Extreme Gitter

Diplomarbeit an der Universität Ulm

– Fakultät für Mathematik und Wirtschaftswissenschaften –

März 2006

Extreme Gitter

Diplomarbeit an der Universität Ulm

– Fakultät für Mathematik und Wirtschaftswissenschaften –

vorgelegt von

Cordian Benedikt Riener

März 2006

Gutachter:

Prof. Dr. Werner Lütkebohmert

Prof. Dr. Gabriele Nebe

Danksagung

Mein Dank gilt allen, die mich während meines Studiums unterstützt haben. An erster Stelle meiner Mutter, die mir während des Studiums oft den Rücken frei gehalten hat, sowie der Konrad-Adenauer-Stiftung der ich ein Stipendium verdanke, das mir ein Studium möglich gemacht hat.

Bedanken möchte ich mich auch bei Frau Prof. Nebe, die mich bei dieser Diplomarbeit betreut hat und mir im Zuge der Erstellung einen Aufentalt in Bordeaux ermöglicht hat. Auch den Professorinnen und Professoren in Bordeaux ganz besonders Christine Bachoc und Jacques Martinet möchte ich ein herzliches Dankeschön aussprechen. Ich fand bei ihnen während meines Aufenthaltes in Bordeaux stets eine offene Tür für alle meine Fragen.

Ulm, im Februar 2006 Cordian Benedikt Riener

Inhaltsverzeichnis

Einführung

Ubi materia ibi geometria

Strena, Seu de Nive Sexangula
Johannes Kepler

Quadratische Formen zählen zu den ältesten Disziplinen der Mathematik. Schon Platos Schüler Menaichmos widmete sein Interesse diesen; damals allerdings nur in Form der Kegelschnitte, welche er an konkreten Modellen studierte. Aber auch die Suche der Pythagoräer nach ganzen Zahlen der Form $a^2 + b^2 = c^2$ motivierte das Studium quadratischer Formen bereits in der Antike.

Pierre de Fermat beschäftigte sich eingehend mit der Darstellbarkeit natürlicher Zahlen durch definite quadratische Formen, jedoch war es erst Lagrange, der ein systematisches Studium der binären quadratischen Formen begann. In seiner Schrift [13] erwähnte Gauß zum ersten mal ein elementares geometrisches Modell für quadratische Formen über den ganzen Zahlen: „ Ein regelmäßiges System von parallelogrammatisch geordneten Punkten, die in den Durchschnitten zweier Systeme von Parallellinien liegen.“ Somit wird Gauß als der Begründer des Studiums der Gitter angesehen. Trotz dieser Anschauung quadratischer Formen als Gitter, bleiben die Arbeiten von Gauß streng abgefasst nur in der Sprache der Formen. Erst Dirichlet wagt zum ersten Mal den Schritt, Sätze über quadratische Fomren ganz in der Sprache der Gitter zu beweisen. Im ersten Kapitel dieser Arbeit beschäftigen wir uns eingehend mit dem Konzept des Gitters und erläutern alle für den Verlauf der Arbeit notwendigen Eigenschaften.

Zeitgleich zu Dirichlet zeigte Hermite, dass es für eine Form $f(x)$ eine von der Determinanten und der Dimension abhängige Konstante γ gibt, so dass die Ungleichung $f(x) \leq \gamma$ wenigstens eine von 0 verschiedene Lösung besitzt. Gut ein halbes Jahrhundert später deutete Minkowski diese Ungleichung als Ellipsoid und die Punkte x mit ganzzahligen Koordinaten als Punkte eines Gitters in einem euklidischen Vektorraum E und konnte somit einen einfacheren Beweis und eine Verschärfung der Ausssage Hermites angeben.

Die entscheidende Leistung Minkowskis ist nun darin zu sehen, dass er diejenige Eigenschaft des Ellipsoids herausgearbeitet hat, die in seinem Beweisverfahren die maßgebliche Rolle gespielt hat: Das Elliposid ist ein konvexer Körper. In Kapitel 2 beschäftigen wir uns daher eingehend mit dem Begriff des Konvexen. Außerdem führen wir kurz in die sogenannte lineare Programmierung ein und zeigen mit dem Simplexverfahren eine Möglichkeit, solche linearen Programme zu lösen.

Die Arbeit Hermites fußt in der sogenannten Reduktionstheorie quadratischer Formen, also der Frage, wie aus den Äquivalenzklassen quadratischer Formen möglichst eindeutige Repräsentanten gewählt

werden können. Diese Arbeit war von Lagrange begonnen worden und von Seeber und Gauss auf Dimension 3 erweitert worden. Gauß zeigte hierbei eine Abschätzung der Determinanten und Hermite verallgemeinerte diese Abschätzung auf alle Dimensionen, indem er zeigte, dass sich der Quotient aus arithmetischem Minimum und n-ter Wurzel der Determinante stehts nach oben begrenzen läßt.

Minkowski erkannte nun einen fundamentalen Zusammenhang dieser Fragestellung mit einer Frage, die von Johannes Kepler 1616 bereits behandelt worden war: Wie lassen sich Kugeln möglichst optimal packen. Zwar hatte Kepler in seiner Abhandlung „Strena, Seu de Nive Sexangula“ eher an die Lagerung von Kanonenkugeln gedacht, doch Minkowski sah, dass die Ungleichung von Hermite, wenn man sie in die Sprache der Gitter übersetzt, genau jener Frage der maximalen Dichte entspricht. Hieraus konnte er dann eine bessere Abschätzung herleiten. Die Geometrie der Zahlen war nun geboren, über die Hermite in einem Brief an Laugel schreibt „Je crois voir la terre promise.“ In Kapitel 3 beschäftigen wir uns daher kurz mit der Reduktionstheorie und zeigen dabei, wie Minkowski hierbei seine geometrische Deutung einbringen konnte.

Die Frage nach der Bestimmung der von Hermite eingeführten Konstante war in den 1870er Jahren von den Russen Korkine und Zolotareff weitergeführt worden. Von ihnen stammt der Begriff *extreme Form*, oder *extremes Gitter*. Darunter verstehen Korkine und Zolotareff eine Form F, deren Hermite-Zahl bei leichten Veränderungen der Koeffizienten der Form nur abnimmt. Betrachtet man die Dichtefunktion auf dem Raum aller Gitter, so kann man auch sagen, dass extreme Gitter genau diejenigen sind, die ein lokales Maximum der Dichtefunktion liefern.

Aufbauend auf den Arbeiten von Korkine und Zolotareff fand der ukrainische Mathematiker Voronoi, dass sich solche extremen Formen durch zwei Eigenschaften ihrer Minimalvektoren charakterisieren lassen. Der ersten Eigenschaft gab er den Namen *Perfektion*, die andere bekam 50 Jahre später von Coxeter den Namen *Eutaxie*. In Kapitel 4 führen wir diese Begriffe ein und zeigen die Eigenschaften extremer Gitter. Als Abschluss dieses Kapitels zeigen wir das Theorem von Voronoi, dass besagt, dass extreme Gitter genau diejenigen sind, die perfekt und eutaktisch sind. Für den Beweis spielt die Theorie der konvexen Kegel eine wichtige Rolle; auch diese läßt sich auf Minkowski und auf Weyl zurückführen.

Das Hauptergebnis dieser Diplomarbeit ist eine neue Methode, um ein Gitter auf Eutaxie zu testen. Diese wird in Kapitel 5 vorgestellt und auf die perfekten Gitter in Dimension 8 angewandt. Durch diesen neuen Algorithmus wird es möglich, aus den 10916 perfekten Gittern diejenigen zu finden, die lokale Maxima der Dichtefunktion sind und den Hauptsatz (Satz 5.1) dieser Arbeit zu beweisen: In Dimension 8 gibt es genau 2408 Ähnlichkeitsklassen extremer Gitter.

Während bis einschließlich Dimension 5 alle perfekten Gitter auch extrem sind tritt in Dimension 6 das erste Gegenbeispiel auf. Dieses Gegenbeispiel scheint bereits Voronoi gekannt zu haben scheint, es wurde allerdings erst 1957 durch Barnes veröffentlicht. In Dimenesion 7 sind nur 3 der 33 perfekten Gitter nicht extrem. In Dimension 8 wendet sich das Blatt: Nur noch 2408 der 10916 perfekten Gitter sind auch extrem. Die Resultate in Dimension 8 bekräftigen somit eine Vermutung, die Martinet in seinem Buch ausgesprochen hat. Er vermutet dort, dass in hohen Dimensionen fast alle perfekten Gitter nicht mehr extrem sind. Die Resultat der Untersuchung werden in [22] erscheinen. Martinets Vermutung wird auch durch weitergehende Untersuchungen in Dimension 9 noch mehr unterstüzt.

1 Gitter

Die positive binäre Form $axx + 2bxy + cyy$ stellt allgemein das Quadrat der Entfernung zweier unbestimmten Puncte in einer Ebene vor, deren Coordinaten in Beziehung auf zwei unter einem Winkel, dessen Cosinus= $\frac{b}{\sqrt{ac}}$ ist, gegeneinander geneigte Axen um $x\sqrt{a}$, $y\sqrt{c}$ verschieden sind. Insofern x und y also nur ganze Zahlen bedeuten sollen,[..] erscheint die Ebene in lauter gleiche Parallelogramme getheilt, deren Endpuncte das Punctesystem ausmachen.

Recensionder Untersuchungen über die Eigenschaften der positiven tenären Formen von Lugwig August Seeber
C.F. Gauss

1.1 Definitionen

Im Folgenden bezeichnen wir mit E einen Euklidischen Vektorraum. Soweit nichts anderes vermerkt ist denken wir uns dabei den $\mathbb{R}^n$ mit dem Skalarprodukt:

$$\begin{array}{rccc} (\,\cdot\,) : & \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & \sum_{i=1}^{n} x_i y_i \end{array}$$

Die euklidische Norm wird durch $\|x\| = \sqrt{(x,x)}$ gegeben. Statt $\|x\|$ betrachten wir der Einfachheit halber oft nur $N(x) := ||x||^2$. Mit $e_1, ..., e_n$ bezeichnen wir eine beliebige orthonormale Basis von E.

Zu beliebigen Vektoren $b_1, b_2, ..., b_m$ sei

$$< b_1, ..., b_m >_{\mathbb{R}} = \sum_{i=1}^{m} b_i \mathbb{R}$$

der von den Vektoren $b_1, ..., b_m$ aufgespannte lineare Raum und

$$< b_1, b_2, ..., b_m >_{\mathbb{R}}^{\perp} := \{y \in \mathbb{R}^n \mid (y, b_i) = 0 \; \textit{für} \; i = 1, 2..., m\}$$

das orthogonale Komplement in $\mathbb{R}^n$.

Ziel dieses ersten Kapitels soll es sein, den zentralen Begriff des Gitters aus verschiedenen Blickwinkeln darzustellen. Wir beginnen mit einer ersten Definition:

Definition 1.1. Sei $\mathcal{B} := \{b_1, \ldots, b_m\}$ eine Menge linear unabhängiger Vektoren von $\mathbb{R}^n$. Ein Gitter $\mathbb{R}^n$ ist eine Untergruppe der Form

$$\Lambda = \Lambda((B)) =< b_1, ..., b_m >_{\mathbb{Z}} .$$

Das m-Tupel $(b_1, ..., b_m)$ heißt eine Basis von Λ. Ist eine Basis eines Gitters auch eine Basis von E, also wenn $n = m$ ist, so sprechen wir von einem vollen Gitter.

Soweit wir das nicht anders vermerken, werden wir mit dem Begriff des Gitters im Folgenden immer ein volles Gitter meinen.

Eine gegebene Basis definiert ein Gitter eindeutig, jedoch ist die Zuordnung von Basis auf Gitter nicht bijektiv. Auch wenn -anders als bei einem Vektorraum- nicht jede Menge von n linear unabhänigen Vektoren eine Basis bildet, besitzt ein Gitter mehrere mögliche Basen.

Für ein $n \in \mathbb{N}$, bezeichnen wir mit $\mathrm{GL}_n(\mathbb{Z}) := \{A \in \mathbb{Z}^{n\times n} \mid \ detA = \pm 1\}$ die Gruppe der ganzzahligen unimodularen Transformationen. Die Zuordnung von Gitter zu Basis ist nun eindeutig modulo $\mathrm{GL}_n(\mathbb{Z})$, mit anderen Worten es gilt folgende Proposition.

Proposition 1.2. *Zwei Basen $\mathcal{A}$ und $\mathcal{B}$ erzeugen das selbe Gitter, genau dann wenn es ein $U \in GL_n(\mathbb{Z})$ gibt, so dass $\mathcal{A} = U\mathcal{B}$.*

Beweis. Sei Λ das durch $\mathcal{A}$ erzeugte Gitter. Damit die Elemente von $\mathcal{B}$ auch in Λ liegen, muss es eine Matrix $U \in \mathbb{Z}^{n\times n}$ geben, so dass $\mathcal{B} = U\mathcal{A}$. Damit $\mathcal{B}$ das selbe Gitter erzeugt, muss diese Gleichung auch nach $\mathcal{A}$ auflösbar sein. Somit muss $detU = \pm 1$ und somit $U \in \mathrm{GL}_n(\mathbb{Z})$. □

Den großen Vorteil, den die Betrachtung von Gittern für die im späteren Verlauf der Arbeit zu diskutierende Problemstellung liefert, ist in einer gewissen geometrischen Anschauung begründet. Um diese Anschauung zu erläutern betrachten wir für eine gegebene Basis $b_1, ..., b_n$ das von den Vektoren aufgespannte Parallelogramm

$$P := P(b_1, ..., b_m) = \left\{ \sum_{i=1}^{n} a_i b_i, \, | a_i \in \mathbb{R}, 0 \leq a_i \leq 1 \right\}.$$

Das Volumen dieser Grundmasche wird durch $|det(b_1, ..., b_n)|$ gegeben und auch als *Diskriminante* des Gitters bezeichnet. In Proposition 1.2 hatten wir gesehen, dass zwei Basen ein und des selben Gitters durch ganzzahlige unimodulare Transformationen ineinander gebracht werden können. Dabei ändert sich das Volumen der Grundmasche nicht. Diese ist also eine geometrisch fassbare Invariante. Wir notieren dies in nachfolgendem Korollar:

Korollar 1.3. *Sei Λ ein Gitter. Das Volumen der Grundmasche ist unabhängig von der Wahl der Basis.*

Ein einfaches Beispiel für ein Gitter erhalten wir, wenn wir $\mathbb{Z}^n$ betrachten. Mögliche Basen sind in diesem Fall n linear orthogonale Einheitsvektoren. Das Volumen der Grundmasche ist eins.

Dieses Beispiel zeigt auch, dass die Struktur als freier $\mathbb{Z}$-Modul nicht ausreichend ist, um ein Gitter sinnvoll zu definieren. Alle freien $\mathbb{Z}$-Moduln vom Rang n sind isomorph zu $\mathbb{Z}^n$. Die Unterscheidung verschiedener Gitter wird also erst durch die Einbettung in einen euklidischen Vektorraum möglich.

Es ist also besser, ein Gitter nicht durch Angabe einer Basis sondern durch die Skalarprodukte der Basisvektoren anzugeben. Wir definieren hierzu die sogenannte Gram-Matrix.

Definition 1.4 (Gram-Matrix). Sei $\mathcal{B} = (b_1, .., b_n)$ eine Basis des Gitters Λ. Die Gram-Matrix von Λ wird dann definiert durch:

$$A := ((b_i, b_j))_{i,j=1,\ldots,n}$$

Zusätzlich wird $\det \Lambda := \det A$ als die *Determinante* von Λ bezeichnet.

Betrachten wir die darstellende Matrix $\mathcal{C}$ einer Gitterbasis $\mathcal{B}$ in einer Orthonormalbasis, so ist $A = \mathcal{C}^t\mathcal{C}$. Dann haben wir

$$vol(P) = det\mathcal{C} = \sqrt{detA} = \sqrt{det\Lambda}.$$

Nach Korollar 1.3 ist $\det(\Lambda)$ somit ebenfalls unabhängig von der Wahl einer Basis.

Unterziehen wir ein Gitter einer Drehung oder Spiegelung, so ist das resultierende Gitter der Anschauung nach das selbe. Auch das Volumen der Grundmasche bleibt bei solchen Transformationen erhalten. Transformationen, des $\mathbb{R}^n$, die das Skalarprodukt invariant lassen, werden *Isometrien* genannt.

Definition 1.5. Eine Abbildung $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ heißt Isometrie bezüglich des Skalarprodukts $(\ ,\ )$, wenn

$$(f(x), f(y)) = (x, y) \text{ für alle } x, y \in \mathbb{R}^n.$$

Die Gruppe der Isometrien wird mit $\mathrm{O}(\mathbb{R}^n)$ bezeichnet.

Gitter die auseinander durch lineare Isometrien entstehen wollen wir als äquivalent betrachten. Auch Gitter, die durch Zentrische Streckung ineinander überführt werden können, wollen wir unter dem Begriff der *Ähnlichkeit* zu einer Klasse zusammenfassen.

Definition 1.6. Seien $\Lambda, \Lambda' \subset \mathrm{E}$ zwei Gitter.

1. Gilt $\Lambda' = f(\Lambda)$ für ein $f \in \mathrm{O}(\mathbb{R}^n)$, so sagen wir Λ und Λ' sind *isometrisch* und schreiben $\Lambda \simeq \Lambda'$.
2. Die Gruppe $\mathrm{Aut}(\Lambda) := \{\sigma \in \mathrm{O}(\mathbb{R}^n) : \sigma(\Lambda) = \Lambda\}$ wird als Automorphismen-Gruppe bezeichnet.
3. Wir sagen Λ, Λ' sind ähnlich, wenn sie isometrisch sind, nachdem sie durch zentrische Streckung auf das selbe Minimum (s. Def. 1.10) gebracht wurden.

Bemerkung 1.7. Da $\mathrm{O}(\mathbb{R}^n)$ eine kompakte Gruppe ist, ist die Automorphismengruppe eines Gitters eine endliche Gruppe.

1.2 Minimum und kürzeste Vektoren

Von besonderem Interesse werden die Vektoren in einem Gitter sein, deren Länge minimal unter allen anderen Gittervektoren ist. Anschaulich ist klar, dass es solche immer gibt, doch wir gehen diesem Sachverhalt genauer auf den Grund.

Definition 1.8. Sei $G \subseteq \mathbb{R}^n$. G heißt dann diskrete Untergruppe, wenn alle $x \in G$ isolierte Punkte von $\mathbb{R}^n$ sind. Das bedeutet, dass es für jedes $x \in G$ eine Umgebung $U_x \subset \mathbb{R}^n$ mit

$$U_x \cap G = \{x\}.$$

Lemma 1.9. *Sei $\Lambda \subset \mathbb{R}^n$ ein Gitter. Dann ist Λ eine diskrete Untergruppe*

Beweis. Sei $b_1, ..., b_m$ eine Basis von Λ. Wir betrachten die Abbildung

$$\begin{array}{rccc} \varphi: & \mathbb{R}^m & \longrightarrow & < b_1, .., b_m >_{\mathbb{R}} \\ & t_1, ..., t_m & \longmapsto & \sum_{i=1}^m t_i b_i \end{array}$$

φ ist ein Isomorphismus und $\varphi(\mathbb{Z}^m) = \Lambda$. Da φ^{-1} stetig und $\mathbb{Z}^n$ diskret ist, folgt, dass Λ auch diskret ist. □

Da ein Gitter also diskret ist, besitzt es insbesondere kürzeste Vektoren. Wir werden später noch genauer darlegen, welche besondere Rolle diese Vektoren spielen. Doch zunächst definieren wir:

Definition 1.10 (Minimum und kürzeste Vektoren). Als das Minimum eines Gitters definieren wir

$$\mathrm{N}(\Lambda) := \min\{\mathrm{N}(x), x \in \Lambda,\, x \neq 0\}$$

Die Vektoren, die dieses Minimum realisieren werden als *kürzeste Vektoren* bezeichnet. Mit $S(\Lambda)$ notieren wir diese kürzesten Vektoren. Diese treten immer in Paaren $(x, -x)$ auf und wir schreiben s für die Anzahl der Paare.
Ein Gitter, das n linear unabhängige Minimalvektoren besitzt, wird *well rounded* genannt.

Wir betrachten nun $\mathbb{Z} + \mathbb{Z}\sqrt{2}$. Dieser $\mathbb{Z}$-Modul ist kein Gitter, denn $1, \sqrt{2}$ sind nicht linear unabhänig über $\mathbb{R}$. Stellt man sich dieses Beispiel als Punktmenge der euklidischen Ebene vor, fällt auch gleich auf, dass dieser Modul nicht diskret ist. Die Diskretheit charakterisiert Gitter also hinreichend, wie wir noch zeigen werden.

In Ebilden je n linearunabhängige Vektoren eine Basis von E. Für ein Gitter ist die lineare Unabhängigkeit alleine nicht ausreichend. Wir definieren daher, zuerst, was man unter einem primitiven Vektor versteht.

Definition 1.11. Sei $\Lambda \subset \mathrm{E}$ ein Gitter. Eine Familie $b_1, \ldots b_i$ von linear unabhänigen Vektoren in Λ heißt *primitives System*, wenn $< b_1, \ldots, b_i >_{\mathbb{R}} \cap \Lambda = < b_1, \ldots, b_i >_{\mathbb{Z}}$ gilt.

Zuerst betrachten wir primitive Systeme, die aus einem Vektor bestehen. Solche *primitiven Vektoren* sind genau diejenigen, die wir zu einer Basis ergänzen können.

Satz 1.12. *Ein $x \in \Lambda$ läßt sich genau dann zu einer Basis von Λ ergänzen, wenn x ein primitiver Vektor ist.*

Beweis. Ist $b_1, ..., b_{n-1}, x$ eine Basis von Λ folgt notwendigerweise, dass x primitiv ist, denn sonst gäbe es ein $x^{'} \in x\ \mathbb{R} \cap \Lambda$, dass sich nicht ganzzahlig in dieser Basis darstellen läßt. Das widerspricht aber den Annahmen.

Ist andererseits x ein primitiver Vektor und ist $b_1, .., b_n$ eine Basis von Λ, so dass x eine Darstellung $x = \lambda_1 b_1 + \ldots + \lambda_n b_n$ mit nicht negativen λ_j besitzt. Diese Darstellung kann so umgeordnet werden, dass $\lambda_j \leq \lambda j + 1$ für alle Koeffizienten λ_j gilt. Es sei nun λ_k der erste positive Koeffizient in dieser Darstellung und wir können die Darstellung für x wie folgt umändern:

$$\begin{aligned} x &= \lambda_k(b_k + qb_{k+1}) + (\lambda_{k+1} - q\lambda_k)b_{k+1} + \lambda_{k+2}b_{k+2} + \ldots + \lambda_n b_n \\ &= \lambda_k b_k^{'} + \lambda_{k+1}^{'} b_{k+1} + \lambda_{k+2}b_{k+2} + \ldots + \lambda_n b_n, \end{aligned}$$

wobei $b_k^{'} := (b_k + qb_{k+1})$ und $\lambda_{k+1}^{'} := \lambda_{k+1} - q\lambda_k$.

Nun können wir q so wählen, dass $0 \leq \lambda_{k+1}^{'} < \lambda_k$ gilt, und in selber Weise mit den anderen Koeffizienten verfahren. Die auf diese Weise erhaltene Darstellung für x ordnen wir wieder wie zu Beginn der Größe nach und weden hierauf wieder das eben beschriebene Verfahren an. Nach endliche vielen Schritten erhalten wir auf diese Weise eine Basis $\tilde{b}_1, \ldots, \tilde{b}_n$ für die $x = \mu\tilde{b}_1$ mit $\mu > 0$ gilt. Nun ist x aber nach Voraussetzung ein primitiver Vektor, und somit muss $\mu = 1$ gelten. Also ist $x = \tilde{b}_1$ und läßt sich daher zu einer Basis ergänzen. □

Dieser Satz läßt sich nun auf eine Familie von primitiven Vektoren verallgemeinern.

Satz 1.13. *Eine Familie $x_1, ..., x_k$ von primitiven Vektoren in einem Gitter Λ läßt sich zu einer Basis von Λ ergänzen.*

Beweis. Der Beweis erfolgt in einer Induktion nach k. Der Induktionsanfang mit $k = 1$ folgt nach Satz 1.12. Sei nun $x_1, \ldots, x_{k+1}$ eine Familie von primitiven Vektoren und wir können nach Induktionsvoraussetzung eine Basis $b_1, \ldots, b_n$ von Λ so wählen, dass $b_1 = x_1, \ldots, b_k = x_k$ gilt. Für x_{k+1} finden wir nun eine Darstellung der Form:

$$x_{k+1} = \lambda_1 b_1 + \ldots \lambda_k b_k + \lambda_{k+1} b_{k+1} + \ldots + \lambda_n b_n.$$

Die letzten Koeffizienten $\lambda_{k+1}, \ldots \lambda_n$ sind ohne Einschränkung nicht negativ und mindestens einer von ihnen ist positiv. Wendet man nun das im Beweise von Satz 1.12 eingeführte Verfahren auf diese Koeffizienten an, ändert sich obige Darstellung zu:

$$x_{k+1} = \lambda_1 b_1 + \ldots \lambda_k b_k + \mu\tilde{b}_{k+1}.$$

Da die Vektoren $x_1, ..., x_{k+1}$ nach Voraussetzung primitiv sind, folgt $\mu = 1$ und $\tilde{b}_{k+1}$ kann als Basisvektor durch x_k ersetzt werden. □

Nun zeigen wir den schon angekündigten Satz über diskrete Untergruppen.

Satz 1.14. *Jede additive diskrete Untergruppe $G \subset \mathrm{E}$, ist ein Gitter.*

Beweis. Sei G eine solche Gruppe und m die maximale Anzahl der linear unabhänigen Vektoren in E. Nach Voraussetzung ist G diskret. Somit findet sich ein der Länge nach kleinstes $b_1 \neq 0$. Somit gilt:

$$\lambda b_1 \notin G, \forall\, 0 \leq \lambda \leq 1.$$

Dieses b_1 dient nun als Ausgang für eine induktiv zu konstruierenden Gitterbasis:

Angenommen es wurden bereits eine Menge $\mathcal{B}$ linear unabhäniger $b_1, \ldots b_i\, (0 < i < m)$ gefunden, so dass

$$P(b_1, ..., b_i) \cap G = \{0\}.$$

Wir setzen $\Lambda_i := \Lambda(b_1, \ldots b_i)$ und wählen nun $b_{i+1} \in G \backslash \Lambda_i$ so, dass $b_1, \ldots, b_i, b_{i+1}$ ein primitives System bilden. Das so gefundene b_{i+1} wird zu $\mathcal{B}$ hinzugefügt und dieser Vorgang wird fortgesetzt bis $i = m$ gilt. Nach Satz 1.13 bilden die Vektoren $b_1, \ldots, b_n$ nun eine Gitterbasis von G. □

Man kann also ein Gitter einfach auch als eine diskrete Untergruppe eines euklidischen Vektorraumes definieren.

1.3 Kugelpackungen

Unter einer Kugelpackung in einem euklidischen Vektorraum E verstehen wir eine Menge von gleich großen Kugeln

$$B_v(r) = \{x \in \mathrm{E} : \|x - v\| \leq r\} \text{ für ein } v \in \mathrm{E} \text{ und } r > 0,$$

die höchstens paarweise Randpunkte gemeinsam haben.

Im Folgenden werden wir Kugelpackungen genauer studieren, bei denen die Kugeln die Punkte eines Gitters Λ zum Mittelpunkt haben. Es ist klar, dass eine solche Kugelpackung nur existieren kann, wenn der Kugelradius $r \leq \sqrt{N(\Lambda)}$ gewählt ist. Wir bezeichnen in diesem Fall auch die Vereinigung aller solcher Kugeln als Kugelpackung zum Gitter Λ. Also

$$\mathcal{K}_\Lambda := \bigcup_{v \in \Lambda} B_v(\sqrt{N(\Lambda)}).$$

Bei der Untersuchung von Kugelpackungen sind nun zwei Fragestellungen von besonderem Interesse, die beim ersten Anblick fast schon trivial anmuten (zumindestens für den Anschauungsraum, ist wohl jeder Mensch in der Lage die Antworten zu finden, wenngleich auch meist nur der Intuition folgend.)

Zum Einen die Frage, wieviele Kugeln gleichzeitig eine Kugel berühren können und zum Anderen die Frage nach einer optimalen Packung von unendlich (oder auch endlich) vielen Kugeln.

Der Ursprung des ersten Problems findet sich im sogenannten Gregory-Newton Disput: Im Jahre 1694 stellten die beiden Gelehrten Isaac Newton und David Gregory die Frage, wieviele Kugeln gleichzeitig im dreidimensionalen Raum eine mittlere Kugel berühren können. Für Newton war 12 die Obergrenze, Gregory hielt aber 13 für möglich. Newtons Position konnte erst 1956 durch Schütte und van der

Waerden abschließend bewiesen werden. Schränkt man die möglichen Kugelpackungen auf die Gitter-Kugelpackungen ein, so ist das Gregory-Newton-Problem gleichbedeutend damit zu fragen, wieviele Minimalvektoren ein Gitter maximal hat. Die Anzahl der Minimalvektoren wird deshalb auch als *Kissing Number* des Gitters bezeichnet, denn im Billiard wird der Begriff kiss für Kugel benutzt, die sich berühren.

Fragen nach der optimalen und platzsparenden Lagerung von Kugeln finden sich zum ersten Mal beim Astronom Johannes Kepler. In seiner Neujahrsschrift von 1611 beschrieb er eine Lagerung von Kugeln, von der er behauptete, dass sie am dichtesten sei. Auch wenn diese Frage wohl auch ganz einfach intuitiv mit ein paar Orangen beantwortbar scheint, dauerte es bis 1998, bis Keplers Vermutung bewiesen werden konnte. Einen schönen historischen Zusammenhang zu den Fragestellungen der Kugelpackungen liefert [23].

Wir interessieren uns nun aber nur für Kugelpackungen zu einem Gitter. Diese Gitterpackungen und ihre Bedeutung für die Arithmetik quadratischer Formen findet sich zum ersten Mal in Minkowskis *Geometrie der Zahlen.*

Um die Dichte einer Gitterkugelpackung nun vernünftig definieren zu können, zeigen wir zuerst folgendes Lemma von Minkowski.

Lemma 1.15. *Sei Λ ein Gitter und P die Grundmasche zu einer beliebigen Basis von Λ. Weiterhin sei $B \subset \mathrm{E}$ beschränkt und meßbar und es gelte*

$$(x + B) \cap (y + B) = \emptyset$$

für alle $x, y \in \Lambda$ mit $x \neq y$. Dann gilt:

$$\sum_{x \in \Lambda} \mathrm{vol}((x + B) \cap P) = \mathrm{vol}(B)$$

Beweis. Wir haben $\mathrm{E} = \cup_{x \in \Lambda}(x + P)$, wobei $(x + P) \cap (y + P)$ für alle $x \neq y \in \Lambda$ eine Nullmenge ist. Somit gilt $B = \cup_{x \in \Lambda} B \cap (x + P)$ mit nur endlich vielen $B \cap (x + P)$, die nicht leer sind. Es ist also

$$\mathrm{vol}\, B = \sum_{x \in \Lambda} \mathrm{vol}\, B \cap (x + P).$$

Nun ist $B \cap (x + P) = (B) \cap P)$ und es folgt

$$vol B = \sum_{x \in \Lambda} vol(x + B) \cap P.$$

□

Unter der Dichte einer Kugelpackung wollen wir anschaulich das Verhältnis des mit Kugeln überdeckten Raumes, zum gesamten Raum verstehen. Nach Lemma 1.15 beschreibt die Größe $\varrho(\Lambda)$ aus der folgenden Definition genau die Dichte von K_Λ.

Definition 1.16. Sei $r = \sqrt{N(\Lambda)}$, dann heißt die Zahl

$$\varrho(\Lambda) := \frac{volB_r(0)}{\sqrt{det(\Lambda)}}$$

die Dichte von Λ.

Ein Gitter, dessen Dichte bei kleinen Veränderungen nur abnimmt, wird *extremes Gitter* genannt. Ein n-dimensionales Gitter, das die maximale Packungsdichte aller n-dimensionalen Gitterpackungen realisiert, nennen wir *absolut extrem*. Wir werden diese Begriffe in Kapitel 3 noch einmal genau definieren.

Bemerkung 1.17. Ist ein Gitter Λ nicht well rounded, so ist anschaulich klar, dass eine Gitterpackung zu Λ keine maximale Packungsdichte realisieren kann.

1.4 Unter- und Duale Gitter

Ein Gitter kann andere Gitter enthalten und man spricht dann von einem Untergitter. Untergitter liefern im Allgemeinen wichtige Informationen über das Gitter, das sie enthält. Wir definieren:

Definition 1.18. Seien Λ und $\Lambda^{'}$ Gitter mit $\Lambda^{'} \subseteq \Lambda$, dann heißt $\Lambda^{'}$ *Untergitter* von Λ.

Als ein anschauliches Beispiel betrachten wir im Gitter $\mathbb{Z}^2$ die Gitterpunkte mit geraden Koordinaten. Diese bilden ein Untergitter $\Lambda^{'} \subset \mathbb{Z}^2$.

An diesem einfachen Beispiel kann man sich leicht die Aussage der nächsten Proposition klar machen.

Proposition 1.19. *Sein Λ und $\Lambda^{'}$ (volle) Gitter in* E *mit $\Lambda \subset \Lambda^{'}$. Dann gilt:*

$$det\Lambda = [\Lambda^{'} : \Lambda]^2 det\Lambda'.$$

Beweis. Seien P und $P^{'}$ die Grundmaschen zu Gitterbasen von Λ und $\Lambda^{'}$. Jeder Punkt $v \in$ E läßt sich eindeutig darstellen als $v = x + p$, wobei $p \in P$ und $x \in \Lambda$. Das liefert uns eine Bijektion von P mit E/Λ. Aus der Inklusion in $\Lambda^{'}$ erhalten wir die Bijektion

$$\Lambda^{'} \cap P \leftrightarrow \Lambda^{'}/\Lambda.$$

Da $\Lambda^{'}$ diskret ist, ist $(\Lambda^{'} \cap P)$ eine endliche Menge, und somit ist $k := [\Lambda^{'} : \Lambda] < \infty$.

Sei $x_1, ..., x_k$ ein Repräsentantensystem von $\Lambda^{'}/\Lambda$, also gilt:

$$\Lambda^{'} = \cup_{1 \leq i \leq k} \Lambda + x_i.$$

Hieraus folgt:

$$\mathrm{E} = \cup_{x^{'} \in \Lambda^{'}} x^{'} + P^{'} = \cup_{x \in \Lambda} \cup_{1 \leq \ i \leq \ k} x + x_i + P^{'} = \cup_{x \in \Lambda} x + B,$$

wobei $B = \cup_{1 \leq i \leq k} x_i + P^{'}$. Mit Lemma 1.15 folgt nun

$$\text{vol}\, P = \text{vol} \cup_{x \in \Lambda} (x + B) \cap P = volB = k\ volP^{'}.$$

Insgesamt haben wir also $det\Lambda = vol(P)^2 = k^2(volP^{'})^2 = k^2 det\Lambda^{'}$. □

Im letzten Satz ist natürlich wichtig zu beachten, dass Λ und $\Lambda^{'}$ beides volle Gitter sind.

Manchmal ist aber auch der Fall interessant, wenn ein Untergitter in einem Unterraum F von E liegt.

Sei Λ ein Gitter und φ eine lineare Abbildung. Nicht in jedem Fall ist $\varphi(\Lambda)$ ein Untergitter, wie wir an folgendem Beispiel sehen:

$$\begin{array}{rccl} \varphi : & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & x_1 + \sqrt{2}x_2 \end{array}$$

Eingangs hatten wir nämlich schon darauf hingewiesen, dass $\varphi(\mathbb{Z}^2)$ kein Gitter ist.

Sei nun $\mathrm{F} \subset \mathrm{E}$ ein linearer Unterraum von E. Wir können E als orthogonale Summe schreiben in der Form

$$\mathrm{E} = \mathrm{F} \oplus F^{\perp}$$

Das heißt jedes $x \in \mathrm{E}$ läßt sich eindeutig schreiben als $x = x_1 + x_2$ mit $x_1 \in F$ und $x_2 \in F^{\perp}$.

Definition 1.20. Mit den obigen Bezeichnungen nennen wir

$$\begin{array}{rccl} \varphi : & \mathrm{E} & \longrightarrow & \mathrm{F} \\ & x & \longmapsto & \varphi(x) = x_1 \end{array}$$

die orthogonale Projektion von E nach F

Die Bilder solcher orthogonalen Projektionen sind nun Untergitter.

Proposition 1.21. *Sei $\Lambda \subset \mathrm{E}$ ein volles Gitter und seien $b_1, ..., b_m$ linear unabhängige Vektoren in Λ. Dann ist $\Lambda' := \pi_{<b_1,...,b_m>_{\mathbb{R}}^{\perp}}(\Lambda)$ ein $(n-m)$-dimensionales Gitter*

Beweis. Man sieht, dass $\Lambda^{'}$ eine Untergruppe von E der Dimension $(n-m)$ ist. Aus Satz 1.14 wissen wir, dass diskrete Untergruppen und Gitter das selbe sind. Nehmen wir also an, $\Lambda^{'}$ besitz einen Häufungspunkt $x \in \mathrm{E}$. Wir wählen nun eine Folge von $v_i \in \Lambda$, so dass die Folge $(\tilde{v}_i)_{i \in \mathbb{N}} := \pi_{<b_1,...,b_m>_{\mathbb{R}}^{\perp}}(v_i)$ aus paarweise verschiedenen Elementen besteht und gegen x konvergiert. Dabei gilt

$$\pi_{<b_1,...,b_m>_{\mathbb{R}}^{\perp}}(v_i) = v_i - \sum_{j=1}^{m} \frac{(v_i, b_j)}{(v_i, v_i)} b_j.$$

Wir betrachten nun die weitere Folge $w_i := v_i - \sum_{j=1}^{m} \lfloor \frac{(v_i,b_j)}{(v_i,v_i)} \rfloor$. Die Folge (w_i) liegt in Λ und wir haben für alle Folgeglieder die Abschätzung $\|w_i - \tilde{v}_i\| \leq \sum_{j=1}^{m} \|b_j\|$. Die Folge $(w_i - \tilde{v}_i)_{i \in \mathbb{N}}$ ist also beschränkt und besitzt nach dem Satz von *Bolzano und Weierstraß* eine konvergente Teilfolge. Somit besitzt auch die Foge w_i einen Häufungspunkt. Das widerspricht aber der Diskretheit von Λ. Somit kann $\Lambda^{'}$ keinen Häufungspunkt besitzen und ist nach Satz 1.14 ein Gitter. □

Ein wichtiges mathematisches Konzept, das wir im Laufe der Arbeit in verschiedenen Situationen antreffen werden, ist die *Dualität*:

Sei R ein kommutativer Ring und M ein endlicherzeugter $R-Modul$, dann bezeichnen wir den Modul der R-Homomorphismen von M nach R als den dualen Modul M^* zu M. Besitzen wir auf M eine nicht entartete Bilinearform B, so können wir M eindeutig mit seinem dualen Modul identifizieren.

Zu einem Gitter Λ definiert man das duale Gitter wie folgt:

Definition 1.22. Sei $\Lambda \subset \mathrm{E}$ ein Gitter. Das duale Gitter Λ^* ist definiert durch:

$$\Lambda^* := \{y \in \mathrm{E} : (x,y) \in \mathbb{Z} \text{ für alle } x \in \Lambda\}$$

Weiterhin sagen wir dass:

1. Λ ein ganzes Gitter ist, wenn $(x,y) \in \mathbb{Z}$ für alle $x,y \in \Lambda$ gilt.
2. Λ ein gerades Gitter ist, wenn $(x,y) \in 2\mathbb{Z}$ für alle $x,y \in \Lambda$ gilt.
3. Λ ein unimodulares Gitter ist, wenn Λ ganz ist mit $det\Lambda = 1$.

Aus der sogenannten Polarisationsgleichung für das Skalarprodukt

$$(x,y) = \frac{1}{2}[((x+y),(x+y)) - (x,x) - (y,y)]$$

leitet man sofort ab, dass ein gerades Gitter stets auch ganz ist. Darüberhinaus haben wir die folgenden Eigenschaften:

Proposition 1.23. *Sei Λ ein Gitter, dann gilt:*

1. $det\Lambda^* = (det\Lambda)^{-1}$
2. $u(\Lambda)^* = (u^t)^{-1}(\Lambda^*)$ *für* $u \in \mathrm{GL(E)}$
3. $\Lambda^{**} = \Lambda$
4. *Für* $\Lambda' \subset \Lambda$ *gilt* $\Lambda^* \subset \Lambda'^*$.
5. *Genau dann ist* Λ *ein ganzes Gitter, wenn* $\Lambda \subset \Lambda^*$ *gilt.*
6. *Genau dann ist* Λ *unimodular, wenn* $\Lambda = \Lambda^*$ *gilt.*

Beweis. Sei $b_1, ..., b_n$ eine Basis von Λ. Wir finden dann $b_1^*, ..., b_n^*$ die dazugehörende Dualbasis von E. Diese ist eine Basis von Λ^*. Bezüglich $e_1, ..., e_n$ erhalten wir dann

$$b_i = \sum_{j=1}^{n} a_{ij} e_j \ , \ b_i^* = \sum_{j=1}^{n} c_{ij} e_j .$$

Wir erhalten somit zwei Matrizen $A = (a_{ij})$ und $C = (c_{ij})$.
Es gilt nun $b_j b_k^* = \sum_{i=1}^{n} a_{ij} c_{ik} = \delta_{jk}$ Hieraus folgt $A^t = C^{-1}$. Also gilt $det\Lambda^* = (detC)^2 = (detA)^{-2} = (det\Lambda)^{-2}$. Hieraus folgen leicht die anderen Behauptungen. □

1.5 Quadratische Formen und Gitter

Wie wir eingangs erwähnt hatten, entspringt das Interesse für Gitter dem Studium der quadratischen Formen. Wir legen nun dar, worin der Zusammenhang zwischen Gittern und quadratischen Formen besteht.

Sei $\Lambda \subset \mathrm{E}$ ein Gitter, das von der Basis $(b_1, ..., b_n)$ aufgespannt wird. Hierzu definieren wir folgende quadratische Form.

$$Q_\Lambda(x_1, ..., x_n) := \sum_{1 \leq i,j \leq n} (b_i, b_j)\, x_i x_j = \| \sum_{i=1}^{n} x_i b_i \|^2$$

Diese so dem Gitter zugeordnete Form ist positiv definit, und den Werten dieser Form auf $\mathbb{Z}^n$ entsprechen genau den Längenquadraten der Vektoren im Gitter Λ.

Betrachten wir andererseits eine beliebige positiv definite quadratische Form

$$Q^{'} = \sum_{1 \leq 1, j \leq n} q_{ij} x_i x_j.$$

Wir können dann linear unabhängig Vektoren $b_1, ..., b_n$ finden mit $(b_i, b_j) = q_{ij}$ für alle $1 \leq i, j \leq n$.

Das von diesen Vektoren aufgespannte Gitter $\Lambda_{Q'}$ steht nun zur Quadratischen Form $Q^{'}$ in der selben Beziehung wie das obige Paar. Zwei Quadratische Formen Q und $Q^{'}$ werden als arithmetisch äquivalent bezeichnet, wenn es eine ganzzahlige unimodulare Transformation $A \in GL_n(\mathbb{Z})$ gibt, so daß für alle $x \in \mathbb{Z}^n$ gilt: $Q(x) = Q^{'}(Ax)$. Dies ist also genau dann der Fall, wenn wir Q und $Q^{'}$ das selbe Gitter zuordnen können.

Aus diesen Überlegungen erhalten wir eine kanonische Bijektion zwischen den Isometrieklassen von Gittern und den Äquivalenzklassen von positiv definiten Quadratischen Formen. Wir können also in Zukunft unbekümmert den Sprachgebrauch zwischen diesen beiden Sichtweisen wechseln.

1.6 Folgen von Gittern

Wir hatten im dritten Abschnitt die Dichte für Gitter definiert und auch erwähnt, dass wir am Maximum dieser Dichte interessiert sind. Zwar ist die Dichte nach oben natürlich durch eins beschränkt, ob aber wirklich immer ein Maximum für die Dichte existiert, ist damit nicht gesagt. Um sicher zu stellen, dass dies der Fall ist, müssen wir uns zuerst klar machen, dass die Menge der Gitter kompakt ist.

Mit $\mathcal{L}$ bezeichnen wir die Menge der vollen Gitter in E. Die Menge der Automorphismen von E wird mit GL(E) bezeichnet. Ist $\mathcal{B}$ eine Basis von E, so ist auch das Bild von $\mathcal{B}$ unter GL(E) wieder eine Basis von E. Somit ist dann auch das Bild $u(\Lambda)$ eines Gitters $\Lambda \in \mathcal{L}$ unter einem Element $u \in \mathrm{GL(E)}$ wieder ein Gitter.

Indem wir eine Basis für E wählen, können wir GL(E) mit der Gruppe aller invertierbaren $n \times n$ Matrizen über $\mathbb{R}$, für die wir $\mathrm{GL}_n(\mathbb{R})$ schreiben, identifizieren. Beginnend mit einem bestimmten Gitter Λ_0, erhalten wir nun alle Gitter in E als $\Lambda = u(\Lambda_0)$ für ein bestimmtes $u \in \mathrm{GL}_n(\mathbb{R})$. Weiterhin

haben wir $u(\Lambda_0) = v(\Lambda_0)$ genau dann, wenn $v^{-1}u$ in $\mathrm{GL}_n(\mathbb{Z})$ liegen. Daher können wir $\mathcal{L}$ mit der Menge der Linksnebenklassen in $\mathrm{GL}_n(\mathbb{R})$ modulo $\mathrm{GL}_n(\mathbb{Z})$ identifizieren.

Auf diese Weise erhalten wir eine Topologie auf $\mathcal{L}$; Die Umgebung eines Gitters $\Lambda = u(\Lambda_0)$ ist einfach die Menge der Bilder einer Umgebung der Einheitsmatrix I_n in $GL_n(\mathbb{R})$ unter u. Insbesondere können wir so auch den Abstand zweier Gitter definieren:

Definition 1.24. Seien $\Lambda_1, \Lambda_2 \in \mathcal{L}$ zwei Gitter und $T(\Lambda_1, \Lambda_2) := \{M \in \mathbb{R}^{n\times m} \mid \Lambda_1 = M\Lambda_2\}$, die Menge der Matrizen, die Λ_2 in Λ_1 überführen. Ferner sei

$$\sigma(\Lambda_1, \Lambda_2) := min\{|M - I_n| : M \in T(\Lambda_1, \Lambda_2)\}$$

Dann heißt

$$\delta(\Lambda_1, \Lambda_2) = max\{log(1 + \sigma(\Lambda_1, \Lambda_2)), log(1 + \sigma(\Lambda_2, \Lambda_1))\}$$

der Abstand von Λ_1 und Λ_2.

Diese Definition liefert uns eine Metrik auf $\mathcal{L}$ und es macht Sinn, von konvergenten Folgen von Gittern zu sprechen.

Eine Menge von Gittern kann in zweierlei Weisen entarten. Zum einen kann sie eine Folge von Gittern enthalten, die immer dünner wird. Andererseits kann sie eine Folge von Gittern enthalten, deren Basen immer kleinere Winkel bilden, so dass das Gitter schon fast nicht mehr voll ist. Um dies zu verhindern definieren wir was wir unter einer beschränkten Menge von Gittern verstehen wollen:

Definition 1.25. Sei $\mathcal{M}$ eine Menge von Gittern. $\mathcal{M}$ heiße beschränkt, wenn es $\sigma, \varsigma > 0$ gibt, so dass

1. $N(\Lambda) \geq \sigma$ für alle $\Lambda \in \mathcal{M}$.
2. $det(\Lambda) \leq \varsigma$ für alle $\Lambda \in \mathcal{M}$.

Folgender Satz, der so genannte Auswahlsatz von Mahler, ist ein Analogon zum bekannten Satz von Bolzano-Weierstrass:

Theorem 1.26. *Jede beschränkte Folge von Gittern besitzt eine konvergente Teilfolge*

Beweis. Zwei Beweise finden sich in [17], Kapitel 11, Abschnitt 17.5 Theorem 2 □

2 Konvexität

> Dieser Umstand führte Minkowski zum ersten Mal zu der Erkenntnis, daß überhaupt der Begriff des konvexen Körpers ein fundamentaler Begriff in unserer Wissenschaft ist und zu deren fruchtbarsten Forschungsmitteln gehört
>
> *Nachruf auf Hermann Minkowski*
> David Hilbert

Einer der zentralen Begriffe im mathematischen Werk von Hermann Minkowski, ist der Begriff des Konvexen. Ausgehend von den elementaren Eigenschaften konvexer Mengen, die wir im ersten Abschnitt erläutern werden, führen wir im zweiten Abschnitt konvexe Kegel ein. Im dritten Abschnitt erläutern wir knapp die Prinzipien der linearen Optimierung und des sogenannten Simplex-Algorithmus, mit dem sich lineare Optimierungsprobleme lösen lassen. Im letzten Abschnitt stellen wir noch kurz die konvexen Funktionen vor. Für einen breiteren Einstieg verweisen wir auf [27].

2.1 Konvexe Mengen

Zuerst werden wir den Begriff der konvexen Menge etwas genauer beleuchten. Wir definieren hierzu:

Definition 2.1. Eine Menge $\mathcal{K} \subset \mathrm{E}$ heißt konvex, wenn für $x, y \in \mathcal{K}\, x \neq y$ auch die Verbindungsstrecke

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

ganz in $\mathcal{K}$ liegt.

Ein n-dimensionaler Würfel ist beispielsweise stets konvex. Eine Hyperebene ist ein affiner Unterraum von E mit Codimension 1. Eine Hyperebene schreiben wir als

$$H_{x,\alpha} := \{y \in \mathrm{E} : \; (x, y) = \alpha\}$$

Der Vektor $0 \neq x \in \mathrm{E}$ wird dabei als der Normalenvektor von H bezeichnet. Jede Hyperebene definiert zwei abgeschlossene Halbräume

$$H^+_{x,\alpha} := \{y \in \mathrm{E} : \; (x, y) \geq \alpha\} \text{ und } H^-_{x,\alpha} := \{y \in \mathrm{E} : \; (x, y) \leq \alpha\}$$

Jeder dieser Halbräume ist offensichtlich konvex.
Bevor wir uns weitergehend mit den Eigenschaften konvexer Mengen beschäftigen wollen, machen wir in diesem ersten Abschnitt noch einige Definitionen.

Definition 2.2. Sei $x \in \mathrm{E}$. Wir sagen, dass x eine Konvexkombination der $x_1, ..., x_r \in \mathrm{E}$ ist, wenn wir eine Darstellung der folgenden Form haben:

$$x = \lambda_1 x_1 + ... + \lambda_r x_r \tag{1}$$
$$\lambda_1 + ... + \lambda_r = 1 \tag{2}$$
$$\lambda_1 \geq 0, ..., \lambda_r \geq 0 \tag{3}$$

Wir sagen, dass x eine positive oder konische Kombination, der $x_1, ..., x_r$ ist, wenn nur (1) und (3) erfüllt ist. Erfüllt x nur (1) und (2) sprechen wir von einer affinen Kombination.

Für eine beliebige Teilmenge $M \subset \mathrm{E}$ bezeichnen wir mit $conv(M)$ die sogenannte konvexe Hülle von M. Dies ist die Menge aller Konvexkombinationen aus endlich vielen Elementen von M. Analog definieren wir die die konische bzw. affine Hülle von M und schreiben hierfür $cone(M)$ bzw. $aff(M)$. Folgender Satz ist anschaulich klar und wir verzichten darauf den Beweis wiederzugeben, sondern verweisen auf [12] Theorem 1.5.

Proposition 2.3. *Für $M \subset \mathrm{E}$ ist $conv(M)$ stets eine konvexe Menge. Ist M bereits schon eine konvexe Menge, haben wir $M = conv(M)$*

Wir hatten die konvexe Hülle einer Menge M als Menge aller Konvexkombinationen von endlich vielen Punkten aus m definiert. Eine andere Definition liefert uns die folgende Proposition:

Proposition 2.4. *Sei $M \subset \mathrm{E}$ dann gilt:*

$$conv(M) = \bigcap_{\mathcal{K} \supset M, \mathcal{K}\ konvex} \mathcal{K}$$

Beweis. Sei

$$S := \bigcap_{\mathcal{K} \supset M, \mathcal{K}\ \text{konvex}} \mathcal{K}$$

Der Durchschnitt von konvexen Mengen ist konvex, wie man sich mit der Definition einer konvexen Menge ersichtlich machen kann. S ist also konvex und $M \subset S$ nach Konstruktion. Somit haben wir

$$conv(M) \subset conv(S) = S.$$

Da nun $conv(M)$ eine konvexe Menge ist mit $M \subset conv(M)$, folgt $S \subset conv(M)$. Hieraus folgt die Behauptung. □

Die konvexe Hülle ist also die kleinste konvexe Menge, die M enthält. Der nächste Satz geht auf Carathéodory zurück.

Theorem 2.5 (Carathéodory). *Für $M \subset \mathrm{E}$ gilt*

$$conv(M) = \{\sum_{i=1}^{n+1} \lambda_i x_i : \ x_i \in M, \ \lambda_i \geq 0, \ \sum_{i=1}^{n+1} \lambda_i = 1\}$$

Beweis. Sei $x \in conv(M)$. Somit finden sich $x_i \in M$ und $\lambda_i > 0$ mit $\sum_{i=1}^{m} \lambda_i = 1$ und $x = \sum_{i=1}^{m} \lambda_i x_i$. Ist $m \leq n+1$ ist nichts zu zeigen. Also sei nun $m \geq n+2$. Die Vektoren $x_m - x_1, ..., x_2 - x_1$ sind dann linear abhängig. Wir haben also eine Darstellung

$$\sum_{i=2}^{m} \alpha_i (x_i - x_1) = 0 \text{ wobei } \alpha_j \neq 0 \text{ für mindestens ein } 2 \leq j \leq m.$$

Wir setzten nun $\alpha_1 := -\sum_{i=2}^{m} \alpha_i$. Dann ist $\sum_{i=1}^{m} \alpha_i x_i = 0$ woraus wir folgern können, dass $x = \sum_{i=1}^{m} (\lambda_i - \lambda \alpha_i) x_i$ wobei wir λ aus $\mathbb{R}$ beliebig wählen können.

Wir wählen nun $\lambda^* := \min_i \{\frac{\lambda_i}{\alpha_i}\}$. Es gilt nun:

$$\lambda_i - \lambda^* \alpha_i \geq 0, \ \sum_{i=1}^{m} (\lambda_i - \lambda^* \alpha_i) = 1, \text{ und } \lambda_j - \lambda^* \alpha_j = 0 \text{ für mindestens ein } j.$$

Somit findet sich eine Kombination mit weniger strikt positiven Koeffizienten. Durch Iteration folgt hieraus die Behauptung. □

Als Korollar aus diesem Satz erhalten wir folgende Behauptung.

Korollar 2.6. *Ist $M \subset \mathrm{E}$ kompakt, so ist auch $conv(M)$ kompakt.*

Beweis. Wir betrachten die Abbildung $f : \mathbb{R}^{n+1} \times M^{n+1} \mapsto \mathbb{R}^n$, definiert durch

$$f(\lambda_1, ..., \lambda_{n+1}, x_1, ..., x_{n+1}) := \sum_{i=1}^{n+1} \lambda_i x_i.$$

Offensichtlich ist f stetig und mit dem Satz von Carathéodory gilt

$$f(K \times M^{n+1}) = conv(M)$$

wobei $K := \{(\lambda_1, ..., \lambda_{n+1}) \in \mathbb{R}^{n+1} \lambda_i \geq 0, \ \sum_{i=1}^{n+1} \lambda_i = 1\}$. Hieraus erhalten wir die Behauptung, denn mit K ist auch $K \times M^{n+1}$ kompakt. □

Wir möchten nun genau das relative Innere einer konvexen Menge betrachten. Wir definieren dazu:

Definition 2.7. Sei $M \subset \mathrm{E}$ eine konvexe Menge. Ein $x \in M$ heißt innerer Punkt von M, wenn es ein $\epsilon > 0$ gibt, so dass

$$B_\epsilon(x) \cap \mathrm{aff}(M) \neq \emptyset$$

gilt. Die Menge aller inneren Punkte bezeichnen wir mit $\mathrm{ri}\, M$. Die Menge $M \setminus \mathrm{ri}\, M$ wird der Rand von M genannt und mit $\mathrm{bd}\, M$ bezeichnet.

Bemerkung 2.8. Einen Randpunkt x einer konvexen Menge M kann man auch so charakterisieren, dass für alle $\epsilon > 0$ $B_\epsilon(x) \cap \text{aff}(M)$ stets einen Punkt aus M und einen Punkt aus $\text{E}\backslash M$ enthält.

In folgendem Satz zeigen wir, wann ein x ein innerer Punkt ist.

Satz 2.9. *Sei $x \in M \subset \text{E}$ und M konvex. Genau dann ist $x \in riM$, wenn wir für alle $y \in M$ ein $\delta > 1$ finden können, so dass $(1-\delta)y + \delta x \in M$ gilt.*

Beweis. Wenn x bereits ein innerer Punk ist, so folgt die Behauptung leicht aus der Definition des relativen Inneren.
Also erfülle nun x die genannte Bedingung. Zu einem $y \in riM$ finden wir nun ein $\delta > 1$, so dass $z := (1-\delta)y + (\delta\, x)$ in M liegt. Mit $0 < \lambda := 1 - \delta^{-1} < 1$ haben wir eine Darstellung $x = \lambda y + (1-\lambda)z$. Angenommen x wäre nun ein Randpunkt. Aus Bemerkung 2.8 leitet man leicht her, dass dann entweder y und z auch Randpunkte sind, oder z nicht in M liegt. Beides widerspricht den Annahmen und somit ist $x \in riM$. □

Für den Spezialfall, $cone(a_1, ..., a_m)$ und $conv(a_1, ..., a_m)$ gilt folgende Aussage:

Korollar 2.10. *Genau dann ist x in $\text{ri}\, M = cone(a_1, ..., a_m)$ (bzw. in $\text{ri}\, M' := conv(a_1, ..., a_m)$) wenn x als eine konische (bzw. konvexe) Kombination der $a_1, ..., a_m$ mit strikt positiven Koeffizienten geschrieben werden kann.*

Beweis. Wir zeigen nur den Fall der konischen Hülle. Der Fall der konvexen Hülle verläuft analog.
Sei zuerst $x := \lambda_1 a_1 + ... + \lambda_m a_m$ so dass alle $\lambda_i > 0$ gewählt sind. Dann ist $x \in M$. Sei nun $y := \delta_1 a_1 + ...\delta_m a_m$ mit $\delta_i \geq 0$ Wir können nun ein $\delta^* > 0$ so bestimmen, dass gilt:

$$\delta^*\lambda_1 + (1-\delta^*)\delta_1 \geq 0, ..., \delta^*\lambda_m + (1-\delta^*)\delta_m \geq 0$$

Nun haben wir $(1-\delta^*)y + \delta^* x \in M$. Nun folgt die Behauptung aus dem Theorem.
Nehmen wir nun an, dass $x \in riM$ gilt. Wir setzten $a^* := \frac{1}{m}(a_1 + ... + a_m)$. Aus aus Theorem 2.9 folgt nun die Existenz eines $\delta > 1$ und eines $y \in M$, so dass

$$y = (1-\delta)a^* + \delta x$$

Da y in M liegt gilt $y = \delta_1 a_1 + ... + \delta_m a_m$ mit positiven $\delta_i \geq 0$. Hieraus erhalten wir für x die Darstellung:

$$x = ((\delta_1 + (\delta - 1)/m)/\delta)a_1 + ... + ((\delta_m + (\delta - 1)/m)/\delta)a_m$$

Somit ist x eine konische Kombination mit strikt positiven Koeffizienten. □

Nun wollen wie erklären, was man unter den Ecken einer abgeschlossenen konvexen Menge versteht. Hierzu folgende Definition:

Definition 2.11. Sei $M \neq \emptyset$ eine abgeschlossene konvexe Menge. $x \in M$ heißt Ecke oder Extrempunkt von M,wenn

$$x = \lambda(x_1) + (1-\lambda)x_2 \text{ mit } x_1, x_2 \in M,\ 0 \leq \lambda \leq 1 \implies \lambda \in \{0, 1\}$$

Die Menge der Ecken bezeichnen wir mit $ext(M)$

Man macht sich leicht klar, dass die Ecken genau die Punkte x sind, für die $M\backslash\{x\}$ immer noch konvex sind. Die Bedeutung der Ecken ergibt sich aus folgendem Satz von Minkowski:

Theorem 2.12 (Minkowski). *Sei M eine kompakte konvexe Menge. Dann gilt:*

$$M = conv(ext(M))$$

Beweis. Siehe [4] Theorem 5.10. □

Im Laufe der Arbeit werden wir entscheiden müssen, ob ein Punkt $y \in E$ in einer gegebenen konvexen Menge M liegt. Für solch einen beliebigen Punkt $y \in$ E suchen wir dafür nun einen Punkt x in einer abgeschlossenen konvexen Menge M, so dass der Abstand $\|x - y\|$ minimal ist. Wie definieren:

Definition 2.13. Sei M eine beliebige abgeschlossene und konvexe Menge in E. Für $y \in$ E heißt der Punkt $p(y, M)$, mit der Eigenschaft

$$\|y - p(y, M)\| \leq \|y - x\| \textit{ für alle } x \in M$$

die metrische Projektion von y auf M.

Dass $p(y, M)$ wohldefiniert ist und sogar eindeutig bestimmt ist beweisen wir an dieser Stelle nicht, sondern verweisen auf [4]. Ist y bereits in M so gilt offensichtlich $p(y, M) = y$. Mit $u(y, M)$ bezeichnen wir im folgenden die Verbindungsstrecke von y zu $p(y, M)$ formal:

$$u(y, M) := \frac{y - p(y, M)}{\|y - p(y, M)\|}$$

Wir hatten eingangs gesehen, dass eine Hyperebene eine Halbraum definiert, der eine konvexe Menge darstellt. Nun wollen wir kurz das Konzept von trennenden und stützenden Hyperebenen erklären.

Wir sagen, dass die Hyperebene $H_{x,\alpha}$ die konvexe Menge M im Punkt x stützt, wenn $x \in M \cap H_{x,\alpha}$ und $M \subset H^+_{x,\alpha}$. Wer nennen x dann auch einen äußeren Normalenvektor von M. Ist $M \cap H_{x,\alpha} \neq M$ so sprechen wir von einer echten Stütze.

Die Stützebenen geben nun eine „äußere“ Beschreibung von konvexen Mengen, wie wir in den folgenden zwei Sätzen sehen werden.

Satz 2.14. *Sei $M \subset$ E konvex und abgeschlossen. Dann geht durch jeden Randpunkt von M eine Stützebene.*

Beweis. Nehmen wir zuerst an, dass M beschränkt ist. In diesem Fall macht man sich recht einfach klar, dass zu jedem x auf dem Rand von M mindestens ein $y \in$ E$\backslash M$ gibt mit $x = p(y, M)$. Die zu $u(y, M)$ orthogonale Hyperebene stützt dann M in x.
Sei nun M unbeschränkt. Wir betrachten die beschränkte konvexe Menge $B_1(x) \cap M$.

Nach dem ersten Teil finden wir zu x eine Hyperebene H die $B_1(x) \cap M$ in x stützt. Angenommen es gibt nun ein $z \in H^- \cap M$. Da M konvex ist, ist somit $[z, x] \in M$ und $[z, x) \cap B(x, 1) \cap M \not\subset H^+$. Das Widerspricht aber der Tatsache, dass H eine Stützebene an x für $B(x, 1) \cap M$ ist. Also muss H auch M in x stützen. □

Umgekehrt können konvexe Mengen mit nicht leerem Inneren dadurch gekennzeichnet werden, dass durch jeden Randpunkt eine Stützebene geht:

Theorem 2.15. *Sei M eine abgeschlossene Menge, deren Inneres nicht leer ist und mit der Eigenschaft, dass durch jeden Randpunkt von M eine echte Stützebene von M geht. Dann ist M eine konvexe Menge.*

Beweis. Angenommen M erfüllt die Voraussetzungen, ist aber nicht konvex. Es gibt dann $x, y \in M$ mit $[x, y] \not\subset M$. Wir finden also $x^* \in [x, y] \cap \operatorname{bd} M$. Nach Vorasusetzung gibt es eine Stützebene H durch x^*. Und wir haben $x, y \in H^+$. Das ist auf Grund der Lage von x, y nur möglich wenn $x, y \in H$. Das widerspricht der Annahme, dass die Stütze eine echte ist. □

Seinen $A, B \subset \mathrm{E}$ zwei beliebige Mengen. Wir sagen dann, dass eine Hyperebene $H_{x,\alpha}$ diese zwei Mengen trennt, wenn gilt $A \subset H^+_{x,\alpha}$ und $B \subset H^-_{x,\alpha}$ oder umgekehrt. Ist $A \cap H = \emptyset$ oder $B \cap H = \emptyset$ so sprechen wir von einer striken Trennung.

Wie begnügen uns mit dem Fall der Trennung eines Punktes von einer konvexen Menge. Es gilt der folgende Satz:

Satz 2.16 (Trennungssatz). *Sei $M \subset \mathrm{E}$ eine konvexe Menge und $x \in \mathrm{E} \backslash M$. Dann können x und M getrennt werden. Ist M abgeschlossen, gibt es eine strikte Trennung*

Beweis. Sei M eine konvexe Menge. Die Hyperebene durch $p(x, M)$ senkrecht zu $u(x, M)$ trennt x und M. Ist M zusätzlich abgeschlossen, so ist $\|p(x, M)\| \neq 0$. Somit liegt x nicht in der trennenden Hyperebene. □

2.2 Konvexe Kegel

In diesem Abschnitt wollen wir die konvexen Kegel genauer studieren. Diese treten beispielsweise als Lösungsmenge von homogenen Ungleichungssytemen auf. Wir definieren:

Definition 2.17. Sei $\mathcal{C}$ eine nicht leere konvexe Menge. Folgt aus $x \in \mathcal{C}$ stets auch $\lambda x \in \mathcal{C}$ für alle $\lambda \geq 0$, so heißt $\mathcal{C}$ ein *konvexer Kegel.*

Für eine beliebige Menge $M \subset \mathrm{E}$ ist die Menge $\operatorname{cone}(M)$ ein konvexer Kegel und wir haben wieder, dass M genau dann ein Kegel ist, wenn $cone(M) = M$ gilt.

Wir nennen einen Kegel $\mathcal{C}$ polyhedral, wenn es eine endliche Menge M gibt mit $\mathcal{C} = cone(M)$.

Ist $\mathcal{C}$ ein konvexer Kegel, so ist

$$\mathcal{C} + (-\mathcal{C}) := \{x - y | x, y \in \mathcal{C}\}$$

ein Unterraum von E, ebenso auch

$$\mathcal{C} \cap (-\mathcal{C}) := \{x \in \mathrm{E} | x \in \mathcal{C},\ -x \in \mathcal{C}\}.$$

Wir nennen $\mathcal{C} \cap (-\mathcal{C})$ den *Linearitätsraum* von $\mathcal{C}$. Ein Kegel $\mathcal{C}$ heißt spitz, falls sein Linearitätsraum nur aus dem 0 Vektor besteht.

Wir definieren nun, was man unter dem dualen oder polaren Kegel versteht:

Definition 2.18. Sei $\mathcal{C} \subset \mathrm{E}$ eine beliebige Teilmenge. Die Menge

$$\mathcal{C}^* := \{y \in \mathrm{E} | x \cdot y \geq 0 \forall x \in \mathcal{C}\}$$

heißt der duale Kegel zu $\mathcal{C}$. Die Menge

$$\mathcal{C}^\perp := \{y \in \mathrm{E} | x \cdot y = 0 \forall x \in \mathcal{C}\} = \mathcal{C}^* \cap (-\mathcal{C}^*)$$

heißt das orthogonale Komplement zu $\mathcal{C}$ in E.

Aus der Definition ist sofort ersichtlich, dass $\mathcal{C}^*$ ein Kegel ist. Dieser ist aber sogar abgeschlossen, wie man sich auch recht einfach klar macht.

Beispiel 2.19. Sei $M := \{x_1, ..., x_m\} \subset \mathrm{E}$ und sei $\mathcal{K} := cone(M)$. Für den dualen Kegel haben wir dann:

$$\begin{aligned}\mathcal{K}^* &:= \{y \in \mathrm{E} : y \cdot x \geq 0 \ \forall x \in \mathcal{K}\} \\ &= \{y \in \mathrm{E} : \sum_{i=1}^{m} y \cdot \lambda_i x_i \geq 0 \ \forall \lambda_i \geq 0\} \\ &= \{y \in \mathrm{E} : y \cdot x_j \geq 0 \ \forall 1 \leq j \leq m\}\end{aligned}$$

Man macht sich einfach klar, dass Dualisieren eine inklusionsumkehernde Abbildung ist, d.h. es gilt

Proposition 2.20. *Seien $\mathcal{K}$ und $\mathcal{C}$ zwei Kegel in* E. *Genau dann ist $\mathcal{K} \subset \mathcal{C}$, wenn $\mathcal{C}^* \subset \mathcal{K}^*$. Insbesondere haben wir $(\mathcal{C} \cup \mathcal{K})^* = \mathcal{C}^* \cap \mathcal{K}^*$.*

Aus Proposition 2.20 leiten wir noch folgende Aussage über die Summe von Kegeln her.

Proposition 2.21. *Seien $\mathcal{C}$ und $\mathcal{K}$ zwei Kegel in* E. *Es gilt dann*

$$(\mathcal{C} + \mathcal{K})^* = \mathcal{C}^* \cap \mathcal{K}^*.$$

Beweis. Man sieht leicht, dass $\mathcal{C} \cup \mathcal{K} \subseteq (\mathcal{C} + \mathcal{K})$ gilt. Somit haben wir mit Proposition 2.20: $\mathcal{C}^* \cap \mathcal{K}^* = (\mathcal{C} \cup \mathcal{K})^* \supseteq (\mathcal{C} + \mathcal{K})^*$. Somit bleibt also noch zu zeigen, dass $\mathcal{C}^* \cap \mathcal{K}^* \subseteq (\mathcal{C} + \mathcal{K})^*$ gilt. Sei $u \in \mathrm{E}$ mit $0 \leq u \cdot x$ und $0 \leq u \cdot y$ für alle $x \in \mathcal{C}$ und $y \in \mathcal{K}$. Für $z := x + y$ gilt dann $u \cdot z = u \cdot x + u \cdot y \geq 0$. □

Mit dem Begriff des dualen Kegel leiten wir aus dem Trennungssatz für konvexe Mengen folgenden Trennungssatz für konvexe Kegeln her:

Satz 2.22. *Sei $\mathcal{C} \subset \mathrm{E}$ ein abgeschlossener Kegel und $x \in \mathrm{E}$. Genau dann ist $x \in \mathrm{E} \backslash \mathcal{C}$, wenn es ein $y \in \mathcal{C}^*$ gibt mit*

$$x \cdot y < 0.$$

Beweis. Sei H eine Stützebene an $\mathcal{C}$. Es existiert also ein $y \in H \cap \mathcal{C}$. Auf Grund der Kegeleigenschaft haben wir $\lambda y \in \mathcal{C}$ für alle $\lambda \geq 0$ und insbesondere $0 = 0y \in H$. Jede Stützebene läßt sich also als $H_{y,0}$ für ein $y \in \mathcal{C}^*$ schreiben. Ist $x \notin \mathcal{C}$ gibt es nach dem Trennungssatz (Satz2.16) für konvexe Mengen nun also ein $y \in \mathcal{C}^*$ mit $x \cdot y < 0$. Somit ist die Hinrichtung gezeigt. Aus der Definition des dualen Kegels folgt, dass $x \cdot y < 0$ für ein $y \in \mathcal{C}^*$ nur dann gelten kann, wenn $x \notin \mathcal{C}$. □

Die Definition des Dualkegels legt nahe, dass $\mathcal{K} \subseteq (\mathcal{K}^*)^*$. Aus dem Trennungssatz können wir folgern, dass bei abgeschlossenen Kegel $\mathcal{K} = \mathcal{K}^{**}$ gilt:

Satz 2.23. *Sei $\mathcal{K}$ ein abgeschlossener Kegel. Dann ist $\mathcal{K} = \mathcal{K}^{**}$.*

Beweis. Angenommen es gibt ein $x \in \mathcal{K}^{**}$ mit $x \notin \mathcal{K}$. Dann gibt es nach dem Trennungssatz ein $y \in \mathcal{K}^*$ mit $x \cdot y < 0$. Da x nach Definition in $\mathcal{K}^{**}$ muss aber gelten $x \cdot y \geq 0$ für alle $y \in \mathcal{C}^*$. Also haben wir einen Widerspruch. □

Zusammen mit Proposition 2.21 erhalten wir das folgende Korollar:

Korollar 2.24. *Seien $\mathcal{C}$ und $\mathcal{K}$ abgeschlossene Kegel in* E. *Wir haben dann:*

$$(\mathcal{C} \cap \mathcal{K})^* = \mathcal{C}^* + \mathcal{K}^*$$
$$(\mathcal{C} + \mathcal{K})^\perp = \mathcal{C}^\perp \cap \mathcal{D}^\perp$$

Beweis. Wir dualisieren die Identität aus Proposition 2.21 und erhalten:

$$(\mathcal{C}^* + \mathcal{K}^*)^* = \mathcal{C}^{**} \cap \mathcal{K}^{**} = \mathcal{C} \cap \mathcal{K}.$$

Also haben wir

$$(\mathcal{C} \cap \mathcal{K})^* = (\mathcal{C}^* + \mathcal{K}^*)^{**} = \mathcal{C}^* + \mathcal{K}^*.$$

Somit ist die erste Aussage gezeigt.
Da $\mathcal{C} \cup \mathcal{K} \subseteq \mathcal{C} + \mathcal{K}$ gilt, ist $(\mathcal{C} + \mathcal{K})^\perp \subseteq (\mathcal{C} \cup \mathcal{K})^\perp = \mathcal{C}^\perp \cap \mathcal{K}^\perp$. Sei nun $u \in (\mathcal{C} \cup \mathcal{K})^\perp$, so gilt für beliebige $x \in \mathcal{C}$ und $y \in \mathcal{K}$ $u \cdot (x + y) = 0$, also $u \in (\mathcal{C} + \mathcal{K})^\perp$. □

Das folgende Korollar ist als Lemma von Farkas bekannt:

Korollar 2.25. *Sei A eine $n \times n$ Matrix und $b \in \mathbb{R}^m$. Genau eine der beiden Alternativen richtig:*

1. *Die Gleichung $Ax = b$ hat eine Lösung mit $x \in \mathbb{R}^n_{\geq 0}$*
2. *$A^t y \geq 0,\ b^t y < 0$ hat eine Lösung $y \in \mathbb{R}^m$*

Beweis. Seien a_i die Spalten der Matrix A. Die erste Alternative bedeutet dann $b \in cone((a_i))$. Es ist $cone((a_i))^{**} = \{z \in \mathbb{R}^m : z^t \cdot y \geq 0 \text{ für alle } y \text{ mit } A^t y \geq 0\}$. Nach Satz 2.22 und 2.16 erfüllt ein b genau dann die zweite Alternative, wenn $b \notin cone((a_j))^{**}$. Nach Satz 2.23 ist $cone((a_j))^{**} = cone((a_j))$. Die zweite Alternative liest sich also als $b \notin cone((a_i))$. Damit ist klar, dass sich beide Alternativen ausschließen und jedes $b \in \mathbb{R}^m$ eine davon erfüllt.

□

Wir definieren nun was man unter den Seiten eines Kegels versteht.

Definition 2.26. Sei x auf dem Rand eines abgeschlossenen Kegels $\mathcal{C}$ und sei H eine echte Stützebene an $\mathcal{C}$ durch x. Die Menge $F := \mathcal{C} \cap H$ ist dann wieder ein Kegel und heißt dann Seite von $\mathcal{C}$.

Wir nennen $dim(F + (-F))$ die Dimension der Seite. Seiten mit Codimension 1 werden als Facetten bezeichnet, Seiten mit Dimension 1 als Kanten.

Es ist klar, dass den Facetten eines Kegels $\mathcal{K}$ genau die Kanten des dualen Kegels $\mathcal{K}^*$ zugeordnet werden können. Wenn F eine Facette von $\mathcal{K}$ ist, schreiben wir F^Δ für die entsprechende Kante in $\mathcal{C}$.

Wir hatten im letzten Abschnitt das Innere einer konvexen Menge betrachtet. Aus Theorem 2.9 folgt, dass wenn y im Inneren eines konvexen Kegels, es durch y keine echte Stützebene geben kann. Also gilt:

Proposition 2.27. *Sei $\mathcal{K}$ ein konvexer Kegel: Genau dann ist y in $ri\mathcal{K}$, wenn gilt:*

$$\mathcal{K}^* \cap H_{y,0} = \mathcal{K}^\perp$$

Bemerkung 2.28. Der Begriff des Inneren eines Kegels $\mathcal{C}$ ist immer bezüglich der Topologie von $(\mathcal{C} + (-\mathcal{C})$ zu denken.

Im weiteren Verlauf werden uns maßgeblich polyhedrale Kegel interessieren. Die inneren Pukte solcher Kegel können wir noch besonders auszeichnen.

Proposition 2.29. *Sei $\mathcal{C} \subset \mathrm{E}$ ein polyhedraler Kegel. Genau dann ist $x \in ri\mathcal{C}$, wenn gilt:*

$$\mathcal{C} + (-\mathcal{C}) = \mathcal{C} + (-\operatorname{cone}(x)).$$

Beweis. Sei x ein innerer Punkt. Dann ist $\operatorname{cone}(x) \subseteq \mathcal{C}$ und somit auch $\mathcal{C} = \mathcal{C} + (\operatorname{cone} x)$. Wir haben also

$$\begin{aligned}
\mathcal{C} + (-\operatorname{cone}(x)) &= \big(\mathcal{C} + \operatorname{cone}(x)\big) + \big(-\operatorname{cone}(x)\big) \\
&= \mathcal{C} + \big(\operatorname{cone}(x) + (-\operatorname{cone}(x))\big) \\
&= \mathcal{C}^{**} + \big(\operatorname{cone}(x)^\perp\big)^* \\
&= \big(\mathcal{C}^* \cap (\operatorname{cone}(x)^\perp)\big).
\end{aligned}$$

Aus Proposition 2.27 haben wir $\mathcal{C}^* \cap (\operatorname{cone}(x)^\perp) = \mathcal{C}^\perp$. Somit folgt insgesamt also

$$\mathcal{C} + (-\operatorname{cone}(x)) = \big(\mathcal{C}^\perp\big)^* = \mathcal{C} + (-\mathcal{C}).$$

Nun wollen wir die Rückrichtung zeigen. Sei also $x \in \mathcal{C}$ ein Punkt in $\mathcal{C}$, so dass $\mathcal{C} + (-\operatorname{cone}(x)) = \mathcal{C} + (-\mathcal{C})$ gilt. Zuerst machen wir klar, dass wir für jedes $z \in \mathcal{C}$ eine Darstellung der Form $z + y = \alpha x$ mit $y \in \mathcal{C}$ und $\alpha \geq 0$ finden können. Sei also z ein beliebiger Punkt in $\mathcal{C}$. Nun ist $-z \in \mathcal{C} + (-\mathcal{C}) = \mathcal{C} + (-\operatorname{cone}(x))$. Also gibt es ein $y \in \mathcal{C}$ und ein $\alpha \geq 0$, sodass $-z = y + (-\alpha x)$ gilt und wir haben $y + z = \alpha x$. Mit dieser Darstellung für alle Punkte in $\mathcal{C}$ wollen wir nun zeigen, dass x im Inneren liegt. Nach Voraussetzung ist $\mathcal{C}$ ein polyhedraler Kegel. Also gibt es eine endliche Menge M, mit $\mathcal{C} = \operatorname{cone}(M)$. Aus dieser Menge

M wählen wir nun eine Basis $e_1, ..., e_k$ für $\mathcal{C} + (-\mathcal{C})$ aus. Da dann insbesondere für alle $i = 1..k$ $e_i \in \mathcal{C}$ liegt, haben wir eine Darstellung $\alpha_i x = y_i + e_i$ mit geeignet gewählten $y_i \in \mathcal{C}$ und $\alpha_i \geq 0$. Wir setzen nun

$$\epsilon := \min\{\frac{1}{k\alpha_i} \mid \ i = 1..k, \alpha_i \geq 0\}$$

und betrachten $B := B_\epsilon(x)$. Um zu zeigen, dass x ein innerer Punkt ist, wollen wir nun klar machen, dass $B \subset \mathcal{C}$ gilt. Dazu sei $z := (\lambda_1 e_i + ... + \lambda_k e_k) + x$ ein beliebiger Punkt in B. Also ist $|\lambda_i| < \epsilon \leq \frac{1}{k\alpha_i}$ und somit $|\alpha_i \lambda_i| < \frac{1}{r}$. Mit $e_i = \alpha_i x - y_i$ haben wir dann

$$\begin{aligned} z &= \sum_{\lambda_i \geq 0} \lambda_i e_i + \sum_{\lambda_i < 0} \lambda_i (\alpha_i x - y_i) + x \\ &= \sum_{\lambda_i \geq 0} \lambda_i e_i + \big(\sum_{\lambda_i < 0} \alpha_i \lambda_i + 1\big) x + \sum_{\lambda_i < 0} (-\lambda_i) y_i \end{aligned}$$

Da alle $y_i \in \mathcal{C}$ sind und ebenso alle $e_i \in \mathcal{C}$. Somit ist der erste und der letzte Term ebenfalls in $\mathcal{C}$. Da sich die $\alpha_i \lambda_i$ betragsmäßig durch $\frac{1}{r}$ abschätzen lassen, ist die mittlere Summe positiv und somit ist auch der mittlere Term in $\mathcal{C}$. Also ist $z \in \mathcal{C}$ und somit $B \subset \mathcal{C}$ und x ist also im Inneren. □

Abschließend wollen wir uns noch klar machen, wann es Punkte im Schnitt des Inneren zweier Kegel gibt. Wir beschränken uns dabei auf den Fall zweier polyhedraler Kegel

Proposition 2.30. *Seien M, M' nicht leere Teilmengen von* E *und $\mathcal{C} = \text{cone}(M)$, $\mathcal{K} = \text{cone}(M'$ zwei polyhedrale Kegel. Äquivalent sind dann:*

1. $ri\mathcal{C} \cap ri\mathcal{K} \neq \emptyset$
2. $\mathcal{C}^* \cap (-\mathcal{K}^*) = \mathcal{C}^\perp \cap \mathcal{K}^\perp$.

Beweis. „1. $\Rightarrow$ 2. " Die Inklusion $\mathcal{K}^\perp \cap \mathcal{K}^\perp \subseteq \mathcal{C}^* \cap \mathcal{K}^*$ ist sofort klar. Wir müssen also noch zeigen, dass unter den Voraussetzungen auch $\mathcal{K}^\perp \cap \mathcal{K}^\perp \supseteq \mathcal{C}^* \cap \mathcal{K}^*$ gilt. Sei nun $x \in ri\mathcal{C}$ und $x \in ri\mathcal{K}$. Mit Proposition 2.24 haben wir $\mathcal{C}^* \cap (-\mathcal{K}^* = (\mathcal{C} + (-\mathcal{K}))^*$. Da $x \in ri\mathcal{C}$ und $x \in ri\mathcal{K}$ haben wir mit 2.29

$$\begin{aligned} (\mathcal{C} + (-\mathcal{K}))^* &= \big(\mathcal{C} + \text{cone}(x) + (-\mathcal{K} - \text{cone}(x))\big)^* \\ &= \big(\mathcal{C} + (-\text{cone}(x)) + (-\mathcal{K} + \text{cone}(x))\big)^* \\ &= \big(\mathcal{C} + (-\mathcal{C}) + \mathcal{K} + (-\mathcal{K})\big)^* \\ &= \mathcal{C}^\perp \cap \mathcal{K}^\perp \end{aligned}$$

„2. $\Rightarrow$ 1. " Wir haben $(\mathcal{C} + (-\mathcal{K})^* = \mathcal{C}^* \cap (-\mathcal{K}^*) = \mathcal{C}^\perp \cap \mathcal{K}^\perp$. Somit ist $(\mathcal{C} + (-\mathcal{K}))$ ein Vektorraum und insbesondere ist $0 \in ri(\mathcal{C} + (-\mathcal{K}))$. Also $0 = \sum_{z_i \in M} \lambda_i z_i - \sum_{y_i \in M'} \delta_i y_i$ mit strikt positiven λ_i und δ_i. Somit ist $x = \sum_{z_i \in M} \lambda_i z_i = \sum_{y_i \in M'} \delta_i y_i$ sowohl in $ri\mathcal{C}$ wie auch in $ri\mathcal{K}$ und $ri\mathcal{C} \cap ri\mathcal{K}$ ist nicht leer. □

Wir wollen nun noch den Begriff des Polyeders klären, diesen benötigen wir im nächsten Abschnitt, um kurz die Ideen der sogenannten Linearen Optimierung darzustellen.

Definition 2.31. Eine Menge $P \subset \mathbb{R}^n$ heißt *Polyeder*, wenn sie als Durchschnitt von endlich vielen abgeschlossenen affinen Halbräumen dargestellt werden kann. Ein beschränktes Polyeder wird als Polytop bezeichnet.

Für den folgendes Darstellungssatz definieren wir die sogenannte Minkowski Summe.

Definition 2.32. Seien $A, B \subset \mathbb{R}^n$. Dann ist die *Minkowski-Summe* von A und B definiert als

$$A + B := \{a + b \mid \ a \in A,\, b \in B\}.$$

Theorem 2.33. *Jedes Polyeder $P \subset \mathbb{R}^n$ kann als Minkowski-Summe*

$$P = convV + coneY$$

mit endlichen Mengen V, Y geschrieben werden.

Beweis. Für den Beweis verweisen wir auf [28] Theorem 1.2. □

2.3 Lineare Optimierung

In diesem Abschnitt wollen wir kurz das Prinzip der *linearen Optimierung* vorstellen. Hierbei soll eine lineare Zielfunktion über einem durch Ungleichungen gegebenen Polyeder P maximiert (bzw. minimiert) werden. Wer werden zeigen, dass der Optimalwert, wenn ein solcher existiert, immer auch in einer Ecke von P angenommen wird. Wir werden im letzten Kapitel zeigen, wie man die Frage, ob ein Gitter extrem ist, mit Hilfe von linearer Optimierung lösen kann.

Im folgenden bezeichnen wir mit $x \leq y$ für zwei Vektoren x, y die komponentenweise Ungleichungsrelation. Unter einem linearen Optimierungsproblem (kurz: *lineares Programm*, *LP*) verstehen wir das folgende:

Definition 2.34. Für eine Matrix $A \in \mathbb{R}^{m \times n}$ und Vektoren $c \in \mathbb{R}^n$ und $b \in \mathbb{R}^m$ bezeichnen wir mit (LP) dein Optmierungsproblem folgender Gestalt:

$$\begin{array}{lrcl} & c \cdot x & \rightarrow & \max \\ \text{wobei} & Ax & \leq & b \end{array} \qquad \text{(LP)}$$

Ein Vektor x mit $Ax \leq b$ nennen wir *zulässige Lösung* des linearen Programms. Das Polyeder der zulässigen Vektoren nennen wir zulässigen Bereich und schreiben $P(A, b)$. Eine zulässige Lösung, die das Maximum annimmt, heißt eine *Optimallösung.*

Es gibt zwei Möglichkeiten, in denen das LP keine Optimallösung besitzt:
Das Problem kann *unzulässig* sein (d.h. $P(A, b) = \emptyset$), oder *unbeschränkt* (d.h. für alle $\alpha \in \mathbb{R}$ existiert ein $x \in P(A, b)$ so dass $c^T x > \alpha$). Für ein LP, das weder unzulässig noch unbeschränkt ist, werden wir sehen, dass es eine optimale Lösung hat.

Unter den Punkten des konvexen Polyeders $P(A, b)$ werden uns im Folgenden die Ecken besonders interessieren. In Satz 2.33 hatten wir gesehen, dass $P(A, b)$ eine Darstellung der Form $\text{conv}(v_1, ..., v_r) +$

$\text{cone}(y_1, ..., y_m)$ besitzt. Die Ecken des Polyeders sind somit die $v_1, ..., v_r$. Die Bedeutung dieser Ecken zeigt nun der nächste Satz:

Satz 2.35. *Hat das lineare Problem (LP) eine Optimallösung, dann gibt es auch eine Ecke von $P(A, b)$, die optimal ist.*

Beweis. Sei nun x^* eine Optimallösung des linearen Problems, d.h. $c \cdot x^* \geq c \cdot x$ für alle $x \in P(A, b)$. Wir nehmen nun an, dass keine der Ecken $v_1, ..., v_n$ Optimallösung ist, also ist $c \cdot x^* > c \cdot v_i$. Auf Grund von Satz 2.33 haben wir eine Darstellung der Form

$$x^* = \sum_{i=1}^{r} \lambda_i v_i + \sum_{j=1}^{m} \delta_j y_j, \text{ mit } 0 \leq \lambda_i \leq 1, \ \sum_{i=1}^{n} \lambda_i = 1, \ \delta_j \geq 0.$$

Dann ist $c \cdot x^* = \sum_{i=1}^{r} \lambda_i c \cdot v_i + \sum_{j=1}^{m} \delta_j c \cdot y_j$. Da nun nach den Annahmen $c \cdot v_i < c \cdot x^* \forall 1 \leq i \leq r$ gilt haben wir $\sum_{i=1}^{r} \lambda_i c \cdot v_i < c \cdot x^*$. Somit muss $\sum_{j=1}^{m} \delta_j c \cdot y_j > 0$ sein. Insbesondere gibt es ein j mit $c \cdot y_j > 0$. Für alle $\delta \geq 0$ ist aber $\tilde{x} := x^* + \delta y_j$ in $P(A, b)$, aber $c \cdot \tilde{x} > c \cdot x^*$, was zum Einen der Optimalität von x^* als auch überhaupt der Existenz einer (endlichen) Optimallösung widerspricht. □

Wir wollen nun noch etwas weitergehend die geometrische Struktur des Randes eines Polyeders $P(A, b)$ erleuchten. Für einen Randpunkt $v \in \text{bd}\, P$ sind einige Ungleichungen mit Gleichheit definiert. Diese Ungleichungen nennen wir *aktiv*. Wir können nun die Matrix A so umordnen, dass wir folgende Darstellung erhalten:

$$A_1 x = b_1 , \tag{2.1}$$
$$A_2 x < b_2 , \tag{2.2}$$

Die Gleichungen der Matrix $A_1(v)$ definieren die affinen Stützebenen, auf denen der Randpunkt v liegt. Die Normalenvektoren dieser Stützebenen erhalten wir also aus den transponierten Zeilen der Matrix $A_1(v)$. Wir nennen daher den von den tarnsponierten Zeilen der Matrix A_1 erzeugte Kegel

$$\mathcal{N}(v) := \text{cone}\{a_1(v)^T, \ldots, a_k(v)^T\}$$

den *Kegel der äußeren Normalen* in v. Für diesen haben wir folgende Darstellung

Lemma 2.36. *Es sei $P = \{x \in \mathbb{R}^n : A_1 x \leq b_1\}$. Wir haben*

$$\mathcal{N}(v) = \bigcap_{x \in P - v} \{ c_0 \in \mathbb{R}^n : c_0^T x \leq 0\} .$$

Beweis. Sei $c_0 \in N(v)$ und $a_1^T \ldots, a_k^T$ die Zeilen von A_1. Dann gibt es $\lambda_1, \ldots, \lambda_k \geq 0$ mit $c_0 = \sum_{j=1}^{k} \lambda_j a_j$. Wegen $P - v = \{x : A_1 x \leq 0\}$ folgt somit für jeden Punkt $x \in P - v$

$$c_0^T x = \sum_{i=1}^{k} \lambda_j a_j^T x \leq 0 .$$

Also Sei $c_0 \notin \mathcal{N}(v)$. Dann gibt es nach dem Trennungssatz einen Vektor w mit $w^T c_0 > 0$, aber $w^T z \leq 0$ für alle $z \in \mathcal{N}(v)$. Insbesondere gilt daher $w^T a_j \leq 0$ für alle $j \in \{1, \ldots, k\}$ und somit $A_1 w \leq 0$. Es

folgt $w \in C - v$, aber $c_0^T w > 0$, d.h.

$$c_0 \notin \bigcap_{x \in P - v} \{c_1 \in \mathbb{R}^n \, : c_1^T x \leq 0\} \, .$$

□

Hieraus erhalten wir eine Charakterisierung für optimale Lösungen:

Korollar 2.37. *Es sei (LP) ein lineares Problem wie in Definition 2.34 und v ein Ranpunkt der zulässigen Bereiches $P(A, b)$. Genau dann ist der Punkt v eine Optimallösung des LPs, wenn c im Kegel $\mathcal{N}(v)$ der äußeren Normalen an v enthalten ist.*

Beweis. Der Punkt $v \in \operatorname{bd} P$ ist optimal für (LP) genau dann, wenn für alle $x \in P$ gilt $c^T x \leq c^T v$ bzw. $c^T(x - v) \leq 0$. Damit folgt die Behauptung aus dem voranstehenden Lemma. □

Ein Punkt v ist also genau dann eine Optimallösung für das LP, wenn das Ungleichungssystem

$$\begin{aligned} A^T y &= c \\ y &\geq 0 \end{aligned}$$

in $y = (y_1, \ldots, y_m)$ eine Lösung besitzt, für die höchstens solche Komponenten von 0 verschieden sind, die zu den aktiven Nebenbedingungen (für v) gehören.

Wie wir eingangs schon gezeigt haben, sind die Ecken des Polyeders von besonderem Interesse. Eine Ecke liegt immer im Durchschnitt von n affinen Hyperebenen. Wir haben folgende Proposition:

Proposition 2.38. *Für $v \in \operatorname{bd} P$ sind die folgenden Aussagen äquivalent:*

1. *Der Punkt v ist eine Ecke von P.*
2. *Die Matrix A_1 der aktiven Nebenbedingungen hat vollen Rang n.*
3. *Der Kegel $\mathcal{N}(v)$ ist spitz.*

Beweis. „1. $\Rightarrow$ 2. “: Sei v eine Ecke von P. Angenommen, die Matrix A_1 hätte nicht den vollen Rang, dann haben wir ein v' mit $v' \neq v$ und $A_1 v' = b_1$. Sei nun $y_{1/2} := v \pm \delta v - v'$. Nun ist $A_1 y_{1/2} = b_1$ und wir können ein $\delta > 0$ so klein wählen, dass auch $A_2 y_{1/2} < b_2$ gilt. Nun ist $v = \frac{1}{2}(y_1 + y_2)$. Das widerspricht aber der der Annahme, dass v Ecke war.
„2. $\Rightarrow$ 3. “: Der duale Kegel zu $\mathcal{N}(v)$ wird von den Spalten der Matrix A_1 aufgespannt. Da Zeilenrang gleich Spaltenrang ist, hat auch $\mathcal{N}(v)^*$ volle Dimension. Also $\{0\} = \mathcal{N}(v)^{*\perp} = \mathcal{N}(v) \cap (-\mathcal{N}(v))$, also ist $\mathcal{N}(v)$ ein spitzer Kegel.
„3. $\Rightarrow$ 1. “: Sei $\mathcal{N}(v)$ ein spitzer Kegel. Dann liegt v im Durchschnitt von n affinen Hyperebenen. Somit ist v ein 0 dimensionaler affiner Raum, also eine Ecke. □

Aus dieser Proposition erhalten wir noch eine wichtige Überlegung. Damit v eine Ecke sein kann, muss es ein Teilmatrix A_1 mit vollem Rang n geben. Also kann es maximal so viele Ecken geben, wie Teilmatrizen mit vollem Rang gebildet werden können. Also es gilt

Korollar 2.39. *Ein Polyeder $P(A,b)$ hat höchsten $\binom{m}{n}$ Ecken*

Nun erklären wir, was wir unter dem sogenannten dualen Programm verstehen werden.

Definition 2.40. Zu jedem linearen Programm (LP) ist das *duale Programm* als

$$\begin{array}{rrcll} & b \cdot y & \to & \min & \\ \text{wobei} & A^T y & = & c & (LP^*) \\ & y & \geq & 0 & \end{array}$$

definiert. Das ursprüngliche Problem $\max\{c^T x : Ax \leq b\}$ wird auch das *primale Programm* genannt.

Folgender Satz zeigt uns den wichtigen Zusammenhang der Lösungen der beiden zueinander dualen Probleme:

Satz 2.41. (Schwacher Dualitätssatz.)

1. *Seien x und y zulässige Lösungen der dualen LPs $\max\{c \cdot x : Ax \leq b\}$ und $\min\{b \cdot y : A^T y = c,\ y \geq 0\}$. Dann gilt $c \cdot x \leq b \cdot y$.*
2. *Ist $c \cdot x = b \cdot y$, so sind x, y optimale Lösungen der jeweiligen Probleme.*

Beweis. 1. Für zulässige Lösungen x und y gilt

$$c \cdot x \;=\; (y \cdot A)x \;=\; y \cdot (Ax) \;\leq\; y \cdot b \;=\; b \cdot y\,.$$

Also $c \cdot x \leq b \cdot y$.

2. Angenommen x wäre keine optimale Lösung, dann existiert $\tilde{x} \in P(A,b)$ mit $c \cdot \tilde{x} > c \cdot x$. Nun gilt aber nach dem ersten Teil $b \cdot y \geq c \cdot \tilde{x} > c \cdot x$ was aber im Widerspruch zu den Voraussetzungen steht.

□

Nun können wir uns der Frage zuwenden, wie wir eine optimale Lösung charakterisieren können. Wir formulieren hierzu den sogenannten Komplementaritätssatz.

Satz 2.42 (Komplementaritätssatz). *Sei (x,y) ein Paar zulässiger Lösungen der dualen LPs $\max\{c \cdot x : Ax \leq b\}$ und $\min\{b \cdot y : A^T y = c, y \geq 0\}$. Genau dann sind x und y optimale Lösungen der jeweiligen Programme, wenn sie den sogenannte* Komplementaritätsbedingungen *genügen, d.h wenn gilt:*

$$y \cdot (b - Ax) = 0, x \cdot (c - A^T y) = 0$$

Beweis. Für einen Beweis siehe [21] Satz 2.8 □

Ein solches Paar nennen wir ein *primal-duales Paar.*

Nach dieser kurzen Einführung in den theoretischen Hintergrund der linearen Optimierung wollen wir einen Algorithmus zur Bestimmung einer optimalen Lösung vorstellen. Schon Fourier hatte einen Algorithmus zu Behandlung von linearen Problemen gekannt (vgl. [28]), dennoch ist der heute bekannteste Algorithmus zur linearen Programmierung der *Simplex-Algorithmus*, der von G. Dantzig im Jahre 1947 entwickelt wurde.

Wir hatten eingangs gezeigt, dass falls das Problem eine Optimallösung besitz, stets auch eine optimale Ecke existiert. Die grundlegende Idee des Simplex-Algorithmuses besteht nun darin, die Ecken des Polyeders zu durchlaufen. In jedem Schritt soll dabei geprüft werden, ob die aktuelle Ecke schon optimal ist, oder nicht. Im zweiten Fall werden die von der Ecke ausgehenden Kanten ermittelt, in deren Richtung die Zielfunktion ansteigt. Entweder ist diese Kante unbeschränkt und somit hat dann das Problem keine endliche Optimallösung, oder man findet so eine bessere Ecke des Polyeders.

Da nach Korollar 2.39 nur endlich viele Ecken vorhanden sind, ist dieses so beschriebene Verfahren endlich.

Wir nehmen nun an, dass wir bereits eine Ecke v kennen. Diese Annahme können wir ohne Probleme treffen, denn wir werden abschließend zeigen, wie man das Simplexverfahren in zwei Phasen dazu benutzen kann, eine solche Startecke zu finden.

Jede Basis der Ecke v definiert einen spitzen affinen Kegel mit Scheitel v, der das Polyeder P enthält. Dieser Kegel enspricht lokal einem Simplex. Daher stammt der Name Simplex-Algorithmus.

Für eine Ecke v von P sei I die Menge der Zeilenindizes zu einer Basis von v. Die von I induzierte Teilmatrix von A bezeichnen wir mit A_I, entsprechend für den Vektor b.

Wir wissen, dass A_I regulär ist, und es gilt $A_I v = b_I$. In v können wir nun P durch $K_I = \{x \in \mathbb{R}^n \mid A_I x \leq b_I\}$ aproximieren. Die Kanten von K_I entsprechen genau den Kanten von P, die v enthalten. Jede dieser Kanten ist der Schnitt von genau $n-1$ Facetten von K_I. Für jede Zeile $i \in I$ definiert betrachten wir $K_{I\setminus\{i\}}$. nach Konstruktion enthält der eindimensionale Linealitätsraum von $K_{I\setminus\{i\}}$ eine Kante von K_I, also entspricht einer möglichen Richtung.

Lemma 2.43. *Die Menge $L = \{x \in \mathbb{R}^n \mid A_{I\setminus\{i\}} x = b_{I\setminus\{i\}}\}$ ist eine affine Gerade in $\mathbb{R}^n$, die v enthält. Ferner ist die Spalte von $-(A_I)^{-1}$ mit Index i ein Richtungsvektor von L.*

Beweis. Aus der Regularität von A_I folgt, dass L eine Gerade ist, und offensichtlich gilt $v \in L$. Sei s die Spalte von $-(A_I)^{-1}$ mit Index i. Dann gilt

$$A_{I\setminus\{i\}} s = 0 \quad \text{und} \quad a_i s = -1\,. \tag{2.3}$$

Folglich ist s ein von Null verschiedener Vektor mit $v + s \in L$. □

Sei s nun die Spalte von $-(A_I)^{-1}$ mit Index i. Wir können nun von der Ecke v in Richtung s gehen,ohne das Polyeder zu verlassen. Ob sich das lohnt, hängt davon ab, ob sich der Zielfunktionswert in Richtung s verbessert, d.h. ob $cs > 0$ ist. Dies kann sehr einfach charakterisiert werden, wenn eine bestimmte Art von Lösung für das duale Programm bekannt ist.

Lemma 2.44. *Sei y eine zulässige Lösung des dualen Programms mit $y_j = 0$ für alle $j \notin I$. Es gilt genau dann $cs > 0$, wenn $y_i < 0$.*

Beweis. Sei y eine solche dual zulässige Lösung, dann gilt nach Definition des dualen Programms sowie nach (2.3)

$$cs \;=\; yAs \;=\; y_I A_I s_I \;=\; -y_i\,.$$

□

Nun können wir aus den möglichen Richtungen diejenigen finden, für die wir eine bessere Lösung erhalten. Haben wir solch eine Richtung ausgewählt, muss noch entschieden werden, ob die entsprechende Kante unbeschränkt ist oder ob wir so eine neue Ecke finden können.

Die Idee für das weitere Vorgehen ist nun, von der Ecke v in Richtung s einer geeigneten Kante von K_I soweit zu laufen, wie noch keine der Zulässigkeitsbedingungen verletzt ist. Dabei können zwei Fälle auftreten.

Lemma 2.45. *1. Im Fall $As \leq 0$ ist $v + \lambda s$ für alle $\lambda \geq 0$ zulässig.*

2. Andernfalls ist für

$$\lambda_s := \min\{\frac{\beta_j - a_j v}{a_j s} \mid a_j s > 0\} \tag{2.4}$$

der Punkt $v + \lambda_s s$ ein zulässiger Punkt, und λ_s ist maximal mit dieser Eigenschaft.

Beweis. Wir betrachten hierzu zunächst einen beliebigen Zeilenindex j und die zugehörige Nebenbedingung $a_j v \leq \beta_j$.

Behauptung: Es gilt

$$\max\{\lambda \geq 0 \mid v + \lambda s \in \{x \in \mathbb{R}^n \mid a_j x \leq \beta_j\}\} \;=\; \begin{cases} \frac{\beta_j - a_j v}{a_j s} & \text{falls } a_j s > 0\,, \\ \infty & \text{falls } a_j s \leq 0\,. \end{cases}$$

Da v ein zulässiger Punkt ist, gilt $a_j v \leq \beta_j$. Ist $a_j s \leq 0$, dann folgt für alle $\lambda \geq 0$ die Ungleichung $a_j(v + \lambda s) \;\leq\; a_j v \;\leq\; \beta_j$. Anderenfalls gilt genau dann $a_j(v + \lambda s) \;\leq a_j v \;\leq\; \beta_j$, wenn $\lambda \leq (\beta_j - a_j v)/(a_j s)$.

Um festzustellen, wann $v + \lambda s$ den Zulässigkeitsbereich P verlässt, betrachten wir alle Ungleichungen simultan und erhalten so die zu zeigende Aussage des Lemmas. □

Da das im Fall $As > 0$ in Lemma 2.45 gewählte λ_s maximal ist, wird für $\lambda_s > 0$ im Punkt $v + \lambda_s s$ eine Ungleichung aktiv, die vorher nicht aktiv war. Wir erhalten eine neue Ecke:

Lemma 2.46. *Sei j ein Zeilenindex der Matrix A mit $\lambda_s = (\beta_j - a_j v)/(a_j s)$. Dann ist $v' := v + \lambda_s s$ eine Ecke von P und $(I \setminus \{i\}) \cup \{j\}$ die Indexmenge einer Basis für v'.*

Simplex-Algorithmus(A, b, c, v)

Eingabe: Eine Matrix $A \in \mathbb{R}^{m \times n}$ und Vektoren $b \in \mathbb{R}^m$, $c \in (\mathbb{R}^n)^*$,
eine Ecke v von $P = \{x \in \mathbb{R}^n : Ax \leq b\}$.
Ausgabe: Eine Ecke v von P, die $\max\{c^T x : x \in P\}$ annimmt,
oder ein Vektor $w \in \mathbb{R}^n$ mit $Aw \leq 0$ und $c^T w > 0$ (d.h. das LP ist unbeschränkt).

1 Sei I Indexmenge einer Basis für v.
2 Bestimme ein $y \in (\mathbb{R}^m)^*$ mit $yA = c$ und $y_i = 0$ für alle $i \notin I$.
Falls $y \geq 0$, dann Ende mit Ausgabe v und dem dualen Vektor y.
3 Sei i der minimale Index mit $y_i < 0$.
Sei s die Spalte von $-(A_I)^{-1}$ mit Index i, so dass $A_{I\setminus\{i\}} s = 0$ und $a_i s = -1$.
Falls $As \leq 0$ dann Ende mit Ausgabe s.
4 Setze $\lambda := \min\{\frac{\beta_j - a_j v}{a_j w} \mid a_j w > 0\}$, und sei j der kleinste Zeilenindex, der dieses Minimum annimmt.
5 Setze $I := (I \setminus \{i\}) \cup \{j\}$ und $v := v + \lambda w$. Gehe zu Schritt 2.

Abbildung 2.1: Simplex-Algorithmus

Beweis. Sei $I' = (I \setminus \{i\}) \cup \{j\}$. Es ist zu zeigen, dass $A_{I'}$ regulär ist und dass $A_{I'} v' = b_{I'}$. Wegen $A_{I\setminus\{i\}} s = 0$ und $a_j s > 0$ liegt a_j nicht im Zeilenraum der $(n-1)$-zeiligen Matrix $A_{I\setminus\{i\}}$. Folglich ist $A_{I'}$ regulär. Aus $A_{I\setminus\{i\}} s = 0$ und $\lambda_s = (\beta_j - a_j v)(a_j s)$ folgt

$$A_{I\setminus\{i\}}(v + \lambda_s s) \;=\; A_{I\setminus\{i\}} v \;=\; b_{I\setminus\{i\}}$$

und

$$a_j(v + \lambda_s s) \;=\; a_j v + a_j s \frac{\beta_j - a_j v}{a_j s} \;=\; \beta_j \,.$$

Also gilt $A_{I'} v' = b_{I'}$. □

Das in (2.4) bestimmte λ_s ist nicht eindeutig bestimmt und kann durchaus verschwinden. Jedoch nur für $\lambda_s > 0$ erhalten wir in der Richtung s von v eine bessere Ecke. Das liegt daran, dass im Allgemeinen eine Ecke mehrer mögliche Basen besitzt. Man spricht dann von *entarteten* Ecken. Betrachten wir beispielsweise im dreidimensionalen Anschauungsraum eine Pyramide mit quadratischer Grundfläche. In der Ecke auf der Spitze der Pyramide treffen dann 4 Kanten zusammen, von denen immer nur drei eine Basis bilden.

Dies kann nun dazu führen, dass $\lambda_s = 0$ ist und wir nur von einer Basis der Ecke in eine neue Basis der selben Ecke wechseln. Im Extremfall kann das Simplexverfahren dann eine unendliche periodische Folge von entarteten Ecken liefern und bricht nie ab. Durch geeignete Auswahl der Indizes i und j, der sogenannten *Pivotwahl* kann jedoch gewährleistet werden, dass eine nicht-optimalen Ecke nach endlich vielen Schritten tatsächlich wieder verlassen wird. Die bekannteste solche Auswahlregel („Pivotregel“) ist die *Regel von Bland*. Hierbei werden i und j im Falle mehrerer Möglichkeiten jeweils kleinst möglich gewählt.

Das so beschriebene Verfahren heißt *Simplexverfahren.* Wir geben in Abbildung 2.1 eine formale Darstellung des Simplexalgorithmus mit der Pivotregel von Bland.

Im 2. Schritt des Algorithmus wird eine Lösung des dualen Zulässigkeitsbereiches bestimmt. Terminiert der Algorithmus hier, so bilden v und y ein primal-duales Paar: Es gilt $cv = (yA)v = y(Av) = yb = b^T y$, da die Komponenten von y außerhalb der Indexmenge I Null sind. Aus dem schwachen Dualitätssatz folgt daher die Optimalität von v und y.

Falls die momentane Ecke nicht optimal ist, wird in Schritt 3 eine Suchrichtung s gemäß Lemma 2.43 gewählt. Wir haben nach Lemma 2.44 $cs > 0$ und somit wir der Zielfunktionswert dieser Richtung verbessert. Terminiert der Algorithmuss nach Schritt 3, so ist $v + \lambda s \in P$ für alle $\lambda \geq 0$ und das Problem ist somit unbeschränkt.

Bricht der Algorithmus weder in Schritt 2, noch in Schritt 3 ab, so muss mittels Lemma 2.45 die maximal mögliche Schrittweite λ_s berechnet werden. Dies geschieht in Schritt 4.

Anschließend wird in Schritt 5 gemäß Lemma 2.46 die neue Basis bestimmt.

Hat man eine geeignete Pivotregel implementiert, so wechselt der Algorithmus in jedem Schritt die Ecken. Somit haben wir folgenden Satz.

Satz 2.47. *Der Simplex-Algorithmus terminiert nach höchstens* $\binom{m}{n}$ *Iterationen.*

Das Ergebnis dieses Satzes ist auf den ersten Blick unbefriedigend, denn die Anzahl der Ecken wächst exponentiell mit der Dimension. Besser wäre es wenn man eine polynomiale Abschätzung für die Laufzeit des Algorithmuses hätte.

Wie wir erläutert hatten, ist die Laufzeit des Algorithmuses abhänig von der implementierten Pivotregel. Für die bisher gängigen Pivotregeln existieren Beispiele von exponentieller Laufzeit. Ob eine Pivotregel existiert, die auf einen Polynomialzeitalgorithmus führt, ist noch nicht geklärt. Dennoch hat der Simplex-Algorithmus in den meisten Anwendungen eine gute Laufzeit und ist das Verfahren der Wahl.

Wir zeigen nun, wie allgemeine lineare Programme mit dem Simplex-Algorithmus gelöst werden können. Hierzu ist noch zu klären, wie eine Startecke gefunden wird. Da es Polyeder gibt, die überhaupt keine Ecke besitzen (z.B. $\{x \in \mathbb{R}^2 : x_1 \leq 0\}$), überführen wir im Folgenden LPs in die Form

$$\begin{array}{rrcl} & c^T x & \rightarrow & \max \\ \text{wobei} & Ax & \leq & b \\ & x & \geq & 0 \end{array} \tag{2.5}$$

Diese besitzen immer den Ursprung als Ecke. Eien solche Form ist keine Einschränkung, da jedes LP der Form $\max\{c\dot{x} : Ax \leq b\}$ wie folgt in die Form (2.5) überführt werden kann.

Jeder Vektor$x \in \mathbb{R}^n$ besitzt eine Darstellung der Form $x = y - z$ mit $y, z \in \mathbb{R}^n_+$. Wir ersetzen x durch $y - z$ und schreiben das LP in der Form

$$\begin{array}{rcl} (c^T\ ,\ -c^T)\begin{pmatrix} y \\ z \end{pmatrix} & \mapsto & \max \\ (A\ ,\ -A)\begin{pmatrix} y \\ z \end{pmatrix} & \leq & b\,, \\ y, z & \geq & 0\,. \end{array} \quad . \tag{2.6}$$

Dieses LP ist genau dann zulässig wenn das Ausgangsproblem zulässig ist, und es besitzt die gleiche Optimallösung.

Im folgenden können wir daher von einem LP der Form (2.5) ausgehen. Mit den Bezeichnungen $I = \{i : b_i \geq 0\}$ und $J = \{i : b_i < 0\}$ betrachten wir das Hilfsproblem

$$\begin{array}{rcl} (\mathbf{1}^T A_J)x + \mathbf{1}^T y & \mapsto & \min \\ A_I x & \leq & b'\,, \\ A_J x + y & \geq & b_J\,, \\ x, y & \geq & 0\,, \end{array} \tag{2.7}$$

wobei $\mathbf{1}$ den aus lauter Einsen bestehenden Vektor bezeichnet. Sei P' der Zulässigkeitsbereich des Hilfsproblems.

Lemma 2.48. *Der Punkt* $\binom{x}{y} = 0$ *ist eine Ecke von* P'*. Der Minimalwert* z^* *des Hilfsproblems ist endlich, und es gilt* $z^* \geq \mathbf{1}^T b_J$*. Ist* $z^* > \mathbf{1}^T b_J$*, dann ist* (2.5) *unzulässig. Ist* $z^* = \mathbf{1}^T b_I$*, dann ist jede optimale Ecke* x *des Hilfsproblems eine Ecke des Zulässigkeitsbereiches von* (2.5)*.*

Durch Anwendung des Simplex-Algorithmus auf das Hilfsproblem mit der Startecke 0 kann also entschieden werden, ob das Ausgangsproblem zulässig ist. Im Falle der Zulässigkeit kann x als Startecke für die Anwendung des Simplex-Algorithmus auf das LP (2.5) verwendet werden.

Beweis. Da im Punkt 0 sowohl die Ungleichungen $x_i \geq 0$ als auch die Ungleichungen $y_i \geq 0$ aktiv sind, ist $\binom{x}{y} = 0$ eine Ecke von P'. Die Zielfunktion des Hilfsproblems ist durch $\mathbf{1}^T b_J$ nach unten beschränkt. Für jede zulässige Lösung x von (2.5) ist

$$\begin{pmatrix} x \\ b_J - A_J x \end{pmatrix}$$

eine optimale Lösung von (2.7). Ist daher das Minimum von (2.7) größer als $\mathbf{1}^T b_J$, dann ist (2.5) unzulässig.

Anderenfalls sei $\binom{x}{y}$ eine optimale Ecke von (2.7). Es gilt $A_J x + y = b_J$. Bezeichnen n und m die Dimensionen von x und y, dann gibt es eine Menge S von $n + m$ Ungleichungen von (2.7), die mit Gleichheit erfüllt sind, so dass die zu diesen $n + m$ Ungleichungen korrespondierende Untermatrix regulär ist.

Sei S_I die Menge der Ungleichungen von $A_I x \leq b_I$ und von $x \geq 0$, die zu S gehören. Sei S_J die Menge der Ungleichungen von $A_J x \leq b_J$, für die die korrespondierenden Ungleichungen von $A_J x + y \leq b_J$ und

$y \geq 0$ beide zu S gehören. Wegen $A_J x + y = b_J$ gilt $|S_I \cup S_J| \geq |S| - m = n$, und die Ungleichungen von $S_I \cup S_J$ sind linear unabhängig und an der Stelle x mit Gleichheit erfüllt. Daher erfüllt x mindestens n linear unabhängige Ungleichungen von (2.5) mit Gleichheit; x ist also eine Ecke. □

Wie wir bereits erwähnt hatten, ist nicht bekannt, ob der Simplex-Algorithmus mit einer geeigneten Pivot-Regel ein Polynomialzeit-Algorithmus ist. Numerisch-effektive Varianten des hier vorgestellten Grundalgorithmus für das Simplex-Verfahren sind jedoch in der Praxis zum Lösen linearer Programme am besten geeignet.

Es gibt Polynomialzeit-Algorithmen zur Lösung linearer Optimierungsprobleme: den Ellipsoid-Algorithmus (Khachiyan, 1979), der aber nicht praktikabel ist, sowie Innere-Punkt-Verfahren (Karmarkar, 1984).(vgl. [21] Kapitel 8)

2.4 Konvexe Funktionen

Wir wollen nun den Begriff der Konvexität auf Funktionen übertragen. Wir definieren dazu:

Definition 2.49. Sei $I \subset \mathbb{R}$ ein Interval. Eine Funktion $f : I \mapsto \mathbb{R}$ heißt konvex, wenn

$$f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y)$$

für alle $x, y \in I$ und $\lambda, \mu \geq 0$ mit $\lambda + \mu = 1$.
Wir sagen, dass f konkav ist, wenn $-f$ konvex ist.

Sind $p_1 := (x, f(x))$ und $p_2 := (y, f(y))$ zwei Punkte auf dem Graphen von f. Die Definition einer konvexen Funktion fordert dann, dass die Gerade durch p_1 und p_2 über dem Graphen von f liegt. Mit dieser geometrischen Anschauung kann man sich recht einfach klar machen, dass eine konvexe Funktion f höchstens an abzählbar vielen Stellen nicht differenzierbar ist. Wir setzten nun voraus, dass f auf ganz I differenzierbar ist. Dann können wir folgende Eigenschaft über konvexe Funktionen zeigen.

Theorem 2.50. *Sei $f : I \mapsto \mathbb{R}$ eine differenzierbare Funktion. Genau dann ist f eine konvexe Funktion, wenn f' monoton steigend ist.*

Beweis. Dass für eine konvexe Funktion f stets $f'(x) \geq f'(x + \epsilon)$ $\forall \epsilon > 0$ gelten muss, macht man sich klar. Wir zeigen also die Rückrichtung.
Sei also f' monoton steigend. Wir wählen $a, b \in I$ mit $a < b$ und $\lambda, \mu > 0$ mit $\lambda + \mu = 1$. Der Zwischenwertsatz liefert uns $c, d \in \mathbb{R}$ mit $a < c < \lambda a + \mu b < d < b$ für die gilt:

$$\frac{f(\lambda a + \mu b) - f(a)}{\lambda a + \mu b - a} = f'(c) \geq f'(d) = \frac{f(b) - f(\lambda a + \mu b)}{b - \lambda a - \mu b}$$

Hieraus folgt:

$$f(\lambda a + \mu b) \leq \lambda f(a) + \mu f(b).$$

Also ist f konvex. □

Für zweimal differenzierbare Funktionen erhalten wir damit sofort:

Korollar 2.51. *Sei $f : I \mapsto \mathbb{R}$ eine zweimal differenzierbare Funktion. Genau dann ist f konvex, wenn $f'' \geq 0$ für alle $x \in I$.*

Abschließend zeigen wir noch einige Propositionen, die im späteren Verlauf nützlich sein werden.

Proposition 2.52. *Seien $f, g : I \mapsto$*
R konvexe Funktionen und sei g monoton steigend. Dann ist $g \circ f$ konvex.

Beweis. Seien $x, y \in$
R und $\lambda, \mu \geq 0$ mit $\lambda + \mu = 1$. Wir haben dann

$$g \circ f(\lambda x + \mu y) = g(f(\lambda x + \mu y)) \leq g(\lambda f(x) + \mu f(y)) \leq \lambda(g \circ f)(x) + \mu(g \circ f)(y)$$

somit ist $g \circ f$ eine konvexe Funktion. □

Korollar 2.53. *Sei h :*
$R \mapsto$
R eine positive Funktion. Ist $log(h)$ konvex, so ist auch h eine konvexe Funktion.

Beweis. Die Behauptung folgt direkt aus Proposition 2.52, denn die Exponential Funktion ist monoton steigend und konvex. Also $h = \exp(\log h)$ konvex. □

Proposition 2.54. *Sei $g : I \mapsto$*
$R_{>0}$ eine positive konkave Funktion, dann ist $f := 1/g$ eine konvexe Funktion.

Beweis. Seien $x, y \in I, \lambda, \mu \geq 0, \lambda + \mu = 1$. Setze $c := g(x)g(y)g(\lambda x + \mu y)$. Nach Voraussetzung ist $c \geq o$. Nun ist

$$\begin{aligned} c(\lambda f(x) + \mu f(y) - f(\lambda x + \mu y)) &= (\lambda g(y) + \mu g(x))g(\lambda x + \mu y) - g(x)g(y) \\ &\geq (\lambda g(y) + \mu g(x))(\lambda g(x) + \mu g(y)) - g(x)g(y) \\ &= \lambda\mu(g(x) - g(y))^2 \\ &\geq 0 \end{aligned}$$

Also haben wir $\lambda f(x) + \mu f(y) \geq f(\lambda x + \mu y)$. □

3 Gitterreduktion und Hermite-Konstante

> Bekanntlich hat Lagrange zuerst gezeigt, dass jede binäre quadratische Form reducir werden kann. Zu leichterer Bezeichnung dieses neuen Verfahren von überraschender Einfachheit, wird es zweckmäßig sein, die Sache in ein geometrisches Gewand zu kleiden
>
> *Über die Reduction der positiven quadratischen Formen*
> P.G.L. Dirichlet

Um 1773 begann Lagrange ein systematisches Studium der binären Formen. Hierbei stellte er zum ersten Mal das Problem, in welcher Weise man aus den Äquivalenzklassen der positiv definiten quadratischen Formen je einen eindeutigen Repräsentanten, eine so genannte reduzierte Form, auswählen kann. Gut 50 Jahre später begann Seeber sich des tenären Falles anzunehmen. Er zeigte dass eine Form reduziert ist, wenn sie gewisse Bedingungen an die Koeffizienten erfüllt. Hierzu musste er zeigen, dass das Produkt der Koeffizienten durch den dreifachen Wert der Determinante nach oben beschränkt ist.

In seinem Werk [13] greift Gauß diese Abschätzung erneut auf und verbessert sie. Abschließend gibt er seiner Abschätzung eine geometrische Deutung; Hierbei taucht, wenn auch versteckt, zum ersten Mal der Begriff des Gitters in der Arithmetik auf. Diese geometrische Deutung greift um 1850 Dirichlet erneut auf und leitet in seiner Abhandlung [9] die Reduktionsbedingungen von Lagrange und Gauß-Seeber nur anhand der Geometrie von Gittern her.

In der Arbeit Dirichlets lassen sich somit schon erste Ansätze der, von Minkowski gut 50 Jahre später entwickelten, *Geometrie der Zahlen* erblicken. Minkowski erkannte auch den Zusammenhang dieser Fragestellungen mit dem im ersten Kapitel dargestellten Problem dichter Gitterpackungen. Eine schöne Darstellung der Reduktionstheorie und ihrer Entwicklung bietet [26].

3.1 Dichteste Kreis-und Kugelpackungen und Minkowskireduktion

Wir hatten im ersten Kapitel gesehen, dass die Basis eines Gitters nicht eindeutig bestimmt ist. Das Problem der Reduktion quadratischer Formen läßt sich also nach der Suche von eindeutigen Basen für Gitter verstehen. Die eingangs erwähnten Ansätze suchen hierbei immer möglichst Basen, die aus

kürzest möglichen Vektoren gebildet werden. Minkowski verallgemeinerte diesen Ansatz und definierte hierzu den Begriff der sukzessiven Minima:

Definition 3.1. Sei $\Lambda \subset \mathrm{E}$ ein Gitter. Für $i = 1, .., n$ bezeichnen wir mit λ_i das i-te sukzessive Minima. Darunter verstehen wir den Minimalen Radius einer Kugel, die i linear unabhänigen Vektoren in Λ enthält.

Das Minimum von Λ entspricht also dem Quadrat des 1. sukzessiven Minimums eines Gitters.

Wir sagen, dass eine Gitterbasis die sukzessiven Minima realisiert, wenn sie so geordnet werden kann, dass $\lambda_1 = \|b_1\|, ..., \lambda_n = \|b_n\|$ gilt.

Wie Dirichlet in seinem Artikel darstellt, entspricht die Reduktion von Lagrange und Gauß-Seeber dem Versuch, eine Basis zu finden, welche die sukzessiven Minima realisiert. Dies gelingt aber nicht immer, wie folgendes Beispiel zeigt.

Beispiel 3.2. Wir betrachten beispielsweise das Gitter im $\mathbb{R}^5$, das von folgender Basis aufgespannt wird:

$$b_1 = (2,0,0,0,0), b_2 = (0,2,0,0,0), b_3 = (0,0,2,0,0), b_4 = (0,0,0,2,0), b_5 = (1,1,1,1,1)$$

man sieht sofort, dass $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 2$ gilt. Dass auch $\lambda_5 = 2$ ist folgt, da $x = (0,0,0,0,2)$ in Λ liegt. Allerdings ist $b_5 \notin < b_1, b_2, b_3, b_4, x >_{\mathbb{Z}}$.

Wir sehen also, dass es im Allgemeinen keine Basis von Vektoren gibt, welche die sukzessiven Minima realisiert. Dirichlet zeigt aber mit elementarer Geometrie, dass dies in Dimension 2 und 3 möglich ist. Gegenbeispiele gibt es erst ab Dimension 5 und es gilt der folgende Satz

Satz 3.3. *Sei $\Lambda \subset \mathbb{R}^n$ ein Gitter. Dann gilt:*

1. *Realisieren $b_1, .., b_r$ die ersten r sukzessiven Minima und ist $r \leq 3$ so lassen sich die $b_1, ..., b_r$ zu einer Basis von Λ ergänzen.*
2. *Λ besitzt stets eine Basis $b_1, ..., b_m$ mit $\|b_j\| = \lambda_j$ für $j \leq \min\{m, 4\}$.*

Beweis. Für einen Beweis verweisen wir auf [20] Theorem 3.32. □

Zu einem Gitter eine Basis zu finden, welche die sukzessiven Minima realisiert, ist nun für die Dimension 2 (Lagrange) und 3 (Seeber, Gauß, Dirichlet) möglich. Wir wollen nun kurz zeigen, wie man mit dieser Reduktion die dichteste gitterförmige Kreispackung finden kann und geben das Ergebnis für die dichteste gitterförmige Kugelpackung in Dimension 3.

Theorem 3.4 (Lagrange). *Im $\mathbb{R}^2$ ist die maximale Dichte einer Gitterpackung*

$$\varrho_{max}(S_2) = \frac{\pi}{2\sqrt{3}}$$

Beweis. Wir wählen für Λ eine Basis der Form b_1, b_2 mit $1 = \|b_1\| \leq \|b_2\|$, $\|b_2 - b_1\| \geq \|b_2\|$ sowie $b_1 \cdot b_2 \geq 0$. Weiterhin haben wir $\|b_2 - b_1\|^2 = \|b_1\|^2 + \|b_2\|^2 - 2(b_1 \cdot b_2)$ Dies formen wir zu $b_1 \cdot b_2 = \frac{\|b_1\|^2}{2} + \frac{\|b_2\|^2}{2} - \frac{\|b_2 - b_2\|^2}{2}$. Da nun $\|b_2 - b_1\|^2 \geq \|b_2\|^2$ nach Voraussetzung, erhalten wir die Abschätzung $b_1 \cdot b_2 \leq \frac{\|b_1\|}{2}^2$. Somit erhalten wir dann für die Determinante:

$$\begin{aligned} \det \Lambda = vol(P(b_1, b_2))^2 = \|b_1 \times b_2\|^2 & \\ = \|b_1\|^2 \|b_2\|^2 - (b_1 \cdot b_2)^2 & \\ \geq \|b_1\|^2 \|b_1\|^2 - \left(\frac{\|b_1\|^2}{2}\right)^2 = \frac{3}{4} & \end{aligned}$$

Weiterhin sieht man, dass Gleichheit genau dann eintritt, wenn $b_2 = b_1$ und $b_1 \cdot b_2 = \frac{1}{2}$ gilt. Das Minimum ist also eindeutig definiert. Es wird vom sogenannten hexagonalen Gitter angenommen. □

Einen ähnlich kurzen Beweis erhält auf diese Weise auch für das folgende Theorem, das die Dichte der absolut extremen Kugelpackungen liefert.

Theorem 3.5 (Gauß). *Die maximale Kugelpackungsdichte für Gitter in* $\mathbb{R}^3$ *ist*

$$\varrho_{max}(S_3) = \frac{\pi}{\sqrt{18}}.$$

Es gibt ein bis auf Ähnlichkeit eindeutig bestimmtes Gitter, welches diese Dichte realisiert.

Beweis. Der Beweis beruht wieder auf der Annahme einer Basis, die die sukzessiven Minima realisiert. Die Rechnung wird hierbei allerdings etwas komplizierter, so dass wir auf [29] Theorem 2.1 verweisen. □

Die beiden Sätze von Lagrange und Gauß hatten in ihrer ursprünglichen Fassung nur die Abschätzung der Determinanten als Aussage und waren nur in der Sprache der quadratischen Formen abgefaßt. Dirichlet führte die Beweise zum ersten Mal ganz in einem geometrischen Gewand (siehe [9]), wobei er allerdings auch noch nicht auf den Aspekt der Kugelpackung einging.

Dieser findet sich zuerst bei Minkowski, der den Ansatz dieser Reduktionsmethode auf höhere Dimensionen verallgemeinerte. Da es ja ab Dimension 5 keine Basen mehr geben muss, welche die sukzessiven Minima realisieren, wählte Minkowski folgende Definition:

Definition 3.6. Sei $\Lambda \subset \mathrm{R}$ ein Gitter. Eine Basis $b_1, ..., b_n$ von Λ heißt Minkowski-reduziert, falls für $i = 1, ..., n$:

1. $\{b_1, ..., b_i\}$ bilden ein primitives System (vgl. Definition 1.11)
2. $\|b_i\| = min\{b \in \mathrm{E} \mid (b_1, ..., b_{i-1}, b) \text{ primitiv}\}$

Geht man induktiv wie im Beweis zu Satz 1.14 vor, erhält man für jedes Gitter eine Minkowski-reduzierte Basis.

Der große Nachteil an Minkowskis Form der Reduktion ist die praktische Durchführung. Es müssen ja in jedem Schritt minimale Vektoren gefunden werden. Das Problem einer Beschränkung der Determinanten läßt sich mit dieser Form der Reduktion nicht praktisch lösen. Im nächsten Abschnitt wollen wir daher die Reduktion im Sinne von Hermite darstellen.

3.2 Hermite-Reduktion und Hermite-Konstante

Im ersten Kapitel hatten wir erwähnt, dass das Volumen der Grundmasche für alle Basen eines gegeben Gitters Λ das selbe ist. Hieraus macht man sich leicht klar, dass die Basisvektoren immer länger werden, je stumpfer bzw. spitzer die Winkel zwischen ihnen sind. Kurze Basen sind dagegen eher orthogonal. Daher ist die Abweichung von der Orthogonalität ein gutes Maß für die Güte der Basis. Wir definieren:

Definition 3.7. Sei $\Lambda \subset \mathrm{E}$ ein Gitter mit Basis $b_1, ..., b_n$. Der Quotient

$$\frac{N(b_1) \cdots N(b_n)}{det\Lambda}$$

wird der orthogonale Defekt genannt.

Zu einer gegebenen Basis $b_1, ..., b_n$ von E existiert nach dem Verfahren von Gram-Schmidt immer auch eine Orthogonalbasis. Diese wird rekursiv wie folgt definiert:

$$b_1^* = b_1$$

$$b_i^* = b_i - \sum_{j=1}^{i-1} \mu_{ij} b^* \text{ für } i = 2, .., n \text{ mit } \mu_{ij} = \frac{b_i \cdot b_j^*}{b_j^* \cdot b_j^*}$$

In folgender Proposition halten wir die wichtigen Eigejnschaften fest:

Proposition 3.8. *Für die Gram-Schmidt reduzierten Vektoren gilt:*

1. $b_i^* \cdot b_j^* = 0$ *für* $i \neq j$
2. $< b_1^*, ..., b_n^* >_{\mathbb{R}} = < b_1, ..., b_n >_{\mathbb{R}}$
3. $det(Gram(b_1, ..., b_n)) = det(Gram(b_1^*, ..., b_n^*)) = N(b_1^*) \cdots N(b_n^*) \leq N(b_1) \cdots N(b_n)$

Beweis. Alle Behauptungen ergeben sich durch Nachrechnen □

Für den Spezialfall, wenn $b_1, ..., b_n$ die Basis eine Gitters Λ bilden, erhalten wir sofort folgende Ungleichung von Hadamard:

Korollar 3.9 (Hadamard Ungleichung). *Sei* Λ *ein Gitter in* E *und* $e_1, ..., e_n$ *n linear unabhängige Vektoren in* Λ. *Es gilt dann die folgende Abschätzung:*

$$det(\Lambda) \leq N(e_1) \cdots N(e_n)$$

Beweis. Es sei $\Lambda^{'}$ das von den Vektoren $e_1, ..., e_n$ aufgespannte Gitter. Mit der vorherigen Proposition haben wir $det(\Lambda^{'}) \leq N(e_1) \cdots N(e_n)$. $\Lambda^{'}$ ist ein Untergitter von Λ. Wir haben also:

$$det(\Lambda) = \frac{det(\Lambda^{'})}{[\Lambda : \Lambda']}$$

und somit $\det(\Lambda) \leq \det(\Lambda^{'}) \leq N(e_1) \cdots N(e_n)$. □

Im letzten Abschnitt hatten wir die Reduktion von Minkowski, als Verallgemeinerung des Ansatzes von Lagrange und Gauß eingeführt. Fast 20 Jahre nach Gauß beschäftigte sich der junge Charles de Hermite mit dem Problem, ob auch in höheren Dimensionen, die Determinante einer quadratischen Form in einer Weise beschränkt ist.

Es gelang ihm zu beweisen, dass die Determinate stets beschränkt ist. Sein Ansatz wurde von Korkine und Zolotareff praktisch dazu verwendet, eine andere Form der Reduktion zu definieren. Diese nutzt die eingangs gemachte Bemerkung über die Orthogonalität aus. Zunächst definieren wir, was wir unter Hermite-Reduktion verstehen wollen, dabei gelten die Bezeichnungen des Gram-Schmidt-Verfahrens.

Definition 3.10. Die Basis $b_1, ..., b_n$ eines Gitter Λ heißt Hermite-reduziert, wenn gilt:

1. $|\mu_{ij}| \leq \frac{1}{2}$
2. $N(b^*_{i+1}) \geq \frac{3}{4} N(b^*_i)$

Korkine und Zolotareff konnten ein algorithmisches Verfahren finden, dass für jedes Gitter eine Hermite-reduzierte Basis findet:

Theorem 3.11 (Korkine-Zolotareff). *Jedes Gitter Λ besitzt eine Basis, die im Sinne von Hermite reduziert ist.*

Beweis. Siehe [29] Theorem 2.7 □

Hermite nutzte seine Reduktionstheorie nur theoretisch, um folgendes Korollar zu gewinnen:

Korollar 3.12 (Hermite-Ungleichung). *Jedes Gitter besitzt eine Basis $(b_1, ..., b_n)$, für die gilt:*

$$\prod_{i=1}^{n} N(b_i) \leq \frac{4}{3}^{\frac{n(n-1)}{2}} \det \Lambda$$

und somit

$$N(\Lambda) \leq (\frac{4}{3})^{(n-1)/2} \det(\Lambda)^{1/n}$$

Beweis. Sei $b_1, ..., b_n$ eine Basis, die im Sinne von Korkine-Zolotareff/Hermite reduziert ist. Wir haben

$$\det(\Lambda) = \prod_{i=1}^{n} N(b^*_i)$$

Weiterhin haben wir auf Grund des Gram-Schmid-Verfahren $b_i = b_i^* + \sum_{j=1}^{i-1} \mu_{ij} b_i^*$ und somit $b_i^2 = b_i^{*\,2} + \sum_{j=1}^{i-1} \mu_{ij}^2 b_i^{*\,2}$. Hieraus und aus der Definition der Reduktion leiten wir

$$N(b_i) \leq (\frac{4}{3})^i N(b_i^*).$$

Hieraus folgt die erste Behauptung sofort. Wir haben nun $N(\Lambda)^n \leq \prod_{i=1}^{n} (b_i \cdot b_i) \leq (\frac{4}{3}) \det \Lambda$. □

Wir hatten im ersten Kapitel gezeigt, dass $\det \Lambda$ und $N(\Lambda)$ Invarianten der Isometrieklassen sind. Wir definieren nun die Hermite-Invariante:

Definition 3.13. Die Hermite-Invariante eines Gitters Λ ist definiert als

$$\gamma(\Lambda) := \frac{N(\Lambda)}{\sqrt[n]{det(\Lambda)}}.$$

Das Supremum über alle Hermite-Zahlen in Dimension n bezeichnen wir als *Hermite-Konstante* γ_n, also

$$\gamma_n := sup_{\Lambda \subset \mathbb{R}^n} \gamma(\Lambda).$$

Ein Gitter Λ mit $\gamma(\Lambda) = \gamma_n$ heißt *absolut extremes* Gitter.

Wie man leicht nachrechnet, haben ähnliche Gitter die selbe Hermite-Zahl. Wir notieren dies in folgender Bemerkung.

Bemerkung 3.14. Seien Λ und Λ^* zwei ähnliche Gitter, dann haben sie die selbe Hermite-Invariante.

Für die Dimensionen 2 und 3 waren die Hermite-Konstanten schon von Lagrange und Gauß errechnet worden (vgl. Satz 3.4, 3.5). Den Fall der Dimension 4, 5 und 6 konnten Korkine und Zolotareff (1872-1877) lösen und 1925-1929 löste Blichefeldt die Dimensionen 7 und 8. Wir zeigen diese in nachfolgender Tabelle:

$$\begin{array}{llll} \gamma_1 = 1 & \gamma_2 = \sqrt{\frac{4}{3}} & \gamma_3 = \sqrt[3]{2} & \gamma_4 = \sqrt{2} \\ \gamma_5 = \sqrt[5]{8} & \gamma_6 = \sqrt[6]{\frac{64}{3}} & \gamma_7 = \sqrt[7]{64} & \gamma_8 = 2 \end{array}$$

Im Jahre 2004 konnten Henry Cohn und Abhinav Kumar zeigen, dass in Dimension 24 $\gamma_{24} = 4$ gilt.

Minkowski scheint der erste gewesen zu sein, der einen Zusammenhang des Problems der Hermite-Konstante mit dem Problem der dichtesten Kugelpackung gesehen hat.

Bemerkung 3.15. Es ist $\varrho(\Lambda) = \omega_n(\frac{\gamma_n}{4})^{n/2}$(vgl.1.16), wobei ω_n das Volumen der n-dimensionalen Einheitskugel ist. Wir haben $\omega_n = \frac{\pi^{n/2}}{\Gamma(n/2+1)}$.

Im letzten Kapitel hatten wir Minkowskis Arbeiten über konvexe Mengen angesprochen. Für die Theorie der Gitter ist vor allem auch Minkowskis erster Satz von zentraler Bedeutung. David Hilbert schreibt darüber in der oben schon erwähnten Gedächtnisrede:

„ Dieser Beweis eines tiefliegenden zahlentheoretischen Satzes ohne rechnerische Hilfsmittel und wesentlich nur auf Grund einer geometrischen Betrachtung, ist eine Perle Minkowskischer Erfindungskunst."

Hierbei bezieht sich Hilbert auf folgenden Satz

Satz 3.16. *Sei $\Lambda \subset \mathrm{E}$ ein vollständiges Gitter und sei $S \subset \mathrm{E}$ eine kompakte, konvexe und zentralsymmetrische Menge, mit $volS \geq 2^n\sqrt{\det \Lambda}$, so trifft S das Gitter Λ in mindestens 3 Punkten.*

Beweis. Für den Beweis verweisen wir auf [20] Satz 4.2. □

Mit diesem Satz erhalten wir schon eine einfache Abschätzung für das Minimum eines Gitters es gilt nämlich

Korollar 3.17. *Sei $\Lambda \subset \mathrm{E}$. Es ist $N(\Lambda) \leq n \det(\Lambda)^{1/n}$.*

Beweis. Sei $r := \sqrt{n(\det \Lambda)^{1/n}}$ Für das Volumen der Kugel um 0 mit Radius r gilt:

$$vol(B_r(0)) = \omega_n r^n = \frac{\pi^{n/2}}{\Gamma(n/2+1)}(n(\det \Lambda)^{1/n})^{n/2} \geq 2n.$$

Somit enthält $B_n(0)$ mindestens zwei von 0 verschiedene Gitterpunkte. □

Hieraus erhalten wir also auch $\gamma_n \leq n$. Die Betrachtung des Kugelpackungsproblems und der in Bemerkung 3.15 dargestellte Zusammenhang zwischen dem Problem der dichtesten Kugelpackung und der Hermite-Konstante, ermöglichte Minkowski eine noch bessere Abschätzung.

Satz 3.18. *Es gilt*

$$\gamma_n \leq \frac{4}{\pi}\Gamma(n/2+1)^{\frac{2}{n}} \approx 0,2342n + O(1)$$

Beweis. Sei Λ ein Gitter und wir setzen $\det \Lambda = 1$. Offensichtlich ist $\varrho(\Lambda) \leq 1$. Also ist nun $volB_r(0) \geq 1$, so ist $\lambda_1 \leq 2r$. Das minimale r, für das $volB_r(0) \geq 1$ ist, ist somit $\tilde{r} = \frac{\Gamma(n/2+1)^{1/n}}{\sqrt{\pi}}$. Und somit

$$\frac{N(\Lambda)}{\det \Lambda} \leq 4\tilde{r}^2 = \frac{4}{\pi}\Gamma(n/2+1)^{\frac{2}{n}}.$$

□

Diese Einsicht wirkt trivial, steckt hinter ihr letztlich doch nichts anderes als die Tatsache, dass eine Dichte vernünftigerweise stets maximal 1 sein kann. Jedoch war Minkowski der erste, der diesen Zusammenhang erhellt hatte. Wir wollen dieses Kapitel mit der Würdigung der Arbeit durch Hermite beschließen. Dieser gehörte als Minkowski seine Geometrie der Zahlen begann schon zu den Altmeistern der Mathematik und würdigte Minkowskis Fortsetzung seiner Jugendarbeit in einem Brief an Minkowski aus dem Jahre 1892 wie folgt:

„Je me sens rempli d'étonnement et de plaisir devant vos principes et vos résultats, ils m'ouvrent comme un monde arithmétique entièrement nouveau, où les questions fondamentales de notre sience sont traitées avec éclatant succès auquel tous les géomètres rendront hommage. Vous voulez bien [...] rapporter à mes anciennes recherches le point de départ de vos beaux travaux, mais vous les avez tant dépassées qu'elles ne gardent plus d'autre mérite que d'avoir ouvert la voie dans laquelle vous êtes entré"

4 Extreme Gitter

> Chaque forme quadratique extrême est determine par la valuer de son minimum et par touts representations de ce minimum. [...] Ce n'est qt'à Partie des formes positives à six variables que j'ai rencounter des formes quadratiques positives qui jouissent la propriété et ne sont pas des formes extrêmes.
>
> *Nouvelles applications des paramètres continus à la théorie des formes quadratiques*
> GEORGE VORONOÏ

Wir hatten im ersten Kapitel dargestellt, wie man zu einem Gitter eine Kugelpackung definiert und wie hierfür die Packungsdichte definiert ist. Ein Gitter, dessen Packungsdichte bei kleinen Veränderungen der Gitterbasis nur abnimmt, wurde dabei als ein extremes Gitter bezeichnet.

Über die Reduktionstheorie der Quadratischen Formen in Kapitel 3, hatten wir die Hermite-Konstante eingeführt und erwähnt, dass diese Invariante der Ähnlichkeitsklassen proportional zur Packungsdichte ist. Zum Ende des 19. Jahrhunderts begannen die russischen Mathematiker Korkine und Zolotareff, diese maximale Hermite-Konstante für die Dimensionen 4 und 5 zu finden. Dazu betrachten wir die Zuordnung $\Lambda \mapsto \gamma(\Lambda)$ als eine Funktion auf den Ähnlichkeitsklassen von Gittern respektive Quadratischen Formen. Um das globale Maximum dieser Hermite-Funktion zu bestimmen, begannen Korkine und Zolotareff, alle lokalen Maxima zu charakterisieren. Solche Gitter bezeichnen wir wieder als extrem und definieren diesen Begriff nun noch genauer.

Definition 4.1. Ein Gitter Λ in E heißt extrem, wenn es eine Umgebung $\mathcal{U} \subset \mathrm{GL}_n(\mathbb{R})$ der Identitätsabbildung Id gibt, mit

$$\gamma(\Lambda) > \gamma(u(\Lambda)) \ \forall \, \mathrm{Id} \neq u \in \mathcal{U}.$$

Im vorliegenden Kapitel stellen wir nun die von Korkine und Zolotareff begonnene Charakterisierung dieser extremen Gitter vor.

4.1 Symmetrische Endomorphismen

In diesem ersten Abschnitt stellen wir einige Eigenschaften von symmetrischen Endomorphismen dar.[1]Im Folgenden bezeichnen wir den Raum der Endomorphismen vom E mit $End(\mathrm{E}) = \{\phi : \mathrm{E} \mapsto \mathrm{E}\ \mathbb{R}\text{-linear}\}$. Durch die Wahl einer Basis in E erhalten wir einen Isomorphismus zwischen $End(\mathrm{E})$ und dem Raum der $n \times n$ Matrizen über $\mathbb{R}$, der einem Endomorphismus u die Matrix $Mat(u)$ zuordnet.

Definition 4.2. Für ein $u \in End(\mathrm{E})$ bezeichnet u^t den transponierten Endomorphismus. Dieser ist durch die Gleichung

$$(u(x), y) = (x, u^t(y)) \ \forall\ x, y \in \mathrm{E}$$

bestimmt.

Ein $u \in End(\mathrm{E})$ bezeichnen wir als symmetrisch, wenn $u = u^t$ gilt und wir notieren $\mathrm{End}^s(\mathrm{E})$ für den Unterraum dieser symmetrischen Endomorphismen.

Der Name symmetrisch erklärt sich durch folgende Proposition, in der wir auch noch weitere Eigenschaften zusammenfassen:

Proposition 4.3. *1. Wir haben $u \in \mathrm{End}^s(\mathrm{E})$ genau dann wenn $Mat(u) \in \mathrm{Sym}_n := \{X \in \mathbb{R}n \times n \mid X = X^t\}$*

2. Ist $u \in \mathrm{End}^s(E)$, so sind die Eigenwerte von u alle reell.

3. Zu jedem $u \in \mathrm{End}^s(E)$ gibt es eine orthogonale Basis bestehend aus den Eigenvektoren von u.

4. Ist $A \in \mathrm{Sym}_n$, so ist A diagonalisierbar.

Beweis. 1. Sei $A = Mat(u)$. Die Behauptung ergibt sich sofort aus der Äquivalenz der folgenden drei Aussagen:

$$(x, u(y)) = (u(x), y)\ \forall x, y \in \mathrm{E}$$

$$(e_i, u(e_j)) = (u(e_i, e_j) \forall i, j = 1, .., n$$

$$a_{ij} = a_{ji}\ \forall i, j = 1, ..., n$$

2. Wie betrachten $\mathbb{C} \otimes \mathrm{E}$. Dabei wird u zu einem hermiteschen Endomorphismus und es gilt dann für $\lambda \in \mathbb{C}$ einen Eigenwert von u zum Eigenvektor x:

$$\lambda \|v\|^2 = (\lambda x, \bar{x}) = (u(x), \overline{(x)}) = (x, \overline{u(x)}) = (x, \overline{\lambda x}) = \overline{\lambda}\|x\|^2$$

also $\lambda \in \mathbb{R}$.

3. Auf Grund der letzten Behauptung ist u diagonalisierbar. Es bleibt zu zeigen, dass die Eigenräume paarweise orthogonal sind. Sei $u(x_i) = \lambda_i x_i$ für $i = 1, 2$ und $\lambda_1 \neq \lambda_2$. Dann ist

$$(\lambda_1 x_1, x_2) = (u(x_1), x_2) = (x_1, u(x_2)) = (\lambda_2 v_1, v_2),$$

woraus $(x_1, x_2) = 0$ folgt.

[1] mit den Bezeichnungen folgen wir Martinet

4. Wir schreiben in die Spalten der Matrix P die Vektoren einer Orthonormalbasis $(x_i)_{i=1,..,n}$, zu den Eigenwerten von A. Es ist dann:

$$P^t A P e_i = P^t A x_i = \lambda_i P^t x_i = \lambda e_i$$

□

Für eine $n \times n$ Matrix $A = (a_{ij})$ ist die Spur gegeben durch

$$Tr(A) := \sum_{i=1}^{n} a_{ii},$$

ebenso für ein $u \in End(\mathrm{E})$

$$Tr(u) := \sum_{i=1}^{n} (u(e_i), e_i)$$

Mit Hilfe der Spur definieren wir aus $\mathrm{End}^s(\mathrm{E})$ bzw. Sym_n ein Skalarprodukt:

Definition 4.4.

$$\begin{aligned} \langle \cdot, \cdot \rangle : \mathrm{End}^s(\mathrm{E}) \times \mathrm{End}^s(\mathrm{E}) &\longrightarrow \mathbb{R} \\ (u, v) &\longrightarrow \langle u, v \rangle := Tr(uv) \end{aligned}$$

ist ein Skalarprodukt auf $\mathrm{End}^s(\mathrm{E})$.
In gleicher Weise ist durch

$$\begin{aligned} \langle \cdot, \cdot \rangle : \mathrm{Sym}_n \times \mathrm{Sym}_n &\longrightarrow \mathbb{R} \\ (A, B) &\longrightarrow \langle A, B \rangle := Tr(AB) \end{aligned}$$

ein Skalarprodukt auf Sym_n gegeben. Dieses Skalarprodukt auf beiden Räumen wird als Voronoi-Skalarprodukt bezeichnet. Die hieraus resultierende Norm auf den beiden Räumen wird auch Frobenius Norm genannt. Wir schreiben daher $\|A\|_F := \sqrt{\langle A, A \rangle}$.

Bemerkung 4.5. Seien $A, B \in \mathrm{Sym}_n$. Dann ist $Tr(AB) = \sum_{i,j=1}^{n} a_{ij} b_{ij}$. Das Voronoi-Skalarprodukt stimmt also mit dem kanonischen Skalarprodukt des $\mathbb{R}^{n^2}$ überein und ist somit positiv definit.

Eine wichtige Rolle spielen die sogenannten Projektionsmatrizen. Diese definieren wir durch:

$$P_x := xx^t = (x_i x_j)_{i,j} = \begin{bmatrix} x_1^2 & \cdots & x_1 x_n \\ \vdots & \ddots & \vdots \\ x_n x_1 & \cdots & x_n^2 \end{bmatrix}$$

für ein $0 \neq x \in \mathrm{E}$.

Des Weiteren bezeichnen wir für einen Vektor $0 \neq x \in \mathrm{E}$ mit p_x die Orthogonalprojektion auf die Gerade, die von x erzeugt wird. Das ist eine Abbildung

$$\begin{aligned} p_x : \mathrm{E} &\longrightarrow \mathrm{E} \\ y &\longrightarrow \rho_x(y) := \frac{x \cdot y}{x \cdot x} x. \end{aligned}$$

Offenbar ist $P_x \in \mathrm{Sym}_n$. Außerdem gilt $x_i x_j = e_i \cdot (\|x\|^2 p_x(e_j))$, also $P_x = Mat(\|x\|^2 \rho_x)$. Somit ist also auch $p_x \in \mathrm{End}^s(\mathrm{E})$ gezeigt. Während die Orthogonalprojektionen p_x nur von der Richtung des Vektors x und nicht von dessen Länge abhängen, gilt dies für die Projektionsmatrizen nicht. Für den weiteren Verlauf wichtige Eigenschaften dieser Orthogonalprojektionen beweisen wir in der nachfolgenden Proposition:

Proposition 4.6. *Für alle* $u \in \mathrm{End}^s(\mathrm{E})$ *und* $0 \neq x \in \mathrm{E}$ *gilt:*

$$< u, p_x >= \frac{u(x) \cdot x}{x \cdot x}.$$

Darüber hinaus erzeugen die Orthogonalprojektionen auf alle Linien von E *den Raum* $\mathrm{End}^s(\mathrm{E})$.

Beweis. $< u, p_x >= Tr(up_x) = \sum_{i=1}^n u(p_x(e_i)) \cdot e_i = \frac{1}{x \cdot x} \sum_{i=1}^n u((x \cdot e_i)x) \cdot e_i = \frac{1}{x \cdot x} \sum_{i=1}^n x_i u(x) \cdot e_i = \frac{u(x) \cdot x}{x \cdot x}$
Sei W der von den ρ_x mit $0 \neq x \in \mathbb{R}^n$ erzeugte Unterraum. Sei $u \in W^\perp$. Wir haben dann

$$\frac{u(x) \cdot x}{x \cdot x} = \langle u, \rho_x \rangle = 0 \text{ für alle } x \in \mathbb{R}^n.$$

Somit verschwinden alle Eigenwerte von u und folglich gilt $u = 0$. □

Bemerkung 4.7. In Matrixschreibweise erhalten wir

$$\langle A, P_x \rangle = Tr(Axx^t) = Tr(x^t Ax) = x^t Ax \in \mathbb{R}.$$

Für den weiteren Verlauf der Arbeit schreiben wir $\rho_x := N(x)p_x$.

Unter den symmetrischen Endomorphismen interessieren uns im Folgenden vor allem die positiv definiten. Diese sind wie folgt definiert

Definition 4.8. Ein $u \in \mathrm{End}^s(\mathrm{E})$ heißt positiv, wenn alle Eigenvektoren nicht negativ sind, definit, wenn keiner der Eigenwerte null ist, und somit positiv definit, wenn die Eigenwerte echt positiv sind. Mit $\mathrm{End}^{s+}(\mathrm{E})$ bzw. $\mathrm{End}^{s++}(\mathrm{E})$ bezeichnen wir die positiven bzw. positiv definiten Endomorphismen.

Wie aus der linearen Algebra bekannt ist, gibt er zu einem positiven symmetrischen Endomorphismus u stets eine eindeutig definierte positive Wurzel:

Satz 4.9. *Zu einem* $u \in End(\mathrm{E})^+$ *gibt es ein eindeutig bestimmtes* $u^{\frac{1}{2}} \in End(\mathrm{E})^+$ *mit:*

$$(u^{\frac{1}{2}})^2 = u.$$

Beweis. Ein Beweis findet sich beispielsweise in [16] Satz 10.13 □

In folgendem Satz geben wir äquivalente Bedingungen eine symmetrische Matrix positiv ist. Ein Beweis hierzu findet sich in den gängigen Büchern zur Linearen Algebra.

Satz 4.10. *Für eine symmetrische Matrix A sind die folgenden Aussagen äquivalent:*

1. *A ist positiv*
2. *$x^t Ax \geq 0$ für alle $x \in \mathbb{R}^n$.*
3. *Es gibt $L \in \mathbb{R}^{n \times n}$ mit $A = LL^t$ (Cholesky-Zerlegung)*

Beweis. Siehe beispielswiese [16] Satz 10.15 □

Die Summe zweier positiver Endomorphismen ist wieder positiv, sowie auch das Produkt eines positiven Endomorphismuses mit einem nicht negativen Skalar. Die Menge der positiven Endomorphismen bilden also einen Kegel, den wir mit $\mathcal{P}_n$ bezeichnen. Für $n = 2$ hat dieser Kegel die Form eines Rotationskegels. Für $n > 2$ ist die Situation nicht mehr ganz so schön.

Durch das Voronoi-Produkt haben wir einen Euklidischen Vektorraum und können den Dualen-Kegel $\mathcal{P}_n^*$ bilden. Es zeigt sich, dass $\mathcal{P}_n$ selbstdual ist, denn es gilt:

Satz 4.11 (Féjer). *Ein symmetrischer Endomorphismus u bzw eine symmetrische Matrix A ist positiv, genau dann wenn $\langle u, v \rangle \geq 0$ für alle $v \in End^{s+}(\mathrm{E})$ bzw. $\langle A, B \rangle \geq 0$ für alle positiven symmetrischen B.*

Beweis. Die beiden Aussagen folgen auseinander. Wir müssen also nur eine beweisen:
Seien A, B zwei positive Matrizen. Wir haben dann:

$$\begin{aligned}\langle A, B \rangle = Tr(uv) &= Tr(u^{\frac{1}{2}} u^{\frac{1}{2}} v^{\frac{1}{2}} v^{\frac{1}{2}} \\ &= Tr(u^{\frac{1}{2}} v^{\frac{1}{2}} u^{\frac{1}{2}} v^{\frac{1}{2}}) \\ &= \langle u^{\frac{1}{2}} v^{\frac{1}{2}}, u^{\frac{1}{2}} v^{\frac{1}{2}} \rangle \\ &= \| u^{\frac{1}{2}} v^{\frac{1}{2}} \|^2 \\ &\geq 0.\end{aligned}$$

Sei nun A eine symmetrische Matrix, so dass $\langle A, B \rangle \geq 0$ für alle $B \in \mathrm{Sym}_n^+$. Für ein $x \in \mathrm{E}$ sind die $P_x \in \mathrm{Sym}_n^+$. Also gilt:

$$0 \leq \langle A, P_x \rangle = x^t Ax.$$

Also ist $A \in \mathrm{Sym}_n^+$. □

Der Kegel $\mathcal{P}_n$ ist nicht endlich erzeugt. Man kann diesen Kegel jedoch durch endlich erzeugte Kegel $\mathcal{P}_n$ überdecken. Diese Kegel werden mit Hilfe der perfekten Formen gebildet, welchen wir uns nun zuwenden werden.

4.2 Perfektion

In diesem Abschnitt widmen wir uns den sogenannten perfekten Formen, bzw. Gittern. Die Eigenschaft der Perfektion geht auf das Werk von Korkine und Zolotareff zurück. Jedoch war es Voronoi, der in seiner Arbeit von 1908, dieser Eigenschaft den Namen „parfait“ gab. Wir definieren diesen zentralen Begriff wie folgt:

Definition 4.12. Sei $S \neq \emptyset$ eine endliche Familie von Vektoren in E.

1. S heißt perfekt, wenn $\{\rho_x \mid x \in S\}$ den Raum $End^s(\mathrm{E})$ aufspannt.
2. Ein Gitter respektive eine Quadratische Form heißt perfekt, wenn die Minimalvektoren eine perfekte Familie bilden.

Korkine und Zolotareff stellten in Ihrem Werk fest, dass eine extreme Quadratische Form eindeutig festgelegt ist, durch die Auswertung der Form auf ihren Minimalvektoren. Voronoi gab dieser Eigenschaft dann den Namen „Perfection“.

Im Folgenden machen wir klar, dass diese Eigenschaft mit der oben gegebenen Definition einer perfekten Familie übereinstimmt.

Theorem 4.13 (Korkine, Zolotareff). *Eine perfekte quadratische Form ist eindeutig bestimmt durch ihr Minimum und die Menge der Vektoren, die dieses Minimum realisieren.*

Beweis. Seien Q_1 und Q_2 zwei perfekte Formen mit dem selben Minimum m und gleicher Menge an Minimalvektoren S. Wir betrachten die Form $Q := Q_1 - Q_2$ deren Gram-Matrix wir mit A bezeichnen. Wir haben dann für $x \in S$:

$$0 = x^{tr}Ax = Tr(x^{tr}Ax) = Tr((x^{tr}x)A) = Tr(P_xA),$$

also ist $A \in < P_x \mid x \in S >^{\perp}$. Die Matrizen P_x spannen $\mathrm{Sym}_n(\mathrm{E})$ auf, so dass $A = 0$ gelten muss. □

Im ersten Kapitel wurde dargelegt, was man unter einem ganzen Gitter versteht. Diese ganzen Gitter spielen eine besondere Rolle, wie wir aus nachfolgendem Korollar zu Theorem 4.13 folgern können.

Korollar 4.14. *Ein perfektes Gitter Λ ist proportional zu einem ganzen Gitter*

Beweis. Wir können ohne Einschränkung annehmen, dass $N(\Lambda) = 1$ gilt. Sei A die Gram-Matrix einer beliebigen Basis von Λ. A ist nun eindeutig bestimmt durch die minimal Vektoren in $\mathbb{Z}^n$. Die Koeffizienten von A sind also die Lösung eines Systems linearer Gleichungen mit ganzzahligen Koeffizienten und somit sind sie rational. Sei nun $m \in \mathbb{Q}$ so gewählt, dass gilt $mA \in \mathbb{Z}^{n\times n}$, dann ist $\sqrt{m}\Lambda$ ein ganzes Gitter. □

Wir erinnern daran, dass zwei Gitter ähnlich sind, wenn die beiden Gitter äquivalent sind, nachdem sie auf die selbe Norm transformiert wurden. Der nachfolgende Satz garantiert die Endlichkeit der Anzahl von Ähnlichkeitsklassen perfekter Gitter. Es ist also möglich alle perfekten Gitter eine gegebenen Dimension n bis auf Ähnlichkeit zu klassifizieren.

Theorem 4.15 (Voronoi). *Die Anzahl der Ähnlichkeitsklassen von perfekten Gittern einer gegebenen Dimension n ist endlich.*

Bevor wir diese Aussage beweisen, zeigen wir im folgenden Lemma, dass es in einem perfekten Gitter stets n linear unabhängige Minimalvektoren gib:

Lemma 4.16. *Sei Λ ein perfektes Gitter, dann gilt:*

$$\mathrm{E} =< x \mid \ x \in S(\Lambda) >_{\mathbb{R}} .$$

Beweis. Sei $y \in \mathrm{E}$ orthogonal zu allen $x \in S(\Lambda)$. Die ρ_x spannen $End^s(\mathrm{E})$ auf. Es gibt also auch eine Darstellung

$$Id = \sum_{x \in S(\Lambda)} \lambda_x \rho_x .$$

Hieraus folgern wir nun:

$$y = Id(y) = \sum_{x \in S(\Lambda)} \lambda_x \frac{x \cdot y}{N(x)} x = 0$$

□

Nun wenden wir uns dem eigentlichen Beweis des Endlichkeitssatzes zu:

Beweis. Sei Λ ein perfektes n-dimensionales Gitter der Norm 1. Da wir n linear unabhängige Vektoren mit Norm 1 haben, erhalten wir mit der Ungleichung von Hadamard (3.9):

$$\det(\Lambda) \leq 1.$$

Auf Grund der Hermite-Ungleichung(3.12) wissen wir nun, dass es eine Gitterbasis $(b_1, ..., b_n)$ von Λ gibt, mit

$$N(b_1) \cdots N(b_n) \leq \frac{4}{3}^{n(n-1)/2}$$

Da die Norm dieser n Vektoren jeweils mindestens 1 sein muss, gibt es eine Konstante $c_n := (\frac{4}{3})^{n(n-1)/2}$ mit $N(b_i) \leq c_n \ \forall \ i = 1, ..., n$.
Für ein $x = \sum_i \beta_i b_i \in \Lambda$ haben wir somit nach der Cramerschen Regel:

$$|\beta_i|^2 = \Big(\frac{det(b_1, ..., b_{i-1}, x, b_{i+1}, ..., b_n)}{det(b_1, ..., b_n)}\Big)^2$$

Mit der Hadamard-Ungleichung und den obigen Überlegungen erhalten wir für den Zähler

$$det(b_1, ..., b_{i-1}, x, b_{i+1}, ..., b_n)^2 \leq c^{n-1} N(x).$$

Aus der Hermite-Ungleichung haben wir

$$\frac{1}{det(\Lambda)} = \frac{1}{det(b_1, .., b_n)^2} \leq \gamma_n^n .$$

Insgesamt lassen sich die Komponenten von x wie folgt abschätzen:

$$|\beta_i|^2 \leq \gamma_n^n c_n^{n-1} N(x).$$

Insbesondere sind also die Komponenten der Minimalvektoren nach oben beschränkt, also eine beschränkte Teilmenge von $\mathbb{Z}^n$. Somit kann es nur endlich viele Systeme von möglichen Minimalvektoren perfekter Gitter geben. Da perfekte Gitter aber genau diejenigen sind, die eindeutig durch ihre Minimalvektoren bestimmt sind, folgt hieraus, die Endlichkeit der Ähnlichkeitsklassen perfekter Gitter. □

4.3 Eutaxie

Korkine und Zolotareff hatten gesehen, dass extreme Gitter immer auch perfekt sind. Voronoi fand allerdings in Dimension 6 das erste Beispiel für ein Gitter, das zwar perfekt, jedoch nicht extrem ist.[2] Die Eigenschaft der Perfektion ist zwar notwendig, jedoch nicht hinreichend für Extremität eines Gitter.

Voronoi fand noch eine weitere wichtige Eigenschaft extremer Gitter, die von Coxeter gut 50 Jahre später den Namen *Eutaxie* bekam. Wir werden später sehen, dass Gitter extrem sind, wenn sie perfekt und eutaktisch sind. Doch beginnen wir zuerst mit der Definition:

Definition 4.17. Sei $S \subset \mathrm{E}$ eine endliche Familie von Vektoren. Wir sagen dann, dass S schwach eutaktisch ist, wenn es eine Relation der Form:

$$\mathrm{Id} = \sum_{x \in S} \lambda_x \rho_x \text{ wobei } \lambda_x \in \mathbb{R}$$

auf $End^s(\mathrm{E})$ gibt.
Weiterhin nennen wir S

- semi-eutaktisch, wenn $\lambda_x \geq 0 \ \forall x \in S$ gewählt werden können.
- eutaktisch, wenn $\lambda_x > 0 \ \forall x \in S$ gewählt werden können.
- stark eutaktisch, wenn zusätzlich alle λ_x gleich gewählt werden können.

Ein Gitter heißt (semi/stark-) eutaktisch, wenn die Menge der Minimalvektoren (semi/stark-) eutaktisch ist.

Der aus dem Griechischen entlehnte Begriff „$\epsilon\upsilon\tau\alpha\xi\eta$“ bedeutet auf Deutsch in etwa „gut geordnet“ . Es zeigt sich, dass der Begriff einer „guten Ordnung“ besonders auf die Minimalvektoren stark eutaktischer Gitter zutrifft. Diese bilden nämlich sogenannte $2-$Designs (vgl. [24]) und haben somit eine regelmäßige Struktur.

Folgende Proposition gibt weitere mögliche Definitionen:

Proposition 4.18. *Sei Λ ein Gitter mit Gram-Matrix A, dann sind folgende Aussagen äquivalent:*

1. *Λ ist schwach eutaktisch mit den Eutaxiekoeffizienten λ_x.*

[2] siehe dazu das Zitat zu Beginn des Kapitels

2. *Für alle* $y, z \in \mathrm{E}$ *gilt:* $(y,z) = \sum_{x \in S(\Lambda)} \lambda_x (y,x)(z,x)$

3. *Für alle* $y \in \mathrm{E}$ *gilt:*$N(y) = \sum_{x \in S(\Lambda)} \lambda_x (y,x)^2$

4. $A^{-1} = \sum_{x \in S(\Lambda)} \lambda_x P_x$

Beweis. „1. $\Rightarrow$ 2. " Wir haben eine Darstellung der Form $\mathrm{Id} = \sum_{x \in S(\Lambda)} \lambda_x \rho_x$ Also $y = \mathrm{Id}(y) = \sum_{x \in S(\Lambda)} \lambda_x N(x) p_x(y) = \sum_{x \in S(\Lambda)} \lambda_x (x,y)$. Skalare Multiplikation auf beiden Seiten mit z liefert das gewünschte.

„2. $\Leftrightarrow$ 3. " Die Äquivalenz von 2. und 3. ergibt sich aus dem Zusammenhang von Bilinearformen und Quadratischen Formen.

„2. $\Rightarrow$ 4. " Setzt man $y = e_i$ und $z = e_j$ folgt $e_i \cdot e_j = a^*_{ij} = \sum_{x \in S(\Lambda)} \lambda_x (e_i \cdot x)(e_y \cdot y)$, wobei a^*_{ij} die Komponenten von A^{-1} sind.

„4. $\Rightarrow$ 1. " Es sei 4. erfüllt und wir betrachten $y = \sum \beta_i e^*_i$ und $z = \sum \mu_j e^*_j$. Wir haben dann

$$\begin{aligned}(y,z) = \sum_{i,j} \beta_i \mu_j \sum_{x \in S(\Lambda)} \lambda_x (x, e^*_i)(x, e^*_j) &= \sum_{x \in S(\Lambda)} \lambda_x \big((x, \sum_i \beta_i e^*_i)\big)\big((x, \sum_j \mu e^*_j)\big) \\ &= \sum_{x \in S(\Lambda)} \lambda_x (x,y)(x,z)\end{aligned}$$

Hieraus erhalten wir die Relation $y = \sum_{x \in S(\Lambda)} \lambda_x (x,y)$, da die obige Gleichung für alle $z \in E$ gültig ist. □

Wir wollen nun die Eutaxie geometrisch veranschaulichen und dazu die in Kapitel 3 eingeführten Begriffe der Kegel verwenden. Dazu definieren wir:

Definition 4.19. Für ein Gitter Λ nennen wir den endlich erzeugten Kegel

$$D_\Lambda := cone\{\rho_x \mid \ x \in S(\Lambda)\}$$

den *Voronoi Bereich* des Gitters Λ.

Bemerkung 4.20. Der duale Kegel zum Voronoi-Bereich ist

$$D^*_\Lambda = \{v \in End^s(\mathrm{E}) : \ \langle v, u \rangle \geq 0 \, \forall u \in D_\Lambda\} = \{v \in End^s(\mathrm{E}) : \ \langle v, \rho_x \rangle \geq 0 \ \forall x \in S(\Lambda)\}$$

Mit Hilfe des Voronoi Bereichs können wir die Eigenschaft der Eutaxie geometrisch fassen.

Proposition 4.21. *Sei* Λ *mit Gram-Matrix* A.

1. *Genau dann ist* Λ *semi-eutaktisch, wenn* $\mathrm{Id} \in D_\Lambda$. *(bzw.* $A^{-1} \in cone\{P_x, x \in S(\Lambda)\} \subset \mathrm{Sym}_n$ *)*

2. *Genau dann ist* Λ *eutaktisch, wenn* $\mathrm{Id} \in \mathrm{ri}\, D_\Lambda$. *(bzw.* $A^{-1} \in \mathrm{ri}\, cone\{P_x, x \in S(\Lambda)\} \subset \mathrm{Sym}_n$ *)*

Beweis. Die erste Aussage ergibt sich direkt aus den Definitionen von Semi-Eutaxie und Voronoi Bereich (4.17, 4.19). Ebenso die zweite Aussage, wobei wir hierbei noch 2.10 benutzen. Die jeweils in Klammern angegebene äquivalente Bedingung für die Gram-Matrix ist Konsequenz von 4.18. □

Aus der in Kapitel 2 dargestellten Theorie der konvexen Kegel können wir nun auch eine Bedingung ableiten, wann ein Gitter Λ nicht eutaktisch bzw. nur semi-eutaktisch ist. Wir fassen diese in folgender Proposition zusammen.

Proposition 4.22. *Sei Λ ein Gitter. Wir haben*

1. *Λ ist genau dann nicht semi-eutaktisch,wenn es ein $v \in D_\Lambda^*$ gibt, mit $\langle v, Id\rangle < 0$.*
2. *Λ ist genau dann semi-eutaktisch, aber nicht eutaktisch, wenn es ein $v \in D_\Lambda^* \setminus D_\Lambda^\perp$ gibt, mit $\langle v, Id\rangle = 0$.*
3. *Λ ist genau dann nicht eutaktisch, wenn es ein $0 \neq v \in D_\Lambda^* \setminus D_\Lambda^\perp$ gibt, mit $\langle v, Id\rangle \leq 0$*

Beweis.
1. Ein Gitter Λ ist nach 4.21 genau dann nicht semi-eutaktisch, wenn $Id \notin D_\Lambda$. Dies ist nach Satz 2.22 genau dann der Fall, wenn es ein $v \in D_\Lambda^*$ gibt mit $\langle v, Id\rangle < 0$.
2. Mit den beiden Aussagen von 4.21 ist ein Gitter Λ genau dann semi-eutaktisch aber nicht eutaktisch, wenn $Id \in \operatorname{bd} D_\Lambda$ gilt. Nach Proposition 2.27 ist dies genau dann der Fall, wenn ein $v \in D_\Lambda^* \setminus D_\Lambda^\perp$ gibt, mit $\langle v, Id\rangle = 0$.
3. Sei Λ ein nicht eutaktisches Gitter. Entweder ist Λ nicht semi-eutaktisch, greift die erste Aussage. Ist Λ nicht semie-eutaktisch, greift die zweite Aussage.

□

4.4 Charakterisierung extremer Gitter

In diesem Abschnitt wollen wir darlegen, wie sich Gitter, die lokale Extremstellen der Hermite-Funktion sind, charakterisieren lassen.

4.4.1 Lokales Verhalten von $\det(\Lambda)$ und $N(\Lambda)$

Die Hermite-Funktion ist abhänig von der Norm und der Determinante eines Gitters.. Um also das lokale Verhalten der Hermite-Funktion zu studieren interessiert uns das Verhalten dieser beiden Funktionen. Dazu untersuchen wir, wie sich $\det(u(\Lambda))$ und $N(u(\Lambda))$ für ein $u \in \mathrm{GL(E)}$ von $\det(\Lambda)$ bzw. $N(\Lambda)$ unterscheiden. Hierzu erinnern wir zuerst an den aus der linearen Algebra bekannten Polarisationssatz:

Theorem 4.23 (Polarisationssatz). *Jedes $u \in \mathrm{GL(E)}$ ist Produkt eines orthogonalen und eines symmetrischen Endomorphismuses.*

Beweis. Siehe [18] Theorem 3.1.7. □

Wir können also ohne Beschränkung der Allgemeinheit $u \in End^s(\mathrm{E})$ annehmen.

Wir wollen nun zunächst betrachten, wie sich die Norm von Λ verhält. Zunächst stellt sich die Frage, welche Vektoren im Gitter $u(\Lambda)$ die kürzesten sind.

Lemma 4.24. *Es sei* $n_1 := N(\Lambda)$ *und* $n_2 := min\{N(x)|x \in \Lambda \setminus S(\Lambda)\}$ *Ist* $u \in U := \{u \in \mathrm{GL(E)} \mid \|u^{-1}\| < \sqrt{\frac{n_2}{n_1}}\}$, *so ist* $S(u(\Lambda)) \subset u(S(\Lambda))$.

Beweis. Sei $u \in U$. Nun gilt für y mit $N(y) > n_2$ und $x \in S(\Lambda)$ $N(u(y)) > N(u(x))$. Ist also $u(x') \in S(u(\Lambda))$, so ist x' in $S(\Lambda)$ □

Im Folgenden u stets in dieser Umgebung U gewählt.

Wir interessieren uns nun, wann sich die Norm von $u(\Lambda)$ im Vergleich zu Λ nicht ändert. Für die Normen der Vektoren in $u(\Lambda)$ gilt:

$$N(u(x)) = (u(x), u(x)) = (u^t u(x), x). \tag{4.1}$$

Den positiven Endomorphismus $u^t u$ werden wir im Folgenden als $u^t u = Id + v$ schreiben. Mit dieser Schreibweise haben wir folgendes Lemma für die Norm des Gitter.

Lemma 4.25. *Mit den obigen Bezeichnungen gilt genau dann* $N(u(\Lambda)) = N(\Lambda)$, *wenn* $v \in \mathrm{bd}\, D^*_\Lambda$

Beweis. Wir führen Gleichung4.1 weiter fort und erhalten:

$$N(u(x)) = (\mathrm{Id}(x), x) + (v(x), x) = N(x) + (v(x), x).$$

Aus Lemma 4.24 wissen wir, dass mit $u \in V$ die Elemente von $S(u(\Lambda))$ unter den Bildern der $x \in S(\Lambda)$ unter u zu finden sind. Aus Proposition 4.6 wissen wir, dass $(v(x), x) = \langle \rho_x, v \rangle$ gilt. Somit haben wir sofort

$$N(u(\Lambda)) = N(\Lambda) + \min_{x \in S(\Lambda)} \{\langle \rho_x, v \rangle\}.$$

Somit haben wir Gleichheit, wenn $\min_{x \in S(\Lambda)}\{\langle \rho_x, v \rangle\} = 0$ gilt. Dies ist genau dann der Fall, wenn $v \in \mathrm{bd}\, D^*_\Lambda$. □

Nun wenden wir uns der Betrachtung der Determinante zu. Mit dem Determinanten-Multiplikationssatz erhalten wir sofort:

$$\det(u(\Lambda)) = det(\Lambda) det(u)^2$$

Wir behalten die oben gemachte Zerlegung von $u^t u$ bei und betrachten nun die Funktion $\det(Id + tv)$.

Eine grundlegende Eigenschaft dieser Funktion ist die Konkavität. Wir zeigen dies in nachfolgendem Lemma:

Lemma 4.26. *Sei* $0 \neq v \in End^s$ *und seit* $I \subset \mathbb{R}$ *ein Interval, so dass* $t \mapsto 1 + \lambda t$ *positiv ist für alle* $t \in I$ *und alle Eigenwerte* λ *von* v. *Die Abbildung*

$$t \mapsto \det(\mathrm{Id} + tv)$$

ist logarithmisch konkav auf I.

Beweis. Wir betrachten $\det(\mathrm{Id}+tv)$ als Funktion der Eigenwerte von v. Wir haben $\det(\mathrm{Id}+tv) = \prod_i(1+\lambda_i t)$ Also gilt für den Logarithmus der zweiten Ableitung:

$$\frac{\partial^2}{\partial t^2} log \prod_i (1+\lambda_i t) = -\sum \frac{\lambda_i^2}{(1+\lambda_i t)^2}$$

Somit folgt die Behauptung aus 2.51. Aus 2.53 und 2.54 folgt auch, dass

$$t \mapsto \frac{1}{\det(\mathrm{Id}+tv)}$$

konvex ist. □

Nun können wir uns damit beschäftigen, wie sich $\det(\mathrm{Id}+v)$ verhält. Dazu bezeichnen mit V die Identitätsumgebung, die gegeben ist durch $V := \{v \in End^s : |\lambda_i| < 1 \ \forall \lambda_i \text{ Eigenvektor von } v\}$.

Lemma 4.27. *Mit den obigen Bezeichnungen gilt: Ist* $0 \neq v \in V$ *und* $\langle \mathrm{Id}, v\rangle \leq 0$, *so ist* $det(\mathrm{Id}+v) < 1$.

Beweis. Die Eigenwerte $1+\lambda_i$ von $Id+v$ sind positiv. Wir betrachten wieder die Funktion $\psi(t) = log(\det(\mathrm{Id}+tv))$. Für die Ableitung gilt

$$\psi'(t) = \sum_{i=1}^{n} \frac{\lambda_i}{1+\lambda_i t}.$$

Somit haben wir

$$\psi'(0) = \sum_{i=1}^{n} \lambda_i = Tr(v) = \langle \mathrm{Id}, v\rangle \leq 0.$$

Nachdem vorhergehenden Lemma wissen wir, dass ψ streng konkav ist. Somit ist $\psi(t) < 0$ für alle $t \in]0,1]$. Insbesondere ist $\psi(1) < 0$ woraus $\det(\mathrm{Id}+v) < 1$ folgtg. □

Lemma 4.28. *Sei* $\mathcal{C} \subset End^s(\mathrm{E})$ *ein abgeschlossener Kegel, so dass* $\langle \mathrm{Id}, v\rangle > 0$ *für alle* $0 \neq v \in \mathcal{C}$. *Dann gibt es ein* $\alpha > 0$ *so dass für alle* $v \in \mathcal{C}$ *mit* $0 < \|v\| < \alpha$ *folgt* $det(\mathrm{Id}+v) > 1$

Beweis. Wir bezeichnen mit $\Sigma := \{\omega \in End^s(\mathrm{E}) | \|\omega\| = 1\}$ die Einheitskugel in $End^s(\mathrm{E})$. Für ein $\omega_0 \in \Sigma \cap \mathcal{C}$ betrachten wir wieder die Funktion $\psi_{\omega_0}(t) = \log\det(\mathrm{Id}+\omega_0 t)$. Nach den Voraussetzungen is $Tr(\omega_0) = \langle \mathrm{Id}, \omega_0\rangle > 0$ woraus wie im letzten Lemma $\psi'(0) > 0$ folgt.

Somit finden wir ein $t_{\omega_0} > 0$, so dass $\psi_{\omega_0}(t_{\omega_0}) > 0$ gilt. In einer genügend kleinen Umgebung V_{ω_0} von ω_0 auf $\Sigma \cap \mathcal{C}$ gilt somit auch $\psi_\omega(t_{\omega_0}) > 0$. Insbesondere gilt dann auch $\psi_\omega(t) > 0$ für alle $t \in (0, t_{\omega_0}]$. Wir erhalten auf diese Art und Weise eine Überdeckung $\bigcup_{\omega\in\Sigma\cap\mathcal{C}} V(\omega)$ der kompakten Menge $\Sigma \cap \mathcal{C}$.

Diese Überdeckung enthält eine endliche Teilüberdeckung $\bigcup_{i=1}^{r} V(\omega_i)$. Wir setzten $\alpha = \min\{t_{\omega_1}, .., t_{\omega_r}\}$. Sei nun ein $v \in \mathcal{V}$ gewählt mit $0 < \|v\| < \alpha$. Zu diesem gewählten v betrachten wir nun $\omega = \frac{v}{\|v\|} \in \Sigma$. Aus der oben definierten endlichen Überdeckung erhalten wir, dass $\omega \in V(\omega_i)$ für ein $1 \leq i \leq r$ gilt.

Die Funktion $\psi_\omega(t)$ ist somit positiv auf dem Interval $(0, \alpha)$ Insbesondere also auch für $t = \|v\|$. Also haben wir $log(\det(\mathrm{Id}+v)) > 0$ woraus wir folgern können, dass $det(\mathrm{Id}+v) > 1$ gilt. □

Mit diesen Lemmata als Vorbereitung können wir nun im nächsten Abschnitt zeigen, dass extreme Gitter genau die perfekten und eutaktischen sind.

4.4.2 Das Theorem von Voronoi

Korkine und Zolotareff zeigten in ihren Arbeiten, dass extreme Gitter immer perfekt sind. Voronoi erkannte allerdings, dass die Perfektion ein nicht ausreichendes Kriterium für Extremität ist. Wir wollen uns hier nun dem schon erwähnten Theorem Voronois zuwenden.

Theorem 4.29 (Voronoi). *Ein Gitter ist extrem genau dann wenn es perfekt und eutaktisch ist.*

Da die Hermite-Konstanten ähnlicher Gitter gleich ist, können wir ohne Beschränkung der Allgemeinheit auf die u einschränken, für die $N(u(\Lambda)) = N(\Lambda)$ gilt.

Für den Beweis betrachten wir wieder die Aufteilung $u^t u = Id + v$ für ein u in $End^s(\mathrm{E})$ und machen zuerst die folgende auf Korkine und Zolotareff zurückgehende Bemerkung, die sich unmittelbar aus den Lemmata 4.25, 4.27 und 4.28 ergibt:

Proposition 4.30. *Mit den eingeführten Bezeichnungen gilt: Ein Gitter Λ ist extrem genau dann, wenn* $\{v \in \mathrm{bd}\, D_\Lambda^* : \langle v, \mathrm{Id}\rangle \leq 0\} = \{0\}$

Nun wenden wir uns dem eigentlichen Beweis des Theorems von Voronoi zu.

Beweis. Zuerst beweisen wir nun, dass ein Gitter, dass sowohl perfekt als auch eutaktisch ist, ein extremes Gitter ist
Sei also Λ perfekt und eutaktisch. Angenommen Λ ist nicht extrem. Dann gibt es nach 4.30 ein $0 \neq v \in \mathrm{bd}\, D_\Lambda^*$ mit $\langle v, \mathrm{Id}\rangle \leq 0$. Da Λ eutaktisch ist, gibt es $\lambda_x > 0$ mit $\sum_{x \in S(\Lambda)} \lambda_x \rho_x$. Also ist $\langle v, \mathrm{Id}\rangle = \langle v, \sum_{x \in S(\Lambda)} \lambda_x \rho_x\rangle = \sum_{x \in S(\Lambda)} \lambda_x \langle v, \rho_x\rangle \leq 0$. Da nach den Voraussetzungen $v \in D_\Lambda^*$ ist, muss $\langle v, \rho_x\rangle = 0$ für alle $x \in S(\Lambda)$ gelten, also $v \in D_\Lambda^\perp$. Da Λ perfekt ist, ist dies nur mit $v = 0$ möglich.
Nun wollen wir uns klar machen, dass ein extremes Gitter immer auch perfekt und eutaktisch ist.

Sei Λ also ein extremes Gitter. Angenommen Λ wäre nicht perfekt, dann gibt es ein $0 \neq v \in D_\Lambda^\perp$. Da mit v auch $-v$ in $D_\Lambda^\perp$ liegt, können wir ohne Einschränkung annehmen, dass $\langle v, \mathrm{Id}\rangle \leq 0$ gilt. Somit haben wir einen Widerspruch zu 4.30

Nehmen wir nun an, dass Λ nicht eutaktisch ist. Dann gibt es aber nach Proposition 4.22 ein $v \in \mathrm{bd}\, D_\Lambda^*$ mit $\langle v, \mathrm{Id}\rangle \leq 0$ und nach 4.30 ist das Gitter somit nicht extrem. □

Als Korollar erhalten wir sofort:

Korollar 4.31. *Die n-te Potenz der Hermite-Konstante ist eine rationale Zahl.*

Beweis. Ein Gitter mit maximaler Hermite-Invariante ist extrem, also auch perfekt. Ein perfektes Gitter ist proportional zu einem ganzen Gitter. □

Im ersten Kapitel hatten wir auch das Problem der größten Kusszahl erwähnt. Auch hier können uns perfekte Gitter für die Lösung des Problems im Fall von gitterförmigen Kugelkonfigurationen liefern.

Theorem 4.32. *Sei $\Lambda \subset \mathrm{E}$ ein Gitter dessen Anzahl der minimalen Vektoren maximal ist. Dann ist Λ ein perfektes Gitter.*

Beweis. Sei Λ ein solches Gitter und wir wollen annehmen, dass Λ nicht perfekt ist. Es gibt dann also ein $0 \neq v \in D_\Lambda^\perp$ Wir wählen nun ein $\lambda \geq 0$ genügend klein, so dass $\mathrm{Id} + \lambda v$ noch positiv definit ist. Dann ist auch die Wurzel von $\mathrm{Id} - \lambda v$ noch definiert und wir setzten $u_\lambda := \sqrt{Id - \lambda v}$. Nach Lemma 4.24 und 4.25 gilt für genügend kleine λ $N(u(\Lambda)) = N(\Lambda)$ und es ist $S(u_\lambda(\Lambda)) = u_\lambda(S(\Lambda))$. Mit wachsendem λ geht $\det(u_\lambda)$ gegen 0. Somit gibt es ein μ mit $N(u_\mu(\Lambda)) = N(\Lambda)$ aber $N(u_{\mu+\epsilon}(\Lambda)) < N(\Lambda)$ für alle $\epsilon > 0$. Dann hat aber das Gitter $u_\mu(\Lambda)$ zumindest ein Paar von Minimalvektoren, die nicht in $u_\mu(S(\Lambda))$ liegen. □

Wie wir in diesem Kapitel gesehen haben, bilden perfekte Gitter eine wichtige Klasse von Gittern. Auf Grund des Endlichkeitstheorems für Ähnlichkeitsklasse perfekter Gitter können diese alle klassifiziert werden. Dies ist mit dem sogenannten Voronoi-Algorithmus möglich. Und wurde bisher erfolgreich bis in Dimension 8 angewandt. (vgl.[10])

4.4.3 Dual-extreme Gitter

In ihrem Artikel [3] haben Bergé und Martinet die Hermite-Invariante auf Paare von zueinander dualen Gittern (Λ, Λ^*) erweitert.

Definition 4.33. Sei $\Lambda \subset \mathrm{E}$ ein Gitter. Die *Bergé-Martinet-Invariante* von Λ, die mit γ' bezeichnet wird, ist definiert als

$$\gamma' := (\gamma(\Lambda)\gamma(\Lambda^*))^{1/2} = (N(\Lambda)N(\Lambda^*))^{1/2}.$$

Da offensichtlich $\gamma'(\Lambda) \leq \gamma_n$, gibt es $\gamma_n' = \max_{\Lambda \subset \mathrm{E}} \gamma'(\Lambda)$. Ein Gitter Λ soll nun *dual-extrem* heißen, wenn $\gamma'(\Lambda)$ ein lokales Maximum in Λ annimmt. Um nun auch für die Klassifikation der dual-extremen Gitter eine Theorie à la Voronoi zu erhalten, erweitern wir den Begriff der Perfektion und Eutaxie wie folgt.

Definition 4.34. Sei (Λ, Λ^*) ein Paar dualer Gitter. Spannt $\{\rho_x\, x \in S(\Lambda) \cup S(\Lambda^*)$ den Raum $End^s(\mathrm{E})$ auf, so heißt (Λ, Λ^*) ein *dual-perfektes Paar*. Ein Paar dualer Gitter, für das wir eine Relation der Form

$$\sum_{x \in S(\Lambda)} \lambda_x \rho_x = \sum_{y \in S(\Lambda^*)} \lambda_y' \rho_y$$

mit strikt positiven λ_x, λ_y' finden können, heißt *dual-eutaktisches* Paar.

Es ist wieder sinnvoll, diese Definitionen mittels des Voronoibereiches der jeweiligen Gitter (vgl. Def.4.19) in die Geometrie der Kegeln zu übersetzten. Spannt $\{\rho_x\, x \in S(\Lambda) \cup S(\Lambda^*)$ den Raum $End^s(\mathrm{E})$ auf, so ist $D_\Lambda^\perp \cap D_{\Lambda^*}^\perp = \{0\}$.

Die Existenz einer dualen Eutaxierelation $\sum_{x \in S(\Lambda)} \lambda_x \rho_x = \sum_{y \in S(\Lambda^*)} \lambda_y' \rho_y$ ist gleichbedeutend mit $ri D_\Lambda \cap ri D_{\Lambda^*} \neq \emptyset$.

Aus Proposition 2.30 folgt damit sofort:

Proposition 4.35. *Ein Paar dualer Gitter* (Λ, Λ^*) *ist genau dann dual perfekt und dual-eutaktisch, wenn* $D^*_\Lambda \cap (-D^*_{\Lambda^*}) = \{0\}$ *gilt.*

Nun können wir dual-extreme Gitter charakterisieren.

Theorem 4.36 (Bergé-Martinet). *Ein Gitter* Λ *ist genau dann dual-extrem, wenn es dual-perfekt und dual-eutaktisch ist.*

Beweis. Sei Λ ein dual-extremes Gitter und wir wollen annehmen, dass Λ nicht dual-perfekt und bzw. oder nicht dual-eutaktisch ist. Nach Proposition 4.35 gibt es also ein $0 \neq v \in D^*_\Lambda \cap (-D^*_{\Lambda^*})$. Für ein genügend klein gewähltes $\epsilon > 0$ ist $u_\epsilon := \sqrt{\mathrm{Id} + \epsilon v}$ wohldefiniert. Nach Proposition 1.23 ist $u_\epsilon(\Lambda)^* = (u_\epsilon^t)^{-1}(\Lambda)$. Setzten wir $(u_\epsilon^t u_\epsilon)^{-1} = \mathrm{Id} + w$ erhalten wir mit der Definition von u_ϵ die Identität $\mathrm{Id} + w = (\mathrm{Id} + \epsilon v)^{-1}$ und somit für ϵ klein genug gewählt die Reihenentwicklung für w:

$$w = -\epsilon v + \epsilon^2 v^2 + ... + (-1)^m \epsilon^m v^m + ...$$

Aus Lemma 4.24 und 4.25 folgt nun für alle ϵ, die klein genug gewählt werden:

$$N(u(\Lambda))(N(u(\Lambda)^*)) =$$

$$(N(\Lambda) + \min_{x \in S(\Lambda)} \{\langle \rho_x, \epsilon v \rangle\})(N(\Lambda^*) + \min_{x \in S(\Lambda^*)} \{\langle \rho_x, -\epsilon v \rangle + \langle \rho x, \epsilon^2 v^2 \rangle + ... + \langle \rho_x, (-1)^m \epsilon^m v^m \rangle + ...\})$$

Da $v \in D^*_\Lambda \cap (-D^*_{\Lambda^*})$ liegt, ist sowohl $\min_{x \in S(\Lambda)} \{\langle \rho_x, \epsilon v \rangle \geq 0$ wie auch $\min_{x \in S(\Lambda^*)} \{\langle \rho_x, -\epsilon v \rangle + \langle \rho x, \epsilon^2 v^2 \rangle + ... + \langle \rho_x, (-1)^m \epsilon^m v^m \rangle + ...\} \geq 0$. Für alle $\epsilon > 0$ die genügend klein gewählt sind ist also $\gamma'(u_\epsilon(\Lambda)) \geq \gamma'(\Lambda)$. Dies ist aber ein Widerspruch zu Λ dual-extrem.

Nun ist noch zu zeigen, dass ein Gitter, welches sowohl dual-perfekt, wie auch dual-eutaktisch ist, auch dual-extrem ist. Sei also Λ ein solches Gitter. Wir wollen zeigen, dass ein Gitter $u(\Lambda)$, das nahe bei Λ liegt und für dessen Bergé-Martinet-Invariante gilt $\gamma'(u(\Lambda)) \geq \gamma'(\Lambda)$, ein zu Λ ähnliches Gitter ist. Da $\gamma'(\Lambda)$ eine Invariante der Ähnlichkeitsklasse von Λ ist, können wir dabei annehmen, dass $N(u(\Lambda^*) = N(\Lambda^*)$ gilt. Wir nehmen also an $N(u(\Lambda)) \geq N(\Lambda)$ und schreiben wieder $u = \sqrt{\mathrm{Id} + v}$. Nach Wahl von u ist $v \in \mathrm{bd}(-D^*_{\Lambda^*})$. Damit $N(u(\Lambda)) \geq N(\Lambda)$ muss v auch in D^*_Λ. Nach 4.35 ist somit $v = 0$ und $u = \mathrm{Id}$. □

5 Neue Resultate

I realized while writing this book, that eutaxy, which appears at first sight to be just a minor restriction, is just as important as perfection.

Perfect Lattices in euklidean Spaces
JACQUES MARTINET

5.1 Extreme Gitter in Dimension 8

Wir möchten nun die Liste der 10916 perfekten Gitter, auf Eutaxie untersuchen. Dazu benutzen wir die in Kapitel 2 erleuterten Methoden der linearen Optimierung. In Kapitel 3 hatten wir dargestellt, dass ein Gitter genau dann eutaktisch ist, wenn die Inverse einer Grammatrix im Inneren des Voronoi-Bereiches liegt. Wir können den Raum Sym_n mit dem $\mathbb{R}^{n(n+1)/2}$ identifizieren vermöge

$$\begin{array}{rccc} \phi: & \mathrm{Sym}_n & \longrightarrow & \mathbb{R}^{n(n+1)/2} \\ & A=(a_{ij}) & \longmapsto & a=(a_{11},...,a_{1n},a_{22},...,a_{nn}) \end{array}$$

Wir möchten nun ein perfektes Gitter Λ mit Gram-Matrix A auf Eutaxie untersuchen. Wir notieren $a := \phi(A^{-1})$, $p_x := \phi(P_x)$ und $\mathcal{C} := cone(p_x)_{x \in S(\Lambda)}$. Für den Eutaxie-Test gehen wir nun in 3 Schritten vor:

1. Wir entscheiden, ob Λ nur schwach eutaktisch ist. Dies ist genau dann der Fall, wenn $a \notin \mathcal{C}$. Nach Proposition 4.22 (1) ist dies genau dann der Fall, wenn es ein $q \in \mathcal{C}^*$ gibt mit $q \cdot a < 0$.

2. Wenn wir $a \in \mathcal{C}$ haben, möchten wir den Fall $a \in bd\mathcal{C}$ ausschließen. Nach Proposition 4.22 (2) und der Tatsache, dass $\mathcal{C}^\perp = \{0\}$ gilt, ist $a \in bd\mathcal{C}$ genau dann wenn wir ein $0 \neq q \in \mathcal{C}^*$ finden mit $q \cdot a = 0$.

3. Ein Gitter, für das nach 1.) und 2.) weder nur schwach-eutaktisch noch nur semi-eutaktisch ist, sollte eutaktisch sein. Wir wollen dies dann zeigen, indem wir eine Eutaxie-Relation mit strikt positiven Koeffizienten suchen.

Wir geben nun an, wie die in Kapitel 2 vorgestellte Lineare Programmierung dazu benutzt werden kann, den vorgestellten 3-Punkte Test zu absolvieren:
Im ersten Schritt müssen wir ein $q \in \mathcal{C}^*$ finden mit $a \cdot q < 0$. Wir betrachten hierzu folgendes Programm:

$$\begin{array}{rrcl} & q \cdot a & \rightarrow & \min \\ \text{mit} & q \cdot p_x & \geq & 0 \, \forall \, x \in S(\Lambda) \\ & q \cdot a & \geq & -100 \end{array}$$

Wir minimieren also $a \cdot q$ und lassen nur $q \in \mathcal{C}^*$ zu. Die letzte Ungleichung wird benötigt, um das Problem nach unten zu begrenzen. Unter diesen Voraussetzungen gibt es genau zwei mögliche optimale Lösungen: Entweder ist $a \cdot q^* = -100$ und dann ist $a \notin \text{cone}(p_x)$ und das untersuchte Gitter ist nur schwach eutaktisch.

Ist die optimale Lösung $a \cdot q^* = 0$, so sind zwei Fälle möglich: Entweder löst nur ein $q^* \in \mathcal{C}^\perp$ dieses Problem, oder es gibt ein $q^* \notin \mathcal{C}^\perp$, mit $a \cdot q^* = 0$. Um diese beiden Falle zu unterscheiden nutzen wir folgendes Programm:

$$\begin{array}{rrcl} & \sum_{x \in S(\Lambda)} q \cdot p_x & \rightarrow & \max \\ \text{mit} & q \cdot p_x & \geq & 0 \, \forall \, x \in S(\Lambda) \\ & q \cdot a & = & 0 \\ & \sum q \cdot p_x & \leq & 100 \end{array}$$

Wir erhalten also eine Lösung $q \in \mathcal{C}^*/\mathcal{C}^\perp$ mit $q \cdot a = 0$ genau dann, wenn $\sum_{x \in S(\Lambda)} q \cdot p_x > 0$ gewählt werden kann. Die letzten Ungleichung ist wieder angebracht, um dafür zu sorgen, dass das Problem begrenzt ist und somit wirklich eine optimale Lösung existiert.

Hat ein Gitter die ersten beiden Schritte durchlaufen, ohne dass eine trennende Hyperebene oder eine stützende Hyperebene durch a gefunden wurde, wissen wir, dass dieses Gitter eutaktisch ist. Um dies zu bestätigen wird eine Eutaxierelation mit strikt positiven Koeffizienten gesucht.

Wir bezeichnen mit $B := (p_x)_{x \in S(\Lambda)}$ die Matrix, deren Spalten die p_x sind. Eine Eutaxierelation hat somit die Gestalt $a = B\mu$, $\mu = (\mu_1, ..., \mu_s)$ mit $\mu_1 > 0, ..., \mu_s > 0$.

Um nun solch ein μ zu finden, vermindern wir x in jeder Komponente um eine sogenannte Schlupfvariable λ und versuchen in nachfolgendem LP den Wert von λ zu maximieren.

$$\begin{array}{rrcll} & \lambda & \rightarrow & \max & \\ \text{mit} & B\mu & = & a & \text{(ER)} \\ & \mu_i - \lambda & \geq & 0 \, \forall \, i = 1, .., s & \end{array}$$

Finden wir nun für (ER) eine optimale Lösung $(\mu^*, \lambda^*$ mit $\mu^* = (\mu_1^*, ..., \mu_s^*), \lambda^*$ und $\lambda^* > 0$, ist das Gitter eutaktisch und der Vektor μ^* liefert eine mögliche Eutaxierelation.

Für die Untersuchungen der 8-dimensionalen perfekten Gitter auf Eutaxie wurde zur Durchführung des beschriebenen 3 Schritte Tests, das Programmpacket lrs von Davis Avis benutzt [1] und wir können folgenden Satz formulieren.

Satz 5.1. *In Dimension 8 gibt es genau 2408 Ähnlichkeitsklassen extremer Gitter*

In Tabelle 5.1 geben wir die Resultate dieses Tests an, wobei die Ergebisse nach halber Kissing number $s := \frac{1}{2}|S(\Lambda)|$ geordnet dargestellt sind. Auf der Homepage von Herrn Martinet [2] findet sich die genaue Aufstellung der Gitter.

Es zeigt sich also, dass in Dimension 8 weniger als ein Drittel aller perfekten Formen auch extrem sind. 1951 noch stellte Coxeter fest dass,«we do not know whether every perfect form is extrem»[6] Zwar scheint Voronoi bereits eine nicht extreme perfekte Form in Dimension 6 gekannt zu haben, jedoch war es erst Barnes, der 1957 eine solche angegeben hat. In Dimension 7 konnte Jaquet unter den 33 perfekten Formen 3 nicht extreme Formen finden.

Die hier dargestellten Ergebnisse lassen aber mehr vermuten und unterstützen eine Vermutung, die Martinet in seinem Buch äußert: In hohen Dimensionen ist fast jede perfekte Form nicht extrem.

s	Extrem	Semi-Eutaktisch	Nur perfekt	Gesamt
36	858	8	5388	6254
37	513	1	1519	2033
38	471	6	1021	1498
39	212	1	288	501
40	156	6	180	342
41	71	3	47	121
42	44	1	24	69
43	19	0	7	26
44	20	0	4	24
45	11	1	1	13
46	10	1	1	12
47	4	0	0	4
48	4	0	0	4
49	1	0	0	1
50	2	0	0	2
51	3	0	0	3
52	1	0	0	1
54	4	0	0	4
56	1	0	0	1
58	1	0	0	1
71	1	0	0	1
120	1	0	0	1
total	2408	28	8480	10916

Tabelle 5.1: Ergebnisse in Dimension 8

5.2 Einige Resultate in Dimension 9

Eine komplette Klassifikation der perfekten Gitter ist bisher nur bis in Dimension 8 bekannt. In höheren Dimensionen sind bisher nur partielle Ergebnisse vorhanden. Auf Grund der vorhandenen Ergebnisse in Dimension 9, können wir folgenden Satz formulieren.

Satz 5.2. *Es gibt mindestens 524289 Ähnlichkeitsklassen perfekter Gitter in Dimension 9. Von diesen sind nur 12814 extrem.*

Satz 5.2 bekräftigt noch mehr die Vermutung von Martinet. In Tabelle 5.1 sind die Ergebnisse wieder nach halber Kissing number geordnet aufgeführt.

In Kapitel 2 hatten wir dargestellt, dass der Simplex-Algorithmus im Worst Case eine exponentielle Laufzeit besitzt.(vgl. Satz 2.47) Abschließend stellt sich nun noch die Frage, was sich über die Laufzeit des in 5.1 beschriebenen Tests auf Eutaxie sagen läßt.

Am schwersten ist diese Frage für den dritten Teiltest, also der Suche nach einer Eutaxierelation. Die Anzahl der Variablen, also der Dimension des zugrundeliegenden Polyeders, wird durch s gegeben. Da die untersuchten Gitter in unserem Fall stets perfekt waren gilt $s \geq n(n+1)/2$. Über die genaue Strucktur des Polyeders läßt sich aber im Allgemeinen nichts sagen. Im Worst Case sind somit $\binom{s}{n(n+1)/2}$ Iterationen nötig.

Für die ersten beiden Tests sieht die Situation besser aus. Hier hat das zugrundeliegende Polyeder stets eine einfache Strucktur: Die ersten Zeilen definieren einen Kegel, dessen einzige Ecke die 0 ist. Erst durch die Begrenzungsungleichung werden eventuell neue Ecken generiert. Die Zielfunktion nimmt entweder die optimale Lösung 0 im Nullpunkt an, oder die optimale Lösung 100 bzw. -100 in einer der anderen Ecken. Da nach Konstruktion des (LP) die Zielfunktion in den von 0 verschiedenen Ecken stets diesen Optimalwert annimmt, ist also nur ein Eckenwechsel nötig.

Die eben gemachten Überlegungen erklären das unterschiedliche Laufzeitverhalten des Tests in Dimension 8 und 9.

Für die 10916 perfekten Gitter der Dimension 8 benötigte ein handelsüblicher Laptop 11272 Sekunden Rechenzeit. Also im Durchschnitt knapp eine Sekunde pro Gitter. Ohne den letzten Teilschritt waren hingegen nur 1647 Sekunden nötig. Das Bestimmen der Eutaxiekoeffizienten benötige somit durchschnittlich ca. 4 Sekunden.

In Dimension 9 wurden 684179 Sekunden für die 524289 zu testenden Gitter benötigt. Im Mittel also nicht deutlich mehr, als für Dimension 8. Ohne den letzten Teilschritt waren jedoch nur 104425 Sekunden nötig. Das bestimmen der Eutaxiekoeffizienten benötige hier im Durchschnitt also ca. 45 Sekunden.

Da für einen Test auf Eutaxie die ersten beiden Teilschritte bereits ausreichend sind, ist der in 5.1. beschriebene Test also auch in höheren Dimensionen praktikabel. Jedoch scheint die Klassifikation der Perfekten Gitter ab Dimension 9 noch in weiter Ferne zu liegen, da die Anzahl perfekter Gitter mit wachsender Dimension sehr stark zu steigen scheint. Da die Ergebnise dieser Diplomarbeit auch die schon erwähnte Vermutung von Martinet stützen und zu erwarten ist, dass in höheren Dimensionen ein immer kleinerer Anteil der perfekten Gitter auch extrem sein wird, ist eine solche komplette Klassifikation auch nicht anzustreben. Vielmehr ist es interessant, perfekte Gitter zu klassifizieren, die sich durch weitreichendere Eigenschaften auszeichnen. Die von B. Venkov eingeführten *stark perfekten Gitter* (vgl. [24]) sind ein Beispiel für solche Gitter. Sie wurden bisher erfolgreich bis einschließlich Dimension 12 klassifiziert.

s	Extrem	Semi-Eutaktisch	Nur perfekt	Gesamt
45	16	2	262434	262452
46	22	2	63891	63915
47	53	0	74105	74158
48	118	0	23083	23201
49	229	0	21972	22201
50	441	0	13146	13587
51	649	2	11639	12290
52	899	4	8418	9321
53	1151	5	7589	8745
54	1309	7	5954	7270
55	1430	3	4682	6115
56	1348	2	4132	5482
57	1265	5	2659	3929
58	1125	5	2483	3613
59	858	4	1705	2567
60	571	4	1204	1779
61	394	3	762	1159
62	291	4	588	883
63	208	1	295	504
64	119	0	236	355
65	75	2	102	179
66	68	2	99	169
67	44	0	113	157
68	26	0	46	72
69	22	1	21	44
70	21	1	20	42
71	10	0	16	26
72	17	0	4	21
73	4	0	3	7
74	3	0	0	3
75	4	0	0	4
76	3	0	3	6
77	1	0	0	1
78	1	0	0	1
79	1	0	1	2
80	4	1	6	11
81	2	0	0	2
82	3	0	1	4
84	2	0	0	2
85	0	0	2	2
88	0	0	1	1
90	2	0	0	2
91	1	0	0	1
99	1	0	0	1
129	1	0	0	1
136	1	0	0	1
Gesamt:	12814	60	511415	524289

Tabelle 5.2: Ergebnisse in Dimension 9

Literaturverzeichnis

[1] D. Avis *The lrs homepage* http://cgm.cs.mcgill.ca/ avis/C/lrs.html

[2] Ch. Batut, J. Martinet *A2x-Web-Pages on Lattices* http://www.math.u-bordeaux.fr/ martinet/

[3] A.-M. Bergé, J. Martinet *Sur un problème de dualité lié aux sphères en géométrie des nombres* J. Number Theory 32 (1989), no. 1, 14-42.

[4] A. Brondsted *An Introduction to Convex Polytopes*
Springer Verlag Berlin , 1982

[5] V. Chvatal *Linear Programming* W.H. Freeman and Company, 1983

[6] J.H. Conway, N.J.A. Sloane *Low-dimensional lattices III. Perfect forms* Proc. R. Soc. Lond 418 (1988) 43-80

[7] J.H. Conway, N.J.A. Sloane *Sphere Packings, Lattices and Groups* Springer New York, 1995

[8] H.S.M. Coxeter *Extrem Forms* Canad. J.Math. 3 (1951) 391-441

[9] P.G.L Dirichlet Über die Reduktion der positiven quadratischen Formen in drei unbestimmten ganzen Zahlen J. reine angew. Math (1850) 209-227

[10] M. Dutour, A. Schürmann, F. Valentin *Classification of eight-dimensional perfect forms* Preprint (2005).

[11] W. Ebeling *Lattices and Codes* Vieweg Braunschweig, 1994

[12] G. Ewald *Combinatorial Convexity and Algebraic Geometry* Springer New York, 1996

[13] C.F. Gauß *Recensionder Untersuchungen über die Eigenschaften der positiven tenären Formen von Lugwig August Seeber* J. reine angew. Math (1840) 312-320

[14] J. Heinhold, B. Riedmüller *Lineare Algebra und Analytische Geometrie* Carl Hanser Verlag, 1973

[15] D.-O. Jaquet-Chiffelle *Enumération complète des classes de formes parfaites en dimension 7*, Ann. Inst. Fourier **43** (1993), 21-55

[16] E. Lamprecht *Lineare Algebra 2* Birkhäuser Verlag, 1983

[17] C.G. Lekkerkerker *Geometry of Numbers* Wolters-Noordhoff Publishing, 1969

[18] J. Martinet *Les réseaux parfaits des espace euclidiens* Masson Paris, 1996

[19] H. Minkowski *Ausgewählte Arbeiten zur Zahlentheorie und Geometrie* Teubner Verlagsgesellschaft Leipzig, 1989

[20] M. Pohst, H. Zassenhaus *Algorithmic Algebraic Number Theory* Cambridge Univ. Press, 1989

[21] U. Rieder *Einführung in Operations Research* Vorlesungsskript Universität Ulm, 2002

[22] C. Riener *On extreme forms in dimension 8* erscheint in Journal de Théorie de nombres Bordeaux

[23] G. Szpiro *Kepler's Conjecture* Wiley and Sons, 2003

[24] B. Venkov *Réseaux et Designs Sphériques* Notes par Jaques Martinet

[25] G. Voronoï *Nouvelles applications des paramètres continus à la théorie des formes quadratiques* J. reine angew. Nath. **133** (1908) 97-178

[26] B.L. van der Waerden*Die Reduktionstheorie der positiven quadratischen Formen* Acta Math. **96** (1956) 265-309

[27] R. Webster *Convexity* Oxford University Press, 1994

[28] G. Ziegler *Lectures on Polytopes* Springer New York, 1995

[29] C. Zong *Sphere Packings* Springer New York, 1999

Ehrenwörtliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig angefertigt und nur die angegebenen Quellen benutzt habe.

Wörtlich oder inhaltlich übernommenes Gedankengut wurde nach bestem Wissen und Gewissen als solches kenntlich gemacht.

Diese Arbeit wurde bisher keinem anderen Prüfungsgremium vorgelegt und auch noch nicht veröffentlicht.

Ulm, den 20. Februar 2006