\nu}$ zu kennen brauchen). Ist $\beta : \mathbb{E}^n \to \mathbb{E}^n$ eine Bewegung, so gilt $V_n((\beta K)_\nu) = V_n(\beta K + \nu B^n) = V_n(\beta(K + \nu B^n) + t) = V_n(K_\nu)$, und aus (3.16) folgt $V_m(\beta K) = V_m(K)$. Die inneren Volumina sind also bewegungsinvariant. Ebenso sind sie invariant unter uneigentlichen Bewegungen.

Das m-te innere Volumen ist homogen vom Grad m, d.h. es gilt

$$V_m(\lambda K) = \lambda^m V_m(K) \qquad \text{für } \lambda \geq 0.$$

Für Polytope folgt das sofort aus (3.14), denn die äußeren Winkel sind invariant unter Ähnlichkeiten. Für beliebige konvexe Körper K folgt es dann, indem man K durch eine Folge von Polytopen approximiert und die Stetigkeit von V_m benutzt.

Das m-te innere Volumen eines konvexen Körpers ist unabhängig von der Dimension des umgebenden Raumes. Es genügt, einen konvexen Körper $K \subset H_{u,0}$ mit

$u \in S^{n-1}$ zu betrachten. Die inneren Volumina von K in $\mathbb{E}^n$ sind definiert durch die Steiner-Formel

$$V_n(K + \rho B^n) = \sum_{m=0}^{n-1} \rho^{n-m} \kappa_{n-m} V_m(K)$$

(die Summation erstreckt sich wegen $V_n(K) = 0$ nur bis $n-1$). Andererseits können wir $H_{u,0}$ mit $\mathbb{E}^{n-1}$ identifizieren und dort die Steiner-Formel anwenden. Mit $B^{n-1} := B^n \cap H_{u,0}$ lautet sie

$$V_{n-1}(K + \rho B^{n-1}) = \sum_{m=0}^{n-1} \rho^{n-1-m} \kappa_{n-1-m} V'_m(K),$$

wobei V'_m vorübergehend das m-te innere Volumen in $\mathbb{E}^{n-1}$ bezeichnet. Mit dem Satz von Fubini ergibt sich

$$V_n(K + \rho B^n) = \int_{-\rho}^{\rho} V_{n-1}((K + \rho B^n) \cap H_{u,t})\, dt.$$

Es ist

$$(K + \rho B^n) \cap H_{u,t} = (K + tu) + (\rho^2 - t^2)^{1/2} B^{n-1},$$

also

$$V_{n-1}((K + \rho B^n) \cap H_{u,t}) = \sum_{m=0}^{n-1} (\rho^2 - t^2)^{(n-1-m)/2} \kappa_{n-1-m} V'_m(K).$$

Das ergibt

$$\begin{aligned} V_n(K + \rho B^n) &= \sum_{m=0}^{n-1} \kappa_{n-1-m} V'_m(K) \int_{-\rho}^{\rho} (\rho^2 - t^2)^{(n-1-m)/2} dt \\ &= \sum_{m=0}^{n-1} \rho^{n-m} \kappa_{n-m} V'_m(K), \end{aligned}$$

denn es ist

$$\int_{-\rho}^{\rho} (\rho^2 - t^2)^{(n-1-m)/2} dt = \rho^{n-m} \frac{\kappa_{n-m}}{\kappa_{n-1-m}}$$

(eine Substitution ergibt ein Beta-Integral, das durch die Γ-Funktion ausgedrückt werden kann). Koeffizientenvergleich ergibt $V'_m(K) = V_m(K)$. Damit ist gezeigt, dass $V_m(K)$ nicht von der Dimension des umgebenden Raumes abhängt

Eine weitere wichtige Eigenschaft der inneren Volumina ist die Additivität.

Definition. Das Funktional $\varphi : \mathcal{K}^n \to \mathbb{R}$ heißt *additiv*, wenn

$$\varphi(K_1 \cup K_2) + \varphi(K_1 \cap K_2) = \varphi(K_1) + \varphi(K_2)$$

für alle konvexen Körper $K_1, K_2 \in \mathcal{K}^n$ mit $K_1 \cup K_2 \in \mathcal{K}^n$ gilt.

3.2.2 Hilfssatz. *Für konvexe Körper $K, L \in \mathcal{K}^n$ mit $K \cup L \in \mathcal{K}^n$ und für $\rho \geq 0$ gilt*

$$V_n((K \cup L)_\rho) + V_n((K \cap L)_\rho) = V_n(K_\rho) + V_n(L_\rho).$$

Beweis. Sei $x \in \mathbb{E}^n$. Sei $y := p(K \cup L, x)$ und o.B.d.A. $y \in K$. Dann gilt

$$p(K \cup L, x) = p(K, x).$$

Sei $z := p(L, x)$. Da $K \cup L$ konvex ist, gibt es einen Punkt $a \in [z, y]$ mit $a \in K \cap L$. Wegen $y = p(K \cup L, x)$ gilt $\|y - x\| \leq \|z - x\|$, also $\|a - x\| \leq \|z - x\|$. Wegen $a \in L$ und der Definition von z muss $a = z$ gelten und damit $z \in K \cap L$. Es folgt

$$p(K \cap L, x) = p(L, x).$$

Für $A \in \mathcal{K}^n$ bezeichne $\mathbf{1}_\rho(A, \cdot)$ die charakteristische Funktion von A_ρ, also

$$\mathbf{1}_\rho(A, a) := \begin{cases} 1 & \text{für } a \in A_\rho, \\ 0 & \text{für } a \in \mathbb{R}^n \setminus A_\rho. \end{cases}$$

Dann ergibt sich aus dem Gezeigten

$$\mathbf{1}_\rho(K \cup L, x) = \mathbf{1}_\rho(K, x), \qquad \mathbf{1}_\rho(K \cap L, x) = \mathbf{1}_\rho(L, x),$$

denn es gilt

$$x \in (K \cup L)_\rho \Leftrightarrow \|x - p(K \cup L, x)\| \leq \rho \Leftrightarrow \|x - p(K, x)\| \leq \rho \Leftrightarrow x \in K_\rho$$

und eine analoge Aussage für $K \cap L$. Da x beliebig war, gilt also

$$\mathbf{1}_\rho(K \cup L, \cdot) + \mathbf{1}_\rho(K \cap L, \cdot) = \mathbf{1}_\rho(K, \cdot) + \mathbf{1}_\rho(L, \cdot).$$

Durch Integration dieser Gleichung mit dem Lebesgue-Maß folgt nun die Behauptung. ■

Zusammen mit (3.16) ergibt das nach Koeffizientenvergleich

$$V_m(K \cup L) + V_m(K \cap L) = V_m(K) + V_m(L).$$

Die inneren Volumina sind also additiv.

Es lässt sich zeigen, dass die inneren Volumina und ihre Linearkombinationen die einzigen Funktionale auf dem Raum der konvexen Körper mit diesen drei Eigenschaften sind, d.h. es gilt der folgende Satz.

3.2.3 Satz (Funktionalsatz von Hadwiger). *Ist* $\varphi : \mathcal{K}^n \to \mathbb{R}$ *ein Funktional, das bewegungsinvariant, additiv und stetig ist, so gibt es Konstanten* $c_0, \dots, c_n \in \mathbb{R}$ *mit*

$$\varphi(K) = \sum_{m=0}^{n} c_m V_m(K)$$

für alle $K \in \mathcal{K}^n$.

Dieses Resultat stellt die besondere Bedeutung der inneren Volumina heraus und gestattet elegante Anwendungen. Es gilt als ein zentrales Ergebnis der klassischen Konvexgeometrie. Der Satz ist erstmals von H. Hadwiger 1951 bewiesen worden. Der Beweis findet sich auch in Hadwigers Buch von 1957. Er erfordert dort umfangreichere Vorbereitungen. Seit 1995 gibt es einen recht kurzen Beweis von D. Klain; er soll im wesentlichen hier wiedergegeben werden. Was wir nicht ausführen, ist ein elementares Fortsetzungsargument für additive Funktionale sowie ein analytischer Satz, dessen Herleitung spezielle Hilfsmittel (Kugelfunktionen) erfordert.

Satz 3.2.3 wird sich leicht aus dem folgenden ergeben.

3.2.4 Satz. *Sei* $\varphi : \mathcal{K}^n \to \mathbb{R}$ *ein bewegungsinvariantes, additives, stetiges Funktional, das außerdem* $\varphi(K) = 0$ *erfüllt, wenn entweder* $\dim K < n$ *ist oder* K *ein Einheitswürfel ist. Dann gilt* $\varphi = 0$.

Beweis von 3.2.4 $\Rightarrow$ 3.2.3. Wir nehmen zunächst an, Satz 3.2.4 sei bewiesen. Zum Beweis von 3.2.3 benutzen wir Induktion nach der Dimension. Für $n = 0$ ist die Behauptung trivial. Sei also $n > 0$ und die Behauptung bewiesen in Dimensionen kleiner als n. Sei $H \subset \mathbb{R}^n$ eine Hyperebene. Die Einschränkung von φ auf die in H enthaltenen konvexen Körper ist additiv, stetig und invariant unter Bewegungen von H in sich. Nach der Induktionsannahme gibt es also Konstanten $c_0, \dots, c_{n-1}$ mit

$$\varphi(K) = \sum_{i=0}^{n-1} c_i V_i(K) \tag{3.17}$$

für alle konvexen Körper $K \subset H$ (hier wird benutzt, dass die inneren Volumina nicht von der Dimension des umgebenden Raumes abhängen). Wegen der Bewegungsinvarianz von φ gilt (3.17) für alle $K \in \mathcal{K}^n$ mit $\dim K < n$. Wir definieren ψ durch

$$\psi(K) := \varphi(K) - \sum_{i=0}^{n} c_i V_i(K), \qquad K \in \mathcal{K}^n,$$

wobei c_n so gewählt sei, dass ψ auf einem gegebenen Einheitswürfel verschwindet. Das Funktional ψ erfüllt dann die Voraussetzungen von Satz 3.2.4, also ist $\psi = 0$. Daraus folgt die Behauptung. ■

Beweis von Satz 3.2.4. Wir verwenden Induktion nach der Dimension. Für $n = 0$ ist nichts zu beweisen. Ist $n = 1$, so verschwindet φ auf (abgeschlossenen) Strecken der Länge 1, also auch auf Strecken der Länge $1/k$ für $k \in \mathbb{N}$ und daher auf Strecken rationaler Länge. Wegen der Stetigkeit verschwindet φ auf allen Strecken und somit auf $\mathcal{K}^1$.

Nun nehmen wir an, dass $n > 1$ ist und dass die Behauptung in kleineren Dimensionen als n bewiesen ist. Sei $H \subset \mathbb{R}^n$ eine Hyperebene und I eine abgeschlossene Einheitsstrecke senkrecht zu H. Für konvexe Körper $K \subset H$ definieren wir

$$\psi(K) := \varphi(K + I).$$

Dann sieht man sofort, dass ψ relativ zu H die Eigenschaften von φ im Satz hat. Aus der Induktionsannahme folgt also $\psi = 0$. Für gegebenes $K \subset H$ gilt daher $\varphi(K + I) = 0$. Ein analoges Argument wie oben im Fall $n = 1$ zeigt, dass $\varphi(K + S) = 0$ ist für jede abgeschlossene Strecke S senkrecht zu H. Wegen der Bewegungsinvarianz verschwindet φ also auf geraden konvexen Zylindern.

Sei wieder $K \subset H$ ein konvexer Körper, und sei $S = \operatorname{conv}\{0, s\}$ eine Strecke, die nicht parallel zu H ist. Wird $m \in \mathbb{N}$ genügend groß gewählt, so kann der Zylinder $Z := K + mS$ so durch eine zu S senkrechte Hyperebene H' geschnitten werden, dass für die von H' berandeten abgeschlossenen Halbräume H^-, H^+ die Inklusionen

$$K \subset H^-, \qquad K + ms \subset H^+$$

gelten. Dann ist

$$\overline{Z} := [(Z \cap H^-) + ms] \cup (Z \cap H^+)$$

ein gerader Zylinder. Es gilt also $\varphi(\overline{Z}) = 0$ und daher

$$m\varphi(K + S) = \varphi(Z) = \varphi(\overline{Z}) = 0.$$

Damit ist gezeigt, dass φ auf beliebigen Zylindern verschwindet.

Man kann mit elementaren Hilfsmitteln zeigen (was hier aber nicht ausgeführt wird), dass die additive Funktion φ eine additive Fortsetzung (ebenfalls mit φ bezeichnet) auf die Menge der endlichen Vereinigungen von Polytopen besitzt. Es folgt, dass

$$\varphi\left(\bigcup_{i=1}^{k} P_i\right) = \sum_{i=1}^{k} \varphi(P_i)$$

ist, falls $P_1, \dots, P_k$ Polytope mit $\dim(P_i \cap P_j) < n$ für $i \neq j$ sind.

Sei nun P ein konvexes Polytop und S eine Strecke. Die Minkowski-Summe $P + S$ hat eine Zerlegung

$$P + S = \bigcup_{i=1}^{k} P_i,$$

wo $P_1 = P$ ist, jedes Polytop P_i mit $i > 1$ ein konvexer Zylinder ist und $\dim(P_i \cap P_j) < n$ für $i \neq j$ gilt. Es folgt $\varphi(P+S) = \varphi(P)$. Durch Induktion ergibt sich $\varphi(P+Z) = \varphi(P)$, wenn Z eine endliche Summe von Strecken ist. Wegen der Stetigkeit von φ folgt daraus $\varphi(K+Z) = \varphi(K)$, wenn K ein konvexer Körper ist und Z ein konvexer Körper, der durch endliche Summen von Strecken approximiert werden kann. Ein solcher Körper wird als *Zonoid* bezeichnet.

Wir müssen nun einen analytischen Satz benutzen, der im Rahmen dieser Vorlesung nicht bewiesen wird (der Beweis erfordert Kugelfunktionen). Sei G eine gerade Funktion der Klasse C^∞ auf der Sphäre S^{n-1}. Dann gibt es eine stetige gerade Lösung g der Integralgleichung

$$G(u) = \int_{S^{n-1}} |\langle u, v\rangle| g(v)\, d\omega(v), \qquad u \in S^{n-1}.$$

Darin bezeichnet ω das sphärische Lebesgue-Maß. Es sei nun K ein konvexer Körper mit Symmetriezentrum 0 und mit einer Stützfunktion der Klasse C^∞. Dann gibt es also eine gerade stetige Funktion g auf S^{n-1} mit

$$h(K,u) = \int_{S^{n-1}} |\langle u, v\rangle| g(v)\, d\omega(v).$$

Es sei g^+ der Positivteil und g^- der Negativteil von g. Dann gilt

$$h(K,u) + \int_{S^{n-1}} |\langle u, v\rangle| g^-(v)\, d\omega(v) = \int_{S^{n-1}} |\langle u, v\rangle| g^+(v)\, d\omega(v).$$

Die Funktion $u \mapsto |\langle u, v\rangle|$ ist die Stützfunktion der Strecke $\mathrm{conv}\,\{-v, v\}$. Die beiden Integrale sind daher sublineare Funktionen von u, also Stützfunktionen konvexer Körper. Es gibt also konvexe Körper Z_1, Z_2 mit

$$h(K,u) + h(Z_1,u) = h(Z_2,u) \qquad \text{für } u \in S^{n-1},$$

also mit $K + Z_1 = Z_2$. Indem man die Integrale durch Riemann-Summen approximiert, sieht man, dass Z_1 und Z_2 durch endliche Summen von Strecken approximiert werden können; sie sind also Zonoide. Nach dem vorher Bewiesenen folgt daher

$$\varphi(K) = \varphi(K + Z_1) = \varphi(Z_2) = 0.$$

Da jeder zentralsymmetrische konvexe Körper approximiert werden kann durch zentralsymmetrische Körper mit Stützfunktionen der Klasse C^∞ (auch das wollen wir hier allerdings nicht beweisen), folgt $\varphi(K) = 0$ für alle zentralsymmetrischen konvexen Körper.

Nun sei Δ ein Simplex, o.B.d.A. etwa $\Delta = \mathrm{conv}\,\{0, v_1, \dots, v_n\}$. Sei $v := v_1 + \ldots + v_n$ und $\Delta' := \mathrm{conv}\,\{v, v - v_1, \dots, v - v_n\}$. Dann ist $\Delta' = -\Delta + v$. Die Vektoren

$v_1, \dots, v_n$ spannen ein Parallelotop P auf. Es ist die Vereinigung von Δ, Δ' und dem Teil von P, der zwischen den beiden parallelen Hyperebenen liegt, die von $v_1, \dots, v_n$ bzw. von $v - v_1, \dots, v - v_n$ aufgespannt werden. Dieser Teil, etwa Q, ist ein zentralsymmetrisches Polytop, und es gilt $\dim(\Delta \cap Q) = \dim(\Delta' \cap Q) = n-1$. Damit folgt $0 = \varphi(P) = \varphi(\Delta) + \varphi(\Delta')$, also $\varphi(-\Delta) = -\varphi(\Delta)$. Ist die Dimension n gerade, so geht $-\Delta$ aus Δ durch eine (orientierungserhaltende) Bewegung hervor. Aus der Bewegungsinvarianz von φ folgt also $\varphi(\Delta) = 0$. Ist die Dimension $n > 1$ ungerade, so zerlegen wir das Simplex Δ in der folgenden Weise. Sei z der Mittelpunkt der Inkugel von Δ, und sei p_i der Punkt, wo die Inkugel die Facette F_i von Δ berührt ($i = 1, \dots, n+1$). Für $i \neq j$ sei Q_{ij} die konvexe Hülle der Seite $F_i \cap F_j$ und der Punkte z, p_i, p_j. Das Polytop Q_{ij} ist invariant unter der Spiegelung an der Hyperebene, die von $F_i \cap F_j$ und z aufgespannt wird. Seien $Q_1, \dots, Q_m$ die Polytope Q_{ij} $(1 \leq i < j \leq n+1)$ in beliebiger Reihenfolge. Dann gilt $P = \bigcup_{r=1}^m Q_r$ und $\dim(Q_r \cap Q_s) < n$ für $r \neq s$. Da $-Q_r$ das Bild von Q_r unter einer eigentlichen Bewegung ist, gilt $\varphi(-\Delta) = \sum \varphi(-Q_r) = \sum \varphi(Q_r) = \varphi(\Delta)$. Also gilt $\varphi(\Delta) = 0$ für jedes Simplex Δ.

Indem wir ein gegebenes Polytop in Simplexe zerlegen, erhalten wir $\varphi(P) = 0$. Aus der Stetigkeit von φ folgt $\varphi(K) = 0$ für alle konvexen Körper K. Damit ist der Beweis von Satz 3.2.4 abgeschlossen. ∎

Wir gehen nun auf die geometrische Bedeutung der inneren Volumina ein. Zu V_n ist nichts zu sagen; das ist einfach das Volumen. Das Funktional V_{n-1} hat für Polytope P die Darstellung

$$2V_{n-1}(P) = \sum_{F \in \mathcal{F}_{n-1}(P)} \lambda_{n-1}(F).$$

Wir definieren daher jetzt:

Definition. Für $K \in \mathcal{K}^n$ ist $2V_{n-1}(K) =: S(K)$ die *Oberfläche* von K.

Aus der Steiner-Formel bekommt man sofort

$$S(K) = \lim_{\rho \downarrow 0} \frac{V_n(K + \rho B^n) - V_n(K)}{\rho}, \tag{3.18}$$

was eine anschauliche Zurückführung des Oberflächenbegriffs auf den Volumenbegriff zum Ausdruck bringt.

Die Oberfläche eines konvexen Körpers ist hier also in einer zu (3.18) äquivalenten Weise definiert worden. Dies ist eine Definition und kein Satz, da wir hier den Standpunkt einnehmen, dass uns für so allgemeine Hyperflächen, wie es die Ränder konvexer Körper sein können, noch kein Oberflächenbegriff zur Verfügung steht. Es läßt sich aber zeigen, dass diese Definition im Einklang steht mit allen vernünftigen Oberflächenbegriffen, die anderweitig definiert werden (z.B. in der Analysis oder Differentialgeometrie im Fall hinreichend glatter Hyperflächen, oder mit dem (passend normierten) $(n-1)$-dimensionalen Hausdorff-Maß).

Am anderen Ende der Skala sehen wir uns zunächst V_0 an. Nach (3.14) ist für ein Polytop P

$$V_0(P) = \sum_{F \in \mathcal{F}_0(P)} \gamma(F, P).$$

Die Normalenkegel der Ecken eines nichtleeren Polytops überdecken ganz $\mathbb{E}^n$ und haben paarweise keine inneren Punkte gemeinsam. Nach Definition des äußeren Winkels ergibt sich daraus $V_0(P) = 1$. Wegen der Stetigkeit von V_0 gilt

$$V_0(K) = 1 \qquad \text{für alle } K \in \mathcal{K}^n.$$

Man setzt noch $V_0(\emptyset) := 0$ und nennt V_0 die *Eulersche Charakteristik*. (Der Grund für diese Bezeichnung wird allerdings erst klar, wenn man additive Fortsetzungen der Funktionale V_m auf endliche Vereinigungen konvexer Körper betrachtet und dann feststellt, dass V_0 auf solchen Mengen mit der in der Algebraischen Topologie eingeführten Eulerschen Charakteristik übereinstimmt.)

Das Funktional V_1 hat eine besondere Eigenschaft. Um dies zu sehen, wählen wir zunächst zwei stark isomorphe Polytope P_1, P_2. Sei $m \in \{0, \dots, n-1\}$, $F \in \mathcal{F}_m(P_1)$ und $u \in S^{n-1}$ ein Vektor mit $F(P_1, u) = F$. Wegen der starken Isomorphie gilt für die Seite $G := F(P_2, u)$ dann $\dim G = m$ und $N(P_2, G) = N(P_1, F)$. Ferner ist $F(P_1 + P_2, u) = F(P_1, u) + F(P_2, u) = F + G$ und $N(P_1 + P_2, F + G) = N(P_1, F) = N(P_2, G)$. Für $m = 1$ ergibt sich daher aus (3.14) die Gleichung

$$V_1(P_1 + P_2) = V_1(P_1) + V_1(P_2).$$

Durch Approximation folgt $V_1(K + L) = V_1(K) + V_2(L)$ für $K, L \in \mathcal{K}^n$. Allgemein wird ein Funktional $\varphi : \mathcal{K}^n \to \mathbb{R}$ mit der Eigenschaft $\varphi(K + L) = \varphi(K) + \varphi(L)$ für $K, L \in \mathcal{K}^n$ als *Minkowski-additiv* bezeichnet. Das Funktional V_1 ist also bewegungsinvariant, Minkowski-additiv und stetig. Wir wollen nun (ohne Verwendung des Hadwigerschen Funktionalsatzes) zeigen, dass es bis auf einen konstanten Faktor überhaupt nur ein Funktional mit diesen Eigenschaften gibt. Dazu dienen die folgenden Vorbereitungen.

Definition. Der konvexe Körper $K' \in \mathcal{K}^n$ heißt ein *Drehmittel* des Körpers $K \in \mathcal{K}^n$, wenn es eine Zahl $r \in \mathbb{N}$ und Drehungen $\vartheta_1, \dots, \vartheta_r$ des $\mathbb{E}^n$ gibt mit $K' = \frac{1}{r}(\vartheta_1 K + \dots + \vartheta_r K)$.

3.2.5 Satz (Kugelungs-Satz). *Zu jedem konvexen Körper $K \in \mathcal{K}^n$ existiert eine Folge von Drehmitteln von K, die gegen eine Kugel konvergiert.*

Beweis. Für $L \in \mathcal{K}^n$ sei $R(L)$ der Radius der kleinsten Kugel mit Mittelpunkt 0, die L enthält. Die Funktion R ist auf $\mathcal{K}^n$ stetig: Aus $\delta(K_1, K_2) = \alpha$ folgt

$K_1 \subset K_2 + \alpha B^n \subset R(K_2)B^n + \alpha B^n = (R(K_2) + \alpha)B^n$, also $R(K_1) \leq R(K_2) + \alpha$. Vertauschung von K_1 und K_2 ergibt $|R(K_1) - R(K_2)| \leq \delta(K_1, K_2)$.

Sei nun $K \in \mathcal{K}^n$ gegeben. Setze

$$R_0 := \inf \{R(K') : K' \text{ Drehmittel von } K\}.$$

Eine Kugel mit Mittelpunkt 0, die K enthält, enthält auch alle Drehmittel von K. Die Menge der Drehmittel von K ist also beschränkt. Nach dem Auswahlsatz von Blaschke existiert daher eine Folge $(K_j)_{j\in\mathbb{N}}$ von Drehmitteln von K mit

$$\lim_{j\to\infty} R(K_j) = R_0 \qquad \text{und} \qquad \lim_{j\to\infty} K_j = K_0$$

für einen konvexen Körper $K_0 \in \mathcal{K}^n$. Wegen der Stetigkeit von R ist $R(K_0) = R_0$.

Angenommen, K_0 wäre nicht die Kugel mit Mittelpunkt 0 und Radius R_0. Dann gibt es einen Vektor $u_0 \in S^{n-1}$ mit $h(K_0, u_0) < R_0$. Wegen der Stetigkeit der Stützfunktion existiert eine offene Umgebung $U \subset S^{n-1}$ von u_0 mit $h(K_0, u) < R_0$ für alle $u \in U$. Wegen der Kompaktheit von S^{n-1} gibt es endlich viele Drehungen $\vartheta_1, \ldots, \vartheta_r$ mit $\bigcup_{j=1}^r \vartheta_j U = S^{n-1}$. Setze

$$K^* := \frac{1}{r}(\vartheta_1 K_0 + \ldots + \vartheta_r K_0).$$

Sei $u \in S^{n-1}$. Es gibt ein $i \in \{1, \ldots, r\}$ mit $u \in \vartheta_i U$, also $\vartheta_i^{-1} u \in U$ und daher $h(K_0, \vartheta_i^{-1} u) < R_0$. Es folgt

$$h(K^*, u) = \frac{1}{r} \sum_{j=1}^{r} h(\vartheta_j K_0, u) = \frac{1}{r} \sum_{j=1}^{r} h(K_0, \vartheta_j^{-1} u) < R_0.$$

Wegen der Stetigkeit der Stützfunktion existiert ein $\epsilon > 0$ mit $h(K^*, u) \leq R_0 - \epsilon$ für alle $u \in S^{n-1}$, also ist $R(K^*) \leq R_0 - \epsilon$. Aus $K_j \to K_0$ folgt

$$K_j^* := \frac{1}{r}(\vartheta_1 K_j + \ldots + \vartheta_r K_j) \to K^*,$$

also $R(K_j^*) < R_0$ für genügend große j. Da K_j^* ebenfalls ein Drehmittel von K ist, ist das ein Widerspruch. Somit ist K_0 eine Kugel, woraus die Behauptung folgt. ∎

Damit lässt sich nun zeigen, dass das Funktional V_1 bis auf einen konstanten Faktor übereinstimmt mit der mittleren Breite, die folgendermaßen erklärt ist. Ist $K \in \mathcal{K}^n$ und $u \in S^{n-1}$, so ist

$$b(K, u) := h(K, u) + h(K, -u)$$

die *Breite* von K in Richtung u; das ist gerade der Abstand der beiden zu u senkrechten Stützebenen von K. Die *mittlere Breite* von K ist der Mittelwert dieser Breite bezüglich des sphärischen Lebesgue-Maßes auf S^{n-1}, also

$$b(K) := \frac{1}{\omega(S^{n-1})} \int_{S^{n-1}} b(K,u)\, d\omega(u).$$

Wegen $\omega(S^{n-1}) = n\kappa_n$ und der Spiegelungssymmetrie von ω ist auch

$$b(K) := \frac{2}{n\kappa_n} \int_{S^{n-1}} h(K,u)\, d\omega(u).$$

Die mittlere Breite ist offenbar bewegungsinvariant, Minkowski-additiv und stetig. Der folgende Satz zeigt, dass sie durch diese Eigenschaften bis auf einen konstanten Faktor eindeutig bestimmt ist.

3.2.6 Satz. *Ist $\varphi : \mathcal{K}^n \to \mathbb{R}$ bewegungsinvariant, Minkowski-additiv und stetig, so gibt es eine Konstante $c \in \mathbb{R}$ mit $\varphi(K) = cb(K)$ für alle $K \in \mathcal{K}^n$.*

Beweis. Sei $K \in \mathcal{K}^n$. Wegen $2K = K + K$ und der Minkowski-Additivität von φ gilt $\varphi(2K) = 2\varphi(K)$. Mit Induktion erhält man $\varphi(pK) = p\varphi(K)$ für $p \in \mathbb{N}$ und dann für $p, q \in \mathbb{N}$

$$p\varphi(K) = \varphi(pK) = \varphi\left(q\frac{p}{q}K\right) = q\varphi\left(\frac{p}{q}K\right),$$

also $\varphi(rK) = r\varphi(K)$ für rationales $r \geq 0$. Wegen der Stetigkeit von φ folgt $\varphi(\lambda K) = \lambda\varphi(K)$ für $\lambda \geq 0$.

Sei $K \in \mathcal{K}^n$. Nach Satz 3.2.5 gibt es eine Folge $(K_j)_{j\in\mathbb{N}}$ von Drehmitteln von K mit $K_j \to B_r$, wo B_r eine Kugel vom Radius r ist (eventuell $r = 0$). Jedes K_j ist von der Form

$$K_j = \frac{1}{m}(\vartheta_1 K + \ldots + \vartheta_m K),$$

und es folgt $\varphi(K_j) = \varphi(K)$. Da φ stetig ist, gilt $\varphi(K) = \lim \varphi(K_j) = \varphi(B_r) = \varphi(rB^n + t) = r\varphi(B^n)$. Da die mittlere Breite b dieselben Eigenschaften wie φ hat, ist auch $b(K) = rb(B^n) = 2r$, also ist $\varphi(K) = cb(K)$ mit $c = \frac{1}{2}\varphi(B^n)$. ∎

Hieraus ergibt sich nun, weil das Funktional V_1 bewegungsinvariant, Minkowski-additiv und stetig ist, dass

$$V_1(K) = cb(K) \quad \text{für } K \in \mathcal{K}^n$$

mit einer Konstanten c gilt. Diese Konstante findet man, wenn man für K eine Kugel einsetzt. Nach (3.13) ist

$$\sum_{m=0}^{n} \rho^{n-m}\kappa_{n-m}V_m(B^n) \;=\; V_n((B^n)_\rho) = V_n(B^n + \rho B^n) = (1+\rho)^n V_n(B^n)$$

$$= \quad \kappa_n \sum_{m=0}^{n} \binom{n}{m} \rho^{n-m}$$

für $\rho \geq 0$. Koeffizientenvergleich ergibt

$$V_m(B^n) = \binom{n}{m} \frac{\kappa_n}{\kappa_{n-m}}. \tag{3.19}$$

Für $m = 1$ folgt

$$2c = cb(B^n) = V_1(B^n) = \binom{n}{1} \frac{\kappa_n}{\kappa_{n-1}},$$

also

$$V_1 = \frac{n\kappa_n}{2\kappa_{n-1}} b. \tag{3.20}$$

In den Dimensionen 2 und 3 sind damit alle inneren Volumina anschaulich gedeutet. Es gibt weitere, in allen Dimensionen gültige Interpretationen. Eine Möglichkeit ist differentialgeometrischer Art. Für konvexe Körper mit genügend glatter Berandung interpretiert sie die inneren Volumina als Integrale gewisser Krümmungsfunktionen. Darauf können wir hier nicht eingehen.

Wir wollen aber eine integralgeometrische Deutung der inneren Volumina erläutern. Sie verallgemeinert die Darstellung von V_1 als Vielfaches der mittleren Breite; die mittlere Breite kann auch gedeutet werden als Mittelwert von eindimensionalen Projektionsvolumina.

Wir bezeichnen zu einem Einheitsvektor $v \in S^{n-1}$ mit Π_v die Orthogonalprojektion auf den zu v orthogonalen linearen Unterraum $v^\perp$. Sei P ein n-Polytop, seien $F_1, \dots, F_k$ seine Facetten, und sei u_i der äußere Normalenvektor von F_i. Dann gilt

$$V_{n-1}(\Pi_v F_i) = |\langle u_i, v\rangle| V_{n-1}(F_i).$$

Durch die Projektion Π_v werden der „obere Rand“ $\bigcup_{\langle u_i,v\rangle>0} F_i$ und der „untere Rand“ $\bigcup_{\langle u_i,v\rangle<0} F_i$ jeweils bijektiv auf $\Pi_v P$ abgebildet. Daher gilt

$$V_{n-1}(\Pi_v P) = \frac{1}{2} \sum_{i=1}^{k} |\langle u_i, v\rangle| V_{n-1}(F). \tag{3.21}$$

Ehe wir fortfahren, wollen wir bemerken, dass offenbar auch

$$\sum_{\langle u_i,v\rangle>0} \langle u_i, v\rangle V_{n-1}(F_i) = \sum_{\langle u_i,v\rangle<0} -\langle u_i, v\rangle V_{n-1}(F_i)$$

folgt. Da dies für alle $v \in S^{n-1}$ gilt, ergibt sich die Relation

$$\sum_{i=1}^{k} u_i V_{n-1}(F_i) = 0. \tag{3.22}$$

Nach dieser Nebenbemerkung integrieren wir nun (3.21) über alle v mit dem sphärischen Lebesgue-Maß ω und erhalten

$$\begin{aligned}\int\limits_{S^{n-1}} V_{n-1}(\Pi_v P)\, d\omega(v) &= \frac{1}{2}\sum_{i=1}^{k} \int\limits_{S^{n-1}} |\langle u_i, v\rangle|\, d\omega(v) V_{n-1}(F_i) \\ &= cS(P)\end{aligned}$$

mit

$$c = \frac{1}{2} \int\limits_{S^{n-1}} |\langle u_i, v\rangle|\, d\omega(v).$$

Dieses Integral hängt wegen der Drehinvarianz des Maßes ω nicht von dem Einheitsvektor u_i ab.

Ist $K \in \mathcal{K}^n$ ein beliebiger konvexer Körper, so können wir K durch eine Folge $(P_j)_{j\in\mathbb{N}}$ von Polytopen approximieren. Da die Folge $(V_{n-1}(\Pi_v P_j))_{j\in\mathbb{N}}$ gegen $V_{n-1}(\Pi_v K)$ konvergiert und durch eine Konstante nach oben beschränkt ist und da S stetig ist, folgt

$$\int\limits_{S^{n-1}} V_{n-1}(\Pi_v K)\, d\omega(v) = cS(K).$$

Wählen wir $K = B^n$ und beachten

$$S(B^n) = \lim_{\lambda\downarrow 0} \frac{1}{\lambda}[V_n(B^n + \lambda B^n) - V_n(B^n)] = n\kappa_n = \omega(S^{n-1}),$$

so erhalten wir

$$\kappa_{n-1} S(B^n) = \int\limits_{S^{n-1}} V_{n-1}(\Pi_v B^n)\, d\omega(v) = cS(B^n),$$

also $c = \kappa_{n-1}$. Damit haben wir gezeigt:

3.2.7 Satz (Oberflächenformel von Cauchy). *Für die Oberfläche eines konvexen Körpers $K \in \mathcal{K}^n$ gilt*

$$S(K) = \frac{1}{\kappa_{n-1}} \int\limits_{S^{n-1}} V_{n-1}(\Pi_v K)\, d\omega(v). \tag{3.23}$$

Die Oberfläche eines konvexen Körpers ist also bis auf einen konstanten Faktor der Mittelwert seiner Projektionsinhalte.

Die Integraldarstellung (3.23) läßt sich auf die anderen inneren Volumina übertragen. Dazu bedienen wir uns eines einfachen Kunstgriffs. Er besteht in der

Ersetzung von K in (3.23) durch Parallelkörper mit variablem Parameter, Entwicklungen nach der Steiner-Formel und anschließendem Koeffizientenvergleich. Seien $\rho, \lambda > 0$. Wenden wir auf beide Seiten der Gleichung

$$V_n((K_\rho)_\lambda) = V_n(K_{\rho+\lambda})$$

die Steiner-Formel (3.15) an, so ergibt sich

$$\begin{aligned}
\sum_{m=0}^{n} \lambda^{n-m} \kappa_{n-m} V_m(K_\rho) &= \sum_{k=0}^{n} (\rho+\lambda)^{n-k} \kappa_{n-k} V_k(K) \\
&= \sum_{k=0}^{n} \sum_{j=0}^{n-k} \binom{n-k}{j} \rho^{n-k-j} \lambda^j \kappa_{n-k} V_k(K) \\
&= \sum_{k=0}^{n} \sum_{m=k}^{n} \binom{n-k}{n-m} \rho^{m-k} \lambda^{n-m} \kappa_{n-k} V_k(K) \\
&= \sum_{m=0}^{n} \lambda^{n-m} \sum_{k=0}^{m} \binom{n-k}{n-m} \rho^{m-k} \kappa_{n-k} V_k(K).
\end{aligned}$$

Vergleich der Koeffizienten von λ^{n-m} ergibt

$$\kappa_{n-m} V_m(K_\rho) = \sum_{k=0}^{m} \rho^{m-k} \binom{n-k}{n-m} \kappa_{n-k} V_k(K). \tag{3.24}$$

Alle inneren Volumina genügen also Gleichungen vom Typ der Steiner-Formel. Wir benötigen hier jetzt nur den Spezialfall $m = n-1$; er lautet

$$2V_{n-1}(K_\rho) = \sum_{k=0}^{n-1} \rho^{n-1-k} (n-k) \kappa_{n-k} V_k(K). \tag{3.25}$$

Nun ersetzen wir in der Cauchyschen Oberflächenformel (3.23) den Körper K durch K_ρ. Es gilt

$$\Pi_v(K_\rho) = \Pi_v K + \rho \Pi_v B^n = (\Pi_v K)'_\rho,$$

wo der Strich anzeigt, dass sich die Parallelkörperbildung auf den Raum $v^\perp$ bezieht. Auf die aus (3.23) resultierende Gleichung

$$2V_{n-1}(K_\rho) = \frac{1}{\kappa_{n-1}} \int\limits_{S^{n-1}} V_{n-1}((\Pi_v K)'_\rho)\, d\omega(v)$$

wenden wir links die Gleichung (3.25) an und rechts die Steiner-Formel (3.15) im $(n-1)$-dimensionalen Raum $v^\perp$. Das liefert

$$\sum_{m=0}^{n-1} \rho^{n-1-m} (n-m) \kappa_{n-m} V_m(K) = \frac{1}{\kappa_{n-1}} \sum_{m=0}^{n-1} \rho^{n-1-m} \kappa_{n-1-m} \int\limits_{S^{n-1}} V_k(\Pi_v K)\, d\omega(v).$$

Hier haben wir wieder von der Tatsache Gebrauch gemacht, dass die inneren Volumina nicht von der Dimension des umgebenden Raumes abhängen. Koeffizientenvergleich ergibt eine Darstellung für $V_m(K)$, die wir als Satz festhalten.

3.2.8 Satz (Integralrekursion von Kubota). *Für $K \in \mathcal{K}^n$ und $m = 1, \dots, n-1$ gilt*

$$V_m(K) = \frac{\kappa_{n-1-m}}{(n-m)\kappa_{n-1}\kappa_{n-m}} \int\limits_{S^{n-1}} V_m(\Pi_v K)\, d\omega(v). \tag{3.26}$$

Das m-te innere Volumen von K ist also bis auf einen konstanten Faktor der Mittelwert der m-ten inneren Volumina der Projektionen von K auf Hyperebenen. Für festes v kann analog $V_m(\Pi_v K)$ bis auf einen Faktor als Mittelwert aus den Werten von V_m auf Projektionen von $\Pi_v K$ gewonnen werden. So kann man fortfahren, bis schließlich $V_m(K)$ durch mehrfache Integrationen auf m-dimensionale Volumina zurückgeführt ist. Das m-te innere Volumen von K kann also bis auf einen Faktor gedeutet werden als Mittelwert der m-dimensionalen Volumina aller Orthogonalprojektionen von K auf die m-dimensionalen linearen Unterräume. Da man die Volumina von Projektionen auch Quermaße nennt, ist in der Literatur für die inneren Volumina (mit anderer Normierung) auch die Bezeichnung *Quermaßintegrale* gebräuchlich.

Eine wichtige Folgerung aus der Integralrekursion (3.26) (mit Induktion nach der Dimension) ist die Monotonie der inneren Volumina bezüglich der Inklusion: Sind $K, L \in \mathcal{K}^n$ konvexe Körper mit $K \subset L$, so gilt $V_m(K) \le V_m(L)$. Ist außerdem $K \ne L$ und $\dim K \ge m$, so ist $V_m(K) < V_m(L)$. Insbesondere ergibt sich $V_m(K) \ge 0$ für alle $K \in \mathcal{K}^n$ und

$$V_m(K) > 0 \quad \Leftrightarrow \quad \dim K \ge m.$$

Zwischen den inneren Volumina besteht eine Reihe von Ungleichungen; zu ihnen gehört auch die isoperimetrische Ungleichung zwischen Volumen und Oberfläche. Einige dieser Ungleichungen lassen sich aus dem Brunn-Minkowskischen Satz herleiten, vermöge des folgenden Hilfssatzes.

3.2.9 Hilfssatz. *Sei $n \ge 2$ und*

$$f(\lambda) = \left[\sum_{i=0}^{n} \binom{n}{i} a_i (1-\lambda)^{n-i} \lambda^i\right]^{1/n} \qquad \textit{für } 0 \le \lambda \le 1$$

mit positiven Koeffizienten $a_0, \dots, a_n$. Die Funktion f sei konkav. Dann gilt

$$a_1^n \ge a_0^{n-1} a_n$$

mit Gleichheit genau dann, wenn f linear ist, und

$$a_1^2 \geq a_0 a_2.$$

Beweis. Differentiation ergibt

$$\begin{aligned} nf(\lambda)^{n-1}f'(\lambda) &= -\sum_{i=0}^{n-1}\binom{n}{i}a_i(n-i)(1-\lambda)^{n-i-1}\lambda^i \\ &\quad +\sum_{i=0}^{n-1}\binom{n}{i+1}a_{i+1}(1-\lambda)^{n-i-1}(i+1)\lambda^i \\ &= n\sum_{i=0}^{n-1}\binom{n-1}{i}(a_{i+1}-a_i)(1-\lambda)^{n-1-i}\lambda^i \end{aligned}$$

und

$$\begin{aligned} & n(n-1)f(\lambda)^{n-2}(f'(\lambda))^2 + nf(\lambda)^{n-1}f''(\lambda) \\ & = n(n-1)\sum_{i=0}^{n-2}\binom{n-2}{i}(a_{i+2}-2a_{i+1}+a_i)(1-\lambda)^{n-2-i}\lambda^i. \end{aligned}$$

Setzen wir

$$\psi(\lambda) := (1-\lambda)f(0) + \lambda f(1) - f(\lambda) \qquad \text{für } 0 \leq \lambda \leq 1,$$

so ist ψ konvex, und es gilt $\psi(0) = \psi(1) = 0$. Daher ist $\psi'(0) \leq 0$, und es gilt $\psi'(0) = 0$ genau dann, wenn $\psi \equiv 0$ ist. Nun ist

$$\begin{aligned} \psi'(0) &= -f(0) + f(1) - f'(0) \\ &= -a_0^{1/n} + a_n^{1/n} - \frac{a_1 - a_0}{a_0^{(n-1)/n}} \\ &= -\frac{1}{a_0^{(n-1)/n}}\left[a_1 - a_0^{(n-1)/n}a_n^{1/n}\right], \end{aligned}$$

woran man die erste Behauptung abliest.

Aus der Konvexität von ψ folgt ferner $\psi''(0) \geq 0$, also

$$\begin{aligned} 0 &\leq \psi''(0) = -f''(0) \\ &= (n-1)\left[\frac{1}{a_0^{1/n}}\frac{(a_1-a_0)^2}{a_0^{(2n-2)/n}} - \frac{a_2 - 2a_1 + a_0}{a_0^{(n-1)/n}}\right] \\ &= \frac{n-1}{a_0^{(2n-1)/n}}\left[a_1^2 - a_0 a_2\right] \end{aligned}$$

und damit die zweite Ungleichung. ■

Nach dem Brunn-Minkowskischen Satz 3.1.1 (und der anschließenden Folgerung) ist durch

$$f(\lambda) := V_n((1-\lambda)K_0 + \lambda K_1)^{1/n} \quad \text{für } \lambda \in [0,1] \tag{3.27}$$

bei beliebigen konvexen Körpern K_0, K_1 eine konkave Funktion gegeben. Mit $K_0 = K \in \mathcal{K}^n$, $K_1 = B^n$ ist nach der Steiner-Formel (Satz 3.2.1)

$$\begin{aligned} f(\lambda)^n &= \sum_{m=0}^{n} \lambda^{n-m}\kappa_{n-m}V_m((1-\lambda)K) \\ &= \sum_{i=0}^{n}(1-\lambda)^{n-i}\lambda^i\kappa_i V_{n-i}(K). \end{aligned}$$

Aus Hilfssatz 3.2.9 ergeben sich daher die Ungleichungen (mit $V_i = V_i(K)$)

$$\left(\frac{2}{n}V_{n-1}\right)^n \geq \kappa_n V_n^{n-1} \tag{3.28}$$

und

$$\left(\frac{2}{n}V_{n-1}\right)^2 \geq \frac{\kappa_2}{\binom{n}{2}}V_{n-2}V_n. \tag{3.29}$$

Andererseits kann man in (3.27) auch $K_0 = B^n$ und $K_1 = K$ setzen. Dann erhält man in analoger Weise die Ungleichungen

$$\left(\frac{\kappa_{n-1}}{n}V_1\right)^n \geq \kappa_n^{n-1}V_n \tag{3.30}$$

und

$$\left(\frac{\kappa_{n-1}}{n}V_1\right)^2 \geq \frac{\kappa_n\kappa_{n-2}}{\binom{n}{2}}V_2. \tag{3.31}$$

Da $2V_{n-1} = S$ ist, also die Oberfläche, ist (3.28) die *isoperimetrische Ungleichung*

$$S(K)^n \geq n^n\kappa_n V_n(K)^{n-1}. \tag{3.32}$$

Ist $\dim K = n$, so gilt hier nach Hilfssatz 3.2.9 die Gleichheit genau dann, wenn K eine Kugel ist. Das war schon in Abschnitt 3.1 gezeigt worden.

Nach (3.20) ist

$$V_1 = \frac{n\kappa_n}{2\kappa_{n-1}}b,$$

wo b die mittlere Breite bezeichnet. Nach (3.30) gilt also

$$b^n \geq \frac{2^n}{\kappa_n}V_n$$

(Ungleichung von Urysohn). Wegen $D(K) \geq b(K)$ folgt wieder die in Abschnitt 3.1 bewiesene isodiametrische Ungleichung. Der dort gegebene Beweis hätte aber bei passender Abwandlung auch schon die Ungleichung von Urysohn geliefert.

Die zusätzlich erhaltenen quadratischen Ungleichungen schreiben wir für den Fall $n = 3$ noch mit anderen üblichen Bezeichnungen auf. Im dreidimensionalen Raum verwendet man oft die Notationen

$$\begin{aligned} V &:= V_3, \\ S &:= 2V_2, \\ M &:= \pi V_1 = 2\pi b. \end{aligned}$$

Es ist also V das Volumen, S die Oberfläche, und für Polytope P ist

$$M(P) = \pi \sum_{F \in \mathcal{F}_1(P)} V_1(F)\gamma(F, P).$$

Hier wird über die Kanten von P summiert, $V_1(F)$ ist die Länge der Kante F, und $\gamma(F, P)$ ist der Winkel zwischen den äußeren Normalenvektoren der beiden Facetten, die die Kante F enthalten. Man nennt M daher auch die *Kantenkrümmung.* Ein anderer gebräuchlicher Ausdruck für M ist *Integral der mittleren Krümmung*, denn für konvexe Körper mit zweimal stetig differenzierbarer Randfläche läßt sich M als Oberflächenintegral der mittleren Krümmung darstellen.

Wir bemerken noch, dass die Steiner-Formel im $\mathbb{R}^3$ mit diesen Bezeichnungen jetzt

$$V(K + \rho B^3) = V(K) + S(K)\rho + M(K)\rho^2 + \frac{4\pi}{3}\rho^3$$

lautet.

Die obigen quadratischen Ungleichungen (3.31) und (3.29) lauten jetzt

$$M^2 \geq 4\pi S, \tag{3.33}$$

$$S^2 \geq 3MV. \tag{3.34}$$

Man kann zeigen (was wir hier aber nicht tun wollen), dass in der ersten Ungleichung Gleichheit nur für Kugeln gilt, während es in der zweiten auch für die sogenannten Kappenkörper der Kugel (und nur für diese) steht.

Übrigens erhalten wir aus den beiden quadratischen Ungleichungen (3.33), (3.34)

$$\begin{aligned} M^4 &\geq 16\pi^2 S^2 \geq 48\pi^2 VM, \\ S^4 &\geq 9M^2V^2 \geq 36\pi SV^2, \end{aligned}$$

im Fall $\dim K = 3$ also

$$\begin{aligned} M^3 &\geq 48\pi^2 V, \\ S^3 &\geq 36\pi V^2. \end{aligned}$$

Das sind gerade die Ungleichungen (3.30) und (3.28) für $n = 3$. Die wesentlichen und grundlegenden Ungleichungen im dreidimensionalen Fall sind also die quadratischen.

Neben den angegebenen Ungleichungen zwischen den drei Größen V, S, M muss es noch weitere geben. Dies kann man sich folgendermaßen klar machen. Jedem konvexen Körper $K \in \mathcal{K}^3$ mit den inneren Volumina V, S, M ordnen wir den Punkt $(x, y) \in \mathbb{R}^2$ mit den Koordinaten

$$x = \frac{4\pi S}{M^2}, \qquad y = \frac{48\pi^2 V}{M^3}$$

zu. Ähnliche Körper werden also auf denselben Punkt abgebildet. Die im $\mathbb{R}^2$ erhaltene Punktmenge BD wird als *Blaschke-Diagramm* bezeichnet. Die Ungleichungen $M^2 \geq 4\pi S$ und $M^3 \geq 48\pi^2 V$ (zusammen mit $M \geq 0$) ergeben

$$0 \leq x \leq 1, \qquad 0 \leq y \leq 1.$$

Die Ungleichung $S^2 \geq 3MV$ liefert

$$y \leq x^2.$$

Der Parabelbogen $\{(x, y) : y = x^2,\ 0 \leq x \leq 1\}$ gehört zu BD (ihm entsprechen z.B. symmetrische Kappenkörper der Kugel, die eine stetige Verbindung von der Strecke zur Kugel herstellen). Ebenfalls zu BD gehört die Strecke $\{(x, 0); 0 \leq x \leq 8/\pi^2\}$, auf die die ebenen konvexen Körper abgebildet werden. Der Punkt $(1, 0)$ gehört aber nicht zu BD (ein entsprechender Körper müsste $V = 0$ erfüllen und würde dann der ebenen isoperimetrischen Ungleichung widersprechen). Man sieht also, dass zur vollständigen Kenntnis des Blaschke-Diagramms noch eine Begrenzungskurve fehlt; sie entspricht einer noch unbekannten Ungleichung.

3.3 Gemischte Volumina

Die Verbindung des Volumens mit der Minkowskischen Addition, die über die Steiner-Formel zu den inneren Volumina geführt hat, wollen wir nun wesentlich vertiefen. Betrachtet man das Volumen einer beliebigen Minkowskischen Linearkombination von konvexen Körpern mit variablen Koeffizienten, so erhält man ebenfalls ein Polynom. Die Koeffizienten sind die sogenannten gemischten Volumina. Sie sind Gegenstand einer flexiblen Theorie mit vielfältigen Anwendungen.

In diesem Abschnitt werden gemischte Volumina und ihre grundlegenden Eigenschaften eingeführt. Wir tun dies zuerst für stark isomorphe Polytope und dehnen die Ergebnisse dann durch Approximation aus.

3.3.1 Hilfssatz. *Sei $P \in \mathcal{K}^n$ ein n-Polytop, seien $F_1, \dots, F_N$ seine Facetten und $u_1, \dots, u_N$ die zugehörigen äußeren Normaleneinheitsvektoren. Dann gilt*

$$\sum_{i=1}^{N} V_{n-1}(F_i) u_i = 0 \tag{3.35}$$

und

$$V_n(P) = \frac{1}{n} \sum_{i=1}^{N} h(P, u_i) V_{n-1}(F_i). \tag{3.36}$$

Beweis. Die Relation (3.35) ist in (3.22) schon gezeigt worden.

Wegen (3.35) ändert sich die rechte Seite von (3.36) nicht unter Translationen von P. Daher kann o.B.d.A. $0 \in \operatorname{int} P$ angenommen werden. Dann ist P die Vereinigung der Pyramiden $\operatorname{conv}(F_i \cup \{0\})$, $i = 1, \dots, N$, die paarweise keine inneren Punkte gemeinsam haben. Wegen

$$V_n(\operatorname{conv}(F_i \cup \{0\})) = \frac{1}{n} V_{n-1}(F_i) h(P, u_i)$$

(wie leicht mit dem Satz von Fubini folgt) ergibt sich (3.36). ■

Im folgenden machen wir wesentlichen Gebrauch von den in Abschnitt 2.3 eingeführten stark isomorphen Polytopen. Es sei $\mathcal{A}$ eine vorgegebene Äquivalenzklasse von einfachen, stark isomorphen Polytopen. Sei $P \in \mathcal{A}$, und seien $u_1, \dots, u_N$ die Normaleneinheitsvektoren der Facetten von P (sie hängen nur von $\mathcal{A}$ ab). Für gegebenes $P \in \mathcal{A}$ benutzen wir die folgenden Abkürzungen

$$\begin{aligned} F_i &:= F(P, u_i) \\ F_{ij} &:= F_i \cap F_j \\ h_i &:= h(P, u_i). \end{aligned}$$

Die Zahlen $h_1, \dots, h_N$, die P eindeutig bestimmen, heißen die *Stützzahlen* von P. Ferner definieren wir

$$J := \{(i,j) : i, j \in \{1, \dots, N\},\ \dim F_{ij} = n-2\}.$$

Die Menge J hängt ebenfalls nur von der Äquivalenzklasse $\mathcal{A}$ ab. Für $(i,j) \in J$ sei θ_{ij} der Winkel zwischen u_i und u_j, und es sei $v_{ij} \perp u_i$ der Einheitsnormalenvektor des $(n-1)$-Polytops F_i bei seiner $(n-2)$-Seite F_{ij}. Wir definieren

$$h_{ij} := h(F_i, v_{ij})$$

Es gilt

$$u_j = u_i \cos\theta_{ij} + v_{ij} \sin\theta_{ij}.$$

Hier bilden wir das Skalarprodukt mit einem beliebigen Vektor aus F_{ij} und erhalten

$$h_{ij} = h_j / \sin\theta_{ij} - h_i \cot\theta_{ij}. \tag{3.37}$$

Nach diesen Vorbereitungen erhalten wir eine einfache Darstellung für das Volumen der Polytope in $\mathcal{A}$.

3.3.2 Hilfssatz. *Das Volumen eines Polytops $P \in \mathcal{A}$ kann dargestellt werden in der Form*

$$V_n(P) = \sum a_{j_1 \ldots j_n} h_{j_1} \cdots h_{j_n}, \tag{3.38}$$

wobei summiert wird über $j_1, \ldots, j_n \in \{1, \ldots, N\}$ und wo die Koeffizienten $a_{j_1 \ldots j_n}$ symmetrisch sind und nur von der Äquivalenzklasse $\mathcal{A}$ abhängen.

Beweis. Wir benutzen Induktion über n. Für $n = 1$ ist die Behauptung trivial. Sei $n > 1$ und die Behauptung bewiesen in der Dimension $n - 1$. Sei $i \in \{1, \ldots, N\}$. Nach Induktionsannahme gibt es eine Darstellung

$$V_{n-1}(F_i) = \sum_{(i,k_r) \in J} a^{(i)}_{k_1 \ldots k_{n-1}} h_{ik_1} \cdots h_{ik_{n-1}}. \tag{3.39}$$

Hier hängen die Koeffizienten $a^{(i)}_{k_1 \ldots k_{n-1}}$ nur von der Äquivalenzklasse $\mathcal{A}_i$ der Facette F_i ab und daher nach Hilfssatz 2.3.12 nur von der Äquivalenzklasse $\mathcal{A}$. Nach (3.37) gilt

$$h_{ik_r} = h_{k_r} / \sin\theta_{ik_r} - h_i \cot\theta_{ik_r}.$$

Wir setzen dies in (3.39) und dann (3.39) in die Formel (3.36). Das ergibt eine Darstellung der Form

$$V_n(P) = \sum a_{j_1 \ldots j_n} h_{j_1} \cdots h_{j_n}.$$

Hier können wir annehmen (indem wir notfalls zusätzliche Koeffizienten Null einfügen), dass sich die Summation über alle $j_1, \ldots, j_n \in \{1, \ldots, N\}$ erstreckt. Wir können auch (nach Symmetrisierung) annehmen, dass die Koeffizienten symmetrisch in ihren Indizes sind. Da die Zahlen $a^{(i)}_{k_1 \ldots k_{n-1}}$ und die Winkel θ_{ij} nur von der Äquivalenzklasse $\mathcal{A}$ abhängen, gilt dasselbe für die Koeffizienten $a_{j_1 \ldots j_n}$. ■

Nun seien die stark isomorphen Polytope $P_1, \ldots, P_n \in \mathcal{A}$ gegeben. Wir definieren

$$F_i^{(r)}, F_{ij}^{(r)}, h_i^{(r)}, h_{ij}^{(r)}$$

zu dem Polytop P_r in derselben Weise, wie F_i, F_{ij}, h_i, h_{ij} zu P definiert worden sind.

Definition. Das *gemischte Volumen* von $P_1, \ldots, P_n$ ist definiert durch

$$V(P_1, \ldots, P_n) := \sum a_{j_1 \ldots j_n} h_{j_1}^{(1)} \cdots h_{j_n}^{(n)}, \tag{3.40}$$

wo die Koeffizienten diejenigen aus Hilfssatz 3.2.2 sind.

Dann ist also $V(P_1, \ldots, P_n)$ symmetrisch in seinen Argumenten, und es gilt

$$V(P, \ldots, P) = V_n(P)$$

für $P \in \mathcal{A}$.

Nun seien Polytope $P_1, \ldots, P_m \in \mathcal{A}$ und Zahlen $\lambda_1, \ldots, \lambda_m \geq 0$ mit $\sum \lambda_i > 0$ gegeben. Dann folgt aus Hilfssatz 2.3.12 und der Definition der stark isomorphen Polytope, dass auch

$$\lambda_1 P_1 + \ldots + \lambda_m P_m \in \mathcal{A}$$

ist, und es gilt

$$h(\lambda_1 P_1 + \ldots + \lambda_m P_m, u_i) = \lambda_1 h_i^{(1)} + \ldots + \lambda_m h_i^{(m)}.$$

Aus (3.38) und (3.40) ergibt sich also sofort die polynomiale Entwicklung

$$V_n(\lambda_1 P_1 + \ldots + \lambda_m P_m) = \sum_{i_1, \ldots, i_n = 1}^{m} \lambda_{i_1} \cdots \lambda_{i_n} V(P_{i_1}, \ldots, P_{i_n}). \tag{3.41}$$

Das folgende Hilfssatz drückt das gemischte Volumen explizit durch Volumina von Minkowski-Summen aus.

3.3.3 Hilfssatz. *Für* $P_1, \ldots, P_n \in \mathcal{A}$ *gilt*

$$V(P_1, \ldots, P_n) = \frac{1}{n!} \sum_{k=1}^{n} (-1)^{n+k} \sum_{1 \leq i_1 < \ldots < i_k \leq n} V_n(P_{i_1} + \ldots + P_{i_k}). \tag{3.42}$$

Beweis. Wir bezeichnen die rechte Seite von (3.42) mit $f(P_1, \ldots, P_n)$ und lassen bei dieser Definition auch $P_i = \{0\}$ zu. Für $\lambda_1, \ldots, \lambda_n > 0$ folgt aus (3.41), dass $f(\lambda_1 P_1, \ldots, \lambda_n P_n)$ ein homogenes Polynom vom Grad n in $\lambda_1, \ldots, \lambda_n$ ist. Nun gilt

$$\begin{aligned}
&(-1)^{n+1} n! f(\{0\}, P_2, \ldots, P_n) \\
&= \sum_{2 \leq i \leq n} V_n(P_i) - \left[\sum_{2 \leq j \leq n} V_n(\{0\} + P_j) + \sum_{2 \leq i < j \leq n} V_n(P_i + P_j) \right] \\
&\quad + \left[\sum_{2 \leq j < k \leq n} V_n(\{0\} + P_j + P_k) + \sum_{2 \leq i < j < k \leq n} V_n(P_i + P_j + P_k) \right] - \ldots \\
&= 0,
\end{aligned}$$

also ist $f(0P_1, \lambda_2 P_2, \ldots, \lambda_n P_n) \equiv 0$ für alle $\lambda_2, \ldots, \lambda_n$. Daraus folgt, dass in dem Polynom $f(\lambda_1 P_1, \ldots, \lambda_n P_n)$ alle Monome $\lambda_{i_1} \cdots \lambda_{i_n}$ mit $1 \notin \{i_1, \ldots, i_n\}$ mit dem Koeffizienten 0 vorkommen. Da 1 durch jede der Zahlen $2, \ldots, n$ ersetzt werden kann, kommt allein das Monom $\lambda_1 \cdots \lambda_n$ mit einem von Null verschiedenen Koeffizienten vor. Aus der Definition von $f(\lambda_1 P_1, \ldots, \lambda_n P_n)$ und (3.41) folgt, dass dieser Koeffizient gerade $V(P_1, \ldots, P_n)$ ist. ■

Aus (3.42) folgt sofort, dass das gemischte Volumen $V(P_1, \ldots, P_n)$ sich nicht ändert bei beliebigen Verschiebungen der Polytope $P_1, \ldots, P_n$.

Eine zur Volumen-Formel (3.36) analoge Gleichung gilt auch für gemischte Volumina. Seien $P'_1, \ldots, P'_{n-1}$ stark isomorphe $(n-1)$-Polytope. Sie liegen in $(n-1)$-dimensionalen affinen Unterräumen, die zu einem linearen Unterraum H parallel sind. Wir definieren

$$v(P'_1, \ldots, P'_{n-1})$$

als das gemischte Volumen, relativ zu H, der in H liegenden Translate von $P'_1, \ldots, P'_n$. Wegen der oben bemerkten Translationsinvarianz ist dieses gemischte Volumen wohldefiniert. Analog bezeichnen wir mit $v^{(n-2)}$ das $(n-2)$-dimensionale gemischte Volumen von stark isomorphen $(n-2)$-Polytopen.

3.3.4 Hilfssatz. *Für $P_1, \ldots, P_n \in \mathcal{A}$ gilt*

$$V(P_1, \ldots, P_n) = \frac{1}{n} \sum_{i=1}^{N} h_i^{(1)} v(F_i^{(2)}, \ldots, F_i^{(n)}). \tag{3.43}$$

Beweis. Wir verwenden Induktion nach n. Der Fall $n = 1$ ist trivial; sei also $n > 1$ und die Behauptung bewiesen in der Dimension $n-1$.

Wir bezeichnen die rechte Seite von (3.43) mit $W(P_1, \ldots, P_n)$. Wenden wir (3.36) auf $\lambda_1 P_1 + \ldots + \lambda_n P_n$ an und benutzen (3.41) in der Dimension $n-1$, so erhalten wir

$$V_n(\lambda_1 P_1 + \ldots + \lambda_n P_n) = \sum_{i_1, \ldots, i_n = 1}^{n} \lambda_{i_1} \cdots \lambda_{i_n} W(P_{i_1}, \ldots, P_{i_n})$$

für beliebige $\lambda_1, \ldots, \lambda_n > 0$. Es genügt also der Nachweis, dass $W(P_1, \ldots, P_n)$ symmetrisch ist in seinen Argumenten. Nach Induktionsannahme gilt

$$\begin{aligned}
&W(P_1, \ldots, P_n) \\
&= \frac{1}{n} \sum_{i=1}^{N} h_i^{(1)} \frac{1}{n-1} \sum_{(i,j) \in J} h_{ij}^{(2)} v^{(n-2)}(F_{ij}^{(3)}, \ldots, F_{ij}^{(n)}) \\
&= \frac{1}{n(n-1)} \sum_{\substack{(r,s) \in J \\ r < s}} \left[h_r^{(1)} h_{rs}^{(2)} + h_s^{(1)} h_{sr}^{(2)} \right] v^{(n-2)} \left(F_{rs}^{(3)}, \ldots, F_{rs}^{(n)} \right)
\end{aligned}$$

(im Fall $n = 2$ sind die Ausdrücke $v^{(n-2)}$ durch 1 zu ersetzen). Nach (3.37) gilt

$$\begin{aligned} &h_r^{(1)} h_{rs}^{(2)} + h_s^{(1)} h_{sr}^{(2)} \\ &= h_r^{(1)} \left[h_s^{(2)} / \sin\theta_{rs} - h_r^{(2)} \cot\theta_{rs} \right] + h_s^{(1)} \left[h_r^{(2)} / \sin\theta_{rs} - h_s^{(2)} \cot\theta_{rs} \right]. \end{aligned}$$

Dieser Ausdruck ist symmetrisch in den oberen Indizes 1 und 2. Also gilt

$$W(P_1, P_2, P_3, \ldots, P_n) = W(P_2, P_1, P_3, \ldots, P_n),$$

und da $W(P_1, P_2, P_3, \ldots, P_n)$ in den letzten $n-1$ Argumenten symmetrisch ist (wie aus den Definitionen von W und dem gemischten Volumen in der Dimension $n-1$ folgt), ergibt sich die Behauptung des Hilfssatzes. ■

Nun erklären wir das gemischte Volumen für allgemeine konvexe Körper.

3.3.5 Satz und Definition. *Es gibt eine nicht-negative, stetige, symmetrische Funktion $V : (\mathcal{K}^n)^n \to \mathbb{R}$, das* gemischte Volumen, *so dass*

$$V_n(\lambda_1 K_1 + \ldots + \lambda_m K_m) = \sum_{i_1,\ldots,i_n=1}^{m} \lambda_{i_1} \cdots \lambda_{i_n} V(K_{i_1}, \ldots, K_{i_n}) \tag{3.44}$$

für $K_1, \ldots, K_m \in \mathcal{K}^n$ und $\lambda_1, \ldots, \lambda_m \geq 0$ gilt.

Beweis. Für $K_1, \ldots, K_n \in \mathcal{K}^n$ definieren wir

$$V(K_1, \ldots, K_n) := \frac{1}{n!} \sum_{k=1}^{n} (-1)^{n+k} \sum_{i_1 < \ldots < i_k} V_n(K_{i_1} + \ldots + K_{i_k}), \tag{3.45}$$

was nach Hilfssatz 3.3.3 konsistent ist mit der bereits erfolgten Definition des gemischten Volumens für stark isomorphe Polytope. Dann ist V symmetrisch, und aus der Stetigkeit der Minkowski-Addition und des Volumen-Funktionals folgt, dass V auf $(\mathcal{K}^n)^n$ stetig ist. Seien $K_1, \ldots, K_m \in \mathcal{K}^n$ gegeben. Nach Satz 2.3.16 gibt es Folgen $(P_{1i})_{i\in\mathbb{N}}, \ldots, (P_{mi})_{i\in\mathbb{N}}$ von Polytopen mit $P_{ri} \to K_r$ für $i \to \infty$ $(r = 1, \ldots, m)$ und derart, dass für jedes $i \in \mathbb{N}$ die Polytope $P_{1i}, \ldots, P_{mi}$ stark isomorph sind. Aus (3.41) und der Stetigkeit des Volumens und des gemischten Volumens folgt jetzt die Gültigkeit von (3.44).

Dass das gemischte Volumen nichtnegativ ist, zeigt man zuerst für stark isomorphe Polytope, und zwar durch Induktion nach der Dimension unter Verwendung von (3.43) (dort kann $0 \in P_1$ angenommen werden, dann ist $h_i^{(1)} \geq 0$). Das allgemeine Resultat folgt dann durch Approximation. ■

Wir stellen einige Eigenschaften des gemischten Volumens fest. Es gilt

$$V(K, \ldots, K) = V_n(K).$$

Für jede volumentreue affine Abbildung $\alpha : \mathbb{E}^n \to \mathbb{E}^n$ gilt

$$V(\alpha K_1, \ldots, \alpha K_n) = V(K_1, \ldots, K_n).$$

Das gemischte Volumen ist in jedem Argument Minkowski-additiv und monoton. Wegen der Symmetrie genügt es, dies für das erste Argument zu zeigen. Die Behauptung lautet also, dass für $K, L, K_2, \ldots, K_n \in \mathcal{K}^n$ und $\lambda, \mu \geq 0$

$$V(\lambda K + \mu L, K_2, \ldots, K_n) = \lambda V(K, K_2, \ldots, K_n) + \mu V(L, K_2, \ldots, K_n)$$

gilt und dass aus $K \subset L$ die Ungleichung

$$V(K, K_2, \ldots, K_n) \leq V(L, K_2, \ldots, K_n) \tag{3.46}$$

folgt. Beides ergibt sich für stark isomorphe Polytope aus (3.43) und folgt dann allgemein durch Approximation. Es ist allgemein nicht bekannt, wann in (3.46) das Gleichheitszeichen steht. Aus (3.46) folgt auch $V(K_1, \ldots, K_n) > 0$, wenn $K_1, \ldots, K_n$ innere Punkte haben (man wähle in jedem K_i eine gleich große Kugel).

Die polynomiale Entwicklung (3.44) kann auch in anderer Form geschrieben werden. Dazu benutzen wir die Abkürzung

$$V(\underbrace{K_1, \ldots, K_1}_{r_1}, \ldots, \underbrace{K_m, \ldots, K_m}_{r_m}) =: V(K_1[r_1], \ldots, K_m[r_m])$$

und den Multinomialkoeffizienten

$$\binom{n}{r_1 \ldots r_m} := \begin{cases} \dfrac{n!}{r_1! \ldots r_m!}, & \text{falls } \sum r_i = n, r_j \in \{0, \ldots, n\}, \\ 0 \text{ sonst.} & \end{cases}$$

Damit ergibt sich für $K_1, \ldots, K_m \in \mathcal{K}^n$ und $\lambda_1, \ldots, \lambda_m \geq 0$ aus (3.44) mit einem einfachen kombinatorischen Argument

$$\begin{aligned} & V_n(\lambda_1 K_1 + \ldots + \lambda_m K_m) \\ &= \sum_{r_1, \ldots, r_m = 0}^{n} \binom{n}{r_1 \ldots r_m} \lambda_1^{r_1} \cdots \lambda_m^{r_m} V(K_1[r_1], \ldots, K_m[r_m]). \end{aligned} \tag{3.47}$$

Ähnliche polynomiale Entwicklungen gelten für gemischte Volumina, in denen einige Elemente festgehalten werden. Seien eine Zahl $p \in \{1, \ldots, n-1\}$ und konvexe Körper $C_{p+1}, \ldots, C_n \in \mathcal{K}^n$ gegeben. Dann gilt

$$\begin{aligned} & V(\lambda_1 K_1 + \ldots + \lambda_m K_m[p], C_{p+1}, \ldots, C_n) \\ = & \sum_{r_1, \ldots, r_m = 0}^{p} \binom{p}{r_1 \ldots r_m} \lambda_1^{r_1} \cdots \lambda_m^{r_m} V(K_1[r_1], \ldots, K_m[r_m], C_{p+1}, \ldots, C_n). \end{aligned} \tag{3.48}$$

Zum Beweis entwickelt man beide Seiten der Identität

$$\begin{aligned} &V_n(\mu(\lambda_1 K_1 + \ldots + \lambda_m K_m) + \mu_{p+1}C_{p+1} + \ldots + \mu_n C_n) \\ &= V_n(\mu\lambda_1 K_1 + \ldots + \mu\lambda_m K_m + \mu_{p+1}C_{p+1} + \ldots + \mu_n C_n) \end{aligned}$$

gemäß (3.47) und vergleicht dann auf beiden Seiten die Koeffizienten von $\mu^p \mu_{p+1} \cdots \mu_n$.

Als Spezialfall von (3.47) erhalten wir für zwei konvexe Körper K, L die Gleichung

$$V_n(K + \rho L) = \sum_{m=0}^{n} \binom{n}{m} \rho^{n-m} V(K[m], L[n-m]).$$

Ein Vergleich mit der Steiner-Formel (3.15) zeigt also, dass

$$\kappa_{n-m} V_m(K) = \binom{n}{m} V(K[m], B^n[n-m])$$

ist. Bis auf einen Faktor (der die Unabhängigkeit von der Dimension des umgebenden Raumes bewirkt) ist das m-te innere Volumen von K also das gemischte Volumen von m-mal dem Körper K und $(n-m)$-mal der Einheitskugel.

Völlig analog zur Herleitung der Ungleichungen (3.28) – (3.31) aus Hilfssatz 3.2.9 und dem Satz von Brunn-Minkowski ergibt sich die folgende allgemeinere Aussage.

3.3.6 Satz. *Für n-dimensionale konvexe Körper K, L gelten die* Minkowskischen Ungleichungen

$$V(K, L[n-1])^n \geq V_n(K) V_n(L)^{n-1} \tag{3.49}$$

mit Gleichheit genau dann, wenn K und L homothetisch sind, und

$$V(K, L[n-1])^2 \geq V(K[2], L[n-2]) V_n(L). \tag{3.50}$$

Wann in (3.50) das Gleichheitszeichen gilt, ist nicht vollständig bekannt.

3.4 Die Aleksandrov-Fenchelschen Ungleichungen

In Satz 3.3.6 wurden neben den Ungleichungen

$$V(K, L, \ldots, L)^n \geq V_n(K) V_n(L)^{n-1}$$

auch die quadratischen Minkowskischen Ungleichungen

$$V(K, L, L, \ldots, L)^2 \geq V(K, K, L, \ldots, L)V_n(L)$$

notiert. Im dreidimensionalen Fall hatten wir (zumindest für $K = B^n$) gesehen, dass aus den quadratischen Ungleichungen die anderen folgen; die quadratischen Ungleichungen sind also die grundlegenderen. In diesem Abschnitt wollen wir eine ganz allgemeine quadratische Ungleichung für gemischte Volumina zeigen, nämlich

$$V(K, L, K_3, \ldots, K_n)^2 \geq V(K, K, K_3, \ldots, K_n)V(L, L, K_3, \ldots, K_n).$$

Beim Beweis werden wir wesentlichen Gebrauch von stark isomorphen Polytopen machen.

Es sei $\mathcal{A}$ eine gegebene (im folgenden zunächst feste) Äquivalenzklasse von stark isomorphen, einfachen Polytopen. Wir verwenden die in Abschnitt 3.3 eingeführten Bezeichnungen. Es seien $u_1, \ldots, u_N$ die Normaleneinheitsvektoren der Polytope aus $\mathcal{A}$. Für gegebenes $P \in \mathcal{A}$ sei

$$\begin{aligned} h_i &:= h(P, u_i), \quad F_i := F(P, u_i), \quad F_{ij} := F_i \cap F_j, \\ J &:= \{(i, j) : \dim F_{ij} = n - 2\}. \end{aligned}$$

Für $P_r \in \mathcal{A}$ werden die entsprechenden Größen mit $h_i^{(r)}, F_i^{(r)}, F_{ij}^{(r)}$ bezeichnet. Das N-Tupel

$$\overline{P} := (h_1, \ldots, h_N) \in \mathbb{R}^N,$$

durch das das Polytop $P \in \mathcal{A}$ eindeutig bestimmt ist, nennen wir den *Stützvektor* von P. Nach (3.37) sind die Stützzahlen der Facette F_i gegeben durch

$$h_{ij} = h_j / \sin \theta_{ij} - h_i \cot \theta_{ij} \qquad \text{für } (i, j) \in J.$$

Für $P_1, \ldots, P_n \in \mathcal{A}$ gilt

$$V(P_1, \ldots, P_n) = \sum a_{j_1 \ldots j_n} h_{j_1}^{(1)} \cdots h_{j_n}^{(n)}, \tag{3.51}$$

$$v(F_i^{(1)}, \ldots, F_i^{(n-1)}) = \sum a_{k_1 \ldots k_{n-1}}^{(i)} h_{ik_1}^{(1)} \cdots h_{ik_{n-1}}^{(n-1)}, \tag{3.52}$$

wo die Koeffizienten symmetrisch sind und nur von der Äquivalenzklasse $\mathcal{A}$ abhängen. Die Summationen in (3.52) können dabei über alle $k_1, \ldots, k_{n-1}$ von 1 bis N erstreckt werden, wenn zusätzliche Koeffizienten Null eingeführt werden und $h_{ij}^{(r)} := 0$ für $(i, j) \notin J$ vereinbart wird.

Der Ausdruck für $V(P_1, \ldots, P_n)$ in (3.51) hängt nur von den Stützvektoren der Polytope $P_1, \ldots, P_n$ ab. Man kann diese Definition daher formal auf beliebige N-Tupel ausdehnen. Für

$$X_r := (x_1^{(r)}, \ldots, x_N^{(r)}) \in \mathbb{R}^N, \qquad r = 1, \ldots, n$$

setzen wir

$$V(X_1,\dots,X_n) := \sum a_{j_1\dots j_n} x^{(1)}_{j_1}\cdots x^{(n)}_{j_n}.$$

Dann ist V eine n-lineare Funktion auf $\mathbb{R}^N$. Für $X=(x_1,\dots,x_N)\in\mathbb{R}^N$ definieren wir sodann

$$x_{ij} := \begin{cases} x_j/\sin\theta_{ij} - x_i\cot\theta_{ij} & \text{für } (i,j)\in J \\ 0 \text{ für } (i,j)\notin J \end{cases}$$

und

$$\Lambda_i X := (x_{i1},\dots,x_{iN}).$$

Es ist also $\Lambda_i:\mathbb{R}^N\to\mathbb{R}^N$ eine lineare Abbildung. Sie ordnet dem Stützvektor eines Polytops die Stützzahlen seiner i-ten Facette zu, durch Nullen zu einem N-Tupel ergänzt. Wir erweitern die Definition (3.52) durch

$$v(\Lambda_i X_1,\dots,\Lambda_i X_{n-1}) := \sum a^{(i)}_{k_1\dots k_{n-1}} x^{(1)}_{ik_1}\cdots x^{(n-1)}_{ik_{n-1}}. \tag{3.53}$$

Dann gilt eine Verallgemeinerung der Gleichung (3.43), nämlich

$$V(X_1,\dots,X_n) = \frac{1}{n}\sum_{i=1}^N x^{(1)}_i v(\Lambda_i X_2,\dots,\Lambda_i X_n). \tag{3.54}$$

Diese Gleichung ist zunächst richtig für Stützvektoren. Ist $P\in\mathcal{A}$, so ist $\overline{P}+\epsilon X_r$ für alle ϵ aus einer Umgebung von 0 der Stützvektor eines Polytops aus $\mathcal{A}$. Das folgt aus Hilfssatz 2.3.14; hier wird benutzt, dass die Polytope aus $\mathcal{A}$ einfach sind. Da beide Seiten von (3.54) n-lineare Funktionen sind, folgt (3.54) allgemein.

Im folgenden wird ein Polytop $P\in\mathcal{A}$ oft mit seinem Stützvektor $\overline{P}$ identifiziert. Ist in $V(X_1,\dots,X_n)$ oder $v(\Lambda_i X_1,\dots,\Lambda_i X_{n-1})$ eines der Argumente X_r ein Stützvektor $\overline{P}_r$, so ersetzen wir in den Ausdrücken X_r durch P_r und $\Lambda_i X_r$ durch $F^{(r)}_i = F(P_r,u_i)$. Wir bezeichnen $Z=(\zeta_1,\dots,\zeta_N)\in\mathbb{R}^N$ als den Stützvektor eines Punktes, wenn ein Punkt $z\in\mathbb{E}^n$ existiert mit $\zeta_i = h(\{z\},u_i)$ für $i=1,\dots,N$; das ist gleichbedeutend mit

$$Z = (\langle z,u_1\rangle,\dots,\langle z,u_N\rangle).$$

Ziel dieses Abschnitts ist der Beweis des folgenden Satzes.

3.4.1 Satz. *Für stark isomorphe einfache Polytope $P,P_3,\dots,P_n\in\mathcal{A}$ und für $Z\in\mathbb{R}^N$ gilt*

$$V(Z,P,P_3,\dots,P_n)^2 \geq V(Z,Z,P_3,\dots,P_n)V(P,P,P_3,\dots,P_n).$$

Das Gleichheitszeichen gilt genau dann, wenn $Z=\lambda\overline{P}+A$ gilt, wo $\lambda\in\mathbb{R}$ und A der Stützvektor eines Punktes ist.

3.4.2 Korollar. *Für konvexe Körper $K, L, K_3, \dots, K_n$ gilt*

$$V(K, L, K_3, \dots, K_n)^2 \geq V(K, K, K_3, \dots, K_n) V(L, L, K_3, \dots, K_n). \tag{3.55}$$

Man bezeichnet (3.55) als die *Aleksandrov-Fenchelschen Ungleichungen.*

Diese Ungleichung ergibt sich zunächst für stark isomorphe einfache Polytope, indem man in 3.4.1 für Z den Stützvektor eines Polytops aus $\mathcal{A}$ wählt. Der Approximationssatz 2.3.16 und die Stetigkeit der gemischten Volumina ergeben dann die Ungleichung (3.55) für beliebige konvexe Körper. Aus 3.4.1 folgt noch, dass in (3.55) jedenfalls dann Gleichheit gilt, wenn K und L homothetisch sind. Jedoch gilt Gleichheit auch noch in anderen Fällen. Die vollständige Aufklärung, wann Gleichheit gilt, ist ein ungelöstes Problem.

Beweis von Satz 3.4.1. Für $n \geq 2$ führen wir auf $\mathbb{R}^N$ eine symmetrische Bilinearform Φ ein durch

$$\Phi(X, Y) := V(X, Y, P_3, \dots, P_n) \qquad \text{für } X, Y \in \mathbb{R}^N$$

(im Fall $n = 2$ sind $P_3, \dots, P_n$ nicht vorhanden). Es genügt, die folgende Behauptung zu zeigen.

Behauptung 1. *Ist $\Phi(Z, P) = 0$, so ist $\Phi(Z, Z) \leq 0$, und Gleichheit gilt genau dann, wenn Z der Stützvektor eines Punktes ist.*

Nehmen wir an, Behauptung 1 sei bewiesen. Ist $Z \in \mathbb{R}^N$ gegeben, so definieren wir

$$\lambda := \frac{\Phi(Z, P)}{\Phi(P, P)} \quad \text{und} \quad Z' := Z - \lambda \overline{P}.$$

Dies ist möglich wegen $\Phi(P, P) = V(P, P, P_3, \dots, P_n) > 0$. Dann ist $\Phi(Z', P) = 0$, nach Behauptung 1 gilt also $\Phi(Z', Z') \leq 0$, mit Gleichheit genau dann, wenn Z' der Stützvektor eines Punktes ist. Wegen

$$\Phi(Z', Z') = \Phi(Z, Z) - \frac{\Phi(Z, P)^2}{\Phi(P, P)}$$

folgt die Behauptung von Satz 3.4.1.

Zum Beweis der Behauptung 1 zeigen wir zunächst:

Behauptung 2. *Es gilt*

$$V(Z, P, \dots, P)^2 \geq V(Z, Z, P, \dots, P) V_n(P). \tag{3.56}$$

Für $n = 2$ gilt Gleichheit genau dann, wenn $Z = \lambda \overline{P} + A$ ist, wo $\lambda \in \mathbb{R}$ und A der Stützvektor eines Punktes ist.

Zum Beweis bemerken wir, dass nach Hilfssatz 2.3.14 durch $\overline{P} + \epsilon Z$ der Stützvektor eines Polytops $Q \in \mathcal{A}$ gegeben ist, wenn $|\epsilon|$ hinreichend klein ist. Aus der Minkowskischen Ungleichung (3.50) folgt

$$\begin{aligned} 0 &\leq V(Q,P,\ldots,P)^2 - V(Q,Q,P,\ldots,P)V_n(P) \\ &= \epsilon^2[V(Z,P,\ldots,P)^2 - V(Z,Z,P,\ldots,P)V_n(P)]. \end{aligned}$$

Dies beweist die Ungleichung (3.56). Ist $n = 2$, so folgt aus Satz 3.3.6 (erster Teil), dass in (3.56) genau dann Gleichheit gilt, wenn Q und P homothetisch sind. Hieraus ergibt sich die Gleichheitsaussage in Behauptung 2.

Wir beweisen nun Behauptung 1 durch Induktion nach der Dimension. Für $n = 2$ gilt die Behauptung 1 nach Behauptung 2. Wir nehmen an, dass $n \geq 3$ und Behauptung 1 in kleineren Dimensionen bewiesen ist.

Für jedes $i \in \{1,\ldots,N\}$ definieren wir eine symmetrische Bilinearform φ_i auf $\mathbb{R}^N$ durch

$$\varphi_i(X,Y) := v(\Lambda_i X, \Lambda_i Y, F_i^{(4)}, \ldots, F_i^{(n)}) \qquad \text{für } X,Y \in \mathbb{R}^N$$

(für $n = 3$ entfallen die Argumente $F_i^{(k)}$).

Behauptung 3. *$Z \in \mathbb{R}^N$ ist genau dann Eigenvektor der Bilinearform Φ zum Eigenwert 0, wenn Z der Stützvektor eines Punktes ist.*

Zum Beweis stellen wir zunächst fest, dass nach (3.54) die Gleichung

$$\Phi(X,Y) = \frac{1}{n}\sum_{i=1}^{N} x_i \varphi_i(Y, P_3)$$

gilt. Die lineare Funktion $\varphi_i(\cdot, P_3)$ ist von der Form

$$\varphi_i(Y, P_3) = \sum_{j=1}^{N} b_{ij} y_j,$$

also ist

$$\Phi(X,Y) = \frac{1}{n}\sum_{i,j=1}^{N} b_{ij} x_i y_j. \tag{3.57}$$

Hier gilt $b_{ij} = b_{ji}$, weil Φ symmetrisch ist. Daß $Z = (\zeta_1,\ldots,\zeta_N) \neq 0$ ein Eigenvektor von Φ zum Eigenwert 0 ist, bedeutet

$$\sum_{j=1}^{N} b_{ij}\zeta_j = 0 \qquad \text{für } i = 1,\ldots,N.$$

Äquivalent hierzu ist

$$\varphi_i(Z, P_3) = 0 \qquad \text{für } i = 1, \dots, N. \tag{3.58}$$

Ist Z der Stützvektor eines Punktes z, so gilt

$$\varphi_i(Z, P_3) = v(\{z\}, F_i^{(3)}, \dots, F_i^{(n)}) = 0.$$

Umgekehrt nehmen wir nun an, dass (3.58) erfüllt ist. Nach der Induktionsannahme gilt dann $\varphi_i(Z, Z) \le 0$. O.B.d.A. können wir $h(P_3, u_i) > 0$ annehmen; dann folgt

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{i=1}^{N} \zeta_i \varphi_i(Z, P_3) = \Phi(Z, Z) \\ &= V(Z, Z, P_3, \dots, P_n) = V(P_3, Z, Z, P_4, \dots, P_n) \\ &= \frac{1}{n} \sum_{i=1}^{N} h(P_3, u_i) \varphi_i(Z, Z) \le 0 \end{aligned}$$

und daher $\varphi_i(Z, Z) = 0$ für $i = 1, \dots, N$. Nach der Induktionsannahme bedeutet das, dass $\Lambda_i Z$ der Stützvektor (relativ zu der Äquivalenzklasse von F_i) eines Punktes z_i ist. Explizit bedeutet dies, dass

$$\Lambda_i Z = (\langle z_i, v_{i1} \rangle, \dots, \langle z_i, v_{iN} \rangle)$$

mit $v_{ij} := 0$ für $(i, j) \notin J$ gilt. Wir wählen $\epsilon \ne 0$ derart, dass $\overline{P}_3 + \epsilon Z$ der Stützvektor eines Polytops $Q \in \mathcal{A}$ ist. Die Gleichung $\Lambda_i(\overline{P}_3 + \epsilon Z) = \Lambda_i \overline{P}_3 + \epsilon \Lambda_i Z$ ergibt

$$\begin{aligned} h(F(Q, u_i), v_{ij}) &= h(F_i^{(3)}, v_{ij}) + \epsilon \langle z_i, v_{ij} \rangle \\ &= h(F_i^{(3)} + \epsilon z_i, v_{ij}) \end{aligned}$$

und damit $F(Q, u_i) = F_i^{(3)} + t_i$ mit einem Vektor t_i (nämlich $t_i = \epsilon z_i + \alpha_i u_i$ mit $\alpha_i \in \mathbb{R}$). Für $(i, j) \in J$ folgt, dass die $(n-2)$-Seite $G := F(Q, u_i) \cap F(Q, u_j)$ die Gleichung $G = F_{ij}^{(3)} + t_i$ erfüllt und analog $G = F_{ij}^{(3)} + t_j$, also ist $t_i = t_j$. Je zwei Facetten von P_3 können verbunden werden durch eine Folge von Facetten derart, dass je zwei aufeinanderfolgende Facetten einen $(n-2)$-dimensionalen Durchschnitt haben. Es folgt $t_i = t_j$ für alle i, j. Daher ist Q ein Translat von P_3, also Z der Stützvektor eines Punktes. Damit ist Behauptung 3 bewiesen.

Nun führen wir eine zweite symmetrische Bilinearform Ψ auf $\mathbb{R}^N$ ein durch

$$\Psi(X, Y) := \frac{1}{n} \sum_{i=1}^{N} \frac{\varphi_i(P, P_3)}{h(P, u_i)} x_i y_i \qquad \text{für } X, Y \in \mathbb{R}^N.$$

Hier setzen wir o.B.d.A. $h(P, u_i) > 0$ für $i = 1, \dots, N$ voraus. Da $\varphi_i(P, P_3) > 0$ ist, ist die Form Ψ positiv definit.

Wir betrachten die relativen Eigenwerte $\lambda_1 > \lambda_2 > \dots$ der Form Φ bezüglich Ψ. Bekanntlich gilt

$$\begin{aligned} \lambda_1 &= \max\{\Phi(X,X) : \Psi(X,X) = 1\}, \\ \lambda_2 &= \max\{\Phi(X,X) : \Psi(X,X) = 1, \Psi(X,Y) = 0 \\ &\qquad \text{für alle } Y \text{ im } \lambda_1\text{-Eigenraum}\}. \end{aligned} \tag{3.59}$$

In Analogie zu (3.57) schreiben wir

$$\Psi(X,Y) = \frac{1}{n} \sum_{i,j=1}^{N} c_{ij} x_i y_j$$

mit

$$c_{ij} := \begin{cases} \dfrac{\varphi_i(P, P_3)}{h(P, u_i)} & \text{für } i = j, \\ 0 & \text{für } i \neq j. \end{cases}$$

Dann ist $Z = (\zeta_1, \dots, \zeta_N) \in \mathbb{R}^N \setminus \{0\}$ ein Eigenvektor von Φ relativ zu Ψ mit Eigenwert λ genau dann, wenn

$$\sum_{j=1}^{N} (b_{ij} - \lambda c_{ij}) \zeta_j = 0 \qquad \text{für } i = 1, \dots, N$$

gilt; dies ist äquivalent mit

$$\varphi_i(Z, P_3) = \lambda \frac{\varphi_i(P, P_3)}{h(P, u_i)} \zeta_i \qquad \text{für } i = 1, \dots, N.$$

Insbesondere ist $\lambda = 1$ ein Eigenwert mit zugehörigem Eigenvektor $Z = \overline{P}$.

Behauptung 4. *Der einzige positive Eigenwert von Φ relativ zu Ψ ist* 1, *und er ist einfach.*

Zum Beweis nehmen wir zuerst $P = P_3 = \dots = P_n$ an. Angenommen, Behauptung 4 wäre in diesem Fall falsch. Falls es einen positiven Eigenwert $\mu \neq 1$ gibt, dann existiert ein $Z \in \mathbb{R}^N$ mit $\Psi(Z, P) = 0$ und $\Phi(Z, Z) = \mu \Psi(Z, Z) > 0$. Falls 1 ein mehrfacher Eigenwert ist, dann ist der zugehörige Eigenraum mindestens zweidimensional und enthält daher einen Vektor Z mit $\Psi(Z, P) = 0$ und $\Phi(Z, Z) = \Psi(Z, Z) > 0$. In jedem der beiden Fälle ist

$$0 = \Psi(Z, P) = \frac{1}{n} \sum_{i=1}^{N} \frac{\varphi_i(P, P)}{h(P, u_i)} \zeta_i h(P, u_i) = V(Z, P, \dots, P).$$

Wegen $V(Z, Z, P, \dots, P) = \Phi(Z, Z) > 0$ ist das ein Widerspruch zu Behauptung 2.

Nun seien $P_3, \dots, P_n \in \mathcal{A}$ beliebig. Für $\vartheta \in [0, 1]$ sei

$$P_r(\vartheta) := (1 - \vartheta)P + \vartheta P_r, \qquad r = 3, \dots, n.$$

Die Koeffizienten der zugehörigen Formen Φ, Ψ hängen stetig von ϑ ab, daher gilt dasselbe für die relativen Eigenwerte. Nach Behauptung 3 ist 0 stets ein Eigenwert mit Vielfachheit n. Es folgt, dass die Summe der Vielfachheiten der positiven Eigenwerte unabhängig von ϑ ist. Da diese Summe für $\vartheta = 0$ gleich 1 ist, ist sie auch für $\vartheta = 1$ gleich 1. Damit ist Behauptung 4 bewiesen.

Behauptung 4 zeigt nun, dass der Eigenraum zum Eigenwert 1 gleich $\operatorname{lin}\{\overline{P}\}$ ist und dass der zweite Eigenwert nicht positiv ist. Es ist also $\Phi(Z, Z) \le 0$ für alle Z mit $\Psi(Z, P) = 0$, und die Bedingung $\Psi(Z, P) = 0$ ist äquivalent mit $\Phi(Z, P) = 0$. Aus $\Phi(Z, P) = 0$ folgt also $\Phi(Z, Z) \le 0$. Angenommen, hier gilt Gleichheit für ein $Z \ne 0$. Da bei Z das Maximum in (3.59) angenommen wird, ist Z ein Eigenvektor zum Eigenwert 0. Nach Behauptung 3 ist Z dann der Stützvektor eines Punktes. Damit ist der Induktionsbeweis der Behauptung 1 abgeschlossen und somit auch der Beweis von Satz 3.4.1. ■

Als erste Konsequenz aus den Aleksandrov-Fenchelschen Ungleichungen können wir den Satz von Brunn-Minkowski (Fall $m = n$ des nachfolgenden Satzes) wesentlich verallgemeinern.

3.4.3 Satz (allgemeiner Satz von Brunn-Minkowski). *Seien eine Zahl* $m \in \{1, \dots, n\}$ *und konvexe Körper* $K_0, K_1, K_{m+1}, \dots, K_n \in \mathcal{K}^n$ *gegeben, sei* $K_\lambda := (1 - \lambda)K_0 + \lambda K_1$ *und*

$$f(\lambda) := V(K_\lambda[m], K_{m+1}, \dots, K_n)^{\frac{1}{m}}$$

für $0 \le \lambda \le 1$. *Dann ist* f *eine konkave Funktion auf* $[0, 1]$.

Beweis. Wir müssen $f''(\lambda) \le 0$ zeigen. Es genügt, dies für $\lambda = 0$ zu zeigen. Ist das nämlich gezeigt und ist dann $0 < \lambda < 1$, so setzen wir $\overline{K}_\tau := (1-\tau)K_\lambda + \tau K_1$ und

$$h(\tau) := V(\overline{K}_\tau[m], K_{m+1}, \dots, K_n)^{\frac{1}{m}}.$$

Dann ist $f(\lambda + \mu) = h(\mu/(1 - \lambda))$, also folgt $f''(\lambda) \le 0$ aus $h''(0) \le 0$.

Nun gilt

$$f''(0) = (m - 1)V_{(0)}^{(1/m)-2}[V_{(0)}V_{(2)} - V_{(1)}^2]$$

mit

$$V_{(i)} := V(K_0[m - i], K_1[i], K_{m+1}, \dots, K_n).$$

Wir haben zunächst $V_{(0)} > 0$ vorausgesetzt. Aus Satz 3.4.2 folgt jetzt $f''(0) \leq 0$. Durch Approximation ergibt sich dies allgemein. ■

Da das gemischte Volumen $V(K_1, K_2, \ldots, K_n)$ in seinen Argumenten symmetrisch ist, ist klar, dass man aus der Aleksandrov-Fenchel-Ungleichung

$$V(K_1, K_2, K_3, \ldots, K_n)^2 \geq V(K_1, K_1, K_3, \ldots, K_n)V(K_2, K_2, K_3, \ldots, K_n)$$

viele weitere Ungleichungen gewinnen kann, indem man sie wiederholt anwendet. Die folgenden Ausführungen sollen in möglichst übersichtlicher Weise zeigen, wie man eine Serie von weiteren Ungleichungen herleiten kann.

Wir betrachten vorbereitend zunächst eine endliche Folge $(a_0, a_1, \ldots, a_m)$ von reellen Zahlen. Sie heißt *konkav*, wenn

$$a_{i-1} - 2a_i + a_{i+1} \leq 0 \qquad \text{für } i = 1, \ldots, m-1$$

gilt. Äquivalent damit ist

$$a_0 - a_1 \leq a_1 - a_2 \leq \ldots \leq a_{m-1} - a_m.$$

Dies gelte nun. Für $0 \leq i < j < k \leq m$ folgt dann, wenn wir arithmetische Mittel bilden,

$$\frac{(a_i - a_{i+1}) + \ldots + (a_{j-1} - a_j)}{j-i} \leq \frac{(a_j - a_{j+1}) + \ldots + (a_{k-1} - a_k)}{k-j},$$

also

$$\frac{a_i - a_j}{j-i} \leq \frac{a_j - a_k}{k-j}$$

und daher

$$(k-j)a_i + (i-k)a_j + (j-i)a_k \leq 0. \tag{3.60}$$

Gleichheit gilt hier genau dann, wenn

$$a_{r-1} - 2a_r + a_{r+1} = 0 \qquad \text{für } r = i+1, \ldots, k-1$$

gilt.

Die Folge $(a_0, a_1, \ldots, a_m)$ positiver Zahlen heißt *log-konkav*, wenn die Folge

$$(\log a_0, \ldots, \log a_m)$$

konkav ist; äquivalent damit ist

$$a_i^2 \geq a_{i-1}a_{i+1} \qquad \text{für } i = 1, \ldots, m-1.$$

In diesem Fall folgt aus (3.60) für $0 \le i < j < k \le m$

$$a_i^{k-j} a_j^{i-k} a_k^{j-i} \le 1. \tag{3.61}$$

Mit positiven Exponenten umgeschrieben, lautet diese Ungleichung

$$a_j^{k-i} \ge a_i^{k-j} a_k^{j-i}. \tag{3.62}$$

Gleichheit gilt hier genau dann, wenn

$$a_r^2 = a_{r-1} a_{r+1} \qquad \text{für } r = i+1, \dots, k-1$$

gilt.

Wir wenden dies nun an auf konvexe Körper $K_0, K_1, K_{m+1}, \dots, K_n$ mit $m \in \{1, \dots, n\}$. Setzen wir

$$a_i := V_{(i)} := V(K_0[m-i], K_1[i], K_{m+1}, \dots, K_n)$$

für $i = 0, \dots, m$, so gilt nach der Aleksandrov-Fenchelschen Ungleichung

$$a_i^2 \ge a_{i-1} a_{i+1} \qquad \text{für } i = 1, \dots, m-1.$$

Setzen wir zusätzlich voraus, dass alle $a_i > 0$ sind, so ist die Folge $(a_0, \dots, a_m)$ also log-konkav. Daher gilt nach (3.62)

$$V_{(j)}^{k-i} \ge V_{(i)}^{k-j} V_{(k)}^{j-i} \tag{3.63}$$

für $0 \le i < j < k \le m$. Durch Approximation ergibt sich diese Ungleichung auch, wenn zugelassen wird, dass einige der vorkommenden Größen verschwinden.

Wir nehmen jetzt die folgende Spezialisierung vor: $m = n$, $K_0 = B^n$, $K_1 = K$. Dann ist

$$V_{(i)} = V(K[i], B^n[n-i]) = \frac{\kappa_{n-i}}{\binom{n}{i}} V_i(K),$$

wo V_i wieder das i-te innere Volumen bezeichnet. Die Ungleichung (3.63) liefert für $i = 0$ wegen $V_{(0)} = \kappa_n$ insbesondere die Ungleichungen

$$V_{(j)}^k \ge \kappa_n^{k-j} V_{(k)}^j \tag{3.64}$$

für $0 < j < k \le n$. Das Gleichheitszeichen gilt hier jedenfalls dann, wenn K eine Kugel ist. In (3.64) ist also eine ganze Serie von Extremaleigenschaften der Kugel enthalten. Der Fall $k = n$ liefert speziell: Unter allen konvexen Körpern mit gegebenem Volumen haben die Kugeln das kleinste j-te innere Volumen ($j = 1, \dots, n-1$). Der Fall $j = 1$ liefert: Unter allen konvexen Körpern mit gegebener mittlerer Breite haben die Kugeln das größte k-te innere Volumen ($k = 2, \dots, n$).

Leider erhalten wir auf diesem Wege keine Information darüber, ob die Kugeln jeweils die einzigen Extremalkörper sind. Nach der Herleitung gilt in der Ungleichung (3.64) (falls wir $V_{(k)} > 0$ annehmen) genau dann das Gleichheitszeichen, wenn

$$V_{(r)}^2 = V_{(r-1)} V_{(r+1)} \qquad \text{für } r = 1, \dots, k-1$$

gilt. Insbesondere folgt aus Gleichheit in (3.64) also die Gleichheit

$$V_{(1)}^2 = V_{(0)} V_{(2)}.$$

In diesem speziellen Fall läßt sich zeigen, dass die Gleichheit nur für Kugeln gilt. Auf den Beweis können wir hier allerdings nicht mehr eingehen.

3.5 Steiner-Symmetrisierung

Bei Extremalproblemen der Konvexgeometrie treten häufig Kugeln oder Ellipsoide als Extremalkörper auf, also Körper mit einem hohen Maß an Symmetrie. Daher können Beweise für eine Reihe von Extremaleigenschaften der Kugeln oder Ellipsoide durch sogenannte Symmetrisierungsverfahren geführt werden. Ein Beispiel dieser Art haben wir in Abschnitt 3.1 (S. 78) kennengelernt. Es ging dabei um den Nachweis, daß ein konvexer Körper mit gegebenem Durchmesser und maximalem Volumen notwendig eine Kugel sein muss. Der erste Schritt bestand darin, mittels des Auswahlsatzes von Blaschke die Existenz von Extremalkörpern zu zeigen. Der zweite Schritt zeigt dann, dass ein Extremalkörper nur eine Kugel sein kann. Dabei geht man indirekt vor. Angenommen, ein Körper K_0 mit gegebenem Durchmesser und maximalem Volumen wäre keine Kugel. Dann gibt es eine Hyperebene H derart, dass K_0 keine zu H parallele Symmetrieebene besitzt. Man spiegelt nun den Körper K_0 an der Hyperebene H, sei K_1 das Spiegelbild, und bildet den Körper $K' := \frac{1}{2}K_0 + \frac{1}{2}K_1$. Von ihm sagt man, dass er durch *Blaschke-Symmetrisierung* an der Hyperebene H aus K_0 hervorgeht. Wichtig ist nun, dass diese Blaschke-Symmetrisierung die Eigenschaften

$$D(K') \leq D(K_0), \qquad V_n(K') \geq V_n(K_0)$$

hat. Das eine der beiden betrachteten Funktionale wird also bei Blaschke-Symmetrisierung nicht vergrößert, das andere wird nicht verkleinert. Die Ungleichung $V_n(K') \geq V_n(K_0)$ folgt dabei aus dem Satz von Brunn-Minkowski; dieser liefert auch die Information, daß $V_n(K') = V_n(K_0)$ nur dann gilt, wenn K' und K_0 homothetisch sind. Wenn das gilt, hat K_0 eine zu H parallele Symmetrie-Ebene, was aber ausgeschlossen war. Es muß also $V_n(K') > V_n(K_0)$ sein, aber das ist ein Widerspruch zur Extremaleigenschaft von K_0. Also ist K_0 eine Kugel.

Wir wollen nun ein anderes Symmetrisierungsverfahren kennenlernen. Es hat die nützliche Eigenschaft, das Volumen nicht zu verändern. Von einer Reihe anderer

Funktionale läßt sich zeigen, dass sie sich bei dieser Symmetrisierung monoton verhalten. Das gestattet dann die Lösung einiger Extremalprobleme für Körper gegebenen Volumens.

Dieses Verfahren, die *Steiner-Symmetrisierung*, lässt sich allgemein auf kompakte Mengen anwenden. Wir beschränken uns hier aber auf konvexe Körper. Das Verfahren läßt sich anschaulich folgendermaßen beschreiben. Gegeben sei eine Hyperebene H. Sei $K \in \mathcal{K}^n$ ein konvexer Körper. Jede zu H senkrechte Gerade G, die K trifft, schneidet K in einer (eventuell ausgearteten) Strecke. Diese Strecke wird so auf G verschoben, dass ihr Mittelpunkt auf H zu liegen kommt. Die Vereinigung aller so verschobenen Strecken ist ein neuer konvexer Körper $S_H(K)$. Die damit definierte Abbildung $S_H : \mathcal{K}^n \to \mathcal{K}^n$ heißt *Steiner-Symmetrisierung an* H.

Wir beschreiben diese Symmetrisierung jetzt etwas formaler. O.B.d.A. können wir annehmen, dass $H = H_{u,0}$ mit einem Einheitsvektor u gilt. Jeder Punkt $x \in \mathbb{R}^n$ lässt sich eindeutig in der Form

$$x = x_0 + \lambda u \quad \text{mit } x_0 \in H, \lambda \in \mathbb{R}$$

darstellen. Für $x \in \mathbb{E}^n$ ist

$$G_x := \{x + \lambda u : \lambda \in \mathbb{R}\}$$

die zu H senkrechte Gerade durch x. Für $K \in \mathcal{K}^n$ bezeichne $K|H$ das Bild von K unter der Orthogonalprojektion auf H, also

$$K|H = \{x \in H : G_x \cap K \neq \emptyset\}.$$

Für $x \in K|H$ ist $G_x \cap K$ eine (eventuell ausgeartete) Strecke, also von der Form

$$G_x \cap K = \{x + zu : \underline{z}(x) \leq z \leq \overline{z}(x)\}.$$

Hierdurch werden zwei Funktionen

$$\underline{z} : K|H \to \mathbb{R} \quad \text{und} \quad \overline{z} : K|H \to \mathbb{R}$$

definiert. Aus der Konvexität von K folgt, dass $\underline{z}$ konvex und $\overline{z}$ konkav ist. Diese Funktionen sind also auf $\operatorname{relint} K|H$ stetig, aber i.a. nicht auf ganz $K|H$. (Als Beispiel betrachte man im $\mathbb{E}^3$ die konvexe Hülle einer Kreisscheibe in H und einer zu H senkrechten Strecke durch einen Randpunkt der Kreisscheibe.) Es gilt jedoch: Zu $x \in K|H$ und $\epsilon > 0$ existiert eine Umgebung U von x mit

$$\underline{z}(y) \geq \underline{z}(x) - \epsilon \quad \text{für alle } y \in U \cap (K|H).$$

Wäre das nämlich für ein $x \in K|H$ falsch, so gäbe es ein $\epsilon > 0$ und eine Folge $(y_i)_{i\in\mathbb{N}}$ in $K|H$ mit $y_i \to x$ und $\underline{z}(y_i) < \underline{z}(x) - \epsilon$. Man kann o.B.d.A. (d.h. nach

Auswahl einer Teilfolge) annehmen, dass $\underline{z}(y_i) \to a$ für eine reelle Zahl a gilt. Dann gilt $y_i + \underline{z}(y_i)u \in K$, $y_i + \underline{z}(y_i)u \to x + au$, also $x + au \in K$ und daher $\underline{z}(x) \leq a \leq \underline{z}(x) - \epsilon$, ein Widerspruch. Somit ist die Funktion $\underline{z}$ halbstetig nach unten, und es gilt

$$\underline{z}(x) \leq \liminf_{y \to x} \underline{z}(y).$$

Analog gilt

$$\overline{z}(x) \geq \limsup_{y \to x} \overline{z}(y).$$

Die Zahl $\overline{z}(x) - \underline{z}(x)$ ist die Länge der Strecke $G_x \cap K$. Für $x \in K|H$ setzen wir nun

$$s(x) := \left\{x + zu : -\tfrac{1}{2}(\overline{z}(x) - \underline{z}(x)) \leq z \leq \tfrac{1}{2}(\overline{z}(x) - \underline{z}(x))\right\}$$

und dann

$$S_H(K) := \bigcup_{x \in K|H} s(x).$$

3.5.1 Satz und Definition. *Für $K \in \mathcal{K}^n$ gilt $S_H(K) \in \mathcal{K}^n$, $V_n(S_H(K)) = V_n(K)$, und $S_H(K)$ ist symmetrisch bezüglich H. Der Körper $S_H(K)$ heißt die* Steiner-Symmetrisierte *von K bezüglich H.*

Beweis. Die Symmetrie von $S_H(K)$ bezüglich H folgt aus der Definition, die Gleichheit der Volumina folgt aus dem Satz von Fubini. Die Konvexität von K ergibt sich daraus, dass die Funktion $\overline{z} - \underline{z}$ konkav ist. Daß $S_H(K)$ beschränkt ist, ist klar. Zum Nachweis der Abgeschlossenheit sei $(y_i)_{i \in \mathbb{N}}$ eine Folge in $S_H(K)$ mit $\lim_{i \to \infty} y_i = y$. Dann ist $y_i = x_i + z_i u$ mit $x_i \in K|H$ und $z_i \in \mathbb{R}$, und es gilt $x_i \to x$, $z_i \to z$, $y = x + zu$. Da $K|H$ abgeschlossen ist, gilt $x \in K|H$. Ferner ist

$$2|z_i| \leq \overline{z}(x_i) - \underline{z}(x_i)$$

und daher

$$2|z| = \lim_{i \to \infty} 2|z_i| \leq \limsup_{i \to \infty} \overline{z}(x_i) - \liminf_{i \to \infty} \underline{z}(x_i) \leq \overline{z}(x) - \underline{z}(x),$$

also $x + zu \in S_H(K)$. Somit ist $S_H(K)$ abgeschlossen. ∎

Die Steiner-Symmetrisierung $S_H : \mathcal{K}^n \to \mathcal{K}^n$ ist nicht stetig, wohl aber auf dem Teilraum der Körper mit inneren Punkten. Wir zeigen nur die folgende Aussage:

3.5.2 Hilfssatz. *Sei $(K_j)_{j \in \mathbb{N}}$ eine Folge konvexer Körper mit $K_j \to K$ und derart, dass $S_H(K_j) \to \tilde{K}$ für $j \to \infty$ gilt. Dann gilt $\tilde{K} \subset S_H(K)$.*

Beweis. Sei $x \in \tilde{K}$. Nach Satz 1.8.7 gibt es Punkte $x_j \in S_H(K_j)$ mit $x_j \to x$ für $j \to \infty$. Es gibt Punkte $y_j, z_j \in G_{x_j} \cap K_j$ mit $\|y_j - z_j\| \geq \|x_j - \sigma_H x_j\|$, wo σ_H die

Spiegelung an H bezeichnet. Wir können annehmen (Auswahl von Teilfolgen und Umbenennung), dass $y_j \to y$ und $z_j \to z$ für $j \to \infty$ gilt. Dann folgt $y, z \in G_x \cap K$ und $\|y - z\| \geq \|x - \sigma_H x\|$, also $x \in S_H(K)$. ■

Für die Steiner-Symmetrisierung gilt ein Kugelungssatz analog zu Satz 3.2.5 für Drehmittelungen. Sind $H_1, \dots, H_m$ Hyperebenen (o.B.d.A.) durch 0, so wird die Abbildung $S_{H_m} \circ \ldots \circ S_{H_1}$ als eine *Mehrfachsymmetrisierung* bezeichnet.

3.5.3 Satz. *Zu $K \in \mathcal{K}^n$ sei $\mathcal{S}(K)$ die Menge der konvexen Körper, die durch Mehrfachsymmetrisierung aus K hervorgehen. Dann gibt es in $\mathcal{S}(K)$ eine Folge, die gegen eine Kugel konvergiert.*

Beweis. Wie im Beweis von Satz 3.2.5 sei zu $L \in \mathcal{K}^n$ mit $R(L)$ der Radius der kleinsten Kugel mit Mittelpunkt 0 bezeichnet, die L enthält. Sei

$$R_0 := \inf\{R(K') : K' \in \mathcal{S}(K)\}.$$

Da $\mathcal{S}(K)$ beschränkt ist, gibt es eine Folge $(K_j)_{j\in\mathbb{N}}$ in $\mathcal{S}(K)$ mit

$$\lim_{j\to\infty} R(K_j) = R_0 \quad \text{und} \quad \lim_{j\to\infty} K_j = K_0$$

für einen konvexen Körper $K_0 \in \mathcal{K}^n$, und wegen der Stetigkeit von R folgt $R(K_0) = R_0$.

Angenommen, K_0 wäre nicht die Kugel B_0 mit Mittelpunkt 0 und Radius R_0. Dann gibt es einen Punkt $z \in \operatorname{bd} B_0$ mit $z \notin K_0$. Es gibt eine Kugel C mit Mittelpunkt z und mit $\operatorname{bd} B_0 \cap C \cap K_0 = \emptyset$. Ist H eine beliebige Hyperebene durch 0, so gilt offenbar

$$\operatorname{bd} B_0 \cap C \cap S_H(K_0) = \emptyset \quad \text{und} \quad \operatorname{bd} B_0 \cap \sigma_H C \cap S_H(K_0) = \emptyset.$$

Wir überdecken $\operatorname{bd} B_0$ durch endlich viele zu C kongruente Kugeln $C_1, \dots, C_m$ mit Mittelpunkten in $\operatorname{bd} B_0$. Sei H_i die Symmetrieebene von C und C_i. Für den Körper $S^*(K_0)$ mit $S^* := S_{H_m} \circ \ldots \circ S_{H_1}$ gilt dann nach der vorstehenden Bemerkung $S^*(K_0) \cap \operatorname{bd} B_0 = \emptyset$. Da $S^*(K_0)$ kompakt ist, folgt daraus $R(S^*(K_0)) < R_0$. Nun war $K_j \in \mathcal{S}(K)$ und $\lim_{j\to\infty} K_j = K_0$. Wir können o.B.d.A. annehmen (mit der üblichen Schlussweise: Auswahlsatz, Auswahl von Teilfolgen, Umbenennung), dass alle Folgen $(S_{H_r} \circ \ldots \circ S_{H_1}(K_j))_{j\in\mathbb{N}}$ konvergieren $(r = 1, \dots, m)$, insbesondere $S^*(K_j) \to \tilde{K}$. Aus Hilfssatz 3.5.2 folgt dann $\tilde{K} \subset S^*(K_0)$, also $R(\tilde{K}) < R_0$ und daher $R(S^*(K_j)) < R_0$ für genügend große j. Wegen $S^*(K_j) \in \mathcal{S}(K)$ ist das ein Widerspruch. Also ist K_0 die Kugel B_0, und es gilt $\lim_{j\to\infty} K_j = B_0$. ■

Wir wollen nun Beispiele für Anwendungen der Steiner-Symmetrisierung geben. Dabei handelt es sich um Extremaleigenschaften von Ellipsoiden. Für die betrachteten Extremalaussagen sind keine Beweise bekannt, die ohne Steiner-Symmetrisierung auskommen. Dass die Extremwerte nicht nur von Kugeln, sondern von Ellipsoiden angenommen werden, liegt daran, dass es sich um affininvariante Funktionale handelt. Ein Funktional $f : \mathcal{K}_0^n \to \mathbb{R}$ (wo $\mathcal{K}_0^n$ der Raum der konvexen Körper mit inneren Punkten im $\mathbb{E}^n$ ist) heißt *affininvariant*, wenn $f(\alpha K) = f(K)$ für jede nichtausgeartete affine Abbildung $\alpha : \mathbb{E}^n \to \mathbb{E}^n$ gilt. Affininvariante Funktionale konvexer Körper kommen in verschiedenen Zusammenhängen vor und werfen interessante Extremalprobleme auf, die aber zum großen Teil ungelöst sind. Grundsätzlich lässt sich zu diesen Extremalproblemen folgendes sagen.

3.5.4 Satz. *Sei $f : \mathcal{K}_0^n \to \mathbb{R}$ ein stetiges, affininvariantes Funktional. Dann nimmt f auf $\mathcal{K}_0^n$ ein Maximum und ein Minimum an.*

Beweis. Sei $S \subset \mathbb{E}^n$ ein reguläres n-Simplex mit Schwerpunkt 0 und $T = -nS$; dann ist T ein reguläres Simplex, dessen Facetten durch die Ecken von S gehen.

Sei $K \in \mathcal{K}_0^n$ ein konvexer Körper. Es gibt ein in K enthaltenes Simplex Δ mit größtem Volumen. Hierzu gibt es eine affine Transformation α des $\mathbb{E}^n$ mit $\alpha\Delta = S$. Dann ist S ein in αK enthaltenes Simplex von größtem Volumen. Jede Facettenhyperebene von T ist Stützebene von αK, denn andernfalls enthielte αK ein Simplex, das größeres Volumen hat als S (dieselbe Grundfläche, größere Höhe). Also gilt $S \subset \alpha K \subset T$. Die Menge

$$\mathcal{A} := \{K' \in \mathcal{K}_0^n : S \subset K' \subset T\}$$

ist beschränkt und abgeschlossen, nach dem Auswahlsatz also kompakt. Die stetige Funktion f nimmt daher auf $\mathcal{A}$ ein Maximum und ein Minimum an. Dies sind zugleich Maximum bzw. Minimum von f auf $\mathcal{K}_0^n$, da es zu jedem $K \in \mathcal{K}_0^n$ eine Affinität α mit $\alpha K \in \mathcal{A}$ gibt und da $f(\alpha K) = f(K)$ ist. ■

Ein wichtiges affininvariantes Funktional ist das *Volumenprodukt*, das wir hier aber nur für zentralsymmetrische konvexe Körper betrachten wollen. Es bezeichne $\mathcal{K}_c^n$ die Menge der konvexen Körper K in $\mathcal{K}_0^n$ mit $K = -K$. Für $K \in \mathcal{K}_c^n$ ist das Volumenprodukt erklärt durch

$$vp(K) := V_n(K)V_n(K^*),$$

wo K^* der Polarkörper von K ist. Nach Definition ist

$$K^* = \{y \in \mathbb{E}^n : \langle y, x\rangle \le 1 \text{ für alle } x \in K\}.$$

Ist α eine lineare Transformation des $\mathbb{E}^n$, so gilt

$$y \in (\alpha K)^* \quad \Leftrightarrow \quad \langle y, x\rangle \le 1 \,\forall\, x \in \alpha K$$

$$\begin{aligned}
&\Leftrightarrow \quad \langle y, \alpha x\rangle \leq 1 \,\forall\, x \in K \\
&\Leftrightarrow \quad \langle \alpha^* y, x\rangle \leq 1 \,\forall\, x \in K \\
&\Leftrightarrow \quad \alpha^* y \in K^* \\
&\Leftrightarrow \quad y \in \alpha^{*-1} K^*,
\end{aligned}$$

wo α^* die adjungierte lineare Abbildung bezeichnet. Es folgt

$$\begin{aligned}
vp(\alpha K) &= V_n(\alpha K) V_n(\alpha^{*-1} K^*) \\
&= |\det \alpha| V_n(K) |\det \alpha|^{-1} V_n(K^*) \\
&= vp(K).
\end{aligned}$$

Das Volumenprodukt ist also invariant unter linearen Transformationen. Allgemeiner kann man das Volumenprodukt auf $\mathcal{K}_0^n$ erklären durch

$$vp(K) := V_n((K - c_K)) V_n((K - c_K)^*),$$

wo c_K den Schwerpunkt von K bezeichnet; es ist dann affininvariant.

Der folgende Satz zeigt, dass das Volumenprodukt auf $\mathcal{K}_c^n$ sein Maximum auf den Ellipsoiden annimmt.

3.5.5 Satz (Blaschke-Santaló-Ungleichung für zentralsymmetrische Körper). *Für $K \in \mathcal{K}_c^n$ gilt*

$$V_n(K) V_n(K^*) \leq \kappa_n^2.$$

Beweis. Wir wollen das Verhalten des Volumenprodukts unter Steiner-Symmetrisierung untersuchen. Da das Volumen bei Steiner-Symmetrisierung invariant ist, muss $V_n(S_H(K)^*)$ betrachtet werden. Dabei sei H eine Hyperebene durch 0, also $H = u^\perp$ mit einem Vektor $u \in S^{n-1}$. Jedes $x \in \mathbb{E}^n$ hat eine eindeutige Darstellung $x = y + tu$ mit $y \in H$ und $t \in \mathbb{R}$, und wir schreiben $x = (y, t)$.

Sei $K \in \mathcal{K}_c^n$. Es gilt

$$(y, s) \in S_H(K) \Leftrightarrow \exists\, (y, s_j) \in K \ \ (j = 1, 2) \text{ mit } s = \frac{s_1 - s_2}{2}.$$

Für einen konvexen Körper M mit $0 \in \operatorname{int} M$ ist

$$M^* = \{(z, t) : \langle (z, t), (y, s)\rangle \leq 1 \,\forall\, (y, s) \in M\},$$

also gilt

$$(z, t) \in M^* \Leftrightarrow \langle z, y\rangle + ts \leq 1 \,\forall\, (y, s) \in M.$$

Es folgt

$$\begin{aligned}(z,t)\in S_H(K)^* \quad &\Leftrightarrow \quad \langle z,y\rangle + ts \le 1,\ \text{wenn } (y,s)\in S_H(K)\\ &\Leftrightarrow \quad \langle z,y\rangle + ts \le 1,\ \text{wenn } (y,s_j)\in K \text{ und } s=\frac{s_1-s_2}{2}\\ &\Leftrightarrow \quad \langle z,y\rangle + t\frac{s_1-s_2}{2} \le 1,\ \text{wenn } (y,s_j)\in K\ \ (j=1,2).\end{aligned}$$

Für eine Menge $E\subset\mathbb{E}^n$ und für $t\in\mathbb{R}$ setzen wir

$$E_t := \{y\in H : (y,t)\in E\}$$

(in der Maßtheorie als t-Schnitt von E bezeichnet). Unter K_t^* ist $(K^*)_t$ zu verstehen. Sei nun

$$z\in\frac{1}{2}K_t^*+\frac{1}{2}(-K_t^*)=\frac{1}{2}K_t^*+\frac{1}{2}K_{-t}^*.$$

Die Gleichheit folgt, weil $K=-K$ und damit auch $K^*=-K^*$ ist. Es ist also

$$z=\frac{1}{2}p+\frac{1}{2}q \quad \text{mit } p\in K_t^* \text{ und } q\in K_{-t}^*.$$

Für $(y,s_j)\in K$ $(j=1,2)$ folgt

$$\begin{aligned}\langle z,y\rangle + t\frac{s_1-s_2}{2} \quad &= \quad \left\langle \frac{p+q}{2},y\right\rangle + t\frac{s_1-s_2}{2}\\ &= \quad \frac{1}{2}\left(\langle p,y\rangle + ts_1\right) + \frac{1}{2}(\langle q,y\rangle - ts_2) \le 1,\end{aligned}$$

denn wegen $(p,t)\in K^*$ ist $\langle p,y\rangle + ts_1\le 1$, und wegen $(q,-t)\in K^*$ ist $\langle q,y\rangle - ts_2 \le 1$. Also folgt $(z,t)\in S_H(K)^*$ und somit $z\in[S_H(K)^*]_t$. Damit ist

$$\frac{1}{2}K_t^*+\frac{1}{2}(-K_t^*)\subset [S_H(K)^*]_t$$

gezeigt. Mit dem Satz von Brunn-Minkowski folgt

$$V_{n-1}(K_t^*)\le V_{n-1}\left(\frac{1}{2}K_t^*+\frac{1}{2}(-K_t^*)\right)\le V_{n-1}\left([S_H(K)^*]_t\right).$$

Integration nach t und der Satz von Fubini ergeben

$$V_n(K^*)\le V_n(S_H(K)^*).$$

Wegen $V_n(K)=V_n(S_H(K))$ ist also

$$vp(K)\le vp(S_H(K)).$$

Aus Satz 3.5.3 und der Stetigkeit des Volumenprodukts lässt sich nun folgern, dass vp ein Maximum auf der Menge der Kugeln annimmt. Das liefert gerade die behauptete Ungleichung, denn es ist $V_n(rB^n) = r^n\kappa_n$ und $V_n\left((rB^n)^*\right) = V_n\left(\frac{1}{r}B_n\right) = r^{-n}\kappa_n$. ■

Da das Volumenprodukt invariant ist unter linearen Transformationen, nimmt es sein Maximum an auf der Menge der Ellipsoide mit Mittelpunkt 0. Dass das Maximum nur von diesen Ellipsoiden angenommen wird, lässt sich mit zusätzlichem Aufwand ebenfalls zeigen. Es ist ein ungelöstes Problem, auf welchen konvexen Körpern das Volumenprodukt sein Minimum annimmt.

Das nächste Extremalproblem, das wir mit Steinerscher Symmetrisierung behandeln wollen, geht zurück auf eine historische Aufgabe von J.J. Sylvester (um 1865). Er stellte die folgende Frage. In einem ebenen konvexen Körper $K \in \mathcal{K}_0^2$ werden unabhängig vier zufällige Punkte gewählt. Mit $p_4^{(3)}(K)$ sei die Wahrscheinlichkeit dafür bezeichnet, dass die konvexe Hülle dieser Punkte ein Dreieck ist. Was lässt sich über diese Wahrscheinlichkeit sagen? Diese Aufgabe ist natürlich nicht präzise gestellt, wenn wir nicht festlegen, welche Wahrscheinlichkeitsverteilung die zufälligen Punkte haben sollen. Wir wollen die natürlichste Voraussetzung machen, dass jeder der zufälligen Punkte uniform in K verteilt ist, das heißt, seine Wahrscheinlichkeitsverteilung ist gegeben durch das auf K eingeschränkte Lebesgue-Maß, das zu einem Wahrscheinlichkeitsmaß normiert ist. Wegen der vorausgesetzten stochastischen Unabhängigkeit der vier zufälligen Punkte ist ihre simultane Verteilung gleich dem Produkt der einzelnen Verteilungen. Daher kann die fragliche Wahrscheinlichkeit $p_4^{(3)}(K)$ jetzt durch ein mehrfaches Integral ausgedrückt werden. Setzen wir für ebene Polygone P

$$g(P) := \begin{cases} 1, & \text{wenn } P \text{ ein Dreieck ist} \\ 0 & \text{sonst,} \end{cases}$$

so ist

$$p_4^{(3)}(K) = \int_K \cdots \int_K g(\operatorname{conv}\{x_1,\dots,x_4\}) \frac{d\lambda(x_1)}{\lambda(K)} \cdots \frac{d\lambda(x_4)}{\lambda(K)},$$

wo λ das zweidimensionale Lebesgue-Maß bezeichnet. Statt $\lambda(K)$ schreiben wir wieder $V_2(K)$, und $d\lambda(x)$ kürzen wir mit dx ab. Dann ist also

$$p_4^{(3)}(K) = \frac{1}{V_2(K)^4} \int_K \cdots \int_K g(\operatorname{conv}\{x_1,\dots,x_4\})\, dx_1 \cdots dx_4.$$

Der Integrand ist genau dann 1, wenn drei der Punkte $x_1,\dots,x_4$ ein Dreieck bilden und der vierte in diesem Dreieck liegt. Aus Symmetriegründen ist also

$$p_4^{(3)}(K) \quad = \quad \frac{4}{V_2(K)^4} \int_K \int_K \int_K \left(\int_{x_4 \in \operatorname{conv}\{x_1,x_2,x_3\}} dx_4 \right) dx_1 dx_2 dx_3$$

$$= \frac{4}{V_2(K)^4} \int_K \int_K \int_K V_2(\operatorname{conv}\{x_1, x_2, x_3\})\, dx_1 dx_2 dx_3.$$

Mit Hilfe dieser Integraldarstellung ist unschwer nachzuweisen, dass $p_4^{(3)}(K)$, als Funktion des konvexen Körpers K aufgefasst, ein affininvariantes stetiges Funktional auf $\mathcal{K}_0^2$ ist. Es nimmt also nach Satz 3.5.4 ein Maximum und ein Minimum an. Für spezielle Körper K, wie Kreise (und damit für Ellipsen), Dreiecke, Parallelogramme lässt sich das Integral explizit berechnen, im allgemeinen wird das aber nicht der Fall sein. Daher kann die Aufgabe von Sylvester für allgemeine K nur so interpretiert werden, dass man den möglichen Wertebereich der Funktion $p_4^{(3)}$ bestimmen soll, also ihre Extremwerte.

Allgemeiner können wir $m \geq 4$ unabhängige, uniform verteilte zufällige Punkte in K betrachten und nach der Wahrscheinlichkeit $p_m^{(3)}(K)$ fragen, dass ihre konvexe Hülle ein Dreieck ist. Analog wie oben ergibt sich

$$p_m^{(3)}(K) = \frac{\binom{m}{3}}{V_2(K)^m} \int_K \int_K \int_K V_2(\operatorname{conv}\{x_1, x_2, x_3\})^{m-3}\, dx_1 dx_2 dx_3.$$

Entsprechende Fragen kann man natürlich im n-dimensionalen Raum stellen. Wir definieren jetzt allgemeiner für $K \in \mathcal{K}_0^n$, $k \in \mathbb{N}$ mit $k \geq n+1$ und $r \in \mathbb{R}$ mit $r \geq 1$

$$f_r(k, K) := \int_K \cdots \int_K V_n(\operatorname{conv}\{x_1, \ldots, x_k\})^r\, dx_1 \cdots dx_k.$$

Wir können $f_r(k,K)/V_n(K)^k$ auffassen als das r-te Moment des Volumens der konvexen Hülle von k unabhängigen, uniform verteilten zufälligen Punkten in K. Das Funktional

$$K \mapsto \frac{f_r(k, K)}{V_n(K)^{k+r}}$$

ist affininvariant; es ist aber im folgenden bequemer, mit $f_r(k,K)$ zu arbeiten. Für diese Funktionale lässt sich mit Steinerscher Symmetrisierung der folgende Satz zeigen.

3.5.6 Satz. *Für $K \in \mathcal{K}_0^n$, $k \in \mathbb{N}$ mit $k \geq n+1$ und $r \geq 1$ gilt*

$$f_r(k, K) \geq f_r(k, B),$$

wenn B eine Kugel mit $V_n(K) = V_n(B)$ ist. Gleichheit gilt genau dann, wenn K ein Ellipsoid ist.

Für die Wahrscheinlichkeit im ebenen Problem von Sylvester bedeutet das

$$p_4^{(3)}(K) \geq p_4^{(3)}(E),$$

wenn E eine Ellipse ist. Für Kreise und damit für Ellipsen kann man die Wahrscheinlichkeit explizit ausrechnen und erhält

$$p_4^{(3)}(K) \geq \frac{35}{12\pi^2} = 0,29552.$$

Das ist zuerst von W. Blaschke 1917 bewiesen worden. Er hat auch gezeigt, dass in der Ebene

$$p_4^{(3)}(K) \leq \frac{1}{3}$$

gilt, mit Gleichheit genau dann, wenn K ein Dreieck ist.

Im Beweis von Satz 3.5.6 folgen wir einer Arbeit von H. Groemer (1974), werden allerdings die Charakterisierung des Gleichheitsfalls nicht vollständig ausführen.

Warum Steiner-Symmetrisierung hier erfolgreich eingesetzt werden kann, soll zunächst an dem von Blaschke behandelten Spezialfall $n = 2$, $k = 3$, $r = 1$ heuristisch erläutert werden. Zur Abkürzung setzen wir

$$f(K) := \int_K \int_K \int_K V_2(\operatorname{conv}\{x_1, x_2, x_3\})\, dx_1 dx_2 dx_3.$$

Sei $S_H(K)$ die Steiner-Symmetrisierte von K an einer Geraden K. Wir erklären eine Abbildung $\varphi : S_H(K) \to K$ in der folgenden Weise. Sei u ein Einheitsvektor senkrecht zu H. Zu $x \in S_H(K)$ seien die Schnittpunkte $a_1, a_2 \in G_x \cap \operatorname{bd} S_H(K)$ und $b_1, b_2 \in G_x \cap \operatorname{bd} K$ so numeriert, dass $a_2 - a_1$ und $b_2 - b_1$ die Richtung von u haben. Dann ist $x = a_1 + tu$ mit $t \geq 0$, und wir definieren $\varphi(x) := b_1 + tu$. Es folgt aus dem Satz von Fubini, dass die damit erklärte Abbildung φ flächentreu ist. Man kann nun zeigen, dass für $x_1, x_2, x_3 \in S_H(K)$ die Ungleichung

$$\begin{aligned} & V_2(\operatorname{conv}\{x_1, x_2, x_3\}) + V_2(\operatorname{conv}\{\sigma_H x_1, \sigma_H x_2, \sigma_H x_3\}) \\ \leq \; & V_2(\operatorname{conv}\{\varphi(x_1), \varphi(x_2), \varphi(x_3)\}) + V_2(\operatorname{conv}\{\varphi(\sigma_H x_1), \varphi(\sigma_H x_2), \varphi(\sigma_H x_3)\}) \end{aligned}$$

gilt. Wegen der Flächentreue der Abbildung φ kann man hieraus durch Integration die Ungleichung $f(S_H(K)) \leq f(K)$ erhalten. Das Funktional f wird also bei Steinerscher Symmetrisierung nicht vergrößert. Dies ist die Grundidee des Beweises. Im Fall $k > n + 1$ muss man jedoch das entsprechende monotone Verhalten bei Steiner-Symmetrisierung etwas anders nachweisen.

Beweis von Satz 3.5.6 (teilweise).

Im folgenden seien die Zahlen $k \geq n + 1$ und (zunächst) $r > 0$ vorgegeben. Für $K \in \mathcal{K}_0^n$ setzen wir

$$\Phi(K) := \int_K \cdots \int_K V_n(\operatorname{conv}\{x_1, \dots, x_k\})^r \, dx_1 \cdots dx_k.$$

Es ist leicht nachzuweisen, dass Φ stetig ist und dass Φ/V_n^{k+r} affininvariant ist. Nach Satz 3.5.4 gibt es daher einen konvexen Körper B mit $V_n(B) = 1$ und

$$\Phi(K) \geq \Phi(B)$$

für alle konvexen Körper K mit $V_n(K) = 1$. Wir müssen zeigen, dass B ein Ellipsoid ist. Dazu benutzen wir einen geometrischen Hilfssatz, der hier allerdings nicht bewiesen werden soll. Ist K ein konvexer Körper und G eine Gerade im $\mathbb{E}^n$, so bezeichnen wir mit $M(K,G)$ die Menge der Mittelpunkte aller Sehnen $K \cap G'$, wo G' die zu G parallelen Geraden durchläuft.

3.5.7 Hilfssatz. *Der konvexe Körper $B \in \mathcal{K}_0^n$ ist genau dann ein Ellipsoid, wenn für jede Gerade G die Menge $M(K,G)$ in einer Hyperebene liegt.*

Wir wollen nun das Verhalten des Funktionals Φ bei Steinerscher Symmetrisierung untersuchen. Dazu legen wir einige Bezeichnungen fest. Sei H eine Hyperebene durch 0 und u ein zu H senkrechter Einheitsvektor. Wie früher schreiben wir die eindeutige Darstellung $x = y + tu$ mit $y \in H$, $t \in \mathbb{R}$ in der Form $x = (y,t)$. Seien nun Punkte

$$c_1, \dots, c_k \in H$$

vorgegeben. Für $Z = (z_1, \dots, z_k) \in \mathbb{R}^k$ setzen wir

$$\begin{aligned} C(Z) &:= \operatorname{conv}\{(c_1, z_1), \dots, (c_k, z_k)\}, \\ v(Z) &:= V_n(C(Z)). \end{aligned}$$

Wenn Z in $\mathbb{R}^k$ variiert, durchläuft $C(Z)$ Polytope, deren Ecken auf gewissen der zu u parallelen Geraden durch $c_1, \dots, c_k$ liegen.

3.5.8 Hilfssatz. *Die Funktion v ist konvex, d.h. es gilt*

$$v((1-\lambda)Z' + \lambda Z'') \leq (1-\lambda)v(Z') + \lambda v(Z'')$$

für $Z', Z'' \in \mathbb{R}^k$ und $\lambda \in [0,1]$.

Beweis. Zu jedem konvexen Körper $M \in \mathcal{K}_0^n$ gibt es die am Anfang dieses Abschnitts definierten Funktionen $\underline{z}$, $\overline{z}$ auf $M|H$, so dass

$$M = \{(y,z) : y \in M|H,\ \underline{z}(y) \leq z \leq \overline{z}(y)\}$$

gilt. Wir nennen sie die *Unter-* bzw. *Oberfunktion* von M.

Nun seien $Z', Z'' \in \mathbb{R}^k$ und $\lambda \in [0,1]$ gegeben. Es seien $\underline{z}'$, $\overline{z}'$ die Unter- und Oberfunktion des Polytops $C(Z')$, analog seien $\underline{z}''$, $\overline{z}''$ für $C(Z'')$ erklärt. Die Polytope $C(Z')$ und $C(Z'')$ haben dieselbe Projektion auf H, nämlich

$$D := \operatorname{conv}\{c_1, \dots, c_k\}.$$

Daher kann man

$$\begin{aligned} \underline{z}^+ &:= (1-\lambda)\underline{z}' + \lambda\underline{z}'', \\ \overline{z}^+ &:= (1-\lambda)\overline{z}' + \lambda\overline{z}'' \end{aligned}$$

definieren. Da $\underline{z}^+$ konvex und $\overline{z}^+$ konkav ist, wird durch

$$Q := \{(y,z) : y \in D,\ \underline{z}^+(y) \le z \le \overline{z}^+(y)\}$$

ein konvexer Körper Q definiert. Mit $Z' = (z_1', \dots, z_k')$, $Z'' = (z_1'', \dots, z_k'')$ gilt

$$\begin{aligned} (c_i, z_i') \in C(Z') &\Rightarrow \underline{z}'(c_i) \le z_i' \le \overline{z}'(c_i), \\ (c_i, z_i'') \in C(Z'') &\Rightarrow \underline{z}''(c_i) \le z_i'' \le \overline{z}''(c_i). \end{aligned}$$

Es folgt

$$\underline{z}^+(c_i) \le (1-\lambda)z_i' + \lambda z_i'' \le \overline{z}^+(c_i)$$

und damit

$$(c_i, (1-\lambda)z_i' + \lambda z_i'') \in Q.$$

Dies gilt für $i = 1, \dots, k$. Da Q konvex ist, folgt

$$C((1-\lambda)Z' + \lambda Z'') \subset Q$$

und daraus

$$v((1-\lambda)Z' + \lambda Z'') \le V_n(Q).$$

Nun ist (im folgenden bezeichnet dy Integration bezüglich des $(n-1)$-dimensionalen Lebesgue-Maßes in H)

$$\begin{aligned} V_n(Q) &= \int_D [\overline{z}^+(y) - \underline{z}^+(y)]\, dy \\ &= \int_D [(1-\lambda)(\overline{z}'(y) - \underline{z}'(y)) + \lambda(\overline{z}''(y) - \underline{z}''(y))]\, dy \\ &= (1-\lambda)V_n(C(Z')) + \lambda V_n(C(Z'')) \\ &= (1-\lambda)v(Z') + \lambda v(Z''). \end{aligned}$$

Damit ist Hilfssatz 3.5.8 bewiesen.

3.5.9 Hilfssatz. *Sei $r \ge 1$, seien Zahlen $a_1, \dots, a_k > 0$ gegeben. Für $W = (w_1, \dots, w_k) \in \mathbb{R}^k$ sei*

$$M(W) := \int_{w_k - a_k}^{w_k + a_k} \cdots \int_{w_1 - a_1}^{w_1 + a_1} v(Z)^r \, dz_1 \cdots dz_k.$$

Dann gilt

$$M(W) \geq M(O), \quad O = (0, \dots, 0) \in \mathbb{R}^k,$$

und Gleichheit gilt genau im Fall $v(W) = 0$.

Beweis. Mit $u_i := z_i - w_i$, $U := (u_1, \dots, u_k)$ ist

$$M(W) = \int\limits_{-a_k}^{a_k} \cdots \int\limits_{-a_1}^{a_1} v(W+U)^r \, du_1 \cdots du_k.$$

Die Substitution $U \to -U$ ergibt

$$M(W) = \int\limits_{-a_k}^{a_k} \cdots \int\limits_{-a_1}^{a_1} v(W-U)^r \, du_1 \cdots du_k,$$

und da Spiegelung an H das Volumen ungeändert lässt, folgt

$$M(W) = \int\limits_{-a_k}^{a_k} \cdots \int\limits_{-a_1}^{a_1} \frac{1}{2}[v(U+W)^r + v(U-W)^r] \, du_1 \cdots du_k.$$

Nach Hilfssatz 3.5.8 ist

$$v(U) = v\left(\frac{1}{2}(U+W) + \frac{1}{2}(U-W)\right) \leq \frac{1}{2}v(U+W) + \frac{1}{2}v(U-W).$$

Da die Funktion $\zeta \mapsto \zeta^r$ für $\zeta \geq 0$ monoton wachsend und für $r \geq 1$ konvex ist, folgt

$$\begin{aligned} v(U)^r &\leq \left[\frac{1}{2}v(U+W) + \frac{1}{2}v(U-W)\right]^r \\ &\leq \frac{1}{2}v(U+W)^r + \frac{1}{2}v(U-W)^r. \end{aligned} \tag{3.65}$$

Damit ergibt sich

$$M(W) \geq \int\limits_{-a_k}^{a_k} \cdots \int\limits_{-a_1}^{a_1} v(U)^r \, du_1 \cdots du_r = M(O).$$

Gilt hier Gleichheit, so gilt in (3.65) aus Stetigkeitsgründen Gleichheit im ganzen Integrationsbereich, insbesondere für $U = O$. Da die Funktion $\zeta \mapsto \zeta^r$ für $\zeta \geq 0$ streng monoton wachsend ist, folgt $v(W) = 0$. Damit ist Hilfssatz 3.5.9 bewiesen.

Jetzt können wir den Beweis von Satz 3.5.6 zu Ende führen. Für $\lambda > 0$ gilt $\Phi(\lambda K) = \lambda^{n(r+k)}\Phi(K)$, daher können wir o.B.d.A. $V_n(K) = 1$ annehmen. Ferner sei $n \geq 2$; der Fall $n = 1$ ist trivial. Angenommen, der nach Satz 3.5.4 existierende Körper K mit $V_n(K) = 1$, für den Φ ein Minimum annimmt, wäre kein Ellipsoid. Nach Hilfssatz 3.5.7 gibt es dann eine Gerade G derart, dass die Mittelpunktsmenge $M(K, G)$ nicht in einer Hyperebene liegt. Sei H die Hyperebene durch 0 senkrecht zu G. Wegen $k \geq n+1$ gibt es k zu G parallele Geraden $G_1, \dots, G_k$, die K in Strecken positiver Länge schneiden und derart, dass die Mittelpunkte $(c_1, w_1), \dots, (c_k, w_k)$ dieser Sehnen nicht in einer Hyperebene liegen. Es ist also $v(W) > 0$. Nach Hilfssatz 3.5.9 folgt $M(W) > M(O)$.

Bezeichnet $\underline{z}$ die Unter- und $\overline{z}$ die Oberfunktion von K, so gilt

$$\begin{aligned} \Phi(K) &= \int_K \cdots \int_K V_n(\operatorname{conv}\{x_1, \dots, x_k\})^r \, dx_1 \cdots dx_k \\ &= \int_{K|H} \cdots \int_{K|H} \left[\int_{\underline{z}(y_k)}^{\overline{z}(y_k)} \cdots \int_{\underline{z}(y_1)}^{\overline{z}(y_1)} v(y, Z)^r \, dz_1 \cdots dz_k \right] dy_1 \cdots dy_k, \end{aligned}$$

wo $v(y, Z) = V_n(\operatorname{conv}\{(y_1, z_1), \dots, (y_k, z_k)\})$ gesetzt ist. An der Stelle $(y_1, \dots, y_k) = (c_1, \dots, c_k)$ ist das innere Integral $[\cdot]$ gleich $M(W)$ (gebildet mit $2a_i = \overline{z}(c_i) - \underline{z}(c_i)$), ist also $> M(O)$. Aus Stetigkeitsgründen gilt eine entsprechende strikte Ungleichung in einer ganzen Umgebung. An allen anderen Stellen gilt für die Größe $M(y, W)$, die mit der Integrationsvariablen $(y_1, \dots, y_k)$ statt $(c_1, \dots, c_k)$ gebildet ist, nach Hilfssatz 3.5.9 jedenfalls die Ungleichung $M(y, W) \geq M(y, O)$. Nach Ausführung der äußeren Integrationen erhält man daher gerade die Ungleichung

$$\Phi(K) > \Phi(S_H(K)).$$

Das ist ein Widerspruch, da K unter allen konvexen Körpern vom Volumen 1 den kleinsten Wert für das Funktional Φ liefern sollte. Aus diesem Widerspruch folgt, dass K ein Ellipsoid ist. ∎