universität
innsbruck
Institut für Mathematik

Reelle Algebra und Geometrie

Tim Netzer

Inhaltsverzeichnis

Einleitung: Was ist reelle Algebra und Geometrie?

Eine der grundlegenden Fragen der Mathematik ist die nach den Lösungen von Gleichungssystemen. *Lineare Gleichungssysteme* über Körpern werden in der linearen Algebra behandelt, *polynomiale Gleichungen mit ganzzahligen Koeffizienten* führen zur Zahlentheorie, *Differentialgleichungen* werden in der Analysis untersucht...

In der klassischen *algebraischen Geometrie* geht es um (nichtlineare) polynomiale Gleichungssysteme mit Koeffizienten aus einem Körper, deren Lösungen ebenfalls in einem Körper gesucht werden. Oft wird der Körper dabei als algebraisch abgeschlossen vorausgesetzt. Wenn man Lösungen sucht betreibt man Algebra, deshalb der Name *algebraische* Geometrie. Man kann aber große Fortschritte machen, wenn man die ganze Lösungsmenge als geometrisches Objekt betrachtet (Varietät genannt). Daher kommt der Name algebraische *Geometrie*. Grundlegende Ergebnisse wie der *Hilbert'sche Nullstellensatz* beschreiben die geometrischen Objekte anhand der darauf definierten polynomialen Funktionen. So kann man beispielsweise algebraische Zertifikate für die (Un-)Lösbarkeit des Gleichungssystems erhalten: ein System ist genau dann unlösbar, wenn das von den Gleichungen erzeugte Ideal die 1 enthält. Diese Frage kann man dann auch algorithmisch entscheiden, zum Beispiel mit Gröbnerbasen.

Die reelle algebraische Geometrie behandelt nun den Spezialfall, dass alle polynomialen Gleichungen über den reellen Zahlen $\mathbb{R}$ definiert sind, und man die Menge der reellen Lösungen der Gleichungen sucht (man kann auch allgemeine sogenannte *reell abgeschlossene Körper* betrachten). Was zunächst nur wie ein Untergebiet der klassischen algebraischen Geometrie aussieht, entpuppt sich als interessantes eigenständiges Gebiet, in dem man viele völlig neue Fragen stellen kann. Beispielsweise besitzen die reellen Zahlen eine *Anordnung* $\leq$. Man kann also auch polynomiale *Ungleichungssysteme* betrachten, was über den komplexen Zahlen nicht sinnvoll ist. Die Lösungsmenge eines Ungleichungssystems ist ei-

ne sogenannte *semialgebraische Menge*, die man wiederum als geometrisches Objekt betrachten kann. Diese Mengen kann man nun durch die darauf definierten *positiven polynomialen Funktionen* beschreiben. Solche Ergebnisse nennt man *Positivstellensätze*, und wir werden viele davon in der Vorlesung kennen lernen. Diese Sätze können auch als algebraische Zertifikate für die Lösbarkeit von Ungleichungssystem verstanden werden, und sie erlauben dann ebenfalls algorithmische Ansätze, zum Beispiel mit Hilfe der semidefiniten Optimierung.

Der erste wichtige Positivstellensatz ist in Hilbert's 17tem Problem formuliert: Jedes reelle Polynom $p \in \mathbb{R}[x_1, \ldots, x_n]$, das an jedem Punkt des $\mathbb{R}^n$ einen nichtnegativen Wert annnimmt, ist eine Quadratsumme von rationalen Funktionen. Dieser Satz wurde 1926 von Artin bewiesen, und kann als Ausgangspunkt der modernen reellen algebraischen Geometrie gelten. Wir werden ihn in dieser Vorlesung beweisen. Weitere Positivstellensätze beschreiben Polynome, die nur auf gewissen Teilmengen des $\mathbb{R}^n$ positiv sind. Ein besonders wichtiger solcher Satz ist der Darstellungssatz von Schmüdgen aus dem Jahr 1991. Es liefert zum ersten Mal Quadratsummendarstellungen ohne Nenner, also mit Polynomen anstelle von rationalen Funktionen. Auch diesen Satz, und einige Verallgemeinerungen, werden wir beweisen.

Erfreulicherweise lassen sich viele der Ergebnisse auf vielfältige Weise anwenden. Es gibt Bezüge etwa zur Funktionalanalysis, zur Optimierung und zur Konvexgeometrie. Ein wichtiges Beispiel ist die Methode von Lasserre, die den Satz von Schmüdgen zur effizienten numerischen Lösung von beliebigen polynomialen Optimierungsproblemen verwendet. Einige der Anwendungen werden wir in dieser Vorlesung ebenfalls kennenlernen.

Dieser Text ist als begleitendes Vorlesungsskript zu verstehen. Insbesondere werden nicht an jeder Stelle alle Quellen zitiert. In großen Teilen basiert er aber auf einigen Standardwerken zu reellen algebraischen Geometrie. Dies sind insbesondere die Bücher *Positive Polynomials* von Prestel & Delzell [5], *Positive Polynomials and Sums of Squares* von Marshall [4], *Real Algebraic Geometry* von Bochnak, Coste & Roy [2] und ein unveröffentlichtes Vorlesungsskript von Claus Scheiderer. Eventuelle Fehler gehen dabei aber natürlich auf mein Konto. Für Hinweise auf dieselben bin ich sehr dankbar.

Kapitel 1

Angeordnete Körper

Im ersten Kapitel beschäftigen wir uns mit angeordneten Körpern. Wir führen den Begriff eines *reell abgeschlossenen Körpers* ein, eine Verallgemeinerung der reellen Zahlen $\mathbb{R}$. Wir lernen Methoden kennen, die Nullstellen von Polynomen über reell abgeschlossenen Körpern zu zählen. Wir zeigen die Existenz und Eindeutigkeit des reellen Abschlusses. Dann beweisen wir die sogenannte *Quantorenelimination* für reell abgeschlossene Körper, und das *Transferprinzip von Tarski-Seidenberg*, das der wichtigste Schritt bei der Lösung von Hilbert's 17. Problem ist.

1.1 Anordnungen von Körpern

Definition 1.1.1. Sei M eine nichtleere Menge. Eine *lineare Ordnung* auf M ist eine zweistellige Relation $\leq$, die für alle $a, b, c \in M$ die folgendene Bedingungen erfüllt:

$a \leq a$	Reflexivität
$a \leq b, b \leq c \Rightarrow a \leq c$	Transitivität
$a \leq b, b \leq a \Rightarrow a = b$	Antisymmetrie
$a \leq b$ oder $b \leq a$	Linearität

Wir schreiben $a < b$ falls $a \leq b$ und $a \neq b$. △

Definition 1.1.2. Sei K ein Körper. Eine *Körperanordnung* ist eine lineare Ordnung von K, die zusätzlich für alle $a, b, c \in K$ die folgenden Bedinungen erfüllt:

$a \leq b \Rightarrow a + c \leq b + c$	Verträglichkeit mit Addition
$0 \leq a, 0 \leq b \Rightarrow 0 \leq ab$	Verträglichkeit mit Multiplikation

Ist $\leq$ eine Anordnung auf K, nennt man $(K, \leq)$ einen *angeordneten Körper.* $\triangle$

Lemma 1.1.3. *In einem angeordneten Körper $(K, \leq)$ gilt für alle $a, b, c \in K$*

(i) $0 \leq a^2$

(ii) $a \leq b, 0 \leq c \Rightarrow ac \leq bc$

(iii) $0 < a < b \Rightarrow 0 < b^{-1} < a^{-1}$

(iv) *Falls $b \neq 0$, so $0 \leq ab \Leftrightarrow 0 \leq ab^{-1}$*

(v) $0 < \underbrace{1 + \cdots + 1}_{n}$ *für alle $n \in \mathbb{N}$. Insbesondere ist $\mathrm{char(K)} = 0$, d.h. $\mathbb{Q} \subseteq K$.*

Beweis. (i) Falls $0 \leq a$ folgt das aus der Verträglichkeit mit Multiplikation. Falls $a \leq 0$ folgt durch Addition von $-a$ sofort $0 \leq -a$, und also $a^2 = (-a)^2 \geq 0$. (ii) Aus $a \leq b$ folgt $b - a \geq 0$ durch Addition von $-a$, und also $c(b - a) = cb - ca \geq 0$. Durch Addition von ca folgt das Ergebnis. (iii) Sei $0 < a$. Wäre $a^{-1} \leq 0$, so $1 = a^{-1}a \leq 0$, ein Widerspruch zu (i). Sei nun $0 < a < b$. Dann ist $(a^{-1} - b^{-1})ab = b - a > 0$. Wegen $0 < ab$ ist aber auch $0 < (ab)^{-1}$, und somit insgesamt $b^{-1} < a^{-1}$. (iv) folgt durch Multiplikation mit b^2 bzw. $(b^{-1})^2$. (v) folgt mit durch iteriertes Addieren von 1 zur Ungleichung $0 < 1$, und der Transitivität von $\leq$. $\square$

Beispiel 1.1.4. (i) $\mathbb{R}$ und $\mathbb{Q}$ besitzen die bekannten Anordnungen.
(ii) Sei $\mathbb{R}(t)$ der Körper der rationalen Funktionen in einer Variablen t über $\mathbb{R}$, d.h. der Quotientenkörper des Polynomrings $\mathbb{R}[t]$. Elemente von $\mathbb{R}(t)$ sind also Brüche von Polynomen in t, mit der bekannten Äquivalenzrelation. Wir können Elemente auch als Funktionen auf $\mathbb{R}$ auffassen, die aber endlich viele Pole haben können. Wenn zwei Brüche gleich sind, definieren sie überall dort, wo sie beide definiert sind, dieselbe Funktion.
Wir wollen alle Anordnungen von $\mathbb{R}(t)$ bestimmen. Wir fixieren ein $a \in \mathbb{R}$ und setzen für $f, g \in \mathbb{R}(t)$

$$f \leq_{a+} g \quad :\Leftrightarrow \quad \exists \epsilon > 0 \, \forall r \in (a, a + \epsilon) \quad f(r) \leq g(r).$$

Da f und g nur endlich viele Polstellen haben können, ist das eine sinnvolle Definition, und man überzeugt sich, dass sie eine Körperanordnung definiert, die auf $\mathbb{R}$ mit der bekannten übereinstimmt. Man beachte, dass

$$a <_{a+} t <_{a+} b \text{ für alle } b \in \mathbb{R} \text{ mit } a < b$$

gilt. Die Variable t ist also rechts von a und infinitesimal bezüglich $\mathbb{R}$ eingeordnet. Analog erhält man eine Anordnung $\leq_{a_-}$, bei der t infinitesimal links von a eingeordnet ist. Schließlich gibt es noch zwei weitere Anordnungen, $\leq_\infty$ und $\leq_{-\infty}$. Man definiert

$$f \leq_\infty g \quad :\Leftrightarrow \exists s \in \mathbb{R}\, \forall r \in (s, \infty) \quad f(r) \leq g(r),$$

und $\leq_{-\infty}$ analog mit $(-\infty, s)$. Hier gilt $\mathbb{R} <_\infty t$ bzw. $t <_{-\infty} \mathbb{R}$. Wir haben also die folgenden Anordnungen gefunden:

$$\{\leq_{-\infty}, \leq_\infty, \leq_{a_-}, \leq_{a_+} | \ a \in \mathbb{R}\}.$$

Wir überlegen uns nun, dass es keine weiteren Anordnungen gibt. Dazu überlegt man sich zuerst, dass eine Anordnung schon eindeutig durch die Position von t bezüglich $\mathbb{R}$ bestimmt ist, wenn die Verträglichkeit mit $+$ und $\cdot$ erfüllt ist. Sei nun $\leq_?$ eine Anordnung auf $\mathbb{R}(t)$. Wir setzen

$$I = \{b \in \mathbb{R} \mid b \leq_? t\} \text{ und } J = \{b \in \mathbb{R} \mid t \leq_? b\}.$$

Im Fall $I = \emptyset$ gilt also $t <_? \mathbb{R}$, und man erhält die Anordnung $\leq_{-\infty}$; für $J = \emptyset$ analog $\leq_\infty$. Falls $I \neq \emptyset \neq J$ gibt es ein $a \in \mathbb{R}$ mit $I \leq a \leq J$, da $\mathbb{R}$ Dedekind vollständig ist. Je nachdem ob $t < a$ oder $a < t$ erhält man $\leq_{a_-}$ oder $\leq_{a_+}$.
(iii) Da man $\mathbb{Q}(t)$ in $\mathbb{R}(t)$ einbetten kann, kann man alle Anordnungen von $\mathbb{R}(t)$ auf $\mathbb{Q}(t)$ einschränken. Für transzendente Zahlen a sind $\leq_{a_-}$ und $\leq_{a_+}$ auf $\mathbb{Q}(t)$ identisch, für algebraische Zahlen unterscheiden sie sich (Übungsaufgabe 1). △

Definition 1.1.5. Eine Anordnung $\leq$ auf K heißt *archimedisch*, falls für jedes $a \in K$ ein $n \in \mathbb{N}$ existiert mit $a \leq n$. △

Beispiel 1.1.6. (i) Die bekannten Anordnungen auf $\mathbb{R}$ und $\mathbb{Q}$ sind archimedisch.
(ii) Keine der Anordnungen auf $\mathbb{R}(t)$ sind archimedisch. Es gibt immer ein $f \in \mathbb{R}(t)$ mit $\mathbb{R} < f$. Für $\leq_{a_+}$ nimmt man beispielsweise $f = 1/(t-a)$.
(iii) Ist $a \in \mathbb{R}$ transzendent über $\mathbb{Q}$, so ist auf $\mathbb{Q}(t)$ die induzierte Anordnung $\leq_{a_-}$ $(= \leq_{a_+})$ archimedisch. Die restlichen Anordnungen sind nicht archimedisch (Übungsaufgabe 1). △

Lemma 1.1.7. *Ist $(K, \leq)$ archimedisch, so liegt $\mathbb{Q}$ dicht in K, d.h. für $a, b \in K$ mit $a < b$ gibt es ein $q \in \mathbb{Q}$ mit $a < q < b$.*

Beweis. Wähle $m \in \mathbb{N}$ mit $(b-a)^{-1} < m$. Wir multiplizieren mit dem positiven Element $b - a$ und erhalten $1 < m(b-a)$, bzw. $ma < mb - 1$. Sei nun $n \in \mathbb{Z}$ minimal mit $mb \leq n + 1$. Dann ist $ma < mb - 1 \leq n < mb$, und also $a < n/m < b$. □

Satz 1.1.8 (Einbettungssatz)**.** *Jeder archimedisch angeordnete Körper lässt sich ordnungstreu in $\mathbb{R}$ einbetten.*

Beweis. Sei $(K, \leq)$ archimedisch. Dann liegt $\mathbb{Q}$ nach Lemma 1.1.7 dicht in K, und wir definieren $\varphi\colon K \to \mathbb{R}$ wie folgt. Für jedes $a \in K$ sei

$$I_a = \{r \in \mathbb{Q} \mid r \leq a\} \text{ und } J_a = \{r \in \mathbb{Q} \mid a \leq r\}.$$

In $\mathbb{R}$ gibt es ein x mit $I_a \leq x \leq J_a$, da $\mathbb{R}$ Dedekind vollständig ist. Da $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt, gibt es genau ein solches x, und wir definieren $\varphi(a) = x$. Man überlegt sich nun dass φ additiv und multiplikativ ist, und somit eine Einbettung. Die Ordnungstreue von φ ist ebenfalls leicht zu sehen (Übungsaufgabe 2). □

Beispiel 1.1.9. Für transzendentes $a \in \mathbb{R}$ haben wir die archimedische Anordnung $\leq_{a_-}$ $(=\leq_{a_+})$ auf $\mathbb{Q}(t)$. In der Tat kommt sie von einer Einbettung nach $\mathbb{R}$, und zwar derjenigen, die die Variable t auf a abbildet. Da a transzendent ist, ist diese Abbildung wohldefiniert und injektiv auf $\mathbb{Q}(t)$. △

Bisher ist eine Anordnung eine zweistellige Relation auf dem Körper K. Aufgrund der Verträglichkeit mit $+$ kennt man die Anordnung aber bereits vollständig, wenn man die positiven Elemente kennt, also die a mit $0 \leq a$. Auf diese Weise kann mal also von einer zweistelligen Relation auf eine einstellige Relation, d.h. eine Teilmenge von K reduzieren, was sich häufig als einfacher erweist. Das motiviert die folgende Definition:

Definition 1.1.10. Sei K ein Körper. Eine Teilmenge $T \subseteq K$ heißt *Präpositivkegel* oder *Präordnung*, falls gilt

$$T + T \subseteq T, \quad T \cdot T \subseteq T, \quad K^2 \subseteq T, \quad -1 \notin T.$$

Dabei bezeichnet K^2 die Menge der Quadrate in K. Falls zusätzlich

$$T \cup -T = K$$

gilt, nennt man T einen *Positivkegel* oder eine *Anordnung*. Anordnungen werden meist mit P bezeichnet. △

Wir werden die doppelte Vergabe des Namens *Anordnung* gleich rechtfertigen. Zunächst die folgende einfache Bemerkung:

Bemerkung 1.1.11. (i) Die Menge

$$\Sigma K^2 = \left\{ \sum_{i=1}^{n} a_i^2 \mid n \in \mathbb{N}, a_i \in K \right\}$$

aller Quadratsummen von Elementen von K bildet genau dann eine Präordnung, wenn $-1 \notin \Sigma K^2$. Sie ist dann die kleinste Präordnung, d.h. in allen anderen Präordnungen enthalten. Insbesondere enthält jede Präordnung 0 und 1.
(ii) Es gilt

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

für alle $b \in K$, d.h. jedes Element ist eine Differenz von zwei Quadraten. Man sieht damit, dass die Bedingung $-1 \notin T$ für Präordnungen auch ersetzt werden kann durch $T \neq K$.
(iii) Für Präordnungen gilt immer $T \cap -T = \{0\}$. Wäre nämlich $x, -x \in T$ für ein $x \neq 0$, so wäre

$$-1 = x \cdot (-x) \cdot \left(\frac{1}{x}\right)^2 \in T,$$

ein Widerspruch.
(iv) Falls $P \subseteq P'$ beides Anordnungen von K sind, so gilt $P = P'$. Ist nämlich $0 \neq x \in P'$, so kann nicht $-x \in P$ gelten, nach (3). Also ist $x \in P$, da P eine Anordnung ist. △

Satz 1.1.12. *Für jede Anordnung $\leq$ auf K ist*

$$P_{\leq} := \{a \in K \mid 0 \leq a\}$$

eine Anordnung im Sinne von Definition 1.1.10. Ist umgekehrt P eine Anordnung im Sinne von Definition 1.1.10, so definiert

$$a \leq_P b \quad :\Leftrightarrow \quad b - a \in P$$

eine Anordnung $\leq_P$ auf K. Die beiden Konstruktionen sind invers zueinander und stellen also eine Bijektion zwischen Anordnungen $\leq$ und Anordnungen P her.

Beweis. Übungsaufgabe 3. □

Beispiel 1.1.13. Auf $\mathbb{R}(t)$ bilden die beiden Mengen

$$P_1 = \left\{ f/g \mid fg = \sum_{i=k}^{d} a_i t^i, a_d > 0 \right\} \cup \{0\}$$

$$P_2 = \left\{ f/g \mid fg = \sum_{i=k}^{d} a_i t^i, a_k > 0 \right\} \cup \{0\}$$

jeweils Anordnungen (Übungsaufgabe 4). Welchen der in Beispiel 1.1.4 bereits bestimmten Anordnungen entsprechen sie? △

Wir untersuchen nun, wie man Präordnungen zu Anordnungen vergrößern kann.

Lemma 1.1.14. *Sei T eine Präordnung von K, und sei $x \in K \setminus T$. Dann ist*

$$T' := T - xT = \{s - xr \mid s, r \in T\}$$

erneut eine Präordnung, mit $T \subseteq T'$ und $-x \in T$.

Beweis. $T \subseteq T'$ folgt aus $s = s - x \cdot 0$, und $-x \in T'$ folgt aus $-x = 0 - x \cdot 1$. Offensichtlich ist $T' + T' \subseteq T'$ und $T' \cdot T' \subseteq T'$ erfüllt, und $K^2 \subseteq T \subseteq T'$. Es bleibt zu zeigen dass $-1 \notin T'$. Wäre $-1 = s - xr$ mit $s, r \in T$, so folgt $r \neq 0$. Aus der Gleichung

$$r^{-1} = \left(\frac{1}{r}\right)^2 \cdot r \in T$$

folgt dann

$$x = r^{-1} \cdot rx = r^{-1} \cdot (s + 1) \in T,$$

ein Widerspruch. □

Satz 1.1.15. *Jede Präordnung T eines Körpers ist in einer Anordnung P enthalten. Weiter gilt*

$$T = \bigcap_{T \subseteq P} P.$$

Beweis. Sei $T \subseteq K$ eine Präordnung. Die Menge

$$\mathcal{T} = \{T' \subseteq K \mid T \subseteq T' \text{ Präordnung}\}$$

ist nichtleer ($T \in \mathcal{T}$), partiell geordnet durch $\subseteq$, und jede Kette besitzt eine obere Schranke (die Vereinigung der in der Kette vorkommenden Präordnungen). Mit

Zorns Lemma finden wir also ein maximales Element $P \in \mathcal{T}$. Um zu sehen, dass P sogar eine Anordnung ist, sei $x \in K \setminus P$. Mit Lemma 1.1.14 sehen wir dass $P - xP$ wieder in $\mathcal{T}$ liegt. Aus der Maximalität folgt $P = P - xP$, und insbesondere $-x \in P$. Also ist $P \cup -P = K$.
Die Inklusion $T \subseteq \bigcap_{T \subseteq P} P$ ist klar. Sei nun $x \notin T$. Dann gibt es mit Lemma 1.1.14 eine Präordnung $T' \supseteq T$ mit $-x \in T'$. Sei nun P eine Anordnung mit $T' \subseteq P$. Dann ist auch $T \subseteq P$, und wegen $-x \in P$ gilt $x \notin P$. Damit gehört x auch nicht zum Durchschnitt auf der rechten Seite. □

Satz 1.1.16. *Ein Körper K besitzt genau dann eine Anordnung, wenn $-1 \notin \Sigma K^2$. Ein Element ist genau dann in jeder Anordnung positiv, wenn es eine Quadratsumme ist.*

Beweis. Wenn es eine Anordnung P gibt, so gilt $-1 \notin P \supseteq \Sigma K^2$. Ist umgekehrt $-1 \notin \Sigma K^2$, so handelt es sich bei den Quadratsummen um eine Präordnung, die nach Satz 1.1.15 zu einer Anordnung erweitert werden kann. Da ΣK^2 in jeder Anordnung enthalten ist, gilt

$$\Sigma K^2 = \bigcap_{P \text{ Anordnung}} P. \qquad \square$$

Beispiel 1.1.17. (i) Sei $f \in \mathbb{R}(t)$ eine rationale Funktion, die überall dort, wo sie definiert ist, nichtnegative Werte annimmt. Dann ist f offensichtlich in allen Anordnungen auf $\mathbb{R}(t)$ positiv, und deshalb eine Quadratsumme.
(ii) Auf $\mathbb{C}$ gibt es keine Anordnung, da $-1 = i^2$.
(iii) Auf $\mathbb{Q}$ gibt es genau eine Anordnung, und zwar $\Sigma\mathbb{Q}^2$.
(iv) Auch auf $\mathbb{R}$ gibt es nur die Anordnung $\Sigma\mathbb{R}^2$. Hier gilt sogar $\Sigma\mathbb{R}^2 = \mathbb{R}^2$. △

Definition 1.1.18. Ein Körper K heiß *reell*, wenn er (mindestens) eine Anordnung besitzt. Dies ist also äquivalent zu $-1 \notin \Sigma K^2$. Ebenso äquivalent ist es zu der Tatsache, dass $a_1^2 + \cdots + a_n^2 = 0$ immer $a_i = 0$ für alle i impliziert. △

1.2 Ordnungsfortsetzungen und reell abgeschlossene Körper

Sei nun L/K eine Körpererweiterung, und $\leq$ eine Anordnung auf K. Wir interessieren uns für die Frage, ob man die Anordnung auf L fortsetzen kann, es also eine Anordnung $\leq'$ auf L gibt, die für Elemente aus K mit $\leq$ übereinstimmt. In der Formulierung mit Positivkegeln bedeutet das: falls P die Anordnung auf K

ist, sucht man also eine Anordnung P' auf L mit $P' \cap K = P$. In diesem Fall nennt man (L, P') bzw $(L, \leq')$ eine Ordnungserweiterung von (K, P) bzw. $(K, \leq)$.

Lemma 1.2.1. *Sei (K, P) ein angeordneter Körper und L/K eine Körpererweiterung. Die Anordnung P lässt sich auf L genau dann fortsetzen, wenn*

$$-1 \notin T_L(P) := \left\{ \sum_{i=1}^{n} p_i \ell_i^2 \mid n \in \mathbb{N}, \ell_i \in L, p_i \in P \right\},$$

d.h. wenn $T_L(P)$ eine Präordnung auf L ist.

Beweis. Wenn $T_L(P)$ eine Präordnung auf L ist, gibt es eine Anordnung P' von L mit $T_L(P) \subseteq P'$ (Satz 1.1.15). Wegen $P \subseteq T_L(P)$ gilt $P \subseteq P' \cap K$. Da aber $P' \cap K$ offensichtlich eine Anordnung auf K ist, folgt mit Bemerkung 1.1.11 (4) daraus schon $P = P' \cap K$, also ist P' eine Erweiterung von P.
Sei umgekehrt P' eine Anordnung auf L, die P fortsetzt. Aus $-1 \notin P' \supseteq T_L(P)$ folgt die andere Implikation. □

Der nächste Satz besagt, dass man eine Anordnung auf eine quadratische Erweiterung genau dann fortsetzen kann, wenn sie durch Adjunktion eines *positiven* Elements entstand.

Satz 1.2.2. *Sei (K, P) ein angeordneter Körper, $a \in K \setminus K^2$ und*

$$L := K(\sqrt{a}) = K[t]/(t^2 - a).$$

Dann lässt sich P auf L genau dann fortsetzen, wenn $a \in P$.

Beweis. Sei P' eine Fortsetzung von P auf L. In L ist $a = (\sqrt{a})^2 \in P'$, und also $a \in P' \cap K = P$.
Sei umgekehrt $a \in P$. Angenommen es gäbe eine Gleichung

$$-1 = \sum_i p_i (a_i + b_i \sqrt{a})^2$$

mit $p_i \in P, a_i, b_i \in K$. Ausmultiplizieren liefert

$$-1 = \sum_i p_i a_i^2 + p_i b_i^2 a + 2\sqrt{a} \sum_i a_i b_i.$$

Koeffizientenvergleich liefert $-1 = \sum_i p_i a_i^2 + p_i b_i^2 a \in P$, ein Widerspruch. Somit ist $-1 \notin T_L(P)$, und nach Lemma 1.2.1 lässt sich P auf L fortsetzen. □

Satz 1.2.3. *Sei L/K eine endliche Körpererweiterung von ungeradem Grad. Dann lässt sich jede Anordnung von K auf L fortsetzen.*

Beweis. Angenommen der Satz stimmt nicht. Dann gibt es eine Körpererweiterung L/K von ungeradem Grad, und eine Anordnung P von K, die sich nicht auf L fortsetzen lässt. Sei dabei das Gegenbeispiel so gewählt, dass der Grad der Körpererweiterung minimal ist.
Da in Charakteristik Null jede algebraische Erweiterung separabel ist, ist nach dem Satz vom primitiven Element die Erweiterung L/K einfach, d.h.

$$L = K(\alpha) = K[t]/(f),$$

wobei f das Minimalpolynom von α über K ist. Dabei ist $\deg(f) = 2n + 1$ der Grad der Körpererweiterung.
Weil sich P nicht fortsetzen lässt, gibt es nach Lemma 1.2.1 eine Gleichung $-1 = \sum_i p_i \ell_i^2$ mit $\ell_i \in L$. Das übersetzt sich in eine Gleichung

$$1 + \sum_i p_i f_i^2 = h \cdot f, \tag{1.1}$$

mit $f_i, h \in K[t]$. Dabei können wir $\deg(f_i) \leq 2n$ für alle i annehmen. Also ist der Grad auf der linken Seite in (1.1) höchstens $4n$, und er ist gerade. Jeder Term $p_i f_i^2$ hat nämlich geraden Grad, und einen Leitkoeffizienten aus P. Diese Leitkoeffizienten können sich beim Summieren nicht gegenseitig aufheben, da $P \cap -P = \{0\}$.
Also ist $\deg(h) \leq 2n-1$ und ungerade. Sei nun $h_1 \in K[t]$ ein irreduzibler Faktor von h von ungeradem Grad, und β eine Nullstelle von h_1 (aus dem Zerfällungskörper). Wir setzen $L' := K(\beta)$ und erhalten eine Körpererweiterung L'/K von ungeradem Grad $\leq 2n-1$, und durch Einsetzung von β in (1.1) erhalten wir eine Gleichung

$$-1 = \sum_i p_i \delta_i^2,$$

wobei $\delta_i = f_i(\beta) \in L'$. Nach Lemma 1.2.1 lässt sich P also nicht auf L' fortsetzen, ein Widerspruch zur Minimalität von L. □

Satz 1.2.4. *Jede Anordnung von K lässt sich auf $K(t)$ fortsetzen.*

Beweis. Wenn sich die Anordnung P von K nicht fortsetzen ließe, gäbe es eine Gleichung

$$-1 = \sum_i p_i \left(\frac{f_i}{g}\right)^2$$

mit $f_i, g \in K[t]$ und $p_i \in P\backslash\{0\}$ nach Lemma 1.2.1. Dabei können wir annehmen, dass g keinen gemeinsamen Teiler mit allen f_i hat. Wir multiplizieren mit g^2 und setzen 0 ein:

$$-g(0)^2 = \sum_i p_i f_i(0)^2 \in P.$$

Falls $g(0) \neq 0$, erhalten wir $-1 \in P$ durch Multiplikation mit dem Quadrat $g(0)^{-2}$, ein Widerspruch. Falls $g(0) = 0$ folgt $f_i(0) = 0$ für alle i, und alle Polynome werden also von t geteilt. Das ist wiederum ein Widerspruch. □

Definition 1.2.5. Ein Körper K heißt *reell abgeschlossen*, wenn er reell ist, aber keine echte algebraische Erweiterung besitzt, die auch reell ist. △

Beispiel 1.2.6. Die reellen Zahlen $\mathbb{R}$ besitzen eine Anordnung. Die einzige echte algebraische Erweiterung von $\mathbb{R}$ ist aber $\mathbb{C}$, und $\mathbb{C}$ besitzt keine Anordnung. Also ist $\mathbb{R}$ reell abgeschlossen. △

Lemma 1.2.7. *Wenn K reell abgeschlossen ist, gibt es genau eine Anordnung, und zwar gerade $P = K^2$.*

Beweis. Sei P eine Anordnung auf K. Für $a \in P \setminus K^2$ könnte P auf $K(\sqrt{a})$ fortgesetzt werden, nach Satz 1.2.2. Weil aber keine echte algebraische Erweiterung von K noch eine Anordnung besitzt, kann es diesen Fall nicht geben, d.h. $P = K^2$. □

Für den nächsten wichtigen Satz brauchen wir noch das folgende Hilfslemma:

Lemma 1.2.8. *Sei K ein Körper, in dem K^2 eine Anordnung ist. Dann ist jedes Element in $K(\sqrt{-1})$ ein Quadrat.*

Beweis. Sei $z = a + b\sqrt{-1}$ ein Element in $K(\sqrt{-1})$, mit $a, b \in K$. Es gibt $c \in K$ mit $c^2 = a^2 + b^2$, mit einem $c \geq 0$. Wegen

$$(c + a)(c - a) = c^2 - a^2 = b^2 \geq 0$$

und entweder $c + a \geq 0$ oder $c - a \geq 0$, muss beides gleichzeitig gelten. Also gibt es $d, e \geq 0$ mit

$$d^2 = (c + a)/2 \text{ und } e^2 = (c - a)/2.$$

Dann ist $(2de)^2 = (c + a)(c - a) = b^2$, also also $\pm 2de = b$. Nun gilt

$$(d \pm e\sqrt{-1})^2 = d^2 - e^2 \pm 2de\sqrt{-1} = a + b\sqrt{-1} = z.$$ □

Satz 1.2.9 (Artin & Schreier, 1926)**.** *Für einen Körper K sind die folgenden Aussagen äquivalent:*

(i) *K ist reell abgeschlossen.*

(ii) *K^2 ist eine Anordnung auf K, und jedes Polynom $p \in K[t]$ von ungeradem Grad hat eine Nullstelle in K.*

(iii) *$K \neq K(\sqrt{-1})$ und $K(\sqrt{-1})$ ist algebraisch abgeschlossen.*

Beweis. $(i) \Rightarrow (ii)$: Lemma 1.2.7 sagt gerade, dass K^2 eine Anordnung auf K ist. Sei nun also $p \in K[t]$ ein Polynom von ungeradem Grad. Sei p_1 ein irreduzibler Faktor von p von ungeradem Grad. Dann ist $L = K[t]/(p_1)$ eine Körpererweiterung von K, auf die sich die Anordnung nach Satz 1.2.3 fortsetzt. Da K reell abgeschlossen ist folgt $L = K$, d.h. $\deg(p_1) = 1$. Also hat p_1, und damit auch p, eine Nullstelle in K.
$(ii) \Rightarrow (iii)$: Wegen $-1 \notin K^2$ ist $K \neq K(\sqrt{-1})$. Wir müssen nun zeigen, dass $K(\sqrt{-1})$ keine echte algebraische Erweiterung besitzt. Sei dazu L eine endliche algebraische Erweiterung von $K(\sqrt{-1})$. Wir können annehmen, dass die Erweiterung L/K galoissch ist, indem wir notfalls zur normalen Hülle übergehen. Sei $G = \mathrm{Gal}(L/K)$, H eine 2-Sylow-Untergruppe von G, und F der zu H gehörende Fixkörper:

$$\begin{array}{rcl} L & & \{\mathrm{id}\} \\ 2^e \,| & & |\, 2^e \\ F & & H \\ \text{ungerade} \,| & & |\, \text{ungerade} \\ K & & G \end{array}$$

Da es nach (ii) keine echten ungeraden Erweiterungen von K geben kann, folgt $F = K$, und $|G| = 2^e$. Für die Untergruppe $G_1 := \mathrm{Gal}(L/K(\sqrt{-1}))$ folgt also $|G_1| = 2^{e-1}$, und wir wollen $e - 1 = 0$ zeigen, d.h. $L = K(\sqrt{-1})$. Wäre das nicht so, könnten wir eine Untergruppe H_1 von G_1 wählen, mit $|H_1| = 2^{e-2}$. Für den

dazugehörigen Fixkörper F_1 bekämen wir das folgende Bild:

$$\begin{array}{cc} L & \{\mathrm{id}\} \\ 2^{e-2}\,| & |\,2^{e-2} \\ F_1 & H_1 \\ 2\,| & |\,2 \\ K(\sqrt{-1}) & G_1 \\ 2\,| & |\,2 \\ K & G \end{array}$$

Nach Lemma 1.2.8 kann $K(\sqrt{-1})$ aber keine quadratische Erweiterung haben, denn jede solche Erweiterung entsteht durch Adjunktion einer Wurzel. Also gilt $e = 1$, also $L = K(\sqrt{-1})$.
$(iii)\Rightarrow(i)$: Wir zeigen zunächst $\Sigma K^2 = K^2$. Seien dazu $a, b \in K$. Da $K(\sqrt{-1})$ algebraisch abgeschlossen ist, gibt es $c, d \in K$ mit

$$(c + d\sqrt{-1})^2 = a + b\sqrt{-1},$$

also $c^2 - d^2 = a$ und $2cd = b$. Man rechnet nun

$$a^2 + b^2 = c^4 - 2c^2d^2 + d^4 + 4c^2d^2 = c^4 + 2c^2d^2 + d^4 = (c^2 + d^2)^2 \in K^2.$$

Wegen $K \neq K(\sqrt{-1})$ ist also $-1 \notin K^2 = \Sigma K^2$, also ist K reell. Nachdem jede algebraische Erweiterung von K in den algebraischen Abschluss $K(\sqrt{-1})$ einbettet, ist $K(\sqrt{-1})$ die einzige solche Erweiterung. Sie ist nicht reell. □

Korollar 1.2.10. *Sei R ein reell abgeschlossener Körper und $R' \subseteq R$ ein relativ algebraisch abgeschlossener Teilkörper. Dann ist R' ebenfalls reell abgeschlossen.*

Beweis. Übungsaufgabe 6. □

Satz 1.2.11. *Sei R ein reell abgeschlossener Körper. Dann gilt*

(i) Die einzigen irreduziblen normierten Polynome in $R[t]$ sind von der Gestalt $t - a$ und $(t - a)^2 + b^2$, mit $a, b \in R, b \neq 0$.

(ii) (Zwischenwertsatz) Falls für ein $p \in R[t]$ und $a, b \in R$ gilt $p(a) < 0 < p(b)$, so gibt es ein $c \in (a, b)$ mit $p(c) = 0$.

Beweis. (i): Da $R(\sqrt{-1})$ algebraisch abgeschlossen ist, ist jedes irreduzible Polynom über R vom Grad ≤ 2. Wenn wir Normiertheit voraussetzen gibt es also die Möglichkeiten $t - a$ und $t^2 - 2at + c = (t-a)^2 + (c - a^2)$, mit $a, c \in R$. Im zweiten Fall ist das Polynom genau dann irreduzibel, wenn es keine Nullstelle in R hat, also wenn $a^2 - c \notin R^2$. Das ist äquivalent zu $c - a^2 \in R^2 \setminus \{0\}$.
(ii): Zerlege p in irreduzible Faktoren. Ein Vorzeichenwechsel zwischen a und b kann es nur geben, wenn ein Linearfaktor $t - c$ mit $c \in (a, b)$ auftritt. □

Satz 1.2.12 (Satz von Rolle)**.** *Sei R ein reell abgeschlossener Körper und $p \in R[t]$. Falls $p(a) = p(b)$ für $a < b$ in R, so gibt es ein $c \in R$ mit $a < c < b$ mit $p'(c) = 0$. Dabei bezeichnet p' die übliche formale Ableitung von p.*

Beweis. Übungsaufgabe 7. □

1.3 Reelle Nullstellen von Polynomen

In diesem Abschnitt lernen wir eine Methode kennen, um reelle Nullstellen von Polynomen zu zählen. Die Ergebnisse sind für sich selbst interessant, werden aber auch ganz maßgeblich für den wichtigen Satz von Tarski-Seidenberg benötigt, der einen Kernpunkt im Beweis von Hilberts 17. Problem darstellt.
Sei zunächst K ein beliebiger Körper, und

$$p = t^d + a_1 t^{d-1} + a_2 t^{d-2} + \cdots + a_{d-1} t + a_d \in K[t]$$

ein normiertes Polynom über K. Wir bezeichnen die Nullstellen von p im algebraischen Abschluss von K mit $\alpha_1, \ldots, \alpha_d$. Für jedes $r \in \mathbb{N}$ ist die r-te *Newtonsumme* von p definiert als

$$\nu_r(p) := \alpha_1^r + \cdots + \alpha_d^r.$$

Man berechnet also gerade die Summe der r-ten Potenzen der Nullstellen von p. Bekanntermaßen ist es schwierig bis unmöglich, die Nullstellen von p aus den Koeffizienten a_i direkt zu berechnen. Erfreulicherweiserweise muss man die Nullstellen aber gar nicht kennen, um die Newtonsummen zu berechnen. Die Kenntnis der Koeffizienten reicht aus. Das erhält man aus der folgenden Formel:

Satz 1.3.1 (Newton Identitäten)**.** *Für jedes $k \geq 1$ gilt die Gleichung*

$$\nu_k(p) + \nu_{k-1}(p) a_1 + \nu_{k-2}(p) a_2 + \cdots + \nu_1(p) a_{k-1} + k a_k = 0.$$

Dabei setzt man $a_j = 0$ für $j > d$.

Beweis. Wir schreiben $\alpha = (\alpha_1, \ldots, \alpha_d)$ und beginnen mit der Gleichung

$$p = (t - \alpha_1)(t - \alpha_2) \cdots (t - \alpha_d) = \sum_{i=0}^{d} (-1)^{d-i} s_{d-i,d}(\alpha) \cdot t^i. \tag{1.2}$$

Dabei bezeichnet

$$s_{r,d}(\alpha) = \sum_{1 \le i_1 < \cdots < i_r \le d} \alpha_{i_1} \cdots \alpha_{i_r}$$

die r-te elementarsymmetrische Funktion in d Variablen, wobei $s_{0,d} = 1$. Durch Koeffizientenvergleich sehen wir $(-1)^r s_{r,d}(\alpha) = a_r$ für alle $r = 0 \ldots, d$. Wenn wir in (1.2) nun α_j für t einsetzen, erhalten wir

$$0 = \sum_{i=0}^{d} (-1)^{d-i} s_{d-i,d}(\alpha) \alpha_j^i,$$

und wenn wir das über alle j aufsummieren

$$\begin{aligned} 0 = \sum_{j=1}^{d} \sum_{i=0}^{d} (-1)^{d-i} s_{d-i,d}(\alpha) \alpha_j^i &= d \cdot (-1)^d s_{d,d}(\alpha) + \sum_{i=1}^{d} (-1)^{d-i} s_{d-i,d}(\alpha) \nu_i(p) \\ &= d \cdot a_d + \sum_{i=1}^{d} a_{d-i} \cdot \nu_i(p). \end{aligned}$$

Das ist genau die gewünschte Formel im Fall $k = d = \deg(p)$.
Für den allgemeinen Fall überlegen wir uns zuerst, dass sich die linke Seite der Formel nicht ändert, wenn man das Polynom p mit t multipliziert, also 0 als weitere Nullstelle hinzufügen. Die Newtonsummen ändert das offensichtlich nicht, und es entsteht der weiterere Koeffizient $a_{d+1} = 0$, und den hatten wir ja schon für p verwendet, falls in der Formel nötig.
Den Fall $k > d$ erhält man nun, indem man statt p das Polynom $q = t^{k-d}p$ betrachtet, welches Grad k hat. Für q stimmt die Formel, wie eben bewiesen, und andererseits stimmt sie mit der Formel für p überein.
Den Fall $k < d$ beweisen wir mit Induktion über $d - k$. Der Induktionsanfang ist $d - k = 0$, also der schon bewiesene Fall $k = d$. Sei also $k < d$. Wir betrachten das Polynom q mit den Nullstellen $\alpha_1, \ldots, \alpha_{d-1}$. Nach Induktionsannahme, stimmt die Formel mit k für q, denn $(d-1) - k < d - k$. Andererseits ist es die selbe Formel wie für das Polynom mit Nullstellen $\alpha_1, \ldots, \alpha_{d-1}, 0$, wie wir uns oben überlegt haben. Nun sieht man aber, dass die linke Seite der gewünschten

Formel ein (homogenes) Polynom vom Grad k in den Nullstellen $\alpha_1, \ldots, \alpha_d$ ist. Dieses Polynom ist Null, wenn $\alpha_d = 0$ gilt, und also kann man α_d ausklammern. Analog geht das für die anderen Variablen α_i, und wegen $k < d$ muss die linke Seite konstant Null gewesen sein. □

Aus den Newton Identitäten kann man rekursiv alle Newtonsummen $\nu_i(p)$ mit Hilfe der Koeffizienten a_j von p berechnen, z.B.

$$\begin{aligned}
\nu_0(p) &= d \\
\nu_1(p) &= -a_1 \\
\nu_2(p) &= -\nu_1(p)a_1 - 2a_2 = a_1^2 - 2a_2 \\
&\vdots
\end{aligned}$$

Insbesondere sieht man, dass $\nu_i(p)$ ein ganzzahliger polynomialer Ausdruck in den Koeffizienten a_j von p ist, vom Totalgrad i.

Bemerkung 1.3.2. Es gibt noch eine weitere Methode um die Newtonsummen explizit zu berechnen. Dazu definiert man die *Begleitmatrix* des Polynoms p wie folgt

$$C(p) = \begin{pmatrix} 0 & & & -a_d \\ 1 & \ddots & & -a_{d-1} \\ & \ddots & 0 & \vdots \\ & & 1 & -a_1 \end{pmatrix} \in \mathrm{M}_d(K)$$

Man rechnet nun relativ leicht aus, dass

$$\det(tI - C(p)) = p$$

gilt, das ursprüngliche Polynom p ist also das charakteristische Polynom von $C(p)$. Damit stimmen die Eigenwerte von $C(p)$ genau mit den Nullstellen von p überein (alles im algebraischen Abschluss von K). Die Eigenwerte von $C(p)^k$ sind also die k-ten Potenzen der Nullstellen von p. Somit gilt

$$\mathrm{tr}(C(p)^k) = \nu_k(p),$$

wobei tr die Spur bezeichnet, die einseits die Summe der Eigenwerte, und andererseits einfach die Summe der Diagonaleinträge ist. Auch hier sieht man, dass die Newtonsummen ganzzahlige polynomiale Ausdrücke in den Koeffizienten von p sind. △

Definition 1.3.3. Sei K ein Körper und $p \in K[t]$ ein normiertes Polynom vom Grad d. Dann heiß die Matrix

$$\mathcal{H}(p) := (\nu_{i+j}(p))_{i,j=0,\dots,d-1} = \begin{pmatrix} \nu_0(p) & \nu_1(p) & \cdots & \nu_{d-1}(p) \\ \nu_1(p) & \nu_2(p) & \cdots & \nu_d(p) \\ \vdots & & & \vdots \\ \nu_{d-1}(p) & \nu_d(p) & \cdots & \nu_{2d-2}(p) \end{pmatrix}$$

die *Hermite-Matrix von* p. △

Der Eintrag an der Stelle (i, j) von $\mathcal{H}(p)$ hängt nur von $i + j$ ab, ist also konstant entlang der Gegendiagonalen. Eine Matrix von dieser Gestalt nennt man auch *Hankelmatrix*. Der (i, j)-Eintrag von $\mathcal{H}(p)$ ist ein polynomialer Ausdruck vom Grad $i + j$ in den Koeffizienten von p. Man kann $\mathcal{H}(p)$ also direkt aus p berechnen.

Sei $M \in \mathrm{Sym}_d(K)$ eine symmetrische Matrix über einem Körper K mit $\mathrm{char}(K) \neq 2$. Dann gibt es eine invertierbare Matrix $S \in \mathrm{Gl}_d(K)$, so dass

$$S^t M S = \mathrm{diag}(a_1, \dots, a_d)$$

Diagonalgestalt hat. Falls nun (K, P) ein angeordneter Körper ist (insbesondere $\mathrm{char}(K) = 0$), so definieren wir die *Signatur* von M wie folgt:

$$\mathrm{sign}_P M := \sum_{i=1}^{d} \mathrm{sign}_P(a_i).$$

Dabei ist

$$\mathrm{sign}_P(a) := \begin{cases} 1: & a \in P \setminus \{0\} \\ -1: & -a \in P \setminus \{0\} \\ 0: & a = 0 \end{cases}$$

Die Signatur ist also die Anzahl der positiven Elemente minus die Anzahl der negativen Elemente in der Diagonalisierung. Der *Sylvester'sche Trägheitssatz* besagt, dass die Definition der Signatur einer Matrix wohldefiniert ist, also nicht von der gewählten Diagonalisierung abhängt. Der Beweis wird gewöhnlich für $K = \mathbb{R}$ geführt, ist aber für beliebige angeordnete Körper identisch.

Der folgende wichtige Satz liefert eine Methode, um Nullstellen von Polynomen über reell abgeschlossenen Körper zu zählen. Da man die Hermite-Matrix relativ leicht ausrechnen kann, und dafür insbesondere die Nullstellen des Polynoms nicht zu kennen braucht, ist das ein interessantes Ergebnis.

Satz 1.3.4. *Sei R ein reell abgeschlossener Körper und $p \in R[t]$ ein normiertes Polynom vom Grad ≥ 1. Dann gilt*

(i) $\operatorname{rang} \mathcal{H}(p) =$ *Anzahl der verschiedenen Nullstellen von p in $R(\sqrt{-1})$.*

(ii) $\operatorname{sign} \mathcal{H}(p) =$ *Anzahl der verschiedenen Nullstellen von p in R.*

Beweis. Seien $\alpha_1, \ldots, \alpha_d$ die Nullstellen von p in $R(\sqrt{-1})$. Wir setzen

$$\omega_i = (1, \alpha_i, \ldots, \alpha_i^{d-1})^t$$

und sehen dass $\mathcal{H}(p) = \sum_{i=1}^{d} \omega_i \omega_i^t$ gilt. Seien nun o.B.d.A. $\alpha_1, \ldots, \alpha_s$ die verschiedenen Nullstellen, wobei α_i mit Vielfachheit n_i auftrete. Die Vektoren $\omega_1, \ldots, \omega_s$ sind linear unabhängig, denn die Vandermonde-Matrix zu paarweise verschiedenen Zahlen hat vollen Rang. Weiter ist

$$\mathcal{H}(p) = \sum_{i=1}^{s} n_i \cdot \omega_i \omega_i^t, \tag{1.3}$$

und deshalb hat $\mathcal{H}(p)$ gerade Rang s (vergleiche Übungsaufgabe 8; beachte dass $n_i \omega_i \omega_i^t = (\sqrt{n_i}\omega_i)(\sqrt{n_i}\omega_i)^t$ gilt). Das beweist (i). Wegen $p \in R[t]$ kommen Nullstellen immer in komplex konjugierten Paaren vor, d.h. wenn $\alpha = a + b\sqrt{-1}$ eine Nullstelle ist, so ist auch $\overline{\alpha} = a - b\sqrt{-1}$ eine. Wir nehmen also an, dass

$$(\alpha_1, \ldots, \alpha_s) = (\alpha_1, \ldots, \alpha_r, \alpha_{r+1}, \ldots, \alpha_q, \overline{\alpha}_{r+1}, \ldots, \overline{\alpha}_q)$$

mit $\alpha_i \in R$ für $i \leq r$ und $\alpha_i \in R(\sqrt{-1}) \setminus R$ für $i > r$. Gleichung (1.3) ist dann

$$\begin{aligned} \mathcal{H}(p) &= \sum_{i=1}^{r} n_i \cdot \omega_i \omega_i^t + \sum_{i=r+1}^{q} n_i \cdot (\omega_i \omega_i^t + \overline{\omega}_i \overline{\omega}_i^t) \\ &= \sum_{i=1}^{r} n_i \cdot \omega_i \omega_i^t + \sum_{i=r+1}^{q} 2n_i \cdot \operatorname{Re}(\omega_i)\operatorname{Re}(\omega_i)^t - \sum_{i=r+1}^{q} 2n_i \cdot \operatorname{Im}(\omega_i)\operatorname{Im}(\omega_i)^t. \end{aligned}$$

Außerdem sind natürlich mit

$$\omega_1, \ldots, \omega_r, \omega_{r+1}, \ldots, \omega_q, \overline{\omega}_{r+1}, \ldots, \overline{\omega}_q$$

auch

$$\omega_1, \ldots, \omega_r, \operatorname{Re}(\omega_{r+1}), \ldots, \operatorname{Re}(\omega_q), \operatorname{Im}(\omega_{r+1}), \ldots, \operatorname{Im}(\omega_q) \in R^d$$

linear unabhängig. Daraus sieht man, dass $\operatorname{sign} \mathcal{H}(p) = q - (q - r) = r$ gilt, vergleiche wieder Übungsaufgabe 8. Man beachte nochmals, dass die Multiplizitäten n_i positiv sind, also in R eine Wurzel besitzen, und also $n_i \omega_i \omega_i^t$ durch $(\sqrt{n_i}\omega_i)(\sqrt{n_i}\omega_i)^t$ ersetzt werden kann. □

Beispiel 1.3.5. Sei $p = t^2 + 1 \in R[t]$. Man berechnet

$$\mathcal{H}(p) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Diese Matrix hat Rang 2 und Signatur 0. Also hat p zwei verschiedene Nullstellen in $R(\sqrt{-1})$, aber keine in R. Das stimmt hier offensichtlich, denn die Nullstellen sind $\pm\sqrt{-1}$. △

Korollar 1.3.6. *Ein normiertes Polynom $p \in R[t]$ hat genau dann nur Nullstellen in R, wenn $\mathcal{H}(p)$ positiv semidefinit ist, d.h. nach Diagonalisierung keine negativen Einträge hat.*

Beweis. Nach Satz 1.3.4 hat p genau dann nur Nullstellen in R, wenn

$$\operatorname{rang} \mathcal{H}(p) = \operatorname{sign} \mathcal{H}(p)$$

gilt. Das bedeutet aber gerade, dass keine negativen Diagonaleinträge auftreten dürfen. □

Wir wollen auch noch die Nullstellen eines Polynoms unter Nebenbedingungen zählen. Dafür verallgemeinern wir die Hermite-Matrix. Sei wieder $p \in K[t]$ ein normiertes Polynom über einem Körper K, mit Nullstellen $\alpha_1, \ldots, \alpha_d$ im algebraischen Abschluss von K. Sei nun $q \in K[t]$ ein beliebiges weiteres Polynom. Wir definieren die *verallgemeinerten Newtonsummen*

$$\nu_r(p, q) := \alpha_1^r \cdot q(\alpha_1) + \cdots + \alpha_d^r \cdot q(\alpha_d)$$

und die *verallgemeinerte Hermite-Matrix*

$$\mathcal{H}(p, q) = (\nu_{i+j}(p, q))_{i,j=0,\ldots,d-1}.$$

Im Fall $q = 1$ erhält man gerade die bekannten Konstruktionen zurück. Man sieht, dass auch die verallgemeinerten Newtonsummen ganzzahlige polynomiale Ausdrücke in den Koeffizienten von p und q sind, die man leicht ausrechnen kann: wenn $q = \sum_{j=0}^{d'} b_j t^j$ ist, gilt

$$\nu_r(p, q) = \sum_i \alpha_i^r \cdot \sum_j b_j \alpha_i^j = \sum_j b_j \sum_i \alpha_i^{j+r} = \sum_j b_j \nu_{j+r}(p).$$

Wir erhalten die folgende Verallgemeinerung von Satz 1.3.4:

Satz 1.3.7. *Sei R ein reell abgeschlossener Körper und $p \in R[t]$ ein normiertes Polynom vom Grad ≥ 1. Sei $q \in R[t]$ ein beliebiges weiteres Polynom. Dann gilt*

(i) $\operatorname{rang} \mathcal{H}(p,q) =$ *Anzahl der verschiedenen Nullstellen α von p in $R(\sqrt{-1})$ mit $q(\alpha) \neq 0$.*

(ii) $\operatorname{sign} \mathcal{H}(p,q) = \sum_{\alpha \in R, p(\alpha)=0} \operatorname{sign} q(\alpha)$.

Beweis. Mit der Notation aus dem Beweis von Satz 1.3.4 erhalten wir

$$\mathcal{H}(p,q) = \sum_{i=1}^{s} n_i q(\alpha_i) \omega_i \omega_i^t$$

mit linear unabhängigen Vektoren ω_i. Daraus folgt (i), nachdem man wieder den Vorfaktor $n_i q(\alpha_i)$ als Wurzel in $R(\sqrt{-1})$ auf die Vektoren verteilt hat. Für (ii) entstehen wieder Summanden

$$n_i q(\alpha_i) \omega_i \omega_i^t$$

mit $\alpha_i \in R$ und

$$n_i q(\alpha_i) \cdot \omega_i \omega_i^t + n_i \overline{q(\alpha_i)} \cdot \overline{\omega}_i \overline{\omega}_i^t$$

mit $\alpha_i \notin R$. Einen Summanden der zweiten Form kann man aber wieder umschreiben als

$$v_1 v_1^t - v_2 v_2^t$$

mit neuen Vektoren $v_1, v_2 \in R^d$, die den gleichen Raum aufspannen wie $\omega_i, \overline{\omega}_i$. Dazu ersetzen wir erst ω_i durch $\sqrt{n_i q(\alpha_i)} \omega_i$ und gehen dann weiter vor wie im Beweis von Satz 1.3.4. Für die Berechnung der Signatur muss man also wiederum nur die Summanden für $\alpha_i \in R$ betrachten. Dabei kann dann $n_i q(\alpha_i)$ je nach Signatur wieder durch $1, -1$ oder 0 ersetzt werden. □

Man kann nun auch noch die Anzahl der Nebenbedingungen vervielfachen, und auch beliebige Vorzeichen an sie vorgeben. Seien dazu $p, q_1, \ldots, q_m \in R[t]$ mit $p \neq 0$. Wir wollen die Anzahl

$$\#\{\alpha \in R \mid p(\alpha) = 0, q_1(\alpha) >, \ldots, q_m(\alpha) > 0\}$$

bestimmen. Beachte dass durch Übergang zu $-q_i$ damit auch die Bedingung $q_i(\alpha) < 0$ abgedeckt ist. Wenn man p durch $p^2 + q_i^2$ ersetzt, kann man auch $q_i(\alpha) = 0$ erreichen. Für $e \in \{1,2\}^m$ setzen wir

$$q^e := q_1^{e_1} \cdots q_m^{e_m}.$$

Korollar 1.3.8. *Sei R reell abgeschlossen, $p, q_1, \ldots, q_m \in R[t]$, p normiert. Dann gilt*

$$\#\{\alpha \in R \mid p(\alpha) = 0, q_1(\alpha) > 0, \ldots, q_m(\alpha) > 0\} = \frac{1}{2^m} \sum_{e \in \{1,2\}^m} \operatorname{sign} \mathcal{H}(p, q^e).$$

Beweis. Wir wissen aus Satz 1.3.7 dass

$$\operatorname{sign} \mathcal{H}(p, q^e) = \sum_{\alpha \in R, p(\alpha)=0} \operatorname{sign} q^e(\alpha)$$

gilt. Somit ist

$$\begin{aligned} \sum_{e \in \{1,2\}^m} \operatorname{sign} \mathcal{H}(p, q^e) &= \sum_{\alpha \in R, p(\alpha)=0} \sum_{e \in \{1,2\}^m} \operatorname{sign} q_1^{e_1}(\alpha) \cdots q_m^{e_m}(\alpha) \\ &= \sum_{\alpha} \prod_{i=1}^{m} (\operatorname{sign} q_i(\alpha) + \operatorname{sign} q_i(\alpha)^2). \end{aligned}$$

Es ist aber $\operatorname{sign} q_i(\alpha) + \operatorname{sign} q_i(\alpha)^2$ genau dann 2 wenn $q_i(\alpha) > 0$, und sonst immer 0. □

1.4 Der reelle Abschluss

Definition 1.4.1. Sei (K, P) ein angeordneter Körper. Ein Oberkörper R von K heißt *reeller Abschluss von* (K, P), wenn R reell abgeschlossen, R/K algebraisch und die Anordnung auf R eine Fortsetzung von P ist. △

Satz 1.4.2. *Jeder angeordnete Körper besitzt einen reellen Abschluss.*

Beweis. Betrachte einen algebraische Abschluss $\overline{K}$ von K, und die nichtleere Menge

$$\mathcal{M} = \left\{(L, P') \mid K \subseteq L \subseteq \overline{K}, P' \cap K = P\right\}.$$

Sie ist partiell geordnet durch

$$(L_1, P_1) \preceq (L_2, P_2) \;:\Leftrightarrow\; L_1 \subseteq L_2 \text{ und } P_2 \cap L_1 = P_1,$$

und jede Kette $(L_i, P_i)_{i \in I}$ besitzt in $\mathcal{M}$ die obere Schranke $L = \bigcup_{i \in I} L_i$, versehen mit der Anordnung $\bigcup_{i \in I} P_i$. Nach dem Zorn'schen Lemma gibt es also ein maximales Element (R, Q) in $\mathcal{M}$, und wir müssen zeigen dass R reell abgeschlossen ist.

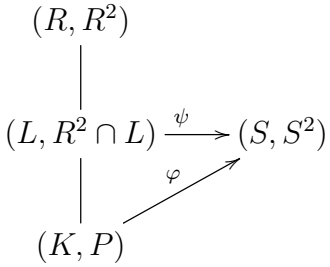
Jedes Element $a \in Q$ ist ein Quadrat in R; ansonsten könnte man die Anordnung weiter auf $R(\sqrt{a})$ fortsetzen (Satz 1.2.2), ein Widerspruch zur Maximalität. Also ist $Q = R^2$, und somit ist jede angeordnete Körperwerweiterung R' von R automatisch eine Ordnungserweiterung (vergleiche Bemerkung 1.1.11 (iv)). Aufgrund der Maximalität von R kann es also keine algebraische angeordnete Körpererweiterung von R geben, und also ist R reell abgeschlossen. □

Unser nächstes Ziel ist zu zeigen, dass der reelle Abschluss eines angeordneten Körpers bis auf Isomorphie eindeutig bestimmt ist. Dafür brauchen wir einige Hilfsaussagen. Das erste Lemma ist bereits eine schwache Version des später zu beweisenden Transferprinzip von Tarski-Seidenberg.

Lemma 1.4.3. *Sei K ein Körper, $p \in K[t]$ und R_1, R_2 zwei reell abgeschlossene Oberkörper von K, die auf K dieselbe Anordnung induzieren. Dann hat p die gleiche Anzahl von Nullstellen in R_1 und R_2.*

Beweis. Nach Satz 1.3.4 ist die Anzahl der Nullstellen in R_1 bzw. R_2 die Signatur der Hermite-Matrix $\mathcal{H}(p)$ über dem jeweiligen Körper. Die Einträge der Hermite-Matrix sind aber polynomial in den Koeffizienten von p, und liegen also in K. Da R_1 und R_2 dieselbe Anordnung auf K induzieren, ist die Signatur über beiden Körpern gleich. □

Lemma 1.4.4. *Sei R ein reeller Abschluss des angeordneten Körpers (K, P) und $K \subseteq L \subseteq R$ ein Zwischenkörper mit $[L : K] < \infty$. Sei $\varphi\colon K \to S$ eine ordnungstreue Einbettung in einen weiteren reell abgeschlossenen Körper S. Dann besitzt φ eine Fortsetzung auf L, d.h. eine Einbettung $\psi\colon L \to S$ die ordnungstreu bezüglich der Anordnung $R^2 \cap L$ auf L ist, mit $\psi = \varphi$ auf K.*



Beweis. Nach dem Satz vom primitiven Element ist $L = K(\alpha)$ für ein $\alpha \in L$. Sei $p \in K[t]$ das Minimalpolynom von α über K. Dann hat p mindestens eine Nullstelle in R, nämlich α. Nach Lemma 1.4.3 hat p dann auch eine Nullstelle in S,

und also gibt es mindestens eine Fortsetzung $\psi\colon L \to S$ von φ. Seien $\psi_1, \dots \psi_m$ alle diese Fortsetzungen. Angenommen, keine davon ist ordnungstreu bezüglich $R^2 \cap L$. Dann gibt es also Elemente $b_1, \dots, b_m \in L$ mit $b_i \in R^2$ und $\psi_i(b_i) < 0$ in S. Wir betrachten den Körper $L' = L(\sqrt{b_1}, \dots, \sqrt{b_m}) \subseteq R$. Keines der ψ_i setzt sich auf L' fort, da Quadrate immer auf positive Elemente in S abgebildet werden. Also setzt sich auch φ nicht auf L' fort, ein Widerspruch zum ersten Teil des Beweises. □

Satz 1.4.5. *Sei (K, P) ein angeordneter Körper mit reellem Abschluss R, und $\varphi\colon K \to S$ ein ordnungstreuer Homomorphismus in einen weiteren reell abgeschlossenen Körper S. Dann gibt es eine eindeutige Fortsetzung $\psi\colon R \to S$ von φ (die automatisch ordnungstreu ist).*

Beweis. Wir betrachten die nichtleere Menge

$$\mathcal{M} = \left\{(L, \psi) \mid K \subseteq L \subseteq R, \psi\colon L \to S, \psi(R^2 \cap L) \subseteq S^2, \psi = \varphi \text{ auf } K\right\}.$$

Sie ist wie üblich partiell geordnet durch die Fortsetzungsrelation, und jede Kette besitzt eine obere Schranke. Also gibt es ein maximales Element in $\mathcal{M}$, und wegen Lemma 1.4.4 muss dafür $L = R$ gelten. Also gibt es eine ordnungstreue Fortsetzung $\psi\colon R \to S$.
Für die Eindeutigkeit sei $p \in K[t]$ ein irreduzibles Polynom. Dann hat p in R einfache Nullstellen $\alpha_1 < \dots < \alpha_r$, und nach Lemma 1.4.3 in S Nullstellen $\beta_1 < \dots < \beta_r$. Die Fortsetzung ψ bildet die α_i auf die β_i ab, und aufgrund der Ordnungstreue muss $\psi(\alpha_i) = \beta_i$ gelten. Da jedes Element $\alpha \in R$ die Nullstelle eines Polynoms über K ist, ist ψ also eindeutig bestimmt. □

Korollar 1.4.6. *Der reelle Abschluss eines angeordneten Körpers (K, P) ist bis auf K-Isomorphie eindeutig bestimmt. Der Isomorphismus ist ebenfalls eindeutig.*

Beweis. Wenn R_1, R_2 zwei reelle Abschlüsse sind, gibt es eindeutig bestimmte K-Homomorphismen $\psi_1\colon R_1 \to R_2$ und $\psi_2\colon R_2 \to R_1$. Aufgrund der Eindeutigkeit aus Satz 1.4.5 muss $\psi_2 \circ \psi_1 = \mathrm{id}_{R_1}$ gelten, und damit sind beides Isomorphismen. □

Bemerkung 1.4.7. Ab jetzt macht es also Sinn, von *dem reellen Abschluss* eines angeordneten Körpers zu sprechen, denn er ist bis auf Isomorphie eindeutig bestimmt. Es ist hier sogar der Isomorphismus eindeutig! Bekanntermaßen ist auch der algebraische Abschluss eines Körpers bis auf Isomorphie eindeutig bestimmt. Dabei kann es aber mehrere Isomorphismen geben. Beispielsweise sind die Identität und die komplexe Konjugation auf $\mathbb{C}$ verschiedene $\mathbb{R}$-Isomorphismen des algebraischen Abschlusses von $\mathbb{R}$. △

Beispiel 1.4.8. Sei $\mathbb{R}_0 = \{\alpha \in \mathbb{R} \mid \alpha \text{ algebraisch über } \mathbb{Q}\}$ der relative algebraische Abschluss von $\mathbb{Q}$ in $\mathbb{R}$. Nach Lemma 1.2.10 ist $\mathbb{R}_0$ reell abgeschlossen, und damit der reelle Abschluss von $\mathbb{Q}$. $\mathbb{R}_0$ ist also der kleinste reell abgeschlossene Körper, d.h. er ist in allen anderen reell abgeschlossenen Körpern auf eindeutige Weise enthalten. △

1.5 Semialgebraische Mengen, der Projektionssatz und das Transferprinzip von Tarski und Seidenberg

Sei im folgenden stets R ein reell abgeschlossener Körper und A ein beliebiger Teilring von R. Wir setzen $x = (x_1, \ldots, x_n)$. Seien $p_1, \ldots, p_m \in R[x] = R[x_1, \ldots, x_n]$ Polynome. Wir setzen

$$O_R(p_1, \ldots, p_m) = \{a \in R^n \mid p_1(a) > 0, \ldots, p_m(a) > 0\},$$

und

$$V_R(p_1, \ldots, p_m) = \{a \in R^n \mid p_1(a) = 0, \ldots, p_m(a) = 0\}.$$

Den Index R lassen wir oft weg, wenn klar ist, welchen Körper wir betrachten.

Definition 1.5.1. (i) Eine Teilmenge von R^n heißt *A-semialgebraisch*, falls sie eine endliche boolesche Kombination (Vereinigungen, Durchschnitte, Komplemente) von Mengen $O(p_1, \ldots, p_m)$ mit $p_i \in A[x]$ ist. Für R-semialgebraisch sagen wir einfach *semialgebraisch.*
(ii) Eine Menge der Gestalt $V(p_1, \ldots, p_m)$ mit $p_i \in A[x]$ heißt *A-algebraisch*. Für R-algebraisch sagen wir einfach *algebraisch.* △

Bemerkung 1.5.2. (i) Jede A-algebraische Menge ist A-semialgebraisch. Die Bedingung $p(a) = 0$ kann man nämlich schreiben als

$$\neg p(a) > 0 \ \wedge \ \neg(-p)(a) > 0.$$

Das übersetzt sich in die boolsche Kombination von Mengen

$$V(p) = O(p)^c \cap O(-p)^c.$$

(ii) Man kann in der Definition einer semialgebraischen Menge also beliebige Bedinungen $p(a) = 0, p(a) \geq 0, p(a) \leq 0, p(a) > 0, p(a) < 0$ und boolesche Kombinationen derselben verwenden. △

Lemma 1.5.3. *(i) Jede A-algebraische Menge ist von der Gestalt $V(p)$ für ein $p \in A[x]$.*
(ii) Jede A-semialgebraische Menge $S \subseteq R^n$ hat eine Beschreibung der Form

$$S = \bigcup_{i=1}^{r} \left(V(p_i) \cap O(q_{i1}, \ldots, q_{im_i})\right).$$

Beweis. (i) folgt aus $V(p_1, \ldots, p_m) = V(p_1^2 + \cdots + p_m^2)$, vergleiche Definition 1.1.18. Für (ii) überlegt man sich, dass das System aller Mengen mit solch einer Beschreibung abgeschlossen unter endlichen Vereinigungen, Durchschnitten und Komplementbildung ist, und alle Mengen $O(p_1, \ldots, p_m)$ enthält. □

Beispiel 1.5.4. (i) Die Teilmenge $\mathbb{Z} \subseteq \mathbb{R}$ ist nicht semialgebraisch. Der Graph der Sinusfunktion $\{(\alpha, \sin\alpha) \mid \alpha \in \mathbb{R}\} \subseteq \mathbb{R}^2$ ist nicht semialgebraisch. Der Graph der Exponentialfunktion $\{(\alpha, e^\alpha) \mid \alpha \in \mathbb{R}\}$ ist nicht semialgebraisch (Übungsaufgabe 16).
(ii) Abzählbare Vereinigungen und Schnitte semialgebraischer Mengen sind im Allgemeinen nicht mehr semialgebraisch (siehe z.B. $\mathbb{Z}$). △

Satz 1.5.5. *Die semialgebraischen Teilmengen von R sind genau die endlichen Vereinigungen von Intervallen (dabei sind abgeschlossene, offene, halboffene, beschränkte und unbeschränkte Intervalle mit Grenzen in R zugelassen; insbesondere auch einzelne Punkte).*

Beweis. Offensichtlich sind alle Intervalle semialgebraisch. Umgekehrt ist die Menge aller endlichen Vereinigungen von Intervallen abgeschlossen unter den booleschen Operationen. Es genügt also zu zeigen, dass jede der Mengen $O(p) = \{\alpha \in R \mid p(\alpha) > 0\}$ eine solche endliche Vereinigung ist. Nach dem Zwischenwertsatz (Satz 1.2.11 (ii)) kann aber ein Polynom zwischen zwei Nullstellen sein Vorzeichen nicht ändern. Bekannterweise hat ein Polynom immer nur endlich viele Nullstellen. Also besteht $O(p)$ aus endlich vielen (offenen) Intervallen. □

Satz 1.5.6. *Seien $p_1, \ldots, p_m \in A[x_1, \ldots, x_n]$ und*

$$\begin{aligned} p\colon R^n &\to R^m \\ a &\mapsto (p_1(a), \ldots, p_m(a)) \end{aligned}$$

die davon induzierte polynomiale Abbildung. Dann ist das Urbild $p^{-1}(T)$ einer A-semialgebraischen (A-algebraischen) Menge $T \subseteq R^m$ wieder A-semi-algebraisch (A-algebraisch).

Beweis. Für $q \in A[x_1, \ldots, x_m]$ ist

$$p^{-1}(O(q)) = \{a \in R^n \mid q(p_1(a), \ldots, p_m(a)) > 0\} = O(h)$$

mit $h = q(p_1, \ldots, p_m) \in A[x_1, \ldots, x_n]$. Der allgemeine semialgebraische Fall folgt aus der Tatsache, dass Urbildnehmen mit den booleschen Operationen verträglich ist. Der algebraische Fall geht genau gleich. □

Bemerkung 1.5.7. (i) Das polynomiale *Bild* $p(V)$ einer algebraischen Menge $V \subseteq R^n$ ist im R^m im Allgemeinen nicht wieder algebraisch, und noch nicht mal eine boolesche Kombination algebraischer Mengen. Beispielsweise erhält man, wenn man die Menge $\{(x, y) \in R^2 \mid y - x^2 = 0\}$ auf die y-Achse projiziert, die Menge $[0, \infty)$. Algebraische Mengen in R sind aber nur endliche Teilmengen und die ganze Gerade R. Die booleschen Kombinationen davon sind also gerade die endlichen und koendlichen Mengen.
(ii) Das polynomiale Bild einer A-semialgebraischen Menge ist allerdings wieder A-semialgebraisch. Diese Aussage folgt aus dem Projektionssatz, den wir nun beweisen wollen. △

Wir beweisen zunächst ein technisches Lemma. Es besagt, dass die Signatur einer Matrix in semialgebraischer Weise von ihren Koeffizienten abhängt:

Lemma 1.5.8. *Zu $n \in \mathbb{N}$ und $k \in \mathbb{Z}$ ist die Menge*

$$\{M \in \mathrm{Sym}_n(R) \mid \operatorname{sign} M = k\}$$

$\mathbb{Z}$-semialgebraisch in $M_n(R) = R^{n^2}$. Dabei kann die semialgebraische Beschreibung sogar unabhängig von R gewählt werden.

Beweis. Wir gehen per Induktion über n vor. Der Fall $n = 1$ ist klar. Wir gehen nun vor wie im Verfahren zur Diagonalisierung von M (als Bilinearform). Wir schreiben $M = (m_{ij})_{i,j}$ und können $M \neq 0$ annehmen. Wir können dann sogar $m_{11} \neq 0$ annehmen, nach einem geeigneten Basiswechsel. Einer von endlich vielen von vorn herein festgelegten Basiswechseln funktioniert dabei immer, je nachdem, welche Einträge von M ungleich null sind. Die Matrixeinträge nach solch einem Basiswechsel kann man $\mathbb{Z}$-polynomial aus den alten Einträgen berechnen. Die verschiedenen Fälle übersetzen sich in eine große Vereinigung von semialgebraischen Mengen. Wir setzen also $m_{11} \neq 0$ voraus. Die Vektoren $e_i' := m_{11}e_i - m_{1i}e_1$ bilden für $i = 2, \ldots, n$ zusammen mit e_1 eine Basis des R^n, bezüglich derer M die Gestalt

$$\begin{pmatrix} m_{11} & 0 \\ 0 & M' \end{pmatrix}$$

mit einer symmetrischen Matrix M' der Größe $n-1$ hat. Man erhält dabei M' $\mathbb{Z}$-polynomial aus M, durch Multiplikation von beiden Seiten mit der Basiswechselmatrix. Für M' stimmt die Behauptung bereits nach Induktionssannahme, und damit auch für M. □

Satz 1.5.9 (Projektionssatz)**.** *Sei $S \subseteq R^m \times R^n$ eine A-semialgebraische Menge, und $\pi\colon R^m \times R^n \to R^n; (x,y) \mapsto y$ die Projektion. Dann ist $\pi(S) \subseteq R^n$ wieder A-semialgebraisch.*

Beweis. Es genügt den Fall $m = 1$ zu betrachten, da man das Ergebnis dann iterieren kann. Da das Bild einer Vereinigung die Vereinigung der Bilder ist, können wir nach Lemma 1.5.3 (ii) annehmen, dass S von der Gestalt

$$\begin{aligned} S &= V(p) \cap O(q_1, \dots, q_m) \\ &= \{(\alpha, a) \in R \times R^n \mid p(\alpha, a) = 0, q_1(\alpha, a) > 0, \dots, q_m(\alpha, a) > 0\} \end{aligned}$$

mit $p, q_i \in A[t,x]$ ist. Wenn wir für ein $a \in R^n$ entscheiden wollen, ob $a \in \pi(S)$ gilt, müssen wir entscheiden, ob ein $\alpha \in R$ existiert mit $(\alpha, a) \in S$. Dies können wir mit Hilfe der Hermite Matrizen tun. Wir fassen dazu alle Polynome als Polynom in t auf, die mit den Variablen x parametrisiert sind:

$$p = \sum_{i=0}^{d} p_i(x) t^i, \qquad q_j = \sum_{i=0}^{d_j} q_{ji}(x) t^i,$$

mit $p_i, q_{ji} \in A[x]$. Für $a \in R^n$ und beliebiges $h \in A[t,x]$ setzen wir $h_a(t) := h(t,a) \in R[t]$. Um die Methode der Hermite Matrizen verwenden zu können, müssen wir eine Fallunterscheidung nach dem Grad von p_a machen. Sei dazu

$$\Sigma_k := \{a \in R^n \mid \deg(p_a) = k\},$$

wobei auch $k = -\infty$ zugelassen ist, wenn $p_a \equiv 0$. Jede der Mengen Σ_k ist A-semialgebraisch im R^n, denn sie ist durch eine Bedingung an das Verschwinden bzw. Nichtverschwinden der p_i definiert. Der R^n zerlegt sich disjunkt in die veschiedenen Σ_k. Wir sind also fertig, wenn wir zeigen können, dass alle Mengen $\pi(S) \cap \Sigma_k$ A-semialgebraisch sind.
Für $k = 0$ ist $\pi(S) \cap \Sigma_k = \emptyset$, da ein Polynom vom Grad 0 keine Nullstelle hat. Die leere Menge ist aber A-semialgebraisch.
Sei nun $k \geq 1$. Für $a \in \Sigma_k$ schreibe $p_a = p_0(a) + p_1(a)t + \cdots + p_k(a)t^k$ mit $p_k(a) \neq 0$. Nach Korollar 1.3.8 gilt

$$a \in \pi(S) \Leftrightarrow \sum_{e \in \{1,2\}^m} \operatorname{sign} \mathcal{H}\left(\frac{1}{p_k(a)} p_a, q_a^e\right) > 0.$$

Die Einträge der einzelnen Hermite-Matrizen rechts sind dabei ganzzahlige polynomiale Ausdrücke in den $\frac{p_i(a)}{p_k(a)}$, und den $q_{ji}(a)$, siehe Abschnitt 1.3. Der Grad ist dabei beschränkt, abhängig vom Grad der auftretenden Polynome p, q_i. Für die Berechnung der Signatur kann man alle Matrizen mit $p_k(a)^N$ multiplizieren, für groß genügendes N, dann sind alle Einträge ganzzahlig-polynomial in den $p_i(a)$ und den $q_{ji}(a)$, und damit A-polynomial in a. Da alle auftretenden Signaturen zwischen $-k$ und $+k$ liegen, gibt es endlich viele Möglichkeiten an die einzelnen Signaturen, um eine positive Summe zu erreichen. Jede Bedingung an die Signatur einer einzelnen Matrix ist aber eine A-semialgebraische Bedingung an a, nach Lemma 1.5.8, und damit sind wir fertig.
Es bleibt noch der Fall $k = -\infty$, d.h. $p_a = 0$. In diesem Fall können wir der Beschreibung von S eine andere Gleichung hinzufügen, ohne $\pi(S) \cap \Sigma_{-\infty}$ zu verändern. Damit sind wir dann im vorigen Fall. Die neue Gleichung erhält man aus dem nachfolgenden Lemma. □

Lemma 1.5.10. *Seien $q_1, \ldots, q_m \in R[t]$ und $q := q_1 \cdots q_m$. Wir setzen*

$$p := (1 - q^2)q'.$$

Falls es einen Punkt $\alpha \in R$ gibt mit $q_1(\alpha) > 0, \ldots, q_m(\alpha) > 0$, so gibt es auch einen solchen mit zusätzlich $p(\alpha) = 0$.

Beweis. Der Fall dass q (und damit alle q_i) konstant sind, ist klar. Wir nehmen also an, dass q nicht konstant ist. Seien $\alpha_1 < \alpha_2 < \cdots < \alpha_r$ die Nullstellen von q in R. In keinem der Intervalle

$$(-\infty, \alpha_1), (\alpha_1, \alpha_2), \ldots, (\alpha_{r-1}, \alpha_r), (\alpha_r, \infty)$$

wechselt ein q_i sein Vorzeichen (Satz 1.2.11 (ii)). Wir sind also fertig wenn wir zeigen, dass p in allen Intervallen eine Nullstelle hat. In den beschränkten Intervallen hat aber q' jeweils eine Nullstelle, nach dem Satz von Rolle (Satz 1.2.12). Weil $1 - q^2$ bei α_1 und α_r den Wert 1 annimmt, für betragsmäßig große Werte aber negativ wird (q nicht konstant!), muss es in den beiden unbeschränkten Intervallen ebenfalls eine Nullstelle haben. □

Korollar 1.5.11. *Jedes polynomiale Bild einer semialgebraischen Menge ist wieder semialgebraisch (vergleiche Bemerkung 1.5.7 (ii)).*

Beweis. Übungsaufgabe 18. □

Bemerkung 1.5.12. Man beachte im Projektionssatz, dass die semialgebraische Formel für $\pi(S)$ nur von den ursprünglichen Formeln für S abhängt, und nicht von R. Das sieht man unmittelbar am Beweis. Das erlaubt uns eine starke Umformulierung des Projektionssatzes, die auch als *Quantorenelimination* bekannt ist. △

Wir führen dazu zunächst nochmals einige Begriffe exakt ein. Sei wieder A ein Ring. Eine *A-Primformel* ist eine Formel der Gestalt $p(x) > 0$ mit einem Polynom $p \in A[x] = A[x_1, \ldots, x_n]$. Allgemeine *$A$-Formeln* erhält man nun iterativ. Jede Primformel ist eine Formel, und wenn φ, ψ Formeln sind, sind auch

$$\varphi \wedge \psi, \quad \neg\varphi, \quad \exists x_i \, \varphi$$

Formeln. Man kann auch die bekannten Verknüpfungen

$$\varphi \vee \psi, \quad \varphi \to \psi, \quad \forall x_i \, \varphi$$

verwenden, als Abkürzungen für

$$\neg((\neg\varphi) \wedge (\neg\psi)), \quad \neg\varphi \vee \psi, \quad \neg(\exists x_i(\neg\varphi)).$$

Insbesondere kann auch $p(x) = 0$ und $p(x) \geq 0$ als Formel aufgefasst werden, und zwar als Abkürzung für $\neg(p(x) > 0) \wedge \neg((-p)(x) > 0)$ bzw. $\neg(-p)(x) > 0$. Der Ausdruck $p(x) = q(x)$ steht als Abkürzung für $(p - q)(x) = 0$.
Eine Variable x_i kommt in einer Formel φ *frei* vor, wenn sie nicht im Wirkungsbereich eines Quantors liegt. Ansonsten heißt ihr Vorkommen *gebunden*. Die Menge aller Variablen, die in einer Formel φ frei vorkommen, bezeichnet man mit $\mathrm{Fr}(\varphi)$. Eine Formel φ mit $\mathrm{Fr}(\varphi) = \emptyset$ nennt man auch *Aussage*.
Sei nun R ein reell abgeschlossener Oberkörper von A. Dann definiert jede A-Formel φ mit $\mathrm{Fr}(\varphi) \subseteq \{x_{i_1}, \ldots, x_{i_r}\}$ eine Teilmenge des R^r. Man nimmt dazu alle Elemente $a \in R^r$, so dass die Formel φ in R wahr ist, wenn man jedes a_j für jedes freie Vorkommen von x_{i_j} in φ einsetzt. Das *Gelten* einer Formel ist dabei genauso definiert, wie man es erwartet. Die so entstehende Menge bezeichnet man dann auch mit $\varphi(R)$:

$$\varphi(R) = \{a \in R^r \mid \varphi(a) \text{ gilt in } R\} .$$

Entsprechend ist eine *Aussage* in einem reell abgeschlossenen Körper entweder richtig oder falsch (je nachdem ob $\varphi(R)$ leer oder alles ist). Semialgebraische Mengen sind gerade die Mengen $\varphi(R)$, für eine Formel φ *ohne Quantoren*.

Beispiel 1.5.13. (i) Sei $A = \mathbb{Z}$ und

$$\varphi\colon \exists x_1\,(x_1x_2 = 1).$$

Es ist $\mathrm{Fr}(\varphi) = \{x_2\}$, und somit definiert φ eine Teilmenge $\varphi(R)$ jedes reell abgeschlossenen Körpers R. Man sieht hier

$$\varphi(R) = R \setminus \{0\}.$$

(ii) Beispiele für Aussagen sind

$$\varphi_1\colon \forall x_1\ \big(x_1 > 0 \to \exists x_2\, x_2^2 = x_1\big)$$

und

$$\varphi_2\colon \exists x_1\,(x_1^2 = -1).$$

Dabei gilt φ_1 in jedem reell abgeschlossenen Körper, und φ_2 in keinem. △

Satz 1.5.14 (Quantorenelimination)**.** *Sei A ein Ring und φ eine A-Formel. Dann gibt es eine quantorenfreie A-Formel γ mit $\mathrm{Fr}(\varphi) = \mathrm{Fr}(\gamma)$, so dass*

$$\varphi(R) = \gamma(R)$$

für jeden reell abgeschlossenen Oberkörper R von A gilt.

Beweis. Wir können induktiv über den Aufbau der Formel vorgehen. Primformeln sind quantorenfrei, also ist hier nichts zu zeigen. Auch die Konstruktionen $\varphi \wedge \psi$ und $\neg\varphi$ fügen keine Quantoren hinzu. Sei also $\varphi = \exists x_i\,\psi$, und ψ sei nach Induktionsvoraussetzung bereits quantorenfrei. Für jeden reell abgeschlossenen Oberkörper R von A ist $\varphi(R)$ die Projektion der Menge $\psi(R)$ entlang der x_i-Achse. Die Menge $\psi(R)$ ist semialgebraisch, da in ψ keine Quantoren auftreten. Nach dem Projektionssatz 1.5.9 ist also auch $\varphi(R)$ eine semialgebraische Menge, lässt sich also durch eine quantorenfreie Formel γ beschreiben. Wir haben uns dabei überlegt, dass die Formel γ unabhängig von R gewählt werden kann. □

Bemerkung 1.5.15. Im Prinzip kann die Quantorenelimination algorithmisch durchgeführt werden. Man geht iterativ über den Formelaufbau vor, und eliminiert jeden Existenzquantor, mit der im Beweis des Projektionssatz beschriebenen Methode. Praktisch ist das aber gewöhnlich nicht durchführbar. △

Beispiel 1.5.16. Betrachte die $\mathbb{Z}$-Formel

$$\varphi\colon \exists t\, xt^2 + yt + z = 0.$$

Es ist $\mathrm{Fr}(\varphi) = \{x, y, z\}$ und die Menge $\varphi(R) \subseteq R^3$ kann als die Menge der quadratischen Polynome mit Nullstelle in R aufgefasst werden. Bekanntlich hat ein (echt) quadratisches Polynom $xt^2 + yt + z$ genau dann eine Nullstelle in R, wenn die Diskriminante $y^2 - 4xz$ eine Quadratwurzel in R besitzt, und das ist genau dann der Fall, wenn sie nichtnegativ ist. Folglich definiert die quantorenfreie Formel

$$\big(x \neq 0 \wedge y^2 - 4xz \geq 0\big) \vee (x = 0 \wedge y \neq 0) \vee (x = 0 \wedge y = 0 \wedge z = 0)$$

über jedem reell abgeschlossenen Körper dieselbe Menge wie φ. △

Die Quantorenelimination ist eine sehr starke Aussage. Das sieht man sehr gut am sogenannten *Transferprinzip von Tarski-Seidenberg*, welches eine unmittelbare Folgerung ist:

Satz 1.5.17 (Transferprinzip von Tarski-Seidenberg)**.** *Sei K ein Körper und R_1, R_2 reell abgeschlossene Oberkörper, die auf K dieselbe Anordnung induzieren. Sei φ eine K-Aussage. Dann ist φ in R_1 und R_2 äquivalent, d.h. wenn φ in einem der beiden Körper gilt, dann gilt φ auch im anderen.*

Beweis. Sei γ eine quantorenfreie K-Aussage, die in jedem reell abgeschlossenen Oberkörper von K zu φ äquivalent ist (Satz 1.5.14). Ob γ in R_i gilt entscheidet sich aber bereits in K, da γ keine Quantoren enthält. □

Beispiel 1.5.18. (i) Wenn ein System aus polynomialen Gleichungen und Ungleichungen (mit Polynomen über K) eine Lösung in einem reell abgeschlossenen Oberkörper von K besitzt, dann in *jedem solchen* (der die gleiche Anordnung induziert). Die Existenz einer Lösung lässt sich ja als A-Aussage formulieren:

$$\exists x_1, \ldots \exists x_n\colon \bigwedge_j p_j(x) \geq 0 \wedge \bigwedge_j q_j(a) \neq 0 \wedge \bigwedge_j f_j(x) = 0.$$

(ii) Aussagen wie der Zwischenwertsatz (Satz 1.2.11 (ii)) können als $\mathbb{Z}$-Aussagen formuliert werden (Übungsaufgabe 17). Da jeder reell abgeschlossene Körper $\mathbb{Q}$ enthält, und es dort nur eine Anordnung gibt, sieht man sofort, dass der Zwischenwertsatz für jeden reell abgeschlossenen Körper gilt (allerdings haben wir ihn für den Beweis der Quantorenelimination schon verwendet). Man kann aber auf diese Weise sehr viele weitere Ergebnisse einfach bekommen. △

Kapitel 2

Global positive Polynome

2.1 Lösung von Hilberts 17. Problem

Sei R ein beliebiger reell abgeschlossener Körper. Wir nennen ein Polynom $p \in R[x] = R[x_1, \ldots, x_n]$ *nichtnegativ*, wenn es an jedem Punkt einen nichtnegativen Wert annimmt:

$$p(a) \geq 0 \quad \forall a \in R^n.$$

Wenn man ein nichtnegatives Polynom angeben möchte, fällt einem gewöhnlich etwas wie $p = x_1^2$ ein. Allgemeiner ist offensichtlich jede Quadratsumme

$$p = q_1^2 + \cdots + q_m^2$$

mit $q_i \in R[x]$ ein nichtnegatives Polynom. Die Frage ist nun, ob es außer den Quadratsummen noch weitere Beispiele gibt. Wenn man versucht, solche explizit zu finden, wird man höchstwahrscheinlich zunächst nicht sehr erfolgreich sein. Allerdings hat Hilbert bereits 1888 gezeigt, dass es nichtnegative Polynome geben muss, die keine Quadratsummen sind. Sein Beweis war allerdings abstrakt, und erst 1967 wurde ein explizites solches Beispiel von Motzkin gefunden. Wir werden diese Beispiele später kennenlernen. Hilbert vermutete allerdings, dass jedes nichtnegative Polynom eine Quadratsumme von *rationalen Funktionen* ist. Diese Vermutung, bekannt als Hilberts 17. Problem, ist in der Tat richtig, und wurde 1926 von Artin bewiesen. Mit Hilfe der bereits entwickelten Theorie können wir nun einen sehr eleganten Beweis dieses Satzes geben:

Satz 2.1.1 (Hilberts 17. Problem)**.** *Ein Polynom $p \in R[x_1, \ldots, x_n]$ ist genau dann nichtnegativ, wenn es eine Quadratsumme von rationalen Funktionen ist, d.h. wenn es*

$q, q_1, \ldots, q_m \in R[x], q \neq 0$ *gibt mit*

$$q^2 p = q_1^2 + \cdots + q_m^2.$$

Beweis. "$\Leftarrow$": Wäre $p(a) < 0$ für ein $a \in R^n$, dann gäbe es auch ein solches mit $q(a) \neq 0$. Dann wäre $(q^2 p)(a) < 0$, und also kann es keine Quadratsumme sein.
"$\Rightarrow$: Sei p nichtnegativ. Wir zeigen, dass p als Element des Körpers $R(x)$ bei jeder Anordnung positiv ist. Die Aussage folgt dann mit Satz 1.1.16.
Angenommen, p ist negativ in einer Anordnung $\leq$ von $R(x)$, d.h. es gilt $p < 0$. Wir bezeichnen mit $\widetilde{R}$ den reellen Abschluss von $R(x)$ bezüglich dieser Anordnung. In $\widetilde{R}$ gilt dann die folgende R-Formel:

$$\exists x_1 \, \exists x_2 \, \cdots \, \exists x_n \quad p(x_1, \ldots, x_n) < 0.$$

Man wählt nämlich für die x_i gerade die Variablen x_i, die in $R(x)$ und damit $\widetilde{R}$ ja als Elemente vorhanden sind. Da R ein Teilring der reell abgeschlossenen Körper R und $\widetilde{R}$ ist, gilt die Formel nach dem Transferprinzip (Satz 1.5.17) auch in R. Somit gibt es ein $a \in R^n$ mit $p(a) < 0$, ein Widerspruch. □

Für jedes nichtnegative Polynom existiert also eine Darstellung, die die Nichtnegativität offensichtlich macht. So eine Darstellung nennt man auch ein *algebraisches Zertifikat* für die geometrische Eigenschaft der Nichtnegativität. Allerdings kommen in der bewiesenen Darstellung *Nenner* vor. Wir werden im nächsten Abschnitt untersuchen, inwiefern solche Nenner notwendig sind.

2.2 Quadratsummen von Polynomen

Sei im ganzen Abschnitt wieder R ein beliebiger reell abgeschlossener Körper. Wir beginnen mit der recht einfachen Tatsache, dass man in einer Variablen keine Nenner für die Darstellung aus Hilberts 17. Problem benötigt:

Satz 2.2.1. *Sei $p \in R[t]$ ein Polynom in einer Variablen. Dann ist p genau dann nichtnegativ, wenn p eine Summe zweier Quadrate von Polynomen ist.*

Beweis. Wir zerlegen p in irreduzible Faktoren, die nach Satz 1.2.11 von der Gestalt $t-a$ oder $(t-a)^2+b^2$ mit $a, b \in R, b \neq 0$ sein müssen. Faktoren des zweiten Typs sind bereits Summen von zwei Quadraten. Jeder Faktor $t - a$ muss aber in gerader Potenz auftreten. Ansonsten wäre p entweder links oder rechts neben a negativ. Da das Produkt von Summen zweier Quadrate wieder eine Summe zweier Quadrate ist (siehe nächste Bemerkung), sind wir fertig. □

Bemerkung 2.2.2. Sei A ein kommutativer Ring. Dann gilt für $a, b, c, d \in A$

$$(a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2. \qquad \triangle$$

Ein weiterer relativ einfacher Fall, in dem nichtnegative Polynome Quadratsummen ohne Nenner sind, ist der von *quadratischen Polynomen*. Dafür benötigen wir zunächst das folgende Lemma:

Lemma 2.2.3. *Sei $M \in \mathrm{Sym}_d(R)$ eine symmetrische Matrix. Dann sind äquivalent:*

(i) $v^t M v \geq 0$ *für alle* $v \in R^n$.

(ii) Es gibt $S \in \mathrm{Gl}_d(R)$ *mit* $S^t M S = \mathrm{diag}(a_1, \ldots, a_d)$ *mit* $a_i \geq 0$ *für alle* i.

(iii) Alle Hauptminoren von M sind nichtnegativ in R.

(iv) $M = \sum_{i=1}^{m} v_i v_i^t$ *für ein* $m \in \mathbb{N}$ *und gewisse* $v_i \in R^d$.

(v) $M = \sum_{i=1}^{\mathrm{rang(M)}} v_i v_i^t$ *für gewisse* $v_i \in R^d$.

Beweis. Diese Aussage ist für $\mathbb{R}$ aus der linearen Algebra bekannt. Der Beweis für beliebige reell abgeschlossenen Körper geht genau gleich. Alternativ folgt die Aussage auch direkt aus dem Transferprinzip. $\square$

Definition 2.2.4. Eine Matrix $M \in \mathrm{Sym}_d(R)$ welche die äquivalenten Bedingungen aus Lemma 2.2.3 erfüllt heiß *positiv semidefinit*. $\triangle$

Satz 2.2.5. *Sei $p \in R[x_1, \ldots, x_n]$ ein nichtnegatives Polynom vom Grad 2. Dann ist p eine Quadratsumme von Polynomen vom Grad 1.*

Beweis. Zunächst können wir annehmen, dass p *homogen* vom Grad 2 ist. Wir können p nämlich mit einer neuen Variablen x_0 homogenisieren, d.h. wir multiplizieren jeden Term so oft mit x_0, bis er Grad 2 hat. Man überlegt sich, dass das neue Polynom dann immer noch nichtnegativ ist. Ist es nun eine Quadratsumme von linearen Polynomen, so erhalten wir eine gewünschte Quadratsummendarstellung von p, indem wir $x_0 = 1$ setzen.
Jedes homogene quadratische Polynom ist aber von der Gestalt

$$p = x^t M x = \sum_{i,j=1}^{n} m_{ij} x_i x_j$$

für eine symmetrische Matrix $M = (m_{ij})_{i,j} \in \mathrm{Sym}_n(R)$. Die Nichtnegativität von p bedeutet dann aber gerade, dass M positiv semidefinit ist.
Wenn also $M = \sum_i v_i v_i^t$ mit gewissen $v_i \in R^n$ ist (Lemma 2.2.3), dann gilt

$$p = x^t \left(\sum_i v_i v_i^t \right) x = \sum_i (v_i^t x)^2.$$

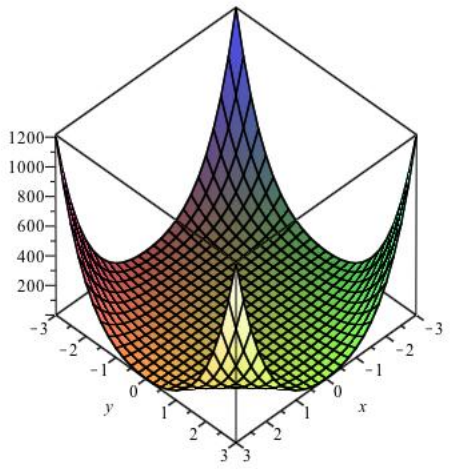
□

Bemerkung 2.2.6. Hilbert hat gezeigt, dass auch jedes nichtnegative Polynom in 2 Variablen vom Grad 4 eine Quadratsumme von Polynomen ist. Dieser Beweis ist relativ schwer, und wir behandeln ihn hier nicht. △

In allen weiteren Fällen ist aber nicht jedes nichtnegative Polynom eine Quadratsumme von Polynomen. Auch das hat Hilbert bereits 1888 abstrakt zeigen können. Das erste explizite Beispiel ist das *Motzkinpolynom* aus dem Jahr 1967:

$$p = x^4y^2 + x^2y^4 - 3x^2y^2 + 1 \in \mathbb{Z}[x, y].$$

Der Graph des Motzkinpolynoms sieht wie folgendermaßen aus:



Satz 2.2.7. *Das Motzkinpolynom p ist nichtnegativ.*

Beweis. 1. Version: Für positive Zahlen $a, b, c \geq 0$ ist das geometrische Mittel stets kleiner als als arithmetische:

$$\sqrt[3]{abc} \leq \frac{1}{3}(a + b + c).$$

Setzt man $a = 1, b = x^4y^2, c = x^2y^4$ folgt daraus die Aussage.

2. Version: Man rechnet leicht die folgende Identität nach:

$$(1 + x^2) \cdot p = (1 - x^2y^2)^2 + x^2(1 - y^2)^2 + x^2y^2(1 - x^2)^2.$$

Dies ist eine Quadratsummendarstellung mit Nennern, und wir haben gesehen, dass daraus die Nichtnegativität folgt.

3. Version: Man rechnet nach, dass die folgende Identität gilt:

$$p(x^3, y^3) = q_1^2 + q_2^2 + q_3^2 + \frac{3}{4}q_4^2 + \frac{3}{4}q_5^2 + \frac{3}{4}q_6^2$$

mit

$$q_1 = x^2y - \frac{1}{2}x^4y^5 - \frac{1}{2}x^6y^3, \quad q_2 = xy^2 - \frac{1}{2}x^3y^6 - \frac{1}{2}x^5y^4,$$

$$q_3 = 1 - \frac{1}{2}x^2y^4 - \frac{1}{2}x^4y^2, \quad q_4 = x^2y^4 - x^4y^2,$$

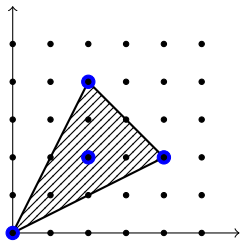
$$g_5 = x^3y^6 - x^5y^4, \quad g_6 = x^4y^5 - x^6y^3.$$

Also ist offensichtlich $p(x^3, y^3)$ nichtnegativ. Da jede reelle Zahl eine dritte Wurzel besitzt, ist damit auch p nichtnegativ. □

Wir wollen nun zeigen, dass das Motzkinpolynom keine Quadratsumme ist.

Definition 2.2.8. Sei $p \in K[x_1, \ldots, x_n]$ ein Polynom über einem Körper K. Schreibe $p = \sum_{\alpha \in \mathbb{N}^n} p_\alpha x^\alpha$, wobei $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ und $p_\alpha \in K$. Das *Newtonpolytop* $\mathcal{N}(p)$ von p ist die konvexe Hülle der Menge $\{\alpha \in \mathbb{N}^n \mid p_\alpha \neq 0\}$ im $\mathbb{R}^n$. △

Beispiel 2.2.9. Das Newton-Polytop des Motzkinpolynoms ist die konvexe Hülle der Punkte $(0,0),(2,2),(4,2),(2,4)$ in der Ebene:



△

Für $v \in \mathbb{R}^n$ und $r \in \mathbb{R}$ betrachten wir den Halbraum

$$H_{v,r} := \{\alpha \in \mathbb{R}^n \mid \langle \alpha, v \rangle \geq r\}.$$

Jedes Polytop ist der Schnitt aller Halbräume, die es enthalten.

Lemma 2.2.10. *Sei R ein reell abgeschlossener Körper und $p \in R[x_1,\ldots,x_n]$. Für $v \in \mathbb{Q}^n$ und $r \in \mathbb{Q}$ sind äquivalent:*

(i) $\mathcal{N}(p) \subseteq H_{v,r}$.

(ii) Für jedes $a \in R^n$ ist

$$\lim_{t \searrow 0} \left| t^{-r} \cdot p(a_1 t^{v_1}, \ldots, a_n t^{v_n}) \right| < \infty.$$

Beweis. Wir schreiben wieder $p = \sum_\alpha p_\alpha x^\alpha$ mit $p_\alpha \in R$.
"$(i) \Rightarrow (ii)$": Die Voraussetzung besagt, dass $\langle \alpha, v \rangle \geq r$ gilt für alle $\alpha \in \mathbb{N}^n$ mit $p_\alpha \neq 0$. Damit ist

$$t^{-r} \cdot p(a_1 t^{v_1}, \ldots, a_n t^{v_n}) = \sum_\alpha p_\alpha \cdot a^\alpha \cdot t^{\langle \alpha, v \rangle - r}.$$

Hier sind also alle auftretenden Exponenten nichtnegativ, und daraus folgt die Behauptung.
"$(ii) \Rightarrow (i)$": Angenommen es gibt einen Exponenten $\alpha \in \mathbb{N}^n$ mit $p_\alpha \neq 0$ und $\langle \alpha, v \rangle = s < r$. Sei dabei s minimal und $\{\alpha^1, \ldots, \alpha^m\}$ die Menge aller dieser α. Es gibt einen Punkt $a \in R^n$ mit $\gamma := \sum_{i=1}^m p_{\alpha^i} a^{\alpha^i} \neq 0$. Es ist dann

$$t^{-r} \cdot p(a_1 t^{v_1}, \ldots, a_n t^{v_n}) = \gamma \cdot t^{s-r} + h$$

wobei alle Terme in h höheren Grad als $s - r$ haben. Da $s - r$ negativ ist, bleibt der Ausdruck nicht beschränkt für $t \searrow 0$. □

Als interessante Folgerung erhalten wir:

Korollar 2.2.11. *Für alle $p, q, q_1, \ldots, q_m \in R[x]$ gilt:*

(i) $\mathcal{N}(p^2) = 2\mathcal{N}(p) \quad (= \{2a \mid a \in \mathcal{N}(p)\})$.

(ii) Sind p, q nichtnegativ, so ist $\mathcal{N}(p) \subseteq \mathcal{N}(p+q)$.

(iii) Ist $p = q_1^2 + \cdots + q_m^2$, so ist $\mathcal{N}(q_i) \subseteq \frac{1}{2}\mathcal{N}(p)$ für alle i.

Beweis. (i): Für $v \in \mathbb{Q}^n$ und $r \in \mathbb{Q}$ gilt $\mathcal{N}(p^2) \subseteq H_{v,r}$ genau dann, wenn $t^{-r}p(a_1t^{v_1}, \ldots, a_nt^{v_n})^2$ beschränkt bleibt für $t \searrow 0$. Das ist äquivalent dazu dass

$$t^{-r/2}p(a_1t^{v_1}, \ldots, a_nt^{v_n})$$

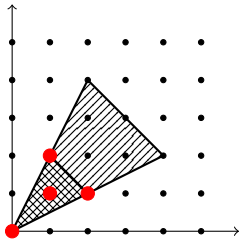
beschränkt bleibt, also zu $\mathcal{N}(p) \subseteq H_{v,r/2} = \frac{1}{2}H_{v,r}$, also zu $2\mathcal{N}(p) \subseteq H_{v,r}$. Zwei Polytope, die in den selben (rationalen) Halbräumen liegen, sind aber gleich.
(ii): Sei $\mathcal{N}(p+q) \subseteq H_{v,r}$, d.h.

$$t^{-r}\left(p(a_1t^{v_1}, \ldots, a_nt^{v_n}) + q(a_1t^{v_1}, \ldots, a_nt^{v_n})\right)$$

bleibt beschränkt für $t \searrow 0$. Da p und q nichtnegativ sind, bleibt auch der Ausdruck mit p allein beschränkt, und also auch $\mathcal{N}(p) \subseteq H_{v,r}$. Daraus folgt wiederum die Aussage. (iii) folgt unmittelbar aus (i) und (ii). □

Satz 2.2.12. *Das Motzkinpolynom p ist keine Quadratsumme.*

Beweis. Wäre $p = q_1^2 + \cdots + q_m^2$, so wäre $\mathcal{N}(q_i) \subseteq \frac{1}{2}\mathcal{N}(p)$ für alle i, nach Korollar 2.2.11. Das bedeutet aber, dass in den q_i höchstens die Monome $1, xy, x^2y, xy^2$ vorkommen können.



Das Monom x^2y^2 entsteht in jedem q_i^2 also auf eindeutige Weise als Quadrat des Monoms xy. Insbesondere hat es einen nichtnegativen Koeffizienten, und somit hat x^2y^2 auch in p einen nichtnegativen Koeffizienten. Dieser Koeffizient ist in p aber -3, ein Widerspruch. □

Bemerkung 2.2.13. Man kann zu p auch eine beliebige positive Zahl addieren, und es ist immer noch keine Quadratsumme. Das Argument benutzt den konstanten Koeffizienten ja überhaupt nicht. Also gibt es auch sehr positive Polynome, die keine Quadratsummen sind. △

Wir wollen noch etwas weiter untersuchen, ob und wie man Polynome als Quadratsummen darstellen kann.

Lemma 2.2.14. *Seien $q_1, \dots, q_m \in R[x_1, \dots, x_n]$ und $p = q_1^2 + \cdots + q_m^2$. Dann gilt*

$$\deg(p) = \max\{2 \cdot \deg(q_i) \mid i = 1, \dots, m\}.$$

Beweis. "$\leq$" ist offensichtlich. Weiter beachte dass

$$\deg(q) = \max_{a \in \mathcal{N}(q)} \langle a, e \rangle \quad \text{mit } e = (1, \dots, 1)$$

gilt, für alle Polynome q. Also gilt mit Korollar 2.2.11 (iii)

$$\deg(p) = \max_{a \in \mathcal{N}(p)} \langle a, e \rangle \geq \max_{a \in 2\mathcal{N}(q_i)} \langle a, e \rangle = 2 \cdot \max_{a \in \mathcal{N}(q_i)} \langle a, e \rangle = 2 \cdot \deg(q_i),$$

für alle i. □

Wir führen nun *Grammatrizen* von Polynomen ein. Wir bezeichnen mit $R[x]_d$ den R-Vektorraum der Polynome in $x = (x_1, \dots, x_n)$ vom Grad $\leq d$. Wir verwenden wieder die Bezeichnung $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ und $|\alpha| := \alpha_1 + \cdots + \alpha_n$. Der Raum $R[x]_d$ besitzt zum Beispiel die monomiale Basis

$$X_d := (x^\alpha)_{\alpha \in \mathbb{N}^n, |\alpha| \leq d} = (1, x_1, x_2, \dots, x_1^2, x_1 x_2, \dots),$$

und seine Dimension ist $\Delta_d := \binom{n+d}{d}$. Wir betrachten nun die folgende lineare Abbildung:

$$G \colon \mathrm{Sym}_{\Delta_d}(R) \to R[x]_{2d}; \quad M \mapsto X_d^t M X_d.$$

Für $M = (m_{\alpha\beta})_{|\alpha|, |\beta| \leq d}$ ist also

$$G(M) = \sum_{|\alpha|, |\beta| \leq d} m_{\alpha\beta} x^\alpha x^\beta = \sum_{|\gamma| \leq 2d} \left(\sum_{\substack{\alpha+\beta=\gamma \\ |\alpha|, |\beta| \leq d}} m_{\alpha\beta} \right) x^\gamma.$$

Offensichtlich ist G surjektiv. Für jedes $p \in R[x]_{2d}$ ist

$$G^{-1}(p) = \{M \mid X_d^t M X_d = p\}$$

also ein nichtleerer affiner Unterraum von $\mathrm{Sym}_{\Delta_d}(R)$.

Definition 2.2.15. Die Elemente von $G^{-1}(p)$ nennt man *Grammatrizen von* p. △

Beispiel 2.2.16. $R[x_1, x_2]_1$ hat die monomiale Basis $X_1 = (1, x_1, x_2)$. Wir haben also die Abbildung

$$G\colon \mathrm{Sym}_3(R) \to R[x_1, x_2]_2,$$

mit

$$\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \mapsto (1, x_1, x_2) \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix}$$
$$= a + 2bx_1 + 2cx_2 + dx_1^2 + 2ex_1x_2 + fx_2^2.$$

Das Polynom $x_1^2 - 2x_1x_2 + x_2^2 + 2x_1 - 2x_2 + 1$ hat also die Grammatrix

$$\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}.$$

In diesem Beispiel hat jedes Polynom genau eine Grammatrix. Ist allerdings $d \geq 2$, stimmt das nicht mehr. △

Satz 2.2.17. *Ein Polynom $p \in R[x]_{2d}$ ist genau dann eine Quadratsumme von Polynomen, wenn p eine positiv semidefinite Grammatrix M besitzt. In diesem Fall ist p eine Summe von* $\mathrm{rang(M)}$ *vielen Quadraten, also maximal* $\binom{n+d}{d}$.

Beweis. Sei $M \in \mathrm{Sym}_{\Delta_d}(R)$ eine positiv semidefinite Grammatrix von p. Nach Lemma 2.2.3 finden wir $\mathrm{rang}(M)$ viele Vektoren $v_i \in R^{\Delta_d}$ mit $M = \sum_i v_i v_i^t$. Damit ist

$$p = X_d^t M X_d = \sum_i X_d^t v_i v_i^t X_d = \sum_i (v_i^t X_d)^2$$

eine Summe von Quadraten, denn $v_i^t X_d \in R[x]_d$.
Sei umgekehrt $p = \sum_i q_i^2$ eine Summe von Quadraten von Polynomen $q_i \in R[x]$. Nach Lemma 2.2.14 liegen alle $q_i \in R[x]_d$. Wir schreiben $q_i = v_i^t X_d$ für ein $v_i \in R^{\Delta_d}$. Es ist v_i also gerade der Koeffizientenvektor von q_i in der monomialen Basis. Also ist

$$p = \sum_i (v_i^t X_d)^2 = \sum_i X_d^t v_i v_i^t X_d = X_d^t \left(\sum_i v_i v_i^t \right) X_d,$$

und p besitzt die positiv semidefinite Grammatrix $\sum_i v_i v_i^t$. □

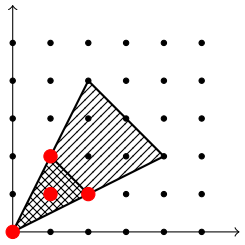
Beispiel 2.2.18. (i) Die Grammatrix von $p = x_1^2 - 2x_1x_2 + x_2^2 + 2x_1 - 2x_2 + 1$ aus Beispiel 2.2.16 ist positiv semidefinit. Man schreibt sie als vv^t mit $v = (1, 1, -1)^t$. Das liefert $p = (x_1 - x_2 + 1)^2$.
(ii) Die (einzige) Grammatrix des Polynoms $p = x_1^2 - 2x_1x_2 + x_2^2 + 2x_1 - 2x_2$ hat im Gegensatz zur Grammatrix aus dem vorigen Beispiel eine 0 in der linken oberen Ecke. Der linke obere 2×2 Minor ist also -1. Also ist p keine Quadratsumme. △

Bemerkung 2.2.19. Wir sehen, dass jede Quadratsummendarstellung von p zu eine positiv semidefiniten Grammatrix führt. Umgekehrt kann eine positiv semidefinite Grammatrix aber zu verschiedenen Quadratsummendarstellungen führen, denn die Zerlegung $M = \sum_i v_i v_i^t$ ist im Allgemeinen nicht eindeutig. △

Satz 2.2.20. *Sei $p \in R[x_1, \ldots, x_n]$ eine Quadratsumme von Polynomen, und sei $r := |\frac{1}{2}\mathcal{N}(p) \cap \mathbb{N}^n|$ die Anzahl der Gitterpunkte in $\frac{1}{2}\mathcal{N}(p)$. Dann ist p eine Summe von r Quadraten von Polynomen.*

Beweis. Wenn $p = q_1^2 + \cdots + q_m^2$ mit $q_i \in R[x]$, so tauchen in den q_i nur Exponenten aus $\frac{1}{2}\mathcal{N}(p)$ auf, nach Satz 2.2.11 (iii). Schreibt man also $q_i = v_i^t X_d$, so können in allen v_i immer nur dieselben r Einträge ungleich null sein. Damit hat $M = \sum_i v_i v_i^t$ aber höchstens Rang r, und da M eine positiv semidefinite Grammatrix von p ist, folgt die Aussage aus Satz 2.2.17. □

Beispiel 2.2.21. Wenn eine Quadratsumme etwa dasselbe Newtonpolytop wie das Motzkinpolynom hat, so ist sie eine Summe von 4 Quadraten:



Die allgemeine Abschätzung aus Satz 2.2.17 wäre hier nur $\binom{2+3}{3} = 10$. △

In den weiteren Kapiteln beschäftigen wir uns mit Polynomen, die zwar nicht unbedingt global, aber auf semialgebraischen Mengen nichtnegativ sind. Dazu entwickeln wir zuerst eine Theorie von angeordneten Ringen.

Kapitel 3

Angeordnete Ringe

In diesem Kapitel sei stets A ein kommutativer Ring mit 1. Ringhomomorphismen zwischen Ringen bilden 1 immer auf 1 ab. Falls A nullteilerfrei ist, bezeichnen wir seinen Quotientenkörper mit $\mathrm{Quot}(A)$. Nach wie vor bezeichne R einen reell abgeschlossenen Körper.

3.1 Präordnungen, Anordnungen und das reelle Spektrum

Definition 3.1.1. Eine *Präordnung* auf A ist eine Teilmenge $T \subseteq A$ mit

$$T + T \subseteq T, \quad T \cdot T \subseteq T, \quad A^2 \subseteq T, \quad -1 \notin T.$$

Die Menge $T \cap -T$ nennt man den *Support* von T und bezeichnet sie auch mit $\mathrm{supp}(T)$. △

Die Definition einer Präordnung ist also genau die gleiche wie bei Körpern. Allerdings gibt es Unterschiede in den Eigenschaften.

Bemerkung 3.1.2. (i) Die Menge ΣA^2 aller Quadratsummen von Elementen aus A ist genau dann eine Präordnung, wenn $-1 \notin \Sigma A^2$. Sie ist dann wiederum die kleinste Präordnung, d.h. in allen anderen Präordnungen enthalten.
(ii) Ist $\frac{1}{2} \in A$, so ist wie in Bemerkung 1.1.11 jedes Element aus A eine Differenz von zwei Quadraten. Die Bedingung $-1 \notin T$ kann also wieder ersetzt werden durch $T \neq A$.

(*iii*) Sei $A = R[x_1, \dots, x_n]$ der Polynomring über dem reell abgeschlossenen Körper R. Für jede nichtleere Teilmenge $S \subseteq R^n$ ist

$$T_S := \{p \in R[x] \mid p(a) \geq 0 \, \forall a \in S\}$$

eine Präordnung. Es gilt

$$\operatorname{supp}(T_S) = T_S \cap -T_S = \{p \in R[x] \mid p(a) = 0 \, \forall a \in S\}.$$

Für gewissen Mengen S enthält der Support also nicht nur die Null. So etwas kann in Körpern nicht passieren, wie wir in Bemerkung 1.1.11 gezeigt haben. Für den Beweis dort mussten wir auch durch Elemente teilen. Das geht in Ringen im allgemeinen nicht. △

Lemma 3.1.3. *Ist $T \subseteq A$ eine Präordnung mit zusätzlich $T \cup -T = A$, so ist $\operatorname{supp}(T)$ ein Ideal in A.*

Beweis. Setze $\mathfrak{p} := T \cap -T$. Die Eigenschaften $0 \in \mathfrak{p}$ und $\mathfrak{p} + \mathfrak{p} \subseteq \mathfrak{p}$ sind klar. Ebenso gilt $\pm T \cdot \mathfrak{p} \subseteq \mathfrak{p}$. Aus $A = T \cup -T$ folgt also $A \cdot \mathfrak{p} \subseteq \mathfrak{p}$. □

Definition 3.1.4. Eine *Anordnung* auf A ist eine Präordnung P, die zusätzlich

$$P \cup -P = A, \quad \operatorname{supp}(P) \text{ ist Primideal von } A$$

erfüllt. △

Beispiel 3.1.5. (*i*) Die Präordnung T_S von $R[x]$ aus Bemerkung 3.1.2 ist genau denn eine Anordnung, wenn $|S| = 1$.
(*ii*) Falls $A = K$ ein Körper ist, stimmt der Begriff einer Anordnung mit dem alten überein. Es ist nämlich $\operatorname{supp}(P) = \{0\}$ ein Primideal in K (das einzige).
(*iii*) Ist $\varphi \colon A \to B$ ein Ringhomomorphismus und $P \subseteq B$ eine Anordnung, so ist $\varphi^{-1}(P)$ eine Anordnung von A. Dabei ist $\operatorname{supp}(\varphi^{-1}(P)) = \varphi^{-1}(\operatorname{supp}(P))$. Insbesondere erhalten wir für nullteilerfreie Ringe durch die Einbettung

$$A \subseteq \operatorname{Quot}(A)$$

aus jeder Körperanordnung von $\operatorname{Quot}(A)$ eine Anordnung von A mit Support $\{0\}$.
(*iv*) Sei $P \subseteq A$ eine Anordnung mit $\mathfrak{p} := \operatorname{supp}(P)$. Sei $A/\mathfrak{p}$ der Restklassenring modulo $\mathfrak{p}$ und $\pi_\mathfrak{p} \colon A \to A/\mathfrak{p}$ die kanonische Projektion. Dann ist $\overline{P} := \pi_\mathfrak{p}(P)$ eine Anordnung auf $A/\mathfrak{p}$ mit $\operatorname{supp}(\overline{P}) = \{0\}$ (Übungsaufgabe 24).

(v) Wir betrachten die Einbettung $R[t] \subseteq R(t)$. Auf $R(t)$ haben wir die Anordnung $P_{a_+}, P_{a_-}, P_\infty, P_{-\infty}$ (siehe Beispiel 1.1.4 (ii)). Sie induzieren also Anordnungen auf $R[t]$ mit Support $\{0\}$. Es ist zum Beispiel

$$P_{a+} = \{p \in R[t] \mid \exists \varepsilon > 0 : \ p > 0 \text{ auf } (a, a+\varepsilon)\} \cup \{0\}.$$

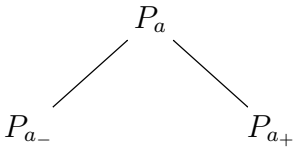
Es gibt aber auf $R[t]$ noch weitere Anordnungen. Ein Beispiel ist

$$P_a = \{p \in R[t] \mid p(a) \geq 0\}$$

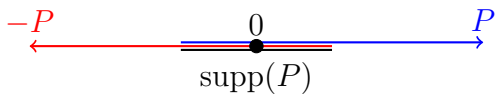
mit

$$\operatorname{supp}(P_a) = \{p \in R[t] \mid p(a) = 0\} = (t-a).$$

Man sieht nun dass sowohl $P_{a_+} \subsetneq P_a$ als auch $P_{a_-} \subsetneq P_a$ gilt. Zum Beispiel ist $t - a \in P_a \setminus P_{a_-}$ und $a - t \in P_a \setminus P_{a_+}$.



In Körpern kann so etwas nicht vorkommen, wie wir in Bemerkung 1.1.11 (iv) gesehen haben. Die Anordnung P_a hat ja auch kein Gegenstück in $R(t)$.
(vi) Man kann P wieder als die Menge aller positiven Elemente einer Ordnungrelation $\leq$ auf A auffassen. Dabei übersetzen sich die Axiome wieder in gewisse Verträglichkeitseigenschaften mit den Ringoperationen, ähnlich wie im Fall von Körpern. Es kann nun allerdings passieren, dass $a \leq 0$ und $a \geq 0$ gilt, ohne dass $a = 0$ ist (nämlich genau dann wenn $a \in \operatorname{supp}(P)$.) Die Relation ist also nicht notwendig antisymmetrisch. Wenn man sich den Ring als Zahlenstrahl vorstellt, sieht das ganze so aus:



△

Lemma 3.1.6. *Seien P, P', P'' Anordnungen von A. Dann gilt:*

(i) $P \subseteq P' \Rightarrow \operatorname{supp}(P) \subseteq \operatorname{supp}(P')$.

(ii) $P \subseteq P', \operatorname{supp}(P) = \operatorname{supp}(P') \Rightarrow P = P'$.

(iii) $P \subseteq P', P \subseteq P'' \Rightarrow P' \subseteq P''$ *oder* $P'' \subseteq P'$.

Beweis. (i) ist klar. Für (ii) sei $a \in P' \setminus P$. Dann ist $-a \in P \subseteq P'$, also $a \in \mathrm{supp}(P') = \mathrm{supp}(P) \subseteq P$, ein Widerspruch. Für (iii) sei $a \in P' \setminus P''$ und $b \in P'' \setminus P'$. Setze $c = a - b$. Wäre $c \in P$ so folgt $c \in P''$ und also $a \in P''$, ein Widerspruch. Wäre $-c \in P$ so folgt $-c \in P'$ und also $b \in P'$, ebenfalls ein Widerspruch. Das ist insgesamt ein Widerspruch zu $P \cup -P = A$. □

Wir wollen wieder untersuchen, wie man Präordnungen zu Anordnungen erweitern kann.

Lemma 3.1.7. *Sei $T \subseteq A$ eine Präordnung und $a, b \in A$ mit $ab \in -T$. Dann ist entweder $T + aT$ oder $T + bT$ wieder eine Präordnung.*

Beweis. Es ist nur zu prüfen ob -1 zu den potentiellen Präordnungen gehört. Angenommen es wäre

$$-1 = t_1 + as_1 \text{ und } -1 = t_2 + bs_2$$

mit $t_1, t_2, s_1, s_2 \in T$. Dann gilt

$$(1 + t_1)(1 + t_2) = abs_1s_2$$

und also

$$-1 = t_1 + t_2 + t_1t_2 - abs_1s_2 \in T,$$

ein Widerspruch. □

Satz 3.1.8. *Jede Präordnung von A ist in einer Anordnung enthalten.*

Beweis. Sei T eine Präordnung. Genau wie im Beweis von Satz 1.1.15 wählen wir mit dem Zorn'schen Lemma eine maximale über T liegende Präordnung P. Sei $a \in A$. Wegen $a(-a) = -a^2 \in -P$ folgt mit Lemma 3.1.7 dass $P + aP$ oder $P - aP$ eine Präordnung ist. Aus der Maximalität von P folgt also $a \in P$ oder $-a \in P$. Also gilt $P \cup -P = A$. Nach Lemma 3.1.3 ist damit $\mathfrak{p} := \mathrm{supp}(P)$ schon ein Ideal, und es bleibt noch die Primidealeigenschaft zu zeigen. Seien also $a, b \in A$ mit $ab \in \mathfrak{p}$ und $a \notin \mathfrak{p}$. Es gilt entweder $a \notin P$ oder $-a \notin P$. Wir nehmen o.B.d.A. $a \notin P$ an (ansonsten ersetze a durch $-a$.) Dann ist $P + aP$ keine Präordnung mehr, aufgrund der Maximalität. Es gilt nun mit Lemma 3.1.7 und der Maximalität von P :

$$\begin{aligned} ab \in -P \quad &\Rightarrow \quad P + bP \text{ ist Präordnung } \Rightarrow \quad P = P + bP \ni b \\ a(-b) \in -P \quad &\Rightarrow \quad P - bP \text{ ist Präordnung } \Rightarrow \quad P = P - bP \ni -b \end{aligned}$$

Das bedeutet aber gerade $b \in \mathfrak{p}$. □

Korollar 3.1.9. *Ein Ring besitzt genau dann eine Anordnung wenn* $-1 \notin \Sigma A^2$.

Definition 3.1.10. (i) Ein Ring heiß *semireell*, wenn $-1 \notin \Sigma A^2$. Er heißt *reell*, wenn aus $a_1^2 + \cdots + a_m^2 = 0$ immer schon $0 = a_1 = \ldots = a_m$ folgt.
(ii) Ein Ideal $I \subseteq A$ heiß *(semi)-reell*, wenn A/I (semi)-reell ist. △

Bemerkung 3.1.11. (i) Reell impliziert semireell.
(ii) Für Körper fallen die Begriffe reell und semireell zusammen, wie wir schon in Definition 1.1.18 gesehen haben.
(iii) Für Ringe sind die beiden Begriffe im Allgemeinen unterschiedlich. Sei $A = \mathbb{R}[x,y]/(x^2+y^2)$. Dann ist A nicht reell, denn es gilt $x^2 + y^2 = 0$ mit $x, y \neq 0$. A ist aber semireell, denn Elemente von A können wohldefiniert im Ursprung ausgewertet werden. Quadratsummen sind dabei nichtnegativ.
(iv) Ein Ring ist semireell genau dann, wenn er eine Anordnung besitzt (Korollar 3.1.9). Ein nullteilerfreier Ring ist genau dann reell, wenn er eine Anordnung P mit $\operatorname{supp}(P) = \{0\}$ besitzt. Sei zunächst P eine solche Anordnung, und sei $a_1^2 + \cdots + a_m^2 = 0$. Auflösen nach a_i^2 liefert $a_i^2 \in \operatorname{supp}(P) = \{0\}$ und $a_i = 0$ für alle i, aus der Nullteilerfreiheit. Sei umgekehrt A reell. Dann ist $\operatorname{Quot}(A)$ aber ein reeller Körper, wie man leicht sieht. Somit besitzt $\operatorname{Quot}(A)$ eine Anordnung (die automatisch Support Null hat), und sie schneidet sich auf A herunter.
(v) Ein Ideal ist genau dann semireell, wenn es im Support einer Anordnung enthalten ist. Ein Primideal $\mathfrak{p}$ ist nach (4) genau dann reell, wenn es eine Anordnung P von A gibt mit $\mathfrak{p} = \operatorname{supp}(P)$. △

Proposition 3.1.12. *Sei A nullteilerfrei und $K = \operatorname{Quot}(A)$. Dann stehen die Körperanordnungen Q von K in Bijektion zu den Ringanordnungen P von A mit* $\operatorname{supp}(P) = \{0\}$. *Die Bijektion ist dabei wie folgt:*

$$Q \mapsto Q \cap A$$

und

$$P \mapsto \operatorname{Quot}(P) := \left\{ \frac{a}{b} \mid ab \in P \right\}.$$

Beweis. Übungsaufgabe 25. □

Für ein beliebiges Primideal $\mathfrak{p}$ des Rings A bezeichen wir den sogenannten *Restklassenkörper* $\operatorname{Quot}(A/\mathfrak{p})$ mit $K_\mathfrak{p}$. Es gibt einen natürlichen Ringhomomorphismus

$$\rho_\mathfrak{p} \colon A \xrightarrow{\pi_\mathfrak{p}} A/\mathfrak{p} \xrightarrow{\iota_\mathfrak{p}} K_\mathfrak{p}.$$

Dabei ist wieder $\pi_\mathfrak{p}$ die kanonische Projektion auf die Restklassen (deren Kern gerade $\mathfrak{p}$ ist), und $\iota_\mathfrak{p}$ die kanonische Inklusion in den Restklassenkörper. Der folgende Satz zeigt, dass man den Begriff einer Ringanordnung vollständig auf Körperanordnungen zurückführen kann.

Satz 3.1.13. *Es gibt eine Bijektion zwischen der Menge aller Ringanordnungen P von A und der Menge aller Tupel $(\mathfrak{p}, Q)$, wobei $\mathfrak{p}$ ein Primideal von A und Q eine Körperanordnung des Restklassenkörpers $K_\mathfrak{p}$ ist. Die Bijektion ist dabei wie folgt (mit der Notation aus Proposition 3.1.12):*

$$\begin{aligned} P &\mapsto (\mathfrak{p} := \operatorname{supp}(P), \operatorname{Quot}(\pi_\mathfrak{p}(P)) \\ (\mathfrak{p}, Q) &\mapsto \rho_\mathfrak{p}^{-1}(Q). \end{aligned}$$

Beweis. Für eine Ringanordnung ist $\operatorname{Quot}(\pi_\mathfrak{p}(P))$ eine Körperanordnung von $K_\mathfrak{p}$. Das folgt aus Beispiel 3.1.5 (iv) und Proposition 3.1.12. Umgekehrt ist für eine Körperanordnung Q von $K_\mathfrak{p}$ das Urbild $\rho_\mathfrak{p}^{-1}(Q)$ eine Anordnung von A, wie in Beispiel 3.1.5 (iii) überlegt.
Wir zeigen nun, dass die beiden Konstruktionen invers zueinander sind. Starten wir zunächst mit P und setzen $\mathfrak{p} := \operatorname{supp}(P)$, so gilt

$$\rho_\mathfrak{p}^{-1}(\operatorname{Quot}(\pi_\mathfrak{p}(P)) = \pi_\mathfrak{p}^{-1}(\pi_\mathfrak{p}(P)) = P + \mathfrak{p} = P.$$

Für die erste Gleichung verwenden wir dabei Proposition 3.1.12. Starten wir umgekehrt mit $(\mathfrak{p}, Q)$, so gilt zunächst $P := \rho_\mathfrak{p}^{-1}(Q) = \pi_\mathfrak{p}^{-1}(\iota_\mathfrak{p}^{-1}(Q))$. Da $\iota_\mathfrak{p}^{-1}(Q)$ nach Proposition 3.1.12 eine Anordnung von $A/\mathfrak{p}$ mit Support $\{0\}$ ist, ist der Support von P gerade $\mathfrak{p}$ (vergleiche Beispiel 3.1.5 (iii)). Wendet man $\pi_\mathfrak{p}$ wieder auf P an, erhält man offensichtlich wieder $\iota_\mathfrak{p}^{-1}(Q)$, und mit Proposition 3.1.12 sind wir fertig. □

Definition 3.1.14. Sei A ein Ring. Die Menge aller Anordnungen nennt man das *reelle Spektrum von A*:

$$\begin{aligned} \operatorname{Sper}(A) &= \{P \subseteq A \mid P \text{ Anordnung}\} \\ &= \{(\mathfrak{p}, Q) \mid \mathfrak{p} \text{ Primideal}, Q \text{ Anordnung von } K_\mathfrak{p}\}. \end{aligned}$$ △

Beispiel 3.1.15. (i) Ein reell abgeschlossener Körper R hat nur eine Anordnung. Also ist $\operatorname{Sper}(R)$ eine einpunktige Menge. Das gleiche gilt etwa für $\operatorname{Sper}(\mathbb{Q})$.
(ii) Sei $A = R[t]$. Wir haben in Beispiel 3.1.5 (v) schon die Anordnungen $P_{-\infty}, P_{a-}, P_a, P_{a+}, P_\infty$ kennengelernt. Sei nun $P = (\mathfrak{p}, Q)$ eine beliebige Anordnung von A. Da A ein

Hauptidealring ist, sind alle Ideale von der Gestalt (p) für ein $p \in A$. Primideale sind dabei von irreduziblen Polynomen erzeugt (bzw. von der Null). Falls $\mathfrak{p} = (0)$, so kommt P von $R(t)$ nach Proposition 3.1.12, und ist also eine der Anordnungen $P_{-\infty}, P_{a-}, P_{a+}, P_\infty$. Falls $\mathfrak{p} = ((t-a)^2 + b^2)$ mit $b \neq 0$, so ist $K_\mathfrak{p} = A/\mathfrak{p} = R[i]$, wie man sich leicht überlegt. Somit gibt es auf $K_\mathfrak{p}$ keine Anordnung. Falls $\mathfrak{p} = (t-a)$ ist $K_\mathfrak{p} = A/\mathfrak{p} = R$, und Q ist eindeutig bestimmt. $\rho_\mathfrak{p} \colon A \to R$ ist dabei einfach die Auswertung in a, und also ist $P = \rho_\mathfrak{p}^{-1}(Q) = P_a$. Also kennen wir hier wirklich schon alle Anordnungen:

$$\operatorname{Sper}(R[t]) = \left\{ P_{-\infty}, P_{a-}, P_a, P_{a+}, P_\infty \mid a \in R \right\}.$$

(iii) Für $A = \mathbb{Z}$ überlegt man entweder direkt dass $\Sigma\mathbb{Z}^2$ die einzige Anordnung ist, oder man geht wie folgt vor. Die Primideale in $\mathbb{Z}$ sind (0) und (p) mit $p \in \mathbb{Z}$ prim. Es ist $K_{(p)} = \mathbb{Z}/(p)$ und hier gibt es keine Anordnung, da $\operatorname{char}(K_{(p)}) \neq 0$. Es ist weiter $K_{(0)} = \mathbb{Q}$, und hier gibt es wiederum nur die Anordnung $\Sigma\mathbb{Q}^2$. Also ist die einzige Anordnung von $\mathbb{Z}$ gerade $\Sigma\mathbb{Q}^2 \cap \mathbb{Z} = \Sigma\mathbb{Z}^2$. △

Man kann nun Elemente von A (auf zunächst vielleicht ungewöhnliche Weise) als Funktionen auf $\operatorname{Sper}(A)$ auffassen. Für $a \in A$ und $P \in \operatorname{Sper}(A)$ definieren wir

$$\hat{a}(P) = \hat{a}(\mathfrak{p}, Q) = \rho_\mathfrak{p}(a) \in K_\mathfrak{p}.$$

Wir bilden das Element $a \in A$ also anhand des kanonischen Morphismus in den Restklassenkörper von $\mathfrak{p} = \operatorname{supp}(P)$ ab. Man beachte, dass das Bild $\hat{a}(P)$ abhängig von P in jeweils einem anderen Körper liegen kann, beziehungsweise

$$\hat{a} \colon \operatorname{Sper}(A) \to \bigcup_{\mathfrak{p} \text{ prim}} K_\mathfrak{p}.$$

Man beachte auch, dass das Element $\hat{a}(P)$ nur von $\mathfrak{p} = \operatorname{supp}(P)$ anhängt.

Beispiel 3.1.16. (i) Für einen reellen Körper K ist $\mathfrak{p} = \operatorname{supp}(P) = \{0\}$ für alle $P \in \operatorname{Sper}(K)$. Also ist immer $K_\mathfrak{p} = K$ und $\rho_\mathfrak{p} = \mathrm{id}$. Also ist

$$\begin{aligned} \hat{a} \colon \operatorname{Sper}(K) &\to K \\ P &\mapsto a \end{aligned}$$

für alle $a \in K$ die konstante Abbildung.
(ii) Für einen reell abgeschlossenen Körper R ist R^n eine Teilmenge von $\operatorname{Sper} R[x_1, \ldots, x_n]$, indem man $a \in \mathbb{R}^n$ identifiziert mit

$$P_a = \{p \in R[x] \mid p(a) \geq 0\}.$$

Dabei ist

$$\mathfrak{p} = \operatorname{supp}(P_a) = \{p \in R[x] \mid p(a) = 0\} = (x_1 - a_1, \ldots, x_n - a_n)$$

und also ist $\rho_{\mathfrak{p}} \colon R[x] \to R$ einfach die Auswertung in a. Also ist

$$\hat{p}(P_a) = p(a),$$

d.h. die Funktion $\hat{p}$ stimmt auf R^n mit der polynomialen Funktion p überein. Man beachte allerdings, dass $\operatorname{Sper} R[x]$ noch mehr Elemente besitzt, auf denen $\hat{p}$ ebenfalls definiert ist, und Werte in anderen Körper annehmen kann.
(*iii*) Im Fall $A = R[t]$ kennen wir $\operatorname{Sper}(A)$ vollständig. Es gibt die Elemente P_a mit $a \in R$, und $\hat{p}(P_a) = p(a)$, wie in (2) gesehen. Für P_{a_+} etwa ist $\mathfrak{p} = \operatorname{supp}(P_{a_+}) = (0)$, also $K_{\mathfrak{p}} = R(t)$ und $\rho_{\mathfrak{p}} \colon R[t] \to R(t)$ ist einfach die Einbettung. Für die Anordnungen $P_{a_-}, P_{-\infty}$ und P_∞ gilt das gleiche. Für $p \in R[t]$ ist also

$$\begin{aligned} \hat{p} \colon \operatorname{Sper}(R[t]) &\to R \cup R(t) \\ P_a &\mapsto p(a) \in R \\ P_{a_+}, P_{a_-}, P_{-\infty}, P_\infty &\mapsto p \in R(t). \end{aligned}$$ △

Anhand der eben definierten Funktionen kann man nun semialgebraische Teilmengen des reellen Spektrums eines Rings definieren. Für $a \in A$ und eine Anordnung $P = (\mathfrak{p}, Q)$ ist ja $\hat{a}(P)$ ein Element von $K_{\mathfrak{p}}$, und Q ist eine Anordnung dieses Körpers. Wir definieren nun

$$\begin{aligned} \hat{a}(P) > 0 \quad &:\Leftrightarrow \quad \hat{a}(P) >_Q 0 \text{ in } K_{\mathfrak{p}} \\ \hat{a}(P) \geq 0 \quad &:\Leftrightarrow \quad \hat{a}(P) \geq_Q 0 \text{ in } K_{\mathfrak{p}} \\ \hat{a}(P) = 0 \quad &:\Leftrightarrow \quad \hat{a}(P) = 0 \text{ in } K_{\mathfrak{p}}. \end{aligned}$$

Hier wird also nun auch die Anordnung Q von $K_{\mathfrak{p}}$ benutzt, während für die Definition von $\hat{a}(P)$ ja nur $\mathfrak{p} = \operatorname{supp}(P)$ eine Rolle spielte. Man kann diese Relationen auch ohne Bezug auf die Restklassenkörper beschreiben. Wenn wir die Äquivalenz aus Satz 3.1.13 verwenden, sehen wir

$$\begin{aligned} \hat{a}(P) > 0 \quad &\Leftrightarrow \quad a \notin -P \\ \hat{a}(P) \geq 0 \quad &\Leftrightarrow \quad a \in P \\ \hat{a}(P) = 0 \quad &\Leftrightarrow \quad a \in \operatorname{supp}(P). \end{aligned}$$

Die erste Sichtweise ist aber oft besser, weil wir uns rechts in einem angeordneten Körper befinden, und dort wie gewohnt rechnen können.
Für $a_1, \ldots, a_m \in A$ setzen wir

$$O(a_1, \ldots, a_m) = \{P \in \operatorname{Sper}(A) \mid \hat{a}_1(P) > 0, \ldots, \hat{a}_m(P) > 0\}$$

$$V(a_1, \ldots, a_m) = \{P \in \operatorname{Sper}(A) \mid \hat{a}_1(P) = 0, \ldots, \hat{a}_m(P) = 0\}\,.$$

Definition 3.1.17. Eine endliche boolsche Kombination von Mengen der Gestalt $O(a_1, \ldots, a_m)$ mit $a_i \in A$ heißt *semialgebraische Teilmenge von* $\operatorname{Sper}(A)$. △

Bemerkung 3.1.18. Man sieht leicht, dass Lemma 1.5.3 hier genauso gilt, also jede semialgebraische Menge ein Standardform

$$\bigcup_i (V(a_i) \cap O(b_{i1}, \ldots, b_{im_i}))$$

mit $a_i, b_{ij} \in A$ hat. △

Beispiel 3.1.19. Fasst man R^n wieder als Teilmenge von $\operatorname{Sper}(R[x])$ auf wie in Beispiel 3.1.16 (ii), so sind die semialgebraischen Teilmengen von $\operatorname{Sper}(R[x])$, schneidet man sie mit R^n, genau die bekannten semialgebraischen Mengen von R^n. △

Definition 3.1.20. Die *spektrale Topologie* auf $\operatorname{Sper}(A)$ ist die Topologie mit den Mengen $O(a_1, \ldots, a_m)$ als Basis. Die offenen Mengen sind also gerade die (beliebigen) Vereinigungen solcher Mengen. △

Definition 3.1.21. Die *konstruierbare Topologie* auf $\operatorname{Sper}(A)$ ist die Topologie mit allen semialgebraischen Mengen als Basis. Sie hat zum Beispiel die Mengen $O(a_1, \ldots, a_m)$ und ihre Komplemente als Subbasis. △

Offensichtlich ist die konstruierbare Topologie feiner als die spektrale Topolgie, d.h sie hat mehr offene Mengen.

Satz 3.1.22. *Die konstruierbare Topologie ist hausdorffsch und quasi-kompakt (d.h. es gilt die endliche Überdeckungseigenschaft mit offenen Mengen) . Die spektrale Topologie ist ebenfalls quasi-kompakt, aber im Allgemeinen nicht hausdorffsch.*

Beweis. Seien $P, Q \in \operatorname{Sper}(A)$ mit $P \neq Q$. Dann gibt es o.B.d.A. ein $a \in P \setminus Q$. Dann ist $Q \in O(-a)$ und $P \in O(-a)^c$, und also lassen sich P und Q durch

zwei offene disjunkte Mengen der konstruierbaren Topologie trennen. Also ist die konstruierbare Topologie hausdorffsch.
Der Raum $\mathrm{Sper}(R[t])$ mit der spektralen Topolgie ist nicht hausdorffsch. Dass eine Menge $O(p)$ die Anordnung P_a enthält bedeutet gerade $p(a) > 0$. Dann ist aber p strikt positiv auf einem Intervall $(a-\epsilon, a+\epsilon)$, und also ist p auch strikt positiv bei P_{a_-} und P_{a_+}, d.h. auch diese beiden Anordnungen liegen in $O(p)$. Damit lassen sie sich nicht durch disjunkte offenen Mengen von P_a trennen.
Die Quasi-Kompaktheit zeigen wir für die konstruierbare Topologie. Für die spektrale Topologie folgt sie dann aus der Tatsache, dass es dort weniger offenen Mengen gibt. Wir fassen nun $\mathrm{Sper}(A)$ als Teilmenge von

$$\{0,1\}^A = \{g\colon A \to \{0,1\}\}$$

auf, indem wir die Teilmenge $P \subseteq A$ mit ihrer charakteristischen Funktion identifizieren. Die Menge $\{0,1\}$ mit der feinstmöglichen Topologie ist kompakt, und nach dem Satz von Tychonoff ist $\{0,1\}^A$ ebenfalls kompakt. Die Produkttopologie ist aber die gröbste Topologie, die alle Projektionen stetig macht, d.h. alle Einsetzungen von Punkten aus a, wenn man $\{0,1\}^A$ als Funktionen auf A auffasst. Auf $\mathrm{Sper}(A)$ wird die induzierte Topologie also gerade durch die Mengen $O(a)$ und deren Komplemente erzeugt. Die induzierte Topolgie ist also gerade die konstruierbare Topologie. Da abgeschlossene Mengen quasi-kompakter Räume wieder quasi-kompakt sind, zeigen wir nur noch, dass $\mathrm{Sper}(A)$ eine abgeschlossene Teilmenge von $\{0,1\}^A$ ist. Sei dazu $M \in \{0,1\}^A \setminus \mathrm{Sper}(A)$, d.h. $M \subseteq A$ ist keine Anordnung. Wir konstruieren eine offene Teilmenge O von $\{0,1\}^A$ die M enthält und disjunkt zu $\mathrm{Sper}(A)$ ist. Dass M keine Anordnung ist kann verschiedene Gründe haben. Beispielsweise könnte es $a, b \in M$ geben mit $a+b \notin M$. In diesem Fall wäre

$$O = \{N \subseteq A \mid a, b \in N, a+b \notin N\}$$

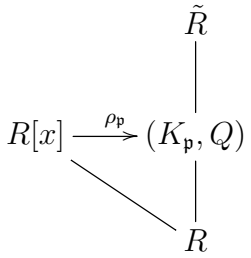
eine solche offene Teilmenge. Alle anderen Möglichkeiten lassen sich ebenso behandeln, wie man sich leicht überlegt (Übungsaufgabe 27). □

Satz 1.1.16 aus Kapitel 1 kann in der eben eingeführten Sprache des reellen Spektrums als abstrakter Positivstellensatz für Körper aufgefasst werden: wenn für ein $a \in K$ die Funktion $\hat{a}$ an jedem Punkt von $\mathrm{Sper}(K)$ nichtnegativ ist, ist a eine Quadratsumme in K. Anhand des Transferprinzips von Tarski-Seidenberg haben wir diesen abstrakten Positivstellensatz dann in einen konkreten umgewandelt: wenn ein Element $p \in R[x]$ an jedem Punkt des R^n nichtnegativ ist, ist

es eine Quadratsumme in $K = R(x)$. Dazu haben wir gezeigt, dass die Nichtnegativität von p auf R^n die Nichtnegativität von $\hat{p}$ auf dem Raum $\mathrm{Sper}(R(x))$ induziert. Eine leichte Verallgemeinerung davon ist der folgende Satz:

Satz 3.1.23. *Ist $p \in R[x]$ nichtnegativ auf dem R^n, so ist $\hat{p}$ nichtnegativ auf $\mathrm{Sper}(R[x])$ (d.h. p liegt in allen Anordnungen von $R[x]$)*

Beweis. Angenommen $\hat{p}(P) < 0$ für ein $P = (\mathfrak{p}, Q) \in \mathrm{Sper}(R[x])$. Sei $\tilde{R}$ der reelle Abschluss von $(K_\mathfrak{p}, Q)$:



Es gilt in $K_\mathfrak{p}$ (und damit in $\tilde{R}$)

$$0 > \rho_\mathfrak{p}(p) = p(\rho_\mathfrak{p}(x_1), \ldots, \rho_\mathfrak{p}(x_n)).$$

Nach dem Transferprinzip von Tarski-Seidenberg gibt es dann aber auch einen Punkt $a \in R^n$ mit $p(a) < 0$. □

Eine leichte Variation des Arguments zeigt sogar:

Korollar 3.1.24. *R^n liegt dicht in $\mathrm{Sper}(R[x])$, bezüglich der konstruierbaren (und damit auch der spektralen) Topologie.*

Beweis. Übungsaufgabe 28. □

Im nächsten Abschnitt beweisen wir zunächst abstrakte Positivstellensätze für Ringe, also Sätze wie 1.1.15 und 1.1.16. Sie können dann wiederum mit dem Transferprinzip in konkrete Positivstellensätze umgewandelt werden.

3.2 Positivstellensätze für Ringe

Für Körper haben wir in Satz 1.1.16 gezeigt, dass

$$\Sigma K^2 = \bigcap_{P \text{ Anordnung}} P$$

gilt. In Ringen stimmt das nicht. Aus Satz 3.1.23 folgt, dass das Motzkinpolynom in jeder Anordnung von $A = R[x, y]$ liegt. Es ist aber keine Quadratsumme in A. Wir müssen Satz 1.1.16 (und Satz 1.1.15) also geeignet anpassen. Seien dazu T eine Präordnung von A, I ein Ideal von A und $G \subseteq A$ eine multiplikativ abgeschlossene Menge mit $1 \in G$.

Proposition 3.2.1. *Die folgenden Aussagen sind äquivalent:*

(i) Es gibt **keine** *Anordnung $P \in \operatorname{Sper}(A)$ mit*

$$\begin{aligned} &\hat{t}(P) \geq 0 \textit{ für alle } t \in T \quad (\textit{d.h. } T \subseteq P) \\ &\hat{i}(P) = 0 \textit{ für alle } i \in I \quad (\textit{d.h. } I \subseteq \operatorname{supp}(P)) \\ &\hat{g}(P) \neq 0 \textit{ für alle } g \in G \quad (\textit{d.h. } G \cap \operatorname{supp}(P) = \emptyset). \end{aligned}$$

(ii) Es gibt $i \in I, g \in G$ und $t \in T$ mit $g^2 + t = i$.

Beweis. $(ii) \Rightarrow (i)$ ist einfach: Für ein Gegenbeispiel $P \in \operatorname{Sper}(A)$ gälte

$$0 = \hat{i}(P) = \widehat{g^2 + t}(P) = \hat{g}(P)^2 + \hat{t}(P) > 0,$$

ein Widerspruch. $(i) \Rightarrow (ii)$: Wir setzen $B := A/I$, betrachten die Projektion auf die Restklassen

$$\pi \colon A \to B$$

und setzen $\overline{G} := \pi(G), \overline{T} = \pi(T)$. Da $\overline{G}$ eine multiplikativ abgeschlossene Teilmenge von B ist, können wir die Lokalisierung nach $\overline{G}$ betrachten, d.h.

$$C := \overline{G}^{-1}B = \left\{ \frac{b}{g} \mid a \in B, b \in \overline{G} \right\},$$

wobei wie üblich eine Äquivalenzrelation definiert ist durch

$$\frac{b}{g} = \frac{c}{h} \quad :\Leftrightarrow \quad f(bh - cg) = 0 \text{ für ein } f \in \overline{G}.$$

Es gibt wieder einen kanonischen Homomorphismus, und zwar

$$\iota \colon B \to C; \quad b \mapsto \frac{b}{1}.$$

Sei nun

$$T' := \left\{ \frac{t}{g^2} \mid t \in \overline{T}, g \in \overline{G} \right\} \subseteq C.$$

1. Fall: $-1 \in T'$, d.h. $f(g^2+t) = 0$ für gewisse $t \in \overline{T}, f, g \in \overline{G}$. Dann ist auch $(fg)^2 + f^2 t = 0$ in B, und nach Zurückziehen mittels π erhalten wir die gewünschte Gleichung in A.
2. Fall: $-1 \notin T'$, d.h. T' ist eine Präordnung in C. Nach Satz 3.1.8 gibt es eine Anordnung P' von C mit $T' \subseteq P'$. Dann ist das Urbild $P := (\iota \circ \pi)^{-1}(P')$ eine Anordnung in A. Offensichtlich gilt $T \subseteq P$ und auch $I \subseteq \operatorname{supp}(P)$, denn bereits $\pi(i) = 0$ gilt für alle $i \in I$. Da für $g \in G$ das Element $\iota(\pi(g))$ in C invertierbar ist, kann es nicht zum Ideal $\operatorname{supp}(P')$ gehört haben. Also ist $g \notin \operatorname{supp}(P)$. Also erfüllt P alle Bedingungen aus (i), ein Widerspruch, und also kann dieser Fall nicht eintreten. □

Für eine Teilmenge $T \subseteq A$ schreiben wir

$$W(T) := \left\{ P \in \operatorname{Sper}(A) \mid \hat{t}(P) \geq 0 \text{ für alle } t \in T \right\}$$

$$V(T) = \left\{ P \in \operatorname{Sper}(A) \mid \hat{t}(P) = 0 \text{ für alle } t \in T \right\}.$$

Satz 3.2.2 (Abstrakter Positivstellensatz)**.** *Sei $T \subseteq A$ eine Präordnung. Dann gilt für jedes $a \in A$*

$$\hat{a} > 0 \text{ auf } W(T) \quad \Leftrightarrow \quad t_1 a = 1 + t_2 \text{ für gewisse } t_1, t_2 \in T.$$

Beweis. "⇐": Für $P \in W(T)$ ist

$$\hat{t}_1(P)\hat{a}(P) = \widehat{t_1 a}(P) = \hat{1}(P) + \hat{t}_2(P) = 1 + \hat{t}_2(P) > 0 \text{ in } K_{\mathfrak{p}}.$$

Nach Teilung durch das positive Element $\hat{t}_1(P)$ erhält man also $\hat{a}(P) > 0$.
"⇒": Verwende Proposition 3.2.1 mit $I = (0), G = \{1\}$ und der Präordnung $T - aT$. □

Satz 3.2.3 (Abstrakter Nichtnegativstellensatz)**.** *Sei $T \subseteq A$ eine Präordnung. Dann gilt für jedes $a \in A$*

$$\hat{a} \geq 0 \text{ auf } W(T) \quad \Leftrightarrow \quad t_1 a = a^{2m} + t_2 \text{ für gewisse } t_1, t_2 \in T, m \in \mathbb{N}.$$

Beweis. Übungsaufgabe 29. □

Satz 3.2.4 (Abstrakter reeller Nullstellensatz)**.** *Sei $I \subseteq A$ ein Ideal. Dann gilt für jedes $a \in A$*

$$\hat{a} = 0 \text{ auf } V(I) \quad \Leftrightarrow \quad a^{2m} + \sigma \in I \text{ für ein } m \in \mathbb{N}, \sigma \in \Sigma A^2.$$

Beweis. Übungsaufgabe 29. □

Definition 3.2.5. Sei A ein Ring und $I \subseteq A$ ein Ideal. Die Menge

$$\mathrm{rrad}(I) = \left\{a \in A \mid a^{2m} + \sigma \in I \text{ für ein } m \in \mathbb{N}, \sigma \in \Sigma A^2\right\}$$

heißt das *reelle Radikal von* I. △

Man beachte dass das gewöhnliche *Radikal* von I entsprechend, aber ohne die Quadratsummen σ definiert ist. Das Radikal von I ist der Durchschnitt aller über I liegenden Primideale. Hier gilt:

Satz 3.2.6. *Für jedes Ideal $I \subseteq A$ gilt*

$$\mathrm{rrad}(I) = \bigcap_{I \subseteq \mathfrak{p} \text{ reelles Primideal}} \mathfrak{p}.$$

Beweis. "$\subseteq$" ist offensichtlich aus der Definition eines reellen Ideals. Für "$\supseteq$" sei a aus jedem reellen Primideal über I. Das impliziert aber, dass $\hat{a}(P) = 0$ ist für jede Anordnung P mit $I \subseteq \mathrm{supp}(P)$. Nach Satz 3.2.4 liegt a dann aber in $\mathrm{rrad}(I)$. □

Man beachte, dass ein Ideal I genau dann reell ist, wenn $I = \mathrm{rrad}(I)$ gilt. Das folgt aus der Tatsache, dass $a_1^2 + \cdots + a_m^2 \in I$ schon $a_i \in I$ für alle i impliziert, für reelle Ideale.

3.3 Positivität auf semialgebraischen Mengen

Wir wollen, ganz wie im Körperfall, die abstrakten Positivstellensätze in konkrete Sätze über Polynome umwandeln. Wie das geht wissen wir im Prinzip schon. Wir verwenden das Transferprinzip.

Satz 3.3.1. *Es sei R ein reell abgeschlossener Körper und $p_1, \ldots, p_r, q_1, \ldots, q_s, f_1, \ldots, f_t \in R[x_1, \ldots, x_n]$. Dann sind äquivalent:*

(i) Es gibt $a \in R^n$ mit

$$\begin{aligned} p_1(a) \geq 0, \ldots, p_r(a) \geq 0 \\ q_1(a) \neq 0, \ldots, q_s(a) \neq 0 \\ f_1(a) = 0, \ldots, f_t(a) = 0. \end{aligned}$$

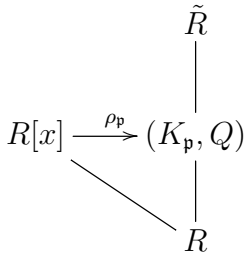
(ii) Es gibt $P \in \mathrm{Sper}(R[x])$ *mit*

$$\begin{aligned}\hat{p}_1(P) \geq 0, \ldots, \hat{p}_r(P) \geq 0\\ \hat{q}_1(P) \neq 0, \ldots, \hat{q}_s(P) \neq 0\\ \hat{f}_1(P) = 0, \ldots, \hat{f}_t(P) = 0.\end{aligned}$$

Beweis. "$(i) \Rightarrow (ii)$" ist wieder klar. Man nehme

$$P = P_a = \{p \in R[x] \mid p(a) \geq 0\}.$$

Für "$(ii) \Rightarrow (i)$" sei $P = (\mathfrak{p}, Q)$, $\rho_\mathfrak{p} \colon R[x] \to K_\mathfrak{p}$ der Restklassenhomomorphismus und $\tilde{R}$ der reelle Abschluss von $K_\mathfrak{p}$ bezüglich Q:



Es gilt in $K_\mathfrak{p}$ und damit in $\tilde{R}$ für alle j:

$$\begin{aligned}0 \leq \hat{p}_j(P) = \rho_\mathfrak{p}(p_j) = p_j(\rho_\mathfrak{p}(x_1), \ldots, \rho_\mathfrak{p}(x_n))\\ 0 \neq \hat{q}_j(P) = \rho_\mathfrak{p}(q_j) = q_j(\rho_\mathfrak{p}(x_1), \ldots, \rho_\mathfrak{p}(x_n))\\ 0 = \hat{f}_j(P) = \rho_\mathfrak{p}(f_j) = f_j(\rho_\mathfrak{p}(x_1), \ldots, \rho_\mathfrak{p}(x_n))\end{aligned}$$

Da es in $\tilde{R}^n$ einen Punkt gibt, der das R-polynomiale System löst, gibt es nach dem Transferprinzip einen solchen Punkt auch in R^n. □

Für Polynome $p_1, \ldots, p_r \in R[x_1, \ldots, x_n]$ betrachten wir die kleinste (potentielle) Präordnung, die diese enthält:

$$T(p_1, \ldots, p_r) = \left\{ \sum_{e \in \{0,1\}^r} \sigma_e p_1^{e_1} \cdots p_r^{e_r} \mid \sigma_e \in \Sigma R[x]^2 \right\},$$

sowie die sogenannte *basisch abgeschlossene Menge* im R^n:

$$W_R(p_1, \ldots, p_r) = \{a \in R^n \mid p_1(a) \geq 0, \ldots, p_r(a) \geq 0\}.$$

Für $f_1, \dots, f_t$ erinnern wir an die Definition

$$V_R(f_1, \dots, f_t) = \{a \in R^n \mid f_1(a) = 0, \dots, f_t(a) = 0\}$$

und

$$I(f_1, \dots, f_t) = \left\{ \sum_i g_i f_i \mid g_i \in R[x] \right\},$$

das von den f_i erzeugte Ideal.

Satz 3.3.2 (Konkreter Positivstellensatz). *Seien $p_1, \dots, p_r \in R[x]$. Für $p \in R[x]$ gilt dann*

$$p > 0 \text{ auf } W_R(p_1, \dots, p_r) \Leftrightarrow t_1 p = 1 + t_2 \text{ mit } t_1, t_2 \in T(p_1, \dots, p_r).$$

Beweis. Unmittelbare Folgerung aus dem abstrakten Positivstellensatz 3.2.2, da die Bedingung $p > 0$ auf $W_R(p_1, \dots, p_r)$ mit Satz 3.3.1 äquivalent zu $\hat{p} > 0$ auf $W(T(p_1, \dots, p_r))$ ist. □

Analog erhält man

Satz 3.3.3 (Konkreter Nichtnegativstellensatz). *Seien $p_1, \dots, p_r \in R[x]$. Für $p \in R[x]$ gilt dann*

$$p \geq 0 \text{ auf } W_R(p_1, \dots, p_r) \Leftrightarrow t_1 p = p^{2m} + t_2 \text{ mit } m \in \mathbb{N}, t_1, t_2 \in T(p_1, \dots, p_r).$$

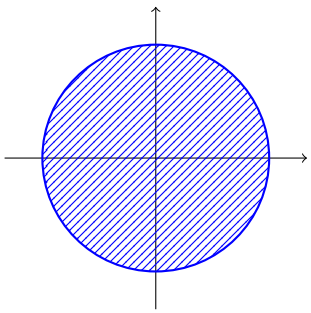
Satz 3.3.4 (Konkreter reeller Nullstellensatz). *Seien $f_1, \dots, f_t \in R[x]$. Für $p \in R[x]$ gilt dann*

$$p = 0 \text{ auf } V_R(f_1, \dots, f_t) \Leftrightarrow p \in \operatorname{rrad}(I(f_1, \dots, f_t)).$$

Beispiel 3.3.5. (*i*) Sei $q = 1 - x^2 - y^2 \in R[x, y]$. Dann ist

$$W_R(q) = \left\{(a, b) \in R^2 \mid a^2 + b^2 \leq 1\right\}$$

die Einheitskreisscheibe:



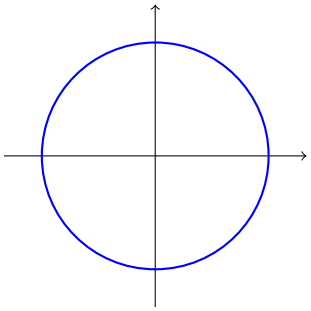
Es ist $T(q) = \{\sigma_1 + \sigma_2 q \mid \sigma_1, \sigma_2 \in \Sigma R[x]^2\}$. Wenn ein Polynom $p \in R[x, y]$ also strikt positiv auf $W_R(q)$ ist, gibt es eine Darstellung

$$(\sigma_1 + \sigma_2 q)p = 1 + \tau_1 + \tau_2 q$$

mit Quadratsummen $\sigma_1, \sigma_2, \tau_1, \tau_2$, und die Positivität kann man daran wiederum ablesen. Falls p nur nichtnegativ auf $W_R(q)$ ist, sieht die Darstellung wie folgt aus:

$$(\sigma_1 + \sigma_2 q)p = p^{2m} + \tau_1 + \tau_2 q.$$

(ii) Sei $f = 1 - x^2 - y^2 \in R[x, y]$. Dann ist $V_R(q)$ der Einheitskreis



und $I(f) = \{gf \mid g \in R[x, y]\}$. Ein Polynom p verschwindet genau dann auf dem Einheitskreis, wenn $p \in \operatorname{rrad}(I(f))$, also $p^{2m} + \sigma \in I(f)$. Man kann zeigen dass das Ideal $I(f)$ reell ist (Übungsaufgabe 30), und somit ist sogar $p \in I(f)$.
(iii) Sei $f = x^2 + y^2 \in R[x, y]$. Dann ist $V_R(f) = \{(0, 0)\}$. Falls $p(0, 0) = 0$, so ist $p^{2m} + \sigma = g(x^2 + y^2)$. Hier ist das Ideal nicht reell, zum Beispiel liegt x nicht in $I(f)$, sondern nur $x^2 + y^2$, also $m = 1$ und $\sigma = y^2$. △

Bemerkung 3.3.6. Der Hilbertsche Nullstellensatz klassifiziert die Polynome, die auf der *komplexen Varietät* $V_C(I)$ verschwinden. Dabei ist $C = R[i]$ der algebraische Abschluss von R. Ein Polynom verschwindet auf $V_C(I)$ genau dann, wenn eine Potenz in I liegt. In Beispiel 3.3.5 (3) sehen wir, dass das mit der kleineren Varietät $V_R(I)$ nicht stimmt. Keine Potenz von x liegt in $(x^2 + y^2)$. Allerdings verschwindet x eben auch nicht auf $V_C(I)$, da zum Beispiel $(1, i) \in V_C(I)$. △

Satz 3.3.7. *Sei $I \subseteq R[x_1, \ldots, x_n]$ ein Ideal. Dann gilt*

(i) I ist semireell genau dann wenn $V_R(I) \neq \emptyset$.

(ii) Ist I reduziert, so ist I reell genau dann wenn $V_R(I)$ R-Zariski-dicht in $V_C(I)$ liegt.

Beweis. (i) "⇒": Schreibe $I = (f_1, \ldots, f_t)$ wähle $P = (\mathfrak{p}, Q) \in \mathrm{Sper}(R[x])$ mit $I \subseteq \mathfrak{p}$. Im reellen Abschluss $\tilde{R}$ von $(K_\mathfrak{p}, Q)$ gibt es wieder ein Tupel $a \in \tilde{R}^n$ mit $f_i(a) = 0$ für alle i. Mit dem Transferprinzip gibt es das also auch in R^n. "⇐" ist klar, da Auswerten in einem Punkt $a \in R^n$ einen Algebrahomomorphismus von $R[x]/I$ nach R liefert.
(ii) "⇒": Sei $p = 0$ auf $V_R(I)$. Dann folgt $p \in \mathrm{rrad}(I) = I$, und also $p = 0$ auf $V_C(I)$. "⇐": Sei $p \in \mathrm{rrad}(I)$. Dann ist $p = 0$ auf $V_R(I)$ und also $p = 0$ auf $V_C(I)$. Mit dem Hilbertschen Nullstellensatz folgt $p^{2m} \in I$, und aus der Reduziertheit also $p \in I$. Somit ist $I = \mathrm{rrad}(I)$, und I ist also reell. □

Bemerkung 3.3.8. Der Positivstellensatz liefert ein algebraisches Zertifikat für die Unlösbarkeit eines polynomialen Ungleichungssystems

$$p_1(x) \geq 0, \ldots, p_r(x) \geq 0.$$

Dies ist offensichtlich genau dann unlösbar, wenn $-1 > 0$ auf $W_R(p_1, \ldots, p_r)$ gilt. Das ist nun äquivalent zu $-t_1 = 1 + t_2$ für gewisse $t_1, t_2 \in T(p_1, \ldots, p_r)$, bzw.

$$-1 \in T(p_1, \ldots, p_r).$$ △

Die bisherigen Positiv- und Nichtnegativstellensätze verwenden immer *Nenner*. Anders gesagt muss man das positive Polynom erst mit etwas multiplizieren, bevor es eine schöne Darstellung bekommt. Der erste *nennerfreie* Positivstellensatz ist der Satz von Schmüdgen, mit dem wir uns im nächsten Kapitel beschäftigen.

Kapitel 4

Der Satz von Schmüdgen

In diesem Kapitel wollen wir den Satz von Schmüdgen über positive Polynome auf kompakten Mengen beweisen. Wir geben dafür nicht Schmüdgens ursprünglichen Beweis von 1991 an, der einen funktionalanalytischen Ansatz verfolgt, sondern einen algebraischen Beweis, der im wesentlichen auf Wörmann zurückgeht.

4.1 Archimedische Präordnungen

Sei wieder A ein kommutativer Ring mit 1.

Definition 4.1.1. Eine Präordnung $T \subseteq A$ heißt *archimedisch*, falls für alle $a \in A$ ein $r \in \mathbb{N}$ existiert mit $r - a \in T$. △

Beachte dass wir den Begriff für Körperanordnungen im ersten Kapitel bereits kennengelernt haben. Unter Annahme der Archimedizität können wir den abstrakten Positivstellensatz 3.2.2 nun deutlich verstärken. Wir werden den Nenner fast vollständig los. Der folgende Satz stimmt auch ohne die Annahme $\mathbb{Q} \subseteq A$, der Beweis wird dann aber etwas technischer.

Satz 4.1.2 (Abstrakter archimedischer Positivstellensatz)**.** *Sei $\mathbb{Q} \subseteq A$ und $T \subseteq A$ eine archimedische Präordnung. Dann gilt für $a \in A$*

$$\hat{a} > 0 \text{ auf } W(T) \quad \Leftrightarrow \quad ka = 1 + t \text{ für gewisse } k \in \mathbb{N}, t \in T.$$

Beweis. "⇐" ist wieder klar. Für "⇒" verwenden wir zunächst den abstrakten Positivstellensatz 3.2.2 und erhalten eine Darstellung

$$t_1 a = 1 + t_2$$

mit $t_1, t_2 \in T$. Nun ist $\ell - t_1 \in T$ für gewisses $\ell \in \mathbb{N}$, aufgrund der Archimedizität von T. Wir betrachten die Identität

$$\ell a + (r\ell - 1) = (\ell - t_1)(a + r) + (t_1 a - 1) + r t_1.$$

Wir sehen daraus, dass falls $a + r \in T$ für ein $r \geq 0$, auch $a + (r - \frac{1}{\ell}) \in T$ gilt (dabei teilen wir durch $\ell \in \mathbb{N}$). Aufgrund der Archimedizität von T gilt aber $a+r \in T$ für ein $r \geq 0$. Durch iterierte Subtraktion von $\frac{1}{\ell}$ erhalten wir schließlich sogar ein negatives solches ℓ und insbesondere

$$a - \tfrac{1}{k} \in T$$

für ein $k \in \mathbb{N}$. Das beweist die Aussage. □

Erstaunlicherweise kann man einen analogen Nichtnegativstellensatz im archimedischen Fall nicht beweisen. Wir werden später noch Gegenbeispiele kennenlernen.

4.2 Der Satz von Schmüdgen

Wir bekommen wieder eine konkrete Version des abstrakten archimedischen Positivstellensatzes, mit dem Transferprinzip. Das besondere ist hier aber noch, dass wir die Archimedizität von T aus der Beschränktheit der semialgebraischen Menge automatisch erhalten, allerdings nur für archimedisch angeordnete reell abgeschlossene Körper, d.h Teilkörper von $\mathbb{R}$. Das wollen wir zunächst beweisen.

Proposition 4.2.1. *Sei R ein archimedischer reell abgeschlossener Körper. Dann ist eine Präordnung $T \subseteq R[x]$ genau dann archimedisch, wenn*

$$r - \sum_{i=1}^{n} x_i^2 \in T$$

für ein $r \in \mathbb{N}$.

Beweis. "⇒" ist klar. Für "⇐" gelte also $r - \sum_i x_i^2 \in T$. Daraus folgt

$$\left(r + \frac{1}{4}\right) \pm x_j = \left(\frac{1}{2} \pm x_j\right)^2 + \left(r - \sum_i x_i^2\right) + \sum_{j \neq i} x_i^2 \in T.$$

Man kann also bezüglich T alle Variablen x_j und alle Koeffizienten aus R durch natürliche Zahlen überschreiten. Für beliebige Polynome zeigen wir es iterativ über ihre Komplexität. Seien dazu $p_1, p_2 \in R[x], r_1, r_2 \in \mathbb{N}$ und $r_1 \pm p_1 \in T, r_2 \pm p_2 \in T$. Dann gilt offensichtlich

$$(r_1 + r_2) \pm (p_1 + p_2) \in T$$

und

$$\begin{aligned} 3r_1r_2 - p_1p_2 &= (r_1 + p_1)(r_2 - p_2) + r_1(r_2 + p_2) + r_2(r_1 - p_1) \in T \\ 3r_1r_2 + p_1p_2 &= (r_1 + p_1)(r_2 + p_2) + r_1(r_2 - p_2) + r_2(r_1 - p_1) \in T. \end{aligned}$$

Damit kann man jedes $p \in R[x]$ geeignet überschreiten. □

Wir nennen eine Menge $S \subseteq R^n$ *beschränkt,* wenn es ein $r \in R$ gibt mit $\|a\|^2 := \sum_i a_i^2 \leq r$ für alle $a \in S$.

Satz 4.2.2. *Sei R ein archimedischer reell abgeschlossener Körper und $p_1, \ldots, p_m \in R[x]$. Dann sind äquivalent:*

(i) $W_R(p_1, \ldots, p_m)$ *ist beschränkt in* R^n.

(ii) $T(p_1, \ldots, p_m)$ *ist archimedisch.*

Beweis. Setze $T := T(p_1, \ldots, p_m)$ und $W := W_R(p_1, \ldots, p_m)$.
"$(ii) \Rightarrow (i)$" ist einfach. Aus der Archimedizität folgt $r - \sum x_i^2 \in T$ für ein $r \in \mathbb{N}$, und da Elemente von T offensichtlich nichtnegativ auf W sind, ist W beschränkt.
Für "$(i) \Rightarrow (ii)$" wählen wir zunächst $r \in \mathbb{N}$ so, dass $p := r - \sum_i x_i^2 > 0$ auf W gilt. Nach dem Positivstellensatz 3.3.2 gibt es dann $t_1, t \in T$ mit $t_1 p = 1 + t$. Dann gilt

$$(1 + t)p = t_1 p^2 \in T. \tag{4.1}$$

Wir betrachten nun

$$T_0 = T(p) = \Sigma R[x]^2 + p\Sigma R[x]^2.$$

Diese Präordnung ist archimedisch, nach Proposition 4.2.1, und es gilt

$$(1 + t)T_0 \subseteq T, \tag{4.2}$$

nach (4.1). Ausserdem gilt

$$p + tr = p + tp + t\sum_i x_i^2 \in T,$$

ebenfalls nach (4.1). Wir wählen nun $s \in \mathbb{N}$ mit $s - t \in T_0$. Dann gilt

$$(1 + s)(s - t) = (1 + t)(s - t) + (s - t)^2 \in T,$$

nach (4.2). Nach Division durch die positive Zahl $1 + s$ folgt also $s - t \in T$. Nun gilt schließlich

$$r(s + 1) - \sum_i x_i^2 = rs + p = (p + tr) + r(s - t) \in T,$$

und also ist T nach Proposition 4.2.1 archimedisch. □

Satz 4.2.3 (Satz von Schmüdgen, konkreter archimedischer Positivstellensatz)**.** *Sei R ein archimedischer reell abgeschlossener Körper, und seien $p_1, \ldots, p_m \in R[x]$ so, dass $W_R(p_1, \ldots, p_m)$ beschränkt ist. Dann gilt für alle $p \in R[x]$*

$$p > 0 \text{ auf } W_R(p_1, \ldots, p_m) \quad \Rightarrow \quad p \in T(p_1, \ldots, p_m).$$

Beweis. Setze $T = T(p_1, \ldots, p_m)$. Wie üblich folgt aus der Bedingung $p > 0$ auf $W_R(p_1, \ldots, p_m)$ schon $\hat{p} > 0$ auf $W(T)$. Nach Satz 4.2.2 ist T aber archimedisch, und wir können den abstrakten archimedischen Positivstellensatz 4.1.2 anwenden. Wir erhalten $kp = 1 + t \in T$, und da wir in $R[x]$ durch die positive Zahl k teilen können folgt daraus schon $p \in T$. □

Beispiel 4.2.4. (i) Wir betrachten erneut die Kreisschreibe in $\mathbb{R}^2$, definiert durch $1 - x^2 - y^2 \geq 0$. Jedes Polynom p, das auf der Kreisscheibe strikt positiv ist, ist von der Gestalt

$$p = \sigma_0 + \sigma_1(1 - x^2 - y^2)$$

mit $\sigma_1, \sigma_2 \in \Sigma\mathbb{R}[x]^2$. Das ist eine deutliche Verstärkung von Beispiel 3.3.5 (i) im positiven Fall.
(ii) Falls die Anzahl m der Polynome p_i steigt, wächst die Anzahl der Summanden in der Darstellung von p exponentiell. In $T(p_1, \ldots, p_m)$ müssen wir ja alle Produkte der p_i berücksichtigen, also 2^m viele. Im nächsten Kapitel werden wir untersuchen, ob man nicht mit weniger auskommt.
(iii) Im Satz von Schmüdgen kann die Bedingung $p > 0$ im Allgemeinen nicht durch $p \geq 0$ ersetzt werden. Sei $p_1 = (1-t^2)^3 \in \mathbb{R}[t]$. Dann ist $W_{\mathbb{R}}(p_1) = [-1, 1]$,

und das Polynom $p = 1 - t^2$ ist darauf nichtnegativ. Angenommen es gäbe eine Darstellung

$$1 - t^2 = \sigma_0 + \sigma_1(1 - t^2)^3.$$

Dann verschwände σ_0 an den Punkten ± 1. Da σ_0 eine Quadratsumme ist, verschwindet es mit gerader Vielfachheit, und also wäre $(1 - t^2)^2$ ein Teiler von σ_0. Nach Kürzen hätten wir eine Darstellung

$$1 = \tilde{\sigma}_0(1 - t^2) + \sigma_1(1 - t^2)^2$$

mit einer neuen Quadratsumme $\tilde{\sigma}_0$. Einsetzen von 1 für t liefert $1 = 0$, ein Widerspruch.
(iv) Der Satz von Schmüdgen stimmt nicht, wenn R nicht archimedisch ist. Ist R eine nichtarchimedische Erweiterung von $\mathbb{R}$, so gibt es dort ein infinitesimales positives Element ε, d.h. es gilt $0 < \varepsilon < r$ für alle $r \in \mathbb{R}, r > 0$. Für p_1, p wie aus (3) ist $p + \varepsilon$ dann strikt positiv auf $W_R(p_1)$. Man kann aber zeigen, dass $p + \varepsilon$ in $R[t]$ nicht zu $T(p_1)$ gehört, siehe Beispiel 5.3.7.
(v) Über die Grade der Quadratsummen σ_i in der Darstellung von p aus dem Satz von Schmüdgen hat man keine Kontrolle. Insbesondere können sie sehr viel höher als der Grad des Polynoms p selbst sein. Ansonsten könnte man nämlich die Aussage als formale Aussage hinschreiben, und dann gälte sie laut Transferprinzip in jedem reell abgeschlossenen Körper.
(vi) Die Beschränktheit der Menge $W_R(p_1, \ldots, p_m)$ kann im Satz von Schmüdgen nicht weggelassen werden. Sei zum Beispiel $p_1 = t^3 \in \mathbb{R}[t]$. Dann ist $W_{\mathbb{R}}(p_1) = [0, \infty)$. Es ist $t + 1 > 0$ auf $W_{\mathbb{R}}(p_1)$. Wäre nun

$$t + 1 = \sigma_0 + \sigma_1 t^3$$

mit Quadratsummen σ_i, so ist der Grad von σ_0 gerade, der von $\sigma_1 t^3$ ungerade. Damit ist der Grad des Ausdrucks auf der rechten Seite entweder gerade, was nicht sein kann, oder ungerade und ≥ 3, was ebenfalls nicht sein kann. $\triangle$

Kapitel 5

Quadratische Moduln und Semiordnungen

In diesem Abschnitt versuchen wir, die Darstellungen von positiven Polynomen zu vereinfachen. Zu gegebenen Polynomen $p_1, \dots, p_m \in R[x]$ bezeichnet ja $T(p_1, \dots, p_m)$ die von den p_i erzeugte Präordnung. Sie besteht genau aus den Elementen der Form

$$\sum_{e \in \{0,1\}^m} \sigma_e p_1^{e_1} \cdots p_m^{e_m},$$

wobei alle σ_e Quadratsummen in $R[x]$ sind. Die Anzahl der Summanden ist dabei 2^m, und das wächst exponentiell mit m. Eine schönere Darstellung wäre gegeben durch einen Ausdruck der Form

$$\sigma_0 + \sigma_1 p_1 + \cdots + \sigma_m p_m.$$

Hier ist die Summenlänge $m + 1$, und eine solche Darstellung ist immer noch ein Zertifikat für die Nichtnegativität auf $W_R(p_1, \dots, p_m)$. Wir haben bisher die kompliziertere Darstellung bekommen, weil Präordnungen unter Multiplikation abgeschlossen sein müssen. Im Folgenden werden wir diese Annahme abschwächen.

5.1 Grundlagen

Sei wieder A ein kommutativer Ring mit 1.

Definition 5.1.1. Eine Teilmenge $M \subseteq A$ heißt *quadratischer Modul,* wenn

$$1 \in M,\ M + M \subseteq M,\ A^2 \cdot M \subseteq M,\ -1 \notin M.$$

Ein quadratischer Modul heiß *Semiordnung,* wenn zusätzlich

$$M \cup -M = A \text{ und } M \cap -M \text{ Primideal von } A$$

gilt. Die Menge

$$\operatorname{supp}(M) = M \cap -M$$

heiß *Support* von M. △

Definition 5.1.2. Ein quadratischer Modul M heiß *archimedisch,* wenn für jedes $a \in A$ ein $r \in \mathbb{N}$ existiert mit $r - a \in M$. △

Bemerkung 5.1.3. (i) Die Quadratsummen ΣA^2 bilden genau dann einen quadratischen Modul, wenn $-1 \notin \Sigma A^2$. In diesem Fall ist es der kleinste quadratische Modul, d.h. er ist in allen anderen enthalten.
(ii) Falls $\frac{1}{2} \in A$, so zeigt die Gleichung

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

wieder, dass $-1 \notin M$ äquivalent zu $M \neq A$ ist.
(iii) Jede Präordnung ist ein quadratischer Modul, und jede Anordnung ist eine Semiordnung.
(iv) Für $a_1, \ldots, a_m \in A$ ist der kleinste quadratische Modul welcher die a_i enthält gerade

$$M(a_1, \ldots, a_m) = \left\{\sigma_0 + \sigma_1 a_1 + \cdots + \sigma_m a_m \mid \sigma_i \in \Sigma A^2\right\},$$

zumindest falls er -1 nicht enthält. △

Beispiel 5.1.4. (i) Im Ring $A = R[x, y]$ (R ein reell abgeschlossener Körper) ist $M(x, y)$ ein quadratischer Modul, der keine Präordnung ist. Man überlegt sich nämlich, dass $xy \notin M(x, y)$ gilt.
(ii) Wir ordnen die Menge $\mathbb{N}^n$ *lexikographisch,* d.h. die Anordnung von α und β entscheidet sich durch den Vergleich der ersten ungleichen Einträge der beiden Vektoren. Wir schreiben Polynome $0 \neq p \in \mathbb{R}[x]$ dann als $p = p_{\alpha_1} x^{\alpha_1} + \cdots + p_{\alpha_r} x^{\alpha_r}$ mit $\alpha_1 < \cdots < \alpha_r$ und $p_{\alpha_i} \neq 0$. Der *Multigrad* $\operatorname{mdeg}(p)$ von p ist dann $\alpha_r \in \mathbb{N}^n$, und der *Leitkoeffizient* $\operatorname{lc}(p)$ ist p_{α_r}.

Wir treffen nun eine Vorzeichenwahl

$$\eta\colon \mathbb{Z}^n/(2\mathbb{Z})^n \to \{-1,1\}$$

mit $\eta(0,\dots,0)=1$. Dann setzen wir

$$S_\eta = \{p \in \mathbb{R}[x]\setminus\{0\} \mid \mathrm{lc}(p)\cdot\eta(\mathrm{mdeg}(p))>0\}\cup\{0\}.$$

Dann ist S_η eine Semiordnung mit $\mathrm{supp}(S_\eta)=(0)$, und S_η ist genau dann eine Anordnung wenn η ein Gruppenhomomorphismus ist. Insbesondere erhält man so (ab $n \geq 2$) Semiordnungen, die keine Anordnungen sind (Übungsaufgabe 38). △

Lemma 5.1.5. *Falls $\frac{1}{2}\in A$, so ist der Support eines quadratischen Moduls ein Ideal von A.*

Beweis. Sei M ein quadratischer Modul und $I=\mathrm{supp}(M)=M\cap -M$. Die Eigenschaft $I+I\subseteq I$ ist klar, ebenso wie $\Sigma A^2\cdot I\subseteq I$. Die Gleichung aus Bemerkung 5.1.3 (ii) zeigt dann $A\cdot I\subseteq I$. □

Satz 5.1.6. *Jeder quadratische Modul ist in einer Semiordnung enthalten.*

Beweis. Sei M quadratischer Modul und S ein maximaler über M liegender quadratischer Modul, den es nach dem Zorn'schen Lemma gibt. Wir zeigen zunächst $S\cup -S=A$. Angenommen es gibt $a\in A\setminus(S\cup -S)$. Aufgrund der Maximalität ist dann

$$\begin{aligned} -1 &= s_1+\sigma_1 a\\ -1 &= s_2-\sigma_2 a\end{aligned}$$

für gewisse $s_1,s_2\in S,\sigma_1,\sigma_2\in\Sigma A^2$. Dann gilt

$$0=\sigma_1(\sigma_2 a)+\sigma_2(-\sigma_1 a)=\sigma_1+\sigma_2+\sigma_1 s_2+\sigma_2 s_1,$$

also $-\sigma_1\in S$. Damit folgt

$$\begin{aligned} -4=4(s_1+\sigma_1 a) &= 4s_1+\sigma_1\left((a+1)^2-(a-1)^2\right)\\ &= 4s_1+\sigma_1(a+1)^2+(-\sigma_1)(a-1)^2\in S,\end{aligned}$$

und also $-1=-4+3\in S$, ein Widerspruch.

Als nächstes zeigen wir dass $S \cap -S$ ein Ideal in A ist. Wie man im Beweis von Lemma 5.1.5 sieht, genügt es zu zeigen dass aus $4a \in \operatorname{supp}(S)$ schon $a \in \operatorname{supp}(S)$ folgt. Dafür reicht es zu zeigen dass aus $4a \in S$ schon $a \in S$ folgt. Sei also $4a \in S$ und $a \notin S$. Dann ist $-a \in S$, wie wir eben gezeigt haben, und also $a = 4a - 3a \in S$, ein Widerspruch.
Wir müssen nun noch zeigen, dass $\operatorname{supp}(S)$ ein Primideal ist. Seien also $a, b \in A$ mit $ab \in \operatorname{supp}(S)$, $b \notin \operatorname{supp}(S)$, und also o.B.d.A. $b \notin S$. Wir müssen $a \in \operatorname{supp}(S)$ zeigen. Angenommen das stimmt nicht, d.h. o.B.d.A. $a \notin S$. Es gilt zunächst $-1 \in S + b \cdot \Sigma A^2$, aufgrund der Maximalität von S. Multiplizieren mit a^2 liefert

$$-a^2 \in S + a(ab) \cdot \Sigma A^2 \subseteq S + \operatorname{supp}(S) \subseteq S.$$

Da auch $a^2 \in S$ folgt also $a^2 \in \operatorname{supp}(S)$. Wegen $a \notin S$ gilt weiter $-1 = s + \sigma a$ für gewisse $s \in S, \sigma \in \Sigma A^2$, wieder aufgrund der Maximalität. Jetzt gilt

$$1 + 2s + s^2 = (1+s)^2 = \sigma^2 a^2 \in \operatorname{supp}(S) \subseteq -S,$$

und deshalb $-1 \in S$, ein Widerspruch. □

Für ein Primideal $\mathfrak{p}$ von A bezeichnen wir wieder mit $\rho_\mathfrak{p}$ den Restklassenhomomorphismus:

$$\rho_\mathfrak{p} \colon A \to A/\mathfrak{p} \to K_\mathfrak{p}.$$

Man kann Satz 3.1.13 ganz analog beweisen: Die Semiordnungen von A stehen in Bijektion zu den Tupeln $(\mathfrak{p}, Q)$, wobei $\mathfrak{p}$ ein Primideal in A und Q eine Semiordnung auf dem Restklassenkörper $K_\mathfrak{p}$ ist. Die Menge aller Semiordnungen von A heiß das *semireelle Spektrum* und wird mit $\operatorname{Semisper}(A)$ bezeichnet. Wir definieren für $a \in A$ und $S = (\mathfrak{p}, Q) \in \operatorname{Semisper}(A)$

$$\begin{aligned}
\hat{a}(S) > 0 \quad &:\Leftrightarrow \quad \rho_\mathfrak{p}(a) >_Q 0 \text{ in } K_\mathfrak{p} \quad \Leftrightarrow \quad a \in S \setminus -S \\
\hat{a}(S) \geq 0 \quad &:\Leftrightarrow \quad \rho_\mathfrak{p}(a) \geq_Q 0 \text{ in } K_\mathfrak{p} \quad \Leftrightarrow \quad a \in S \\
\hat{a}(S) = 0 \quad &:\Leftrightarrow \quad \rho_\mathfrak{p}(a) = 0 \text{ in } K_\mathfrak{p} \quad \Leftrightarrow \quad a \in \operatorname{supp}(S)
\end{aligned}$$

Auf diese Weise fassen wir also Elemente von A als Funktionen auf $\operatorname{Semisper}(A)$ auf, die Werte in den Restklassenkörpern $K_\mathfrak{p}$ annehmen. Ebenso können wir nun wieder spezielle Teilmengen des semireellen Spektrums definieren. Wir setzen für $M \subseteq A$

$$\begin{aligned}
\widetilde{W}(M) :&= \{S \in \operatorname{Semisper}(A) \mid \hat{m}(S) \geq 0 \text{ für alle } m \in M\} \\
&= \{S \in \operatorname{Semisper}(A) \mid M \subseteq S\}.
\end{aligned}$$

$$\begin{aligned}\widetilde{V}(M) :&= \{S \in \operatorname{Semisper}(A) \mid \hat{m}(S) = 0 \text{ für alle } m \in M\} \\ &= \{S \in \operatorname{Semisper}(A) \mid M \subseteq \operatorname{supp}(S)\}.\end{aligned}$$

Beachte nochmals, dass $W(M)$ bzw. $V(M)$ die entsprechende Menge im rellen Spektrum von A bezeichnet, und $W_R(M)$ bzw. $V_R(M)$ die Menge in R^n, falls $A = R[x]$. Es gilt dann

$$W_R(M) \subseteq W(M) \subseteq \widetilde{W}(M),$$

und analog für V.

5.2 Abstrakte Positivstellensätze für quadratische Moduln

Wir wollen in diesem Abschnitt ein Analogon des abstrakten Positivstellensatzes und des abstrakten archimedischen Positivstellensatzes beweisen, dieses Mal für quadratische Moduln anstatt für Präordnungen.

Satz 5.2.1 (Abstrakter Positivstellensatz für quadratische Moduln)**.** *Sei $M \subseteq A$ ein quadratischer Modul. Dann gilt für alle $a \in A$*

$$\hat{a} > 0 \text{ auf } \widetilde{W}(M) \quad \Leftrightarrow \quad \sigma a = 1 + m \text{ für gewisse } \sigma \in \Sigma A^2, m \in M.$$

Beweis. "$\Leftarrow$" ist wieder klar: wäre $-a \in S$ für eine Semiordnung mit $M \subseteq S$, so wäre auch $-\sigma a \in S$, und also $-1 \in S$, ein Widerspruch. "$\Rightarrow$" Falls es keine solche Darstellung gibt, ist $M - a\Sigma A^2$ erneut ein quadratischer Modul. Er ist also in einer Semiordnung S enthalten. Dafür gilt dann $S \in \widetilde{W}(M)$ und $-a \in S$, d.h. $\hat{a}(S) \leq 0$, ein Widerspruch. □

Wir beweisen jetzt den abstrakten archimedischen Positivstellensatz für quadratische Moduln. Um den Beweis zu vereinfachen, nehmen wir ab jetzt an, dass $\mathbb{Q} \subseteq A$ gilt.

Satz 5.2.2 (Abstrakter archimedischer Positivstellensatz für quadratische Moduln)**.** *Sei $\mathbb{Q} \subseteq A$ und $M \subseteq A$ ein archimedischer quadratischer Modul. Dann gilt für alle $a \in A$*

$$\hat{a} > 0 \text{ auf } \widetilde{W}(M) \quad \Leftrightarrow \quad ka = 1 + m \text{ für gewisse } k \in \mathbb{N}, m \in M.$$

Beweis. "$\Leftarrow$" ist erneut klar. Für "$\Rightarrow$" betrachten wir $M' := M - a\Sigma A^2$. Die Voraussetzung besagt, dass über M' keine Semiordnung existiert. Nach Satz 5.1.6 ist also M' kein quadratischer Modul, d.h. $-1 \in M'$, d.h. es gibt eine Gleichung $\sigma a - 1 = m \in M$ mit $\sigma \in \Sigma A^2$. Wegen der Archimedizität von M gibt es ein $k \in \mathbb{N}$ mit $2k - 1 - \sigma^2 a \in M$. Dann gilt

$$2k - \sigma = (2k - 1 - \sigma^2 a) + \sigma(\sigma a - 1) + 1 \in M.$$

Betrachte nun die folgende Identität für $r \in \mathbb{Q}$:

$$k^2 a + k^2 r - 1 = (k - \sigma)^2(a + r) + 2k(\sigma a - 1) + r\sigma(2k - \sigma) + (2k - 1 - \sigma^2 a).$$

Wir sehen nach Teilung durch k^2, dass falls $a + r \in M$ für ein $r \in \mathbb{Q}_{\geq 0}$ gilt, auch $a + (r - \frac{1}{k^2}) \in M$ gilt. Aufgrund der Archimedizität ist aber $a + r \in M$ für ein $r \in \mathbb{N}$, und also gibt es ein rationales $s < 0$ mit $a + s \in M$. Nach Multiplizieren mit dem Nenner von s ist das die gewünschte Aussage. □

Auch hier wollen wir wieder einen konkreten archimedischen Positivstellensatz im Fall $A = R[x]$ erhalten. Dabei tauchen zwei Probleme auf. Zunächst erhalten wir aus der geometrischen Positivität eines Polynoms auf $W_R(M)$ die Positivität auf $W(M) \subseteq \mathrm{Sper}(A)$, wie üblich mit dem Transferprinzip. Wir benötigen aber die Positivität auf der größeren Menge $\widetilde{W}(M) \subseteq \mathrm{Semisper}(A)$, um den abstrakten Positivstellensatz anwenden zu können. Im archimedischen Fall kann man das erreichen, muss dafür aber erst etwas Arbeit investieren. Für nicht-archimedische quadratische Moduln M stimmt das übrigens im Allgemeinen nicht (Übungsaufgabe 39). Zweitens bekommt man im Unterschied zu Schmüdgens Satz aus der Kompaktheit der Menge nicht automatisch die Archimedizität des quadratischen Moduls. Man muss sie hier extra fordern, bzw. sicherstellen.

5.3 Semiordnungen auf Körpern

Wir wollen uns für einen Moment mit Semiordnungen auf Körpern beschäftigen. Die Ergebnisse werden wir später für den konkreten archimedischen Positivstellensatz für quadratische Moduln benötigen. Sei also S eine Semiordnung auf dem Körper K. Dann können wir wieder eine zweistellige Relation $\leq_S$ auf K definieren, durch

$$a \leq_S b \quad :\Leftrightarrow \quad b - a \in S.$$

Den Subskript S lassen wir dabei häufig auch weg. Man sieht leicht dass $\leq$ eine vollständige lineare Ordnung der Menge F ist. Beachte dass $\mathrm{supp}(S) = \{0\}$ in

Körpern immer gilt, also $a \leq b$ und $b \leq a$ immer $a = b$ impliziert. Insbesondere schreiben wir auch $a < b$ für $a \leq b$ und $a \neq b$. Aus den weiteren Eigenschaften von S kann man weitere Eigenschaften von $\leq$ ableiten. Zum Beispiel gilt

$$a \leq b \Rightarrow a + c \leq b + c$$
$$a \leq b \Rightarrow c^2 a \leq c^2 b.$$

Die Eigenschaft $0 \leq a, b \Rightarrow 0 \leq ab$ gilt im Allgemeinen allerdings nicht. Das unterscheidet gerade Semiordnungen von Anordnungen. Weitere Eigenschaften sind im folgenden Lemma zusammengefasst.

Lemma 5.3.1. *Für $a, b \in K$ und $m \in \mathbb{N}$ gilt:*

(i) $0 < a \Rightarrow 0 < \frac{1}{a}$

(ii) $0 < a < b \Rightarrow ba^2 < ab^2$

(iii) $0 < a < b \Rightarrow 0 < \frac{1}{b} < \frac{1}{a}$

(iv) $0 < a < m \Rightarrow a^2 < m^2$ *(und analog für $m < a$).*

Beweis. (i) gilt wegen $\frac{1}{a} = a \cdot \left(\frac{1}{a}\right)^2$. Im Fall von (ii) gilt $0 < a, 0 < b - a$. Daraus folgt $0 < \frac{1}{a}, 0 < \frac{1}{b-a}$ und also $0 < \frac{1}{a} + \frac{1}{b-a}$. Erneutes Inversenbilden liefert

$$0 < \frac{1}{\frac{1}{a} + \frac{1}{b-a}} \cdot b^2 = ab^2 - ba^2,$$

und das ist die Behauptung. Für (iii) folgt zunächst $ba^2 < ab^2$, und wenn wir beide Seiten mit $\left(\frac{1}{ab}\right)^2$ multiplizieren, folgt die Aussage. Im Falle von (iv) schließen wir zuerst $ma^2 < am^2$. Weil

$$\frac{1}{m} = \frac{m}{m^2} = \left(\frac{1}{m}\right)^2 + \cdots + \left(\frac{1}{m}\right)^2$$

in K eine Quadratsumme ist gilt dann $a^2 < am$. Aus $a < m$ folgt aber auch $am < m^2$, und damit die Aussage. □

Satz 5.3.2. *Jede archimedische Semiordnung auf einem Körper ist eine Anordnung.*

Beweis. Sei $\leq$ die von der Semiordnung definierte lineare Ordnung auf dem Körper K. Wir zeigen zunächst, dass $\mathbb{Q}$ dicht in K liegt. Für $a, b \in K$ gilt

$$\begin{aligned} a < b &\Rightarrow 0 < b - a \\ &\Rightarrow 0 < (b-a)^{-1} < m \text{ für ein } m \in \mathbb{N} \\ &\Rightarrow \frac{1}{m} < b - a \\ &\Rightarrow 1 < m(b-a) \quad (\text{da } m \in \Sigma K^2) \\ &\Rightarrow ma < mb - 1. \end{aligned}$$

Dabei haben wir die Archimedizität und mehrfach Lemma 5.3.1 benutzt. Wir wählen nun $n \in \mathbb{Z}$ minimal mit $mb \leq n + 1$. Dann gilt

$$ma < mb - 1 \leq n < mb,$$

und nach Multiplikation mit der Quadratsumme $\frac{1}{m}$ gilt

$$a < \frac{n}{m} < b,$$

die gewünschte Dichtheitsaussage.
Wir zeigen nun die eigentliche Aussage. Seien dazu $0 \leq a, b \in K$. Wir müssen $0 \leq ab$ zeigen. Sei o.B.d.A. $0 < a < b$. Wir setzen $c := b - a$ und $d := b + a$. Dann gilt $0 < c < d$, und wir finden aufgrund der eben bewiesenen Dichtheit gewisse $m, n \in \mathbb{N}, m \neq 0$, mit

$$mc < n < md.$$

Mit Lemma 5.3.1 (iv) folgt

$$m^2c^2 < n^2 < m^2d^2,$$

und nach Multiplikation mit m^{-2} gilt

$$(b-a)^2 = c^2 < d^2 = (b+a)^2.$$

Also gilt $-2ab < 2ab$, und also $0 < ab$. □

Ist eine Semiordnung nicht archimedisch, können wir daraus immer eine archimedische Semiordnung auf einem gewissen anderen Körper konstruieren. Dafür benötigen wir den folgenden Begriff.

Definition 5.3.3. Sei K ein Körper. Ein Teilring $\mathcal{O} \subseteq K$ heißt *Bewertungsring von* K, falls für alle $a \in K$ gilt

$$a \in \mathcal{O} \text{ oder } a^{-1} \in \mathcal{O}. \qquad \triangle$$

Man beachte dass dann automatisch $K = \mathrm{Quot}(\mathcal{O})$ gilt.

Lemma 5.3.4. *Ein Bewertungsring hat genau ein maximales Ideal, und zwar die Menge der Nichteinheiten*

$$\mathfrak{m} = \mathcal{O} \setminus \mathcal{O}^\times = \left\{ a \in \mathcal{O} \mid a^{-1} \notin \mathcal{O} \right\}.$$

Beweis. Wir zeigen zuerst dass $\mathfrak{m}$ ein echtes Ideal von $\mathcal{O}$ ist. $1 \notin \mathfrak{m}$ ist klar, da $1 \in \mathcal{O}$. Seien nun $a, b \in \mathfrak{m}$, d.h. $a^{-1}, b^{-1} \notin \mathcal{O}$. Sei o.B.d.A. $\frac{b}{a} \in \mathcal{O}$. Wäre $(a+b)^{-1} \in \mathcal{O}$, so

$$a^{-1} = \frac{a+b}{a(a+b)} = (a+b)^{-1} + \frac{b}{a}(a+b)^{-1} \in \mathcal{O},$$

ein Widerspruch. Also ist auch $a + b \in \mathfrak{m}$. Sei schließlich $a \in \mathfrak{m}, b \in \mathcal{O}$. Dann ist $(ab)^{-1} = b^{-1}a^{-1} \notin \mathcal{O}$, da sonst auch $a^{-1} \in \mathcal{O}$ wäre. Also ist $ab \in \mathfrak{m}$, und $\mathfrak{m}$ ist somit ein Ideal. Da jedes Element in $\mathcal{O} \setminus \mathfrak{m}$ invertierbar in $\mathcal{O}$ ist, enthält $\mathfrak{m}$ offensichtlich alle echten Ideale, und ist damit das einzige maximale Ideal von $\mathcal{O}$. $\square$

Definition 5.3.5. Der Körper $\mathcal{O}/\mathfrak{m}$ heißt *Restklassenkörper von* $\mathcal{O}$. Die kanonische Projektion $\pi \colon \mathcal{O} \to \mathcal{O}/\mathfrak{m}$ heiß *Restklassenhomomorphismus*. $\triangle$

Der folgende Satz zeigt nun, wie wir aus einer beliebigen Semiordnung eine archimedische konstruieren können.

Satz 5.3.6. *Sei K ein Körper und S eine Semiordnung von K. Dann ist*

$$\begin{aligned} \mathcal{O}(S) :&= \{a \in K \mid |a| \leq_S m \text{ für ein } m \in \mathbb{N}\} \\ &= \{a \in K \mid m \pm a \in S \text{ für ein } m \in \mathbb{N}\} \end{aligned}$$

ein Bewertungsring von K, mit maximalem Ideal

$$\mathfrak{m} = \left\{ a \in K \mid |a| <_S \frac{1}{m} \quad \forall m \in \mathbb{N} \setminus \{0\} \right\}.$$

Die Menge $S \cap \mathcal{O}(S)$ ist eine archimedische Semiordnung auf $\mathcal{O}(S)$, und $\pi(S)$ ist eine archimedische Semiordnung des Restklassenkörpers $\mathcal{O}(S)/\mathfrak{m}$. Insbesondere ist es eine Anordnung, und wir können $\mathcal{O}(S)/\mathfrak{m}$ als angeordneten Teilkörper von $\mathbb{R}$ auffassen.

Beweis. Setze $\mathcal{O} := \mathcal{O}(S)$. Offensichtlich ist $\mathcal{O}$ abgeschlossen unter $+$, enthält additiv inverse Elemente, sowie $\mathbb{Q}$. Aufgrund der Gleichung

$$ab = \left(\frac{a+b}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2$$

erhalten wir die Abgeschlossenheit unter $\cdot$ also wenn wir zeigen, dass $a \in \mathcal{O}$ immer $a^2 \in \mathcal{O}$ impliziert. Das folgt aber aus Lemma 5.3.1 (iv).
Ist nun $a \notin \mathcal{O}$ für ein $a \in K$, so bedeutet das o.B.d.A. $\mathbb{N} < a$. Lemma 5.3.1 (iii) zeigt dann $0 < a^{-1} < \frac{1}{m}$ für alle $m \in \mathbb{N}$. Somit ist $\mathcal{O}$ ein Bewertungsring mit dem angegebenen Ideal $\mathfrak{m}$. Offensichtlich ist nun $S \cap \mathcal{O}$ in $\mathcal{O}$ eine archimedische Semiordnung (mit dem Primideal $\{0\}$ als Support). Wir zeigen nun dass $\pi(S)$ eine Semiordnung von $\mathcal{O}/\mathfrak{m}$ ist. Zu zeigen ist dafür nur $-1 \notin \pi(S)$ und $\operatorname{supp}(\pi(S)) = (0)$. Falls $s, t \in S$ sind mit $\pi(s) = \pi(-t)$, so ist $s + t \in \mathfrak{m}$, d.h. $\frac{1}{m} - s - t \in S$ für alle $m \in \mathbb{N}$. Nach Addition von s bzw. t folgt sofort $s, t \in \mathfrak{m}$, d.h. $\pi(s) = \pi(t) = 0$. Daraus folgen die beiden Aussagen direkt.
Nach Satz 5.3.2 ist $\pi(S)$ als archimedische Semiordnung auf einem Körper eine Anordnung, und jeder archimedisch angeordnete Körper ist ein Teilkörper von $\mathbb{R}$, nach Satz 1.1.8. □

Wir können nun auch zeigen, dass der Satz von Schmüdgen nicht für jeden reell abgeschlossenen Körper gilt.

Beispiel 5.3.7. Sei R ein nichtarchimedisch reell abgeschlossener Erweiterungskörper von $\mathbb{R}$. Dann gibt es ein positives infinitesimales Element ε in R, d.h. es gilt

$$0 < \varepsilon < r \text{ für alle } r \in \mathbb{R}_{>0}.$$

Das Polynom $p = (1 - t^2) + \varepsilon$ ist dann strikt positiv auf

$$[-1, 1] = W_R((1 - t^2)^3) \subseteq R.$$

Angenommen es gäbe eine Darstellung

$$p = \sigma_0 + \sigma_1(1 - t^2)^3 \tag{5.1}$$

mit $\sigma_i \in \Sigma R[t]^2$. Bezeichne mit $\mathcal{O} = \mathcal{O}(R^2)$ die Hülle von $\mathbb{Z}$ (bzw. $\mathbb{R}$) in R bezüglich der Anordnung von R. Wenn $\sigma_i = \sum_j f_{ij}^2$ ist, und jeder Koeffizient von jedem f_{ij} in $\mathcal{O}$ liegt, können wir den Restklassenhomomorphismus $\pi\colon \mathcal{O} \to \mathcal{O}/\mathfrak{m} \subseteq \mathbb{R}$ auf alle Koeffizienten von (5.1) anwenden. Links entsteht dabei $1 - t^2$,

und rechts wieder eine Darstellung wie in (5.1), diesmal über $\mathbb{R}$. Wir haben aber schon in Beispiel 4.2.4 (iii) gezeigt, dass es die nicht geben kann.
Also kann einer der Koeffizienten eines der f_{ij} nicht in $\mathcal{O}$ liegen, er ist also unendlich groß bezüglich $\mathbb{R}$. Wir teilen die Gleichung (5.1) nun durch das Quadrat des betragsmäßig größten solchen Koeffizienten. Dann steht rechts wieder ein solcher Ausdruck, aber diesmal liegen alle Koeffizienten aller f_{ij} in $\mathcal{O}$, und mindestens einer ist genau 1. Links sind alle Koeffizienten nun in $\mathfrak{m}$. Wir wenden wieder den Restklassenhomomorphismus auf alle Koeffizienten an, und erhalten eine neue Gleichung

$$0 = \sigma_0 + \sigma_1(1-t^2)^3, \tag{5.2}$$

mit Quadratsummen $\sigma_i \in \mathbb{R}[t]$. Es sind nicht beide $\sigma_i = 0$, da der Koeffizient 1 in einem f_{ij} vorkommt. Andererseits sind beide σ_i und auch $(1-t^2)^3$ nichtnegativ auf $[-1,1]$, und aus (5.2) folgt also $\sigma_0 = \sigma_1 = 0$. Also haben wir auch hier einen Widerspruch. △

5.4 Positivität auf beschränkten Mengen reloaded

Wir übersetzen nun den abstrakten archimedischen Positivstellensatz für quadratische Moduln in einen konkreten. Zunächst benötigen wir das folgende Ergebnis für beliebige Ringe A:

Satz 5.4.1. *Sei $M \subseteq A$ ein archimedischer quadratischer Modul. Dann ist jede über M liegende maximale Semiordnung eine Anordnung. Insbesondere gilt für alle $a \in A$*

$$\hat{a} > 0 \text{ auf } W(M) \Rightarrow \hat{a} > 0 \text{ auf } \widetilde{W}(M).$$

Beweis. Sei $S = (\mathfrak{p}, Q)$ eine maximale über M liegende Semiordnung von A. Offensichtlich ist auch S archimedisch. Wir betrachten wieder die Restklassenabbildung

$$A \to A/\mathfrak{p} \hookrightarrow K_{\mathfrak{p}}.$$

Auf $A/\mathfrak{p}$ ist die von S induzierte Semiordnung archimedisch, aber die Semiordnung Q auf $K_{\mathfrak{p}}$ muss nicht archimedisch sein. Deshalb betrachten wir $\mathcal{O}(Q)$ in $K_{\mathfrak{p}}$:

$$A \to A/\mathfrak{p} \subseteq \mathcal{O}(Q) \subseteq K_{\mathfrak{p}},$$

und fügen den Restklassenhomomorphismus an

$$A \to A/\mathfrak{p} \subseteq \mathcal{O}(Q) \to \mathcal{O}(Q)/\mathfrak{m} =: K.$$

K trägt nun die von Q induzierte archimedische Semiordnung (vgl. Satz 5.3.6), und diese ist nach Satz 5.3.2 eine Anordnung. Ihr Urbild in A ist also ein Anordnung, die aber offensichtlich S enthält. Aus der Maximalität folgt die Gleichheit, und also ist S eine Anordnung.
Aus $\hat{a} > 0$ auf $W(M)$ folgt also $\hat{a}(S) > 0$ für alle maximalen Semiordnungen über M. Daraus folgt aber $\hat{a} > 0$ auf $\widetilde{W}(M)$, denn ist $-a \in S$ für eine Semiordnung, so gilt das auch für jede darüberliegende. □

Proposition 5.4.2. *Sei R ein archimedischer reell abgeschlossener Körper und $M \subseteq R[x]$ ein quadratischer Modul. Dann ist M genau dann archimedisch, wenn*

$$r - \sum_{i=1}^{n} x_i^2 \in M$$

für ein $r \in \mathbb{N}$ gilt.

Beweis. "⇒" ist klar. Gelte umgekehrt $r - \sum_i x_i^2 \in M$. Betrachte

$$M' := \Sigma R[x]^2 + (r - \sum_i x_i^2)\Sigma R[x]^2 \subseteq M.$$

Da M' von nur einem Element erzeugt wird, ist es sogar eine Präordnung. Nach Satz 4.2.1 ist also bereits M' archimedisch, und also auch M. □

Satz 5.4.3 (Satz von Putinar, Konkreter archimedischer Positivstellensatz für quadratische Moduln)**.** *Sei R ein archimedischer reell abgeschlossener Körper und $p_1, \ldots, p_m \in R[x_1, \ldots, x_n]$, so dass $r - \sum_{i=1}^{n} x_i^2 \in M(p_1, \ldots p_m)$ für ein $r \in \mathbb{N}$ gilt. Dann gilt für alle $p \in R[x]$*

$$p > 0 \text{ auf } W_R(p_1, \ldots, p_m) \Rightarrow p \in M(p_1, \ldots, p_m).$$

Beweis. Setze $M := M(p_1, \ldots, p_m)$. Die Bedingung $r - \sum_{i=1}^{n} x_i^2 \in M$ stellt nach Proposition 5.4.2 die Archimedizität von M sicher. Ist $p > 0$ auf der Menge $W_R(p_1, \ldots, p_m) \subseteq R^n$, so ist $\hat{p} > 0$ auf $W(M) \subseteq \mathrm{Sper}(R[x])$, wie gewöhnlich mit dem Transferprinzip. Nach Satz 5.4.1 ist dann auch $\hat{p} > 0$ auf $\widetilde{W}(M) \subseteq \mathrm{Semisper}(R[x])$, und Satz 5.2.2 liefert die Aussage, nach Teilung durch k. □

Bemerkung 5.4.4. Die Archimedizität von $M(p_1, \ldots, p_m)$ impliziert die Beschränktheit der Menge $W_R(p_1, \ldots, p_m)$, ist aber im Allgemeinen nicht dazu äquivalent. Wir werden gleich ein Beispiel sehen. Allerdings kann man bei Beschränktheit der Menge $W_R(p_1, \ldots, p_m)$ einfach eine weitere Ungleichung $r - \sum x_i^2 \geq 0$ dazunehmen, mit r so groß, dass die Menge sich nicht ändert. Der zusätzliche Erzeuger macht den quadratischen Modul dann archimedisch. △

Beispiel 5.4.5. Betrachte den Einheitswürfel $W = [-1,1]^n \subseteq \mathbb{R}^n$. Er ist definiert durch die Ungleichungen $1 \pm x_i \geq 0$ für $i = 1, \ldots, n$. Der dazugehörige quadratische Modul M besteht aus allen Elementen der Form

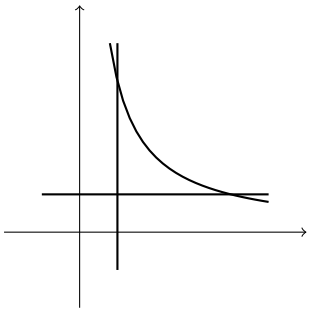
$$\sigma_0 + \sigma_1(1-x_1) + \sigma_2(1+x_1) + \sigma_3(1-x_2) + \cdots + \sigma_{2n-1}(1-x_n) + \sigma_{2n}(1+x_n).$$

Wir überlegen uns zunächst, dass er archimedisch ist. Es gilt

$$(1-x_i)^2(1+x_i) + (1+x_i)^2(1-x_i) = 2(1-x_i^2),$$

und also ist $1-x_i^2 \in M$, und also $n - \sum_i x_i^2 \in M$. Also ist M archimedisch. Somit ist jedes Polynom p, das strikt positiv auf dem Einheitswürfel ist, von der obigen Gestalt. Die Darstellung aus dem Satz von Schmüdgen hätte hier eine Darstellung mit 4^n Summanden geliefert. △

Beispiel 5.4.6. Die Beschränktheit der Menge $W_R(p_1, \ldots, p_m)$ impliziert nicht die Archimedizität des quadratischen Moduls $M = M(p_1, \ldots, p_m)$. Wir wählen $p_1 = x - \frac{1}{2}, p_2 = y - \frac{1}{2}, p_3 = 1 - xy \in \mathbb{R}[x,y]$. Die Menge $W_{\mathbb{R}}(p_1, \ldots, p_m)$ ist beschränkt:



Angenommen es wäre

$$r - x = \sigma_0 + \sigma_1(x - \frac{1}{2}) + \sigma_2(y - \frac{1}{2}) + \sigma_3(1-xy)$$

mit Quadratsummen σ_i und einem $r \in \mathbb{N}$. Die homogenen Summanden höchsten Grades auf der rechten Seite sind

$$\tilde{\sigma}_0,\ \tilde{\sigma}_1 x,\ \tilde{\sigma}_2 y,\ -\tilde{\sigma}_3 xy,$$

wobei die $\tilde{\sigma}_i$ als Terme höchsten Grades der σ wieder Quadratsummen sind. Mindestens einer dieser vier Terme muss Grad ≥ 1 haben. Wir betrachten nun einen

Term mit maximalem Grad. Ist es der erste oder der vierte, muss er sich mit dem jeweils anderen wegheben, denn $n - x$ hat ungeraden Grad. Aus $\tilde{\sigma}_0 - xy\tilde{\sigma}_3 = 0$ folgt aber $\tilde{\sigma}_0 = \tilde{\sigma}_3 = 0,$ da alle Ausdrücke auf dem 2. Orthanten nichtnegativ sind. Dies ist ein Widerspruch zur Maximalität des Grades. Also muss der maximale Grad im zweiten oder dritten Summanden angenommen werden. Da auch $\tilde{\sigma}_1 x + \tilde{\sigma}_2 y$ nicht Null sein kann, diesmal wegen Positivität auf dem 1. Orthanten, muss $\tilde{\sigma}_1 x + \tilde{\sigma}_2 y = -x$ gelten. Setzt man $x = 1$ und $y = 1$ so ist die rechte Seite negativ, die linke Seite nichtnegativ, ein Widerspruch. Also ist M nicht archimedisch. △

Wir wollen in den folgenden Kapiteln verschiedene Anwendungen der bisherigen Ergebnisse kennenlernen. Dabei handelt es sich teilweise um echte Anwendungen, zum Beispiel in der polynomialen Optimierung, und andererseits um Anwendungen in anderen Bereichen der eher theoretischen Mathematik, wie zum Beispiel der Funktionalanalysis.

Kapitel 6

Konvexität und Optimierung

Die Ergebnisse der reellen algebraischen Geometrie die wir bisher kennengelernt haben, lassen sich für die polynomiale Optimierung verwenden. Um diese Anwendungen zu beschreiben, erklären wir zunächst die sogenannte *semidefinite Optimierung*.

6.1 Semidefinite Optimierung

Sei $\mathrm{Sym}_d(\mathbb{R})$ der reelle Vektorraum aller symmetrischen $d \times d$-Matrizen. Für zwei symmetrische Matrizen $A = (a_{ij})_{i,j}$ und $B = (b_{ij})_{i,j}$ setzen wir

$$\langle A, B \rangle := \mathrm{tr}(AB) = \sum_{i,j} a_{ij} b_{ij}.$$

Dabei bezeichnet tr die Spur einer Matrix, d.h. die Summe der Diagonaleinträge, bzw. die Summe der Eigenwerte. $\langle \cdot, \cdot \rangle$ definiert offensichtlich ein Skalarprodukt auf $\mathrm{Sym}_d(\mathbb{R})$.
Wir bezeichnen mit $\mathcal{P}_d$ die Menge der positiv semidefiniten Matrizen in $\mathrm{Sym}_d(\mathbb{R})$ (vergleiche Definition 2.2.4 und Lemma 2.2.3). Die Menge $\mathcal{P}_d$ ist offensichtlich ein abgeschlossener konvexer Kegel in $\mathrm{Sym}_d(\mathbb{R})$. Nach Lemma 2.2.3 ist $\mathcal{P}_d$ sogar basisch abgeschlossen, d.h. von der Gestalt

$$\mathcal{P}_d = W_{\mathbb{R}}(p_1, \ldots, p_r)$$

für gewissen Polynome $p_k \in \mathbb{R}[x_{ij} \mid i, j = 1, \ldots, d]$. Eine Matrix ist ja genau dann positiv semidefinit, wenn alle ihre Hauptminoren nichtnegativ sind. Wir

schreiben $A \succeq 0$ für $A \in \mathcal{P}_d$ und $A \preceq B$ für $B - A \succeq 0$. Analog bezeichnet $\succ$ strikte positive Definitheit, d.h. $A \succ 0$ bedeutet dass alle Eigenwerte von A strikt positiv sind, oder dass $v^t A v > 0$ für alle $v \in \mathbb{R}^d \setminus \{0\}$ gilt.

Proposition 6.1.1. *(i) Für $A \in \mathcal{P}_d$ und $P \in \mathrm{M}_d(\mathbb{R})$ gilt $PAP^t \in \mathcal{P}_d$.*
(ii) $\langle \cdot, \cdot \rangle$ ist invariant unter Konjugation mit orthogonalen Matrizen P, d.h.

$$\langle A, B \rangle = \langle PAP^t, PBP^t \rangle.$$

(iii) Der Kegel $\mathcal{P}_d$ ist selbstdual bezüglich $\langle \cdot, \cdot \rangle$, d.h.

$$A, B \in \mathcal{P}_d \Rightarrow \langle A, B \rangle \geq 0$$

$$\langle A, B \rangle \geq 0 \text{ für alle } B \in \mathcal{P}_d \Rightarrow A \in \mathcal{P}_d.$$

(iv) Falls $A \succeq 0, B \succ 0$ und $\langle A, B \rangle = 0$, so $A = 0$.

Beweis. (i) Es ist PAP^t wieder symmetrisch, und

$$v^t PAP^t v = (P^t v)^t A (P^t v) \geq 0.$$

(ii) Es gilt

$$\begin{aligned} \langle PAP^t, PBP^t \rangle &= \mathrm{tr}(PAP^t PBP^t) = \mathrm{tr}(PABP^t) \\ &= \mathrm{tr}(ABP^t P) = \mathrm{tr}(AB) = \langle A, B \rangle. \end{aligned}$$

Dabei verwenden wir, dass Produkte bei der Berechnung der Spur zyklisch vertauscht werden dürfen, und dass $P^t P = I$ für orthogonale Matrizen gilt.
(iii) Seien $A, B \in \mathcal{P}_d$. Wegen (i) und (ii) können wir annehmen dass A diagonal ist, mit nichtnegativen Diagonaleinträgen. Die Diagonaleinträge der positiv semidefiniten Matrix B sind aber auch nichtnegativ. Damit ist

$$\langle A, B \rangle = \sum_i a_{ii} b_{ii} \geq 0.$$

Mit der gleichen Formel sieht man, dass für Diagonalmatrizen A alle Diagonaleinträge nichtnegativ sein müssen, damit $\langle A, B \rangle \geq 0$ für alle $B \in \mathcal{P}_d$ gilt. Also muss A dann selbst positiv semidefinit sein. Falls A nicht diagonal ist, sondern nur $D := PAP^t$, dann gilt für $B \in \mathcal{P}_d$

$$\langle D, B \rangle = \langle PAP^t, B \rangle = \langle A, P^t B P \rangle \geq 0.$$

Also ist $D \succeq 0$, und also auch $A \succeq 0$.
(iv) Wir können annehmen, dass B diagonal mit strikt positiven Diagonaleinträgen ist. Dann müssen alle Diagonaleinträge von A Null sein, und damit muss $A = 0$ gelten, da A positiv semidefinit ist (man betrachte die 2×2-Hauptminoren). □

Definition 6.1.2. Seien $M, M_1, \ldots, M_m \in \mathrm{Sym}_d(\mathbb{R})$ und $\beta_1, \ldots, \beta_m \in \mathbb{R}$. Das folgende Optimierungsproblem nennt man ein *semidefinites Optimierungsproblem (in primaler Form)*:

$$\begin{array}{ll} \text{finde} & \inf \langle M, A\rangle \\ \text{s.t.} & \langle M_i, A\rangle = \beta_i \text{ für } i = 1, \ldots, m \\ & A \succeq 0 \end{array}$$

Ein *zulässiger Punkt* ist eine Matrix $A \succeq 0$ mit $\langle M_i, A\rangle = \beta_i$ für alle i. Ein *strikt zulässiger Punkt* ist ein zulässiger Punkt A, der zusätzlich positiv definit ist, also $A \succ 0$.
Das dazugehörige Problem in *dualer Form* ist:

$$\begin{array}{ll} \text{finde} & \sup \sum_{i=1}^{m} \lambda_i \beta_i \\ \text{s.t.} & \sum_i \lambda_i M_i \preceq M \end{array}$$

Ein zulässiger Punkt ist hier ein $\lambda \in \mathbb{R}^m$ mit $\sum \lambda_i M_i \preceq M$, und ein strikt zulässiger Punkt ist ein Punkt mit $\sum_i \lambda_i M_i \prec M$. △

Bemerkung 6.1.3. (i) Für festes $M \in \mathrm{Sym}_d(\mathbb{R})$ ist $A \mapsto \langle M, A\rangle$ eine lineare Abbildung auf $\mathrm{Sym}_d(\mathbb{R})$, und jede lineare Abbildung ist von dieser Gestalt (das gilt ja allgemein für Skalarprodukte auf endlich-dimensionalen Räumen). Die Bedingungen $\langle M_i, A\rangle = \beta_i$ definieren also affin-lineare Hyperebenen von $\mathrm{Sym}_d(\mathbb{R})$. Ein semidefinites Optimierungsproblem in primaler Form besteht also darin, eine lineare Funktion über einen affin-linearen Schnitt des Kegels $\mathcal{P}_d$ der positiv semidefiniten Matrizen zu minimieren.
(ii) Im dualen Problem maximiert man die auf dem $\mathbb{R}^m$ definierte lineare Funktion $\lambda \mapsto \beta^t\lambda$. Dabei wird auf die Menge eingeschränkt, die durch die Bedingung

$\sum \lambda_i M_i \preceq M$ definiert ist. Das ist offensichtlich eine abgeschlossene konvexe Menge.
(iii) Das duale Problem lässt sich in die Form eines primalen Problems bringen, und umgekehrt. Im dualen Problem optimiert man ja eine lineare Funktion im affinen Unterraum

$$M + \mathrm{span}_{\mathbb{R}}(M_1, \ldots, M_m) \subseteq \mathrm{Sym}_d(\mathbb{R})$$

über den Schnitt mit $\mathcal{P}_d$. Genau das macht man in einem primalen Problem auch. Umgekehrt, wenn man im primalen Problem eine Basis für den affinen Raum wählt, über den optimiert wird, erhält man ein Problem in dualer Gestalt.
(iv) Für semidefinite Optimierungsprobleme existieren numerische *Innere-Punkte-Methoden*, die in vielen Fällen effizient den Optimalwert und die Optimalstelle berechnen können.
(v) Wählt man die Matrizen M, M_i alle diagonal, so definiert die Bedingung $\sum_i \lambda_i M_i \preceq M$ im dualen Problem einen *Polyeder*, d.h. einen endlichen Durchschnitt von Halbräumen. Man optimiert also eine lineare Funktion über einen Polyeder. Das nennt man auch *lineare Optimierung*. Semidefinite Optimierung ist also eine Verallgemeinerung von linearer Optimierung. △

Satz 6.1.4 (Dualitätssatz der semidefiniten Optimierung)**.** *Sei p^* der Optimalwert des primalen Problems, und d^* der Optimalwert des dualen Problems. Dann gilt*

$$d^* \leq p^*.$$

Falls beide Probleme einen zulässigen Punkt besitzen, und eines der beiden sogar einen strikt zulässigen Punkt, so gilt

$$d^* = p^*.$$

Beweis. Falls das duale Problem überhaupt keinen zulässigen Punkt besitzt, gilt $d^* = -\infty$, und die erste Aussage ist also trivial. Das entsprechende gilt, wenn das primale Problem keinen zulässigen Punkt besitzt.
Seien also $\lambda \in \mathbb{R}^m$ und $A \in \mathcal{P}_d$ dual bzw. primal zulässige Punkte. Wegen $M - \sum_i \lambda_i M_i \in \mathcal{P}_d$ gilt nach Proposition 6.1.1 (iii)

$$0 \leq \langle M - \sum_i \lambda_i M_i, A\rangle = \langle M, A\rangle - \sum_i \lambda_i \langle M_i, A\rangle,$$

d.h.

$$\sum_i \lambda_i \beta_i \leq \langle M, A\rangle.$$

Da d^* das Supremum über alle Ausdrücke links und p^* das Infimum über alle Ausdrücke rechts ist, folgt $d^* \leq p^*$.
Sei nun $\lambda \in \mathbb{R}^m$ ein strikt zulässiger Punkt des dualen Problems, d.h. es gilt $M - \sum \lambda_i M_i \succ 0$. Wir zeigen zunächst dass die Menge

$$K := \{(\langle A, M\rangle, \langle A, M_1\rangle, \ldots, \langle A, M_m\rangle) \mid A \in \mathcal{P}_d\} \subseteq \mathbb{R}^{m+1}$$

ein abgeschlossener konvexer Kegel ist. Die Kegeleigenschaft ist klar. Sei nun $(A_j)_{j\in\mathbb{N}}$ eine Folge in $\mathcal{P}_d$, so dass die Tupel

$$(\langle A_j, M\rangle, \langle A_j, M_1\rangle, \ldots, \langle A_j, M_m\rangle)$$

für $j \to \infty$ gegen $r \in \mathbb{R}^{m+1}$ konvergieren. Wir können annehmen dass $A_j \neq 0$ für alle A_j gilt. Sei $\|\cdot\|$ irgendeine Norm auf $\mathrm{Sym}_d(\mathbb{R})$, z.B. die durch $\langle\cdot,\cdot\rangle$ induzierte. Dann gibt es o.B.d.A. ein $A \in \mathcal{P}_d \setminus \{0\}$ mit

$$\frac{A_j}{\|A_j\|} \overset{j\to\infty}{\to} A,$$

aufgrund des Satzes von Bolzano-Weierstraß und der Abgeschlossenheit von $\mathcal{P}_d$. Nach Proposition 6.1.1 (iv) gilt

$$\begin{aligned} 0 < \langle A, M - \sum_i \lambda_i M_i\rangle &= \lim_j \frac{1}{\|A_j\|}\langle A_j, M - \sum_i \lambda_i M_i\rangle \\ &= \lim_j \frac{1}{\|A_j\|}\left(\langle A_j, M\rangle - \sum_i \lambda_i \langle A_j, M_i\rangle\right). \end{aligned}$$

Der zweite Faktor konvergiert dabei gegen $r_0 - \lambda_1 r_1 - \cdots - \lambda_m r_m$, bleibt also beschränkt. Damit darf der erste Faktor nicht gegen Null gehen, die Normen der A_j sind also beschränkt. Damit können wir also o.B.d.A. annehmen dass die Folge A_j selbst konvergiert, und damit ist $r \in K$. Damit haben wir die Abgeschlossenheit von K gezeigt.
Da beide Probleme einen zulässigen Punkt besitzen, gilt

$$-\infty < d^* \leq p^* < \infty.$$

Sei nun $p < p^*$ beliebig. Dann gehört das Tupel $(p, \beta_1, \ldots, \beta_m)$ nicht zu K. Sonst gäbe es ja einen primal zulässigen Punkt A mit $\langle A, M\rangle = p < p^*$.

Nach dem bekannten Trennungssatz für abgeschlossene konvexe Kegel gibt es also einen Vektor $\gamma \in \mathbb{R}^{m+1}$ mit

$$0 \leq \gamma_0\langle A, M\rangle + \gamma_1\langle A, M_1\rangle + \cdots + \gamma_m\langle A, M_m\rangle \qquad \forall A \in \mathcal{P}_d \tag{6.1}$$

$$\gamma_0 p + \gamma_1\beta_1 + \cdots + \gamma_m\beta_m < 0. \tag{6.2}$$

Durch Einsetzen eines primal zulässigen Punktes A in (6.1) und vergleichen mit (6.2) sehen wir $\gamma_0 > 0$. Wir teilen durch γ_0 und sehen an (6.1) dass $(-\gamma_1/\gamma_0, \ldots, -\gamma_m/\gamma_0)$ dual zulässig ist (wieder mit Proposition 6.1.1 (iii)), und wegen (6.2) dort einen Wert $> p$ liefert. Damit ist $d^* > p$, und da $p < p^*$ beliebig war, folgt $d^* = p^*$.
Den Fall dass das primale Problem strikt zulässig ist erledigt man ähnlich (oder führt es auf den behandelten Fall zurück, indem man das duale Problem geeignet als primales auffasst, und umgekehrt). □

Bemerkung 6.1.5. Der Dualitätssatz zeigt, wie man semidefinite Optimierungsprobleme numerisch *mit Fehlerabschätzung* lösen kann. Man löst numerisch das primale und das duale Problem gleichzeitig. Dabei wird d^* von unten und p^* von oben approximiert. Wegen $d^* \leq p^*$ weiß man, wie weit man höchstens von den eigentlichen Optima entfernt ist. Falls $d^* = p^*$ gilt, konvergieren die approximierenden Folgen sogar gegeneinander. △

Im folgenden Abschnitt schauen wir uns an, wie man den Satz von Schmüdgen oder den archimedischen Positivstellensatz für quadratische Moduln zur Bestimmung des Optimums eines beliebigen polynomialen Optimierungsproblems verwenden kann.

6.2 Die Optimierungsmethode von Lasserre

Gegeben seien Polynome $p_1, \ldots, p_r \in \mathbb{R}[x_1, \ldots, x_n]$. Wieder sei

$$W := W_{\mathbb{R}}(p_1, \ldots, p_r) = \{a \in \mathbb{R}^n \mid p_1(a) \geq 0, \ldots, p_r(a) \geq 0\}$$

die von den p_i definierte basisch abgeschlossene Menge. Für ein weiteres Polynom $p \in \mathbb{R}[x_1, \ldots, x_n]$ interessieren wir uns nun für das Infimum auf W:

$$p_* := \inf\{p(a) \mid a \in W\}.$$

Man beachte, dass W nicht als konvex vorausgesetzt wird, und p nicht als linear. Das Problem ist also kein semidefinites Optimierungsproblem, und auch keins der meisten anderen konvexen Optimierungsprobleme, für die effiziente Algorithmen exisitieren. Die Berechnung von p_* ist im Allgemeinen auch sehr schwer. Bezeichne nun mit $\mathbb{R}[x]_d$ den Raum der Polynome vom Grad $\leq d$. Mit $M_d(p_1, \ldots, p_r)$ bezeichnen wir die Elemente aus dem quadratischen Modul $M(p_1, \ldots, p_r)$, die *offensichtlich* in $\mathbb{R}[x]_d$ liegen. Die genaue Definition ist

$$M_d(p_1, \ldots, p_r) := \{\sigma_0 + \sigma_1 p_1 + \cdots + \sigma_r p_r \mid \deg(\sigma_0), \deg(\sigma_i p_i) \leq d\}.$$

Wir nehmen also alle Elemente die deshalb Grad höchstens d haben, weil alle Summanden Grad höchstens d haben. Das bedeutet auch, dass

$$\deg(\sigma_i) \leq d - \deg(p_i)$$

gilt. Man beachte, dass die Inklusion

$$M_d(p_1, \ldots, p_r) \subseteq M(p_1, \ldots, p_r) \cap \mathbb{R}[x]_d$$

im Allgemeinen eine strikte ist (vergleiche Beispiel 4.2.4(v)). Es gilt

$$M(p_1, \ldots, p_r) = \bigcup_{d \in \mathbb{N}} M_d(p_1, \ldots, p_r).$$

Wir setzen nun für jedes $d \in \mathbb{N}$

$$p_{*,d} := \sup\{s \in \mathbb{R} \mid p - s \in M_d(p_1, \ldots, p_r)\}.$$

Der folgende Satz beschreibt die Optimierungsmethode von Lasserre:

Satz 6.2.1. *Jedes $p_{*,d}$ ist der Optimalwert eines semidefiniten Optimierungsproblems, das man aus $p, p_1, \ldots, p_r$ explizit konstruieren kann. Es gilt $p_{*,d} \leq p_*$ für alle d, und die Folge $(p_{*,d})_{d \in \mathbb{N}}$ ist monoton wachsend. Falls $M(p_1, \ldots, p_r)$ archimedisch ist, konvergiert sie gegen p_*.*

Beweis. Setze $M = M(p_1, \ldots, p_r)$, $M_d = M_d(p_1, \ldots, p_r)$, $W = W_{\mathbb{R}}(p_1, \ldots, p_r)$. Falls $p - s \in M_d$, so natürlich $p - s \in M$, und damit ist $p - s \geq 0$ auf W. Also gilt $s \leq p_*$, und damit ist $p_{*,d} \leq p_*$. Die Folge der $p_{*,d}$ ist offensichtlich monton wachsend, denn $M_d \subseteq M_{d+1}$.

Sei nun M archimedisch, und $\epsilon > 0$ beliebig. Dann ist $p - (p_* - \epsilon)$ strikt positiv auf W, und nach Satz 5.4.3 gilt

$$p - (p_* - \epsilon) \in M = \bigcup_{d \in \mathbb{N}} M_d.$$

Also gibt es ein d mit $p - (p_* - \epsilon) \in M_d$, d.h. $p_{*,d} \geq p_* - \epsilon$. Das beweist die Konvergenz.
Wir müssen nun noch zeigen, dass $p_{*,d}$ der Optimalwert eines semidefiniten Optimierungsproblems ist. Wir verwenden dafür die in Abschnitt 2.2 eingeführten Grammatrizen, und betrachten den endlich-dimensionalen Vektorraum

$$V := \mathbb{R} \times \mathrm{Sym}_{\delta_0}(\mathbb{R}) \times \cdots \times \mathrm{Sym}_{\delta_r}(\mathbb{R}).$$

Dabei ist δ_i so gewählt, dass $G(N_i)$ höchstens Grad $d - \deg(p_i)$ hat, für jedes $N_i \in \mathrm{Sym}_{\delta_i}(\mathbb{R})$. In V betrachten wir nun den affin-linearen Unterraum

$$H := \{(s, N_0, \ldots, N_r) \mid p - s = G(N_0) + G(N_1)p_1 + \cdots + G(N_r)p_r\}.$$

Es ist $p_{*,d}$ das Supremum der linearen Funktion $(s, N_0, \ldots, N_r) \mapsto s$, über den Schnitt von H mit der Menge, die durch die Bedingung $N_i \succeq 0$ für alle i definiert wird (Satz 2.2.17). Das zeigt im Prinzip bereits, dass es sich um ein semidefinites Optimierungsproblem handelt. Wenn man einen affin-linearen Schnitt mit *einem einzigen Kegel* von positiv semidefiniten Matrizen als Definitionsbereich erreichen möchte, kann man V noch in $\mathrm{Sym}_{1+\delta_0+\cdots+\delta_r}(\mathbb{R})$ einbetten, durch

$$(s, N_0. \ldots, N_r) \mapsto \mathrm{diag}(s, N_0, \ldots, N_r). \qquad \square$$

Bemerkung 6.2.2. (i) Man beachte nochmal, dass man die Bedingung

$$''M(p_1, \ldots, p_r) \text{ ist archimedisch}''$$

einfach erreichen kann, indem man ein Polynom $p_{r+1} = N - \sum_i x_i^2$ zu den definierenden Polynomen hinzufügt. Falls $W_{\mathbb{R}}(p_1, \ldots, p_r)$ beschränkt ist, ändert sich diese Menge dabei nicht, vorausgesetzt N ist groß genug.
(ii) Für die Methode von Lasserre gibt es Implementierungen, zum Beispiel im Matlab-Plugin *Yalmip*. Solange die Anzahl der Variablen und der Grad nicht zu groß werden, kann man polynomiale Optimierungsprobleme damit explizit lösen. Dabei kommen noch Verfeinerungen der hier beschriebenen Methode zum Einsatz, mit denen man die Komplexität zu verringern versucht, abhängig von der Struktur von p. Teilweise basieren sie auf verfeinerten Positivstellensätzen. $\triangle$

6.3 Spektraeder

Die Definitionsmengen der semidefiniten Optimierung nennt man *Spektraeder*. Es ist für eine gegebene Menge oft gar nicht einfach zu entscheiden, ob sie ein Spektraeder ist, und sie gegebenenfalls durch explizite positiv semidefinite Matrizen zu beschreiben. Eine Vorstellung davon haben wir schon im Beweis von Satz 6.2.1 bekommen. Andererseits ist es für die Anwendbarkeit der semidefiniten Optimierung wichtig, ihre Definitionsmengen möglichst gut zu kennen. Wir wollen diese Fragen deshalb etwas genauer beschreiben. Die Theorie wird übersichtlicher, wenn wir uns auf spektraedrische *Kegel* beschränken. Außerdem ist das keine wirkliche Einschränkung, der allgemeine Fall entsteht jeweils durch Schnitt des Kegels mit einem affinen Unterraum.

Definition 6.3.1. Ein *spektraedrischer Kegel* ist eine Menge der Gestalt

$$\mathcal{S}(M_1, \ldots, M_n) = \{a \in \mathbb{R}^n \mid a_1 M_1 + \cdots + a_n M_n \succeq 0\},$$

für gewisse Matrizen $M_1, \ldots, M_n \in \mathrm{Sym}_d(\mathbb{R})$. △

Bemerkung 6.3.2. (i) Ein spektraedrischer Kegel ist also gerade das Urbild eines Kegels $\mathcal{P}_d$ von positiv semidefiniten Matrizen unter einer linearen Abbildung $\mathbb{R}^n \to \mathrm{Sym}_d(\mathbb{R})$. Wenn die Matrizen M_i linear unabhängig sind (was man meistens annimmt), kann man es also als Schnitt von $\mathcal{P}_d$ mit einem Unterraum auffassen.
(ii) Spektraedrische Kegel sind konvexe, basisch abgeschlossene (und damit abgeschlossene) Kegel. Das folgt sofort daraus, dass $\mathcal{P}_d$ diese Eigenschaft hat.
(iii) Jeder polyedrische Kegel, also eine Menge der Gestalt

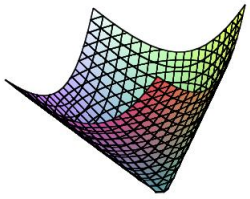
$$\{a \in \mathbb{R}^n \mid v_1^t a \geq 0, \ldots, v_r^t a \geq 0\}$$

für gewisse $v_i \in \mathbb{R}^n$, ist spektraedrisch. Wenn man alle M_i diagonal wählt, entstehen gerade solche Bedingungen.
(iv) Durchschnitte von spektraedrischen Kegeln sind wieder spektraedrisch. Man kann ja zwei Tupel von Matrizen als Blockmatrizen zusammenfügen. △

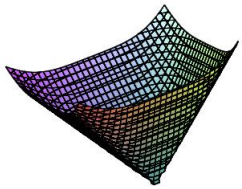
Beispiel 6.3.3. (i) Wählt man $M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, M_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ und $M_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, so entsteht die Bedingungen

$$a_1^2 - a_2^2 - a_3^2 \geq 0, a_1 \geq 0$$

wie man sich leicht überlegt. Es ist $\mathcal{S}(M_1, M_2, M_3) \subseteq \mathbb{R}^3$ also gerade ein einfacher Kreiskegel, der offensichtlich nicht polyedrisch ist:



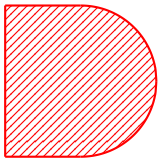
(ii) Der Kegel der durch die Bedingungen $a_1^4 - a_2^4 - a_3^4 \geq 0, a_1 \geq 0$ definiert wird, sieht relativ ähnlich aus:



Er ist allerdings nicht spektraedrisch, wir wir noch sehen werden.
(iii) Der Kegel dessen Querschnitt die Menge

$$[-1, 0] \times [-1, 1] \cup B \subseteq \mathbb{R}^2$$

ist, wobei B die Einheitskreisscheibe bezeichne, ist nicht spektraedrisch. Er ist nämlich nicht basisch semialgebraisch (Übungsaufgabe 33).



△

Beispiel 6.3.4. Sei $\mathbb{R}[x]_d^*$ der *algebraische Dualraum* des endlich-dimensionalen Raums $\mathbb{R}[x]_d = \mathbb{R}[x_1, \ldots, x_n]_d$. Ein Element $\varphi \in \mathbb{R}[x]_d^*$ ist also eine lineare Abbildung $\varphi\colon \mathbb{R}[x]_d \to \mathbb{R}$. Da lineare Abbildungen eindeutig durch die Werte auf einer Basis bestimmt sind, kann man φ identifizieren mit den Werten auf der kanonischen monomialen Basis von $\mathbb{R}[x]_d$, d.h.

$$\varphi = (\varphi(x^\alpha))_{\alpha \in \mathbb{N}^n; |\alpha| \leq d}\,.$$

Auf diese Weise identifiziert sich also $\mathbb{R}[x]_d^*$ mit $\mathbb{R}^{\Delta_d}$.
Seien nun $p_1, \ldots, p_r \in \mathbb{R}[x]$ gegeben. Wir betrachten wieder den trunkierten quadratischen Modul

$$M_d = M_d(p_1, \ldots, p_r) \subseteq \mathbb{R}[x]_d$$

aus dem letzten Abschnitt. Sei nun

$$M_d^\vee = \{\varphi \in \mathbb{R}[x]_d^* \mid \varphi \geq 0 \text{ auf } M_d\}$$

der zu M_d *duale Kegel*. Dann ist $M_d^\vee$ ein Spektraeder in $\mathbb{R}[x]_d^* = \mathbb{R}^{\Delta_d}$. Die Bedingung $\varphi \geq 0$ auf M_d zerlegt sich ja in die Einzelbedingungen

$$\varphi(q^2 p_i) \geq 0 \qquad \forall q \in \mathbb{R}[x]_{k_i}$$

wobei k_i so gewählt ist dass $q^2 p_i \in \mathbb{R}[x]_d$. Schreibt man $q = \sum_\alpha q_\alpha x^\alpha$ so ist

$$\varphi(q^2 p) = \varphi\left(\sum_{\alpha,\beta} q_\alpha q_\beta x^{\alpha+\beta} p\right) = \sum_{\alpha,\beta} q_\alpha q_\beta \left(\varphi(x^{\alpha+\beta} p)\right).$$

Dass dieser Ausdruck für alle Wahlen der Koeffizienten q_α nichtnegativ ist bedeutet aber gerade, dass die Matrix $\left(\varphi(x^{\alpha+\beta} p)\right)_{\alpha,\beta}$ positiv semidefinit ist. Die Einträge sind aber Linearkombinationen der Werte $\varphi(x^\alpha)$, abhängig von p. Das beweist die Aussage. △

Es ist bekannt, dass konvexe Mengen in ihrer affinen Hülle immer nichtleeres Inneres besitzen. Wenn man den umgebenden Raum einer konvexen Menge also durch ihre affine Hülle ersetzt, kann man so immer annehmen, dass die Menge nichtleeres Inneres hat.

Proposition 6.3.5. *Sei $S \subseteq \mathbb{R}^n$ ein spektraedrische Kegel der einen inneren Punkt e besitzt. Dann gibt es symmetrische Matrizen $M_1, \ldots, M_n$ mit $S = \mathcal{S}(M_1, \ldots, M_n)$ und $e_1 M_1 + \cdots + e_n M_n = I$.*

Beweis. Sei zunächst $S = \mathcal{S}(N_1, \ldots, N_n)$ mit symmetrischen Matrizen N_i. Wir setzen

$$a \bullet N := a_1 N_1 + \cdots + a_n N_n.$$

Da e im Innern von S liegt gilt

$$e \bullet N \pm \epsilon N_i \succeq 0$$

für alle i und $\epsilon > 0$ klein genug. Für jeden Vektor $v \in \ker(e \bullet N)$ gilt also

$$0 \leq v^t \left(e \bullet N \pm \epsilon N_i\right) v = \pm \epsilon v^t N_i v.$$

Also muss $v^t N_i v = 0$ und damit $0 = v^t \left(e \bullet N \pm \epsilon N_i\right) v$ gelten. Da $e \bullet N \pm \epsilon N_i$ positiv semidefinit ist, folgt daraus bereits

$$(e \bullet N \pm \epsilon N_i) v = 0,$$

wie man zum Beispiel anhand der Zerlegung in Quadrate vom Rang 1 wie in Lemma 2.2.3 (iv) sieht. Nun folgt daraus $N_i v = 0$, d.h wir haben $\ker(e \bullet N) \subseteq \ker(N_i)$ gezeigt. Nach einem Basiswechsel entsteht in allen N_i oben links ein Block aus Nullen, den wir für die Definition von S einfach weglassen können, und die neuen Matrizen M_i erfüllen dann $\ker(e \bullet M) = \{0\}$, d.h. $e \bullet M \succ 0$. Nach Konjugation mit einer weiteren Matrix können wir also $e \bullet M = I$ erreichen. □

Der folgende Satz stellt einen Zusammenhang zwischen der Konvexgeometrie von Spektraedern und der (reellen) algebraischen Geometrie her:

Satz 6.3.6. *Seien $M_1, \ldots, M_n \in \mathrm{Sym}_d(\mathbb{R})$, und sei $e \in \mathbb{R}^n$ mit $e_1 M_1 + \cdots + e_n M_n = I$. Setze*

$$h := \det\left(x_1 M_1 + \cdots + x_n M_n\right).$$

Dann ist $h \in \mathbb{R}[x_1, \ldots, x_n]$ homogen vom Grad d, es gilt $h(e) = 1$, und für jedes $a \in \mathbb{R}^n$ hat das Polynom

$$h_a(t) := h(a - te) \in \mathbb{R}[t]$$

nur reelle Nullstellen. Es gilt

$$\mathcal{S}(M_1, \ldots, M_n) = \left\{a \in \mathbb{R}^n \mid \textit{alle Nullstellen von } h_a \textit{ sind } \geq 0\right\}.$$

Beweis. Die Homogenität von h ist klar. Es gilt für $a \in \mathbb{R}^n$

$$h_a(t) = \det((a - te) \bullet M) = \det(a \bullet M - tI),$$

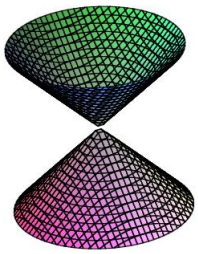
d.h. h_a ist das charakteristische Polynom der symmetrischen Matrix $a \bullet M$. Die Nullstellen von h_a sind also die Eigenwerte von $a \bullet M$, und die sind alle reell. Weiter ist $a \bullet M$ genau dann positiv semidefinit, wenn alle diese Eigenwerte bzw. Nullstellen von h_a nichtnegativ sind. □

Definition 6.3.7. (i) Ein homogenes Polynom $h \in \mathbb{R}[x_1, \ldots, x_n]$ heißt *hyperbolisch in Richtung* $e \in \mathbb{R}^n$, falls $h(e) \neq 0$ und die Nullstellen von $h_a(t) := h(a - te)$ alle reell sind, für alle $a \in \mathbb{R}^n$.
(ii) Ist h hyperbolisch in Richtung e, so nennt man

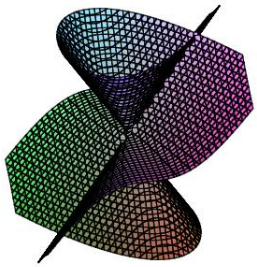
$$\Lambda_e(h) = \{a \in \mathbb{R}^n \mid \text{alle Nullstellen von } h_a \text{ sind } \geq 0\}$$

den *Hyperbolizitätskegel* von h (in Richtung e). △

Beispiel 6.3.8. (i) Das Polynom $h = x_1^2 - x_2^2 - x_3^2$ ist hyperbolisch in Richtung $e = (1, 0, 0)$, wie man zum Beispiel am folgenden Bild sieht:

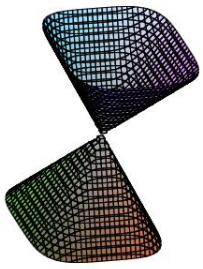


Auf jeder senkrechten Gerade gibt es 2 reelle Schnittpunkte, und da h Grad 2 hat, kann es keine echt komplexen Nullstellen geben. Das gleiche stimmt für das Polynom $h = x_1^3 - x_1^2 x_3 - x_1 x_3^2 + x_1 x_2^2 + x_3^3$, das Grad 3 hat:



Die Hyperbolizitätskegel sind jeweils die ausgefüllten Kegel, die nach oben zeigen.

(*ii*) Das Polynom $h = x_1^4 - x_2^4 - x_3^4$ ist nicht hyperbolisch in Richtung $e = (1, 0, 0)$. Auf jeder senkrechten Geraden gibt es nur 2 reelle Nullstellen. Da h Grad 4 hat, muss es immer zwei echt komplexe Nullstellen geben, die man im Bild nicht sieht. h ist auch in keine andere Richtung hyperbolisch.



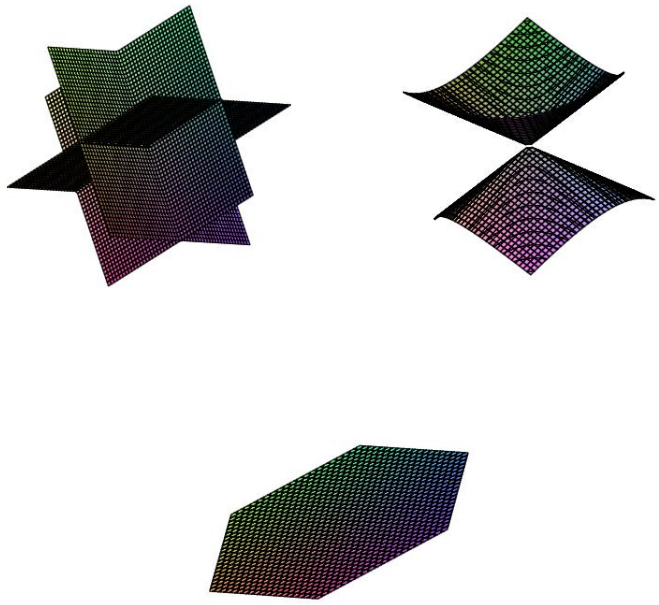
(*iii*) Die elementar-symmetrischen Polynome

$$s_{r,n} = \sum_{1 \le i_1 < \cdots < i_r \le n} x_{i_1} \cdots x_{i_r} \in \mathbb{R}[x_1, \ldots, x_n],$$

die wir in Abschnitt 1.3 schon kennengelernt haben, sind hyperbolisch in Richtung $e = (1, \ldots, 1)$. Am besten sieht man das folgendermaßen. Es ist $s_{n,n} = x_1 \cdots x_n$, und also hat

$$s_{n,n}(a - te) = s_{n,n}(a_1 - t, \ldots, a_n - t) = (a_1 - t) \cdots (a_n - t)$$

die reellen Nullstellen $a_1, \ldots, a_n$. Man sieht nun, dass $s_{r,n}(a - te)$ gerade die r-te Ableitung von $s_{n,n}(a - te)$ nach t ist (bis auf Vorzeichen). Nach dem Satz von Rolle entsteht zwischen zwei reellen Nullstellen eines Polynoms immer eine Nullstelle der Ableitung, also haben alle diese Ableitungen auch nur reelle Nullstellen. Die folgenden Bilder zeigen die reellen Varietäten von $s_{3,3}$, $s_{2,3}$ und $s_{1,3}$:



△

Bemerkung 6.3.9. (i) Hyperbolizitätskegel sind in der Tat immer konvexe Kegel. Das ist allerdings nicht offensichtlich. Man kann es entweder elementar beweisen, allerdings recht technisch, oder man verwendet den Satz von Helton und Vinnikov (Satz 6.3.12) weiter unten, siehe Übungsaufgabe 49.
(ii) Ist h hyperbolisch in Richtung e, so ist für jedes e' im Inneren von $\Lambda_e(h)$ das Polynom h auch hyperbolisch in Richtung e', und $\Lambda_e(h) = \Lambda_{e'}(h)$. Auch das kann man wieder elementar und recht technisch beweisen, oder mit dem Satz von Helton und Vinnikov (siehe Übungsaufgabe 49). Anschaulich ist es sehr plausibel.
(iii) Jeder spektraedrische Kegel ist ein Hyperbolizitätskegel. Das ist die Aussage von Satz 6.3.6 (zusammen mit Proposition 6.3.5). △

Beispiel 6.3.10. Der Kegel K aus Beispiel 6.3.3 (ii), definiert durch die Bedingung $a_1^4 - a_2^4 - a_3^4 \geq 0, a_1 \geq 0$ ist nicht hyperbolisch, und damit auch nicht spektraedrisch. Wäre er nämlich hyperbolisch, so gäbe es ein hyperbolisches Polynom h mit $h = 0$ auf ∂K. Aufgrund der Homogenität wäre $h = 0$ auf der ganzen Varietät $V_{\mathbb{R}}(x_1^4 - x_2^4 - x_3^4) \subseteq \mathbb{R}^3$. Nach dem reellen Nullstellensatz läge also h in $\mathrm{rrad}(I(x_1^4 - x_2^4 - x_3^4))$. Dieses Ideal ist aber reell, wie man sich analog zum Fall $1 - x_1^2 - x_2^2$ überlegt (das war Übungsaufgabe 30). Also enthält h das Polynom $x_1^4 - x_2^4 - x_3^4$ als Faktor. Da dieses Polynom nicht hyperbolisch ist, kann es h auch nicht sein, ein Widerspruch. △

Jeder spektraedrische Kegel ist hyperbolisch. Die Hyperbolizität eines Kegels ist prinzipiell einfacher zu überprüfen als die Spektraedereigenschaft, wie wir in Beispiel 6.3.10 gerade gesehen haben. Das motiviert die folgende Vermutung:

Vermutung 6.3.11 (Allgemeine Lax-Vermutung)**.** *Jeder Hyperbolizitätskegel ist spektraedrisch.*

Die Vermutung ist bisher offen, und es gibt nur Teilresultate. Eines ist der Satz von Helton & Vinnikov, den wir nur zitieren. Der zweite Teil der Aussage folgt dabei direkt aus Satz 6.3.6.

Satz 6.3.12 (Helton & Vinnikov)**.** *Sei $h \in \mathbb{R}[x_1, x_2, x_3]$ hyperbolisch in Richtung $e \in \mathbb{R}^3$ mit $h(e) = 1$. Dann gibt es Matrizen $M_1, M_2, M_3 \in \mathrm{Sym}_d(\mathbb{R})$ mit $e_1M_1 + e_2M_2 + e_3M_3 = I$ und*

$$h = \det\left(x_1M_1 + x_2M_2 + x_3M_3\right).$$

Insbesondere ist jeder Hyperbolizitätskegel in $\mathbb{R}^3$ spektraedrisch.

Die genaue Aussage des Satzes von Helton & Vinnikov stimmt in höheren Dimensionen nicht mehr.

Beispiel 6.3.13. Das Polynom $h = x_1^2 - x_2^2 - x_3^2 - x_4^2 \in \mathbb{R}[x_1, x_2, x_3, x_4]$ ist hyperbolisch in Richtung $e = (1, 0, 0, 0)$. Es ist nämlich

$$h_a(t) = h(a_1 - t, a_2, a_3, a_4) = (a_1 - t)^2 - a_2^2 - a_3^2 - a_4^2,$$

und beide Nullstellen sind reell. Angenommen es gilt

$$h = \det\left(x_1M_1 + x_2M_2 + x_3M_3 + x_4M_4\right)$$

mit symmetrischen Matrizen M_i, die aus Gründen der Homogenität der Größe 2 sein müssen. Die M_i sind im Raum $\mathrm{Sym}_2(\mathbb{R})$ aber linear unabhängig. Könnte man eine nämlich durch eine Linearkombination der anderen ersetzen, so wäre $h = q(Ax)$ für ein Polynom q in drei Variablen und eine Matrix $A \in \mathrm{M}_{3\times 4}(\mathbb{R})$. Für $0 \neq v \in \ker A$ und $\lambda \in \mathbb{R}$ gilt dann aber

$$1 = h(e) = h(e + \lambda v) = \lambda^2 h(v) + 2\lambda v_1 + 1.$$

Daraus folgt $0 = v_1 = h(v) = -v_2^2 - v_3^2 - v_4^2$, also $v = 0$, ein Widerspruch.
In $\mathrm{Sym}_2(\mathbb{R})$ gibt es aber höchstens drei linear unabhängige Matrizen. Also besitzt h keine Determinantendarstellung. △

Bemerkung 6.3.14. Wäre $h^r = \det(x_1 M_1 + \cdots + x_n M_n)$ mit symmetrischen Matrizen und $e \bullet M = I$ für ein $r \in \mathbb{N}$, so wäre

$$\Lambda_e(h) = \Lambda_e(h^r) = \mathcal{S}(M_1, \ldots, M_n)$$

immer noch spektraedrisch. Im Beispiel 6.3.13, und allgemeiner für quadratische hyperbolische Polynome, stimmt das auch wirklich immer.
Es gibt allerdings ein Beispielpolynom h von Brändén, vom Grad 4 und in 4 Variablen, von dem *keine Potenz* eine Determinantendarstellung besitzt. △

Bemerkung 6.3.15. Sehr viel mehr ist über die allgemeine Lax-Vermutung bisher nicht bekannt. Brändén hat noch gezeigt, dass die Hyperbolizitätskegel der elementar-symmetrischen Polynome $s_{r,n}$ alle spektraedrisch sind. Dafür produzierte er keine Determinantendarstellung von Potenzen der Polynome, sondern von anderen *Vielfachen* $h \cdot s_{r,n}$, mit einem Zusatzfaktor h, der den Hyperbolizitätskegel nicht ändert. △

6.4 Spektraedrische Schatten

Definition 6.4.1. Ein *spektraedrischer Schatten* ist das Bild eines spektraedrischen Kegels unter einer linearen Abbildung. △

Bemerkung 6.4.2. (i) Das lineare Bild eines Polyeders ist wieder ein Polyeder. Dasselbe stimmt für Spektraeder nicht. Man betrachte beispielsweise den Kegel

$$K = \left\{(a,b,c,d,e) \in \mathbb{R}^5 \mid b^2 \leq da, c^2 \leq ea, d^2 + e^2 \leq a^2, 0 \leq a, d, e\right\}.$$

Die Bedingungen $b^2 \leq da, 0 \leq a, d$ übersetzen sich zum Beispiel in

$$\begin{pmatrix} a & b \\ b & d \end{pmatrix} \succeq 0.$$

Man sieht so dass K spektraedrisch ist. Projizieren wir K anhand der Abbildung $(a,b,c,d,e) \mapsto (a,b,c)$ in den $\mathbb{R}^3$, erhalten wir den Kegel

$$K' = \left\{(a,b,c) \mid a^4 \geq b^4 + c^4, a \geq 0\right\}.$$

In Beispiel 6.3.10 haben wir gesehen, dass K' nicht spektraedrisch ist.
(ii) Die Klasse der spektraedrischen Schatten ist abgeschlossen unter den meisten bekannten Operationen konvexer Mengen, zum Beispiel unter Bilden von Dualen, Polaren, Minkowskisummen, Abschlüssen und Innerem. △

Es gibt bisher im wesentlichen nur eine Methode, um spektraedrische Schatten systematisch zu konstruieren. Diese Methode geht ebenfalls auf Lasserre zurück, und sie funktioniert folgendermaßen. Seien wieder $p_1, \dots, p_r \in \mathbb{R}[x]$ gegeben. Wir betrachten die Menge

$$W_{\mathbb{R}}(p_1, \dots, p_r) = \{a \in \mathbb{R}^n \mid p_1(a) \geq 0, \dots, p_r(a) \geq 0\}$$

und den trunkierten quadratischen Modul

$$M_d = M_d(p_1, \dots, p_r) \subseteq \mathbb{R}[x]_d.$$

Wir haben in Beispiel 6.3.4 gesehen, dass

$$M_d^\vee = \{\varphi \colon \mathbb{R}[x]_d \to \mathbb{R} \text{ linear } \mid \varphi \geq 0 \text{ auf } M_d\} \subseteq \mathbb{R}[x]_d^* = \mathbb{R}^{\Delta_d}$$

ein Spektraeder ist. Wir projizieren diesen Spektraeder nun anhand der linearen Abbildung

$$\begin{aligned} \pi \colon \mathbb{R}[x]_d^* &\to \mathbb{R}^n \\ \varphi &\mapsto (\varphi(x_1), \dots, \varphi(x_n)) \end{aligned}$$

in den $\mathbb{R}^n$. In Koordinaten formuliert projiziert man das Tupel $\varphi = (\varphi(x^\alpha))_{|\alpha| \leq d}$ auf die Koordinaten, die mit $x_1, \dots, x_n$ indiziert sind.
Für eine Menge $W \subseteq \mathbb{R}^n$ bezeichne $\mathrm{cc}(W)$ deren Kegelhülle, also den kleinsten konvexen Kegel in $\mathbb{R}^n$, der W enthält. Weiter sei $\overline{\mathrm{cc}}(W)$ der Abschluss der Kegelhülle.

Satz 6.4.3. *Seien $p_1, \dots, p_r \in \mathbb{R}[x]$ und $W := W_{\mathbb{R}}(p_1, \dots, p_r)$. Dann gilt:*
(i) Die Menge $\mathcal{L}_d := \pi(M_d^\vee)$ ist ein spektraedrischer Schatten mit

$$\mathrm{cc}\,(W) \subseteq \mathcal{L}_{d+1} \subseteq \mathcal{L}_d$$

für alle $d \in \mathbb{N}$
(ii) Falls für ein $d \in \mathbb{N}$ jedes homogene lineare Polynom $\ell \in \mathbb{R}[x]_1$, welches nichtnegativ auf W ist, zu M_d gehört, so ist

$$\mathrm{cc}\,(W) \subseteq \mathcal{L}_d \subseteq \overline{\mathrm{cc}}\,(W)\,.$$

Beweis. (i): $\mathcal{L}_d$ ist ein spektraedrischer Schatten, da $M_d^\vee$ laut Beispiel 6.3.4 ein spektraedrischer Kegel ist. Die Inklusion $\mathcal{L}_{d+1} \subseteq \mathcal{L}_d$ ist klar, da jedes $\varphi \in M_{d+1}^\vee$ durch Einschränkung auf $\mathbb{R}[x]_d$ ein Element von $M_d^\vee$ definiert. Sein nun $a \in W$.

Dann definiert a ein lineares Funktional $\delta_a \colon \mathbb{R}[x] \to \mathbb{R}$ auf dem gesamten Polynomring (sogar einen Ringhomomorphismus), das auf dem ganzen quadratischen Modul $M(p_1, \ldots, p_r)$ nichtnegativ ist. Durch einschränken auf $\mathbb{R}[x]_d$ erhalten wir $\delta_a \in M_d^\vee$, für alle $d \in \mathbb{N}$. Es ist also

$$a = (\delta_a(x_1), \ldots, \delta_a(x_n)) \in \mathcal{L}_d, \qquad \text{für alle } d.$$

Da mit $M_d^\vee$ auch $\mathcal{L}_d$ ein konvexer Kegel ist, ist also $\mathrm{cc}(W) \subseteq \mathcal{L}_d$ für alle d.
Sei für (ii) also $d \in \mathbb{N}$ wie gefordert. Für $a \notin \overline{\mathrm{cc}}(W)$ gibt es nach dem Trennungssatz für abgeschlossene konvexe Kegel ein homogenes lineares $\ell \in \mathbb{R}[x]_1$ mit $\ell \geq 0$ auf $\overline{\mathrm{cc}}(W)$ und $\ell(a) < 0$. Nach Voraussetzung ist $\ell \in M_d$, und somit für jedes $\varphi \in M_d^\vee$

$$0 \leq \varphi(\ell) = \ell(\varphi(x_1), \ldots, \varphi(x_n)).$$

Das bedeutet $\ell \geq 0$ auf $\mathcal{L}_d$, und also $a \notin \mathcal{L}_d$. Wir haben also $\mathcal{L}_d \subseteq \overline{\mathrm{cc}}(W)$ gezeigt. □

Bemerkung 6.4.4. (i) Der letzte Satz zeigt, dass die Darstellbarkeit aller auf W nichtnegativen lineare Polynome im quadratischen Modul M *mit simultaner Gradschranke* dazu führt, dass die Kegelhülle von W (bis auf Abschluss) ein spektraedrischer Schatten ist.
(ii) Man überzeugt sich, dass es reicht, die strikt positiven linearen Polynome mit Gradschranke in M darzustellen. Der konkrete archimedische Positivstellensatz für quadratische Moduln liefert die Existenz der Darstellungen, im archimedischen Fall. Die Frage nach einer Gradschranke muss extra betrachtet werden.
(iii) Es gibt sowohl Fälle bei denen es eine solche Gradschranke gibt, als auch Fälle, wo es sie nicht gibt, vergleiche Beispiele 6.4.5, 6.4.6 und 6.4.7.
(iv) Es gibt weitere Arbeiten (von Helton & Nie), in denen die Methode von Lasserre lokal angewandt wird. Damit kann für sehr viele Mengen gezeigt werden, dass sie spektraedrische Schatten sind.
(v) Scheiderer hat kürzlich gezeigt, dass *jeder* konvexe semialgebraische Kegel in $\mathbb{R}^3$ ein spektraedrischer Schatten ist. Dabei wird im wesentlichen die hier demonstrierte Methode verwendet, und tiefliegende Ergebnisse über Quadratsummen auf Kurven.
(vi) Ebenso Scheiderer hat gezeigt, dass nicht jede konvexe semialgebraische Menge ein spektraedrischer Schatten ist. Damit wurde die sogenannte *Helton-Nie-Vermutung* widerlegt. △

Wir wollen die Frage nach der Darstellbarkeit aller nichtnegativen linearen Polynome in einem M_d an zwei Beispielen beleuchten. Um die Rechnungen übersichtlich zu halten, betrachten wir dabei konvexe Mengen, die keine Kegel sind.

Man kann die Ergebnisse dann aber auf Kegel übertragen, die die entsprechenden Mengen als Durchschnitte haben.

Beispiel 6.4.5 (Farkas Lemma)**.** Seien $\ell_1, \ldots, \ell_r \in \mathbb{R}[x]_1$ vom Grad 1. Dann ist

$$W = W_{\mathbb{R}}(\ell_1, \ldots, \ell_r) \subseteq \mathbb{R}^n$$

ein Polyeder. Falls ein weiteres $\ell \in \mathbb{R}[x]_1$ auf W nichtnegativ ist, gibt es $\lambda_0, \ldots, \lambda_r \geq 0$ mit

$$\ell = \lambda_0 + \lambda_1\ell_1 + \cdots + \lambda_r\ell_r,$$

zumindest wenn $W \neq \emptyset$. Mit anderen Worten, es gilt $\ell \in M_1(\ell_1, \ldots, \ell_r)$. Das kann man beispielsweise mit dem Dualitätssatz der semidefiniten Optimierung beweisen (Übungsaufgabe 43). △

Beispiel 6.4.6. Wir geben ein weiteres Beispiel an, in dem es die gewünschten Gradschranken gibt. Sei $W = W_{\mathbb{R}}(1 - x_1^4 - x_2^4) \subseteq \mathbb{R}^2$ die Menge, die sich als Durchschnitt des Kegels aus Beispiel 6.3.3 (ii) mit der Ebene $\{x = 1\}$ ergibt. Sei $\ell \in \mathbb{R}[x]_1$ nichtnegativ auf W, und o.B.d.A. $\ell(a) = 0$ für ein $a \in \partial W$. Bis auf Skalierung ist ℓ dann eindeutig bestimmt; mit $a = (r, s)$ gilt

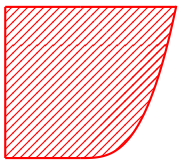
$$\ell = 1 - r^3x_1 - s^3x_2.$$

Das Polynom

$$\ell - \lambda(1 - x_1^4 - x_2^4)$$

ist dann global nichtnegativ, für ein geeignetes $\lambda > 0$. Das sieht man beispielsweise durch Berechnung der kritischen Punkte. Da es Grad 4 und 2 Variablen hat, ist es nach Hilberts Satz eine Quadratsumme von Polynomen vom Grad 2 (siehe Bemerkung 2.2.6). Das zeigt $\ell \in M_4(1 - x_1^4 - x_2^4)$. △

Beispiel 6.4.7. Schließlich geben wir noch ein Beispiel an, in dem es die Gradschranken nicht gibt. Betrachte die Menge $W = W_{\mathbb{R}}(y - x^3, y, 1 - y, x + 1) \subseteq \mathbb{R}^2$:



Für $0 < r < 1$ liegt der Punkt $a = (r, r^3)$ im Rand von W, und das (bis auf Skalierung eindeutige) lineare Polynom ℓ_a, welches auf W nichtnegativ und bei a Null ist, ist

$$\ell_a = 2r^3 - 3r^2x + y.$$

Angenommen es wäre $\ell_a \in M_d$ für alle $a = (r, r^3)$ mit $r > 0$, und einem *festen* d. Wir hätten also Darstellungen

$$\ell_a = \sigma_0^{(a)} + \sigma_1^{(a)}(y - x^3) + \sigma_2^{(a)} y + \sigma_3^{(a)}(1 - y) + \sigma_4^{(a)}(x + 1) \tag{6.3}$$

mit Quadratsummen $\sigma_i^{(a)}$, alle simultan beschränkt im Grad. Durch Einsetzen von a sieht man, dass $\sigma_i^{(a)}(a) = 0$ gelten muss für alle $i \neq 1$. Wir bilden nun den Grenzübergang für $a \to (0, 0)$. Formal macht man das entweder, indem man so skaliert, dass alle Koeffizienten der Quadratsummen simultan beschränkt bleiben, und man dann den Satz von Bolzano-Weierstraß anwenden kann. Alternativ kann man die Existenz einer Darstellung (6.3) für alle $r > 0$ als Formel schreiben, und in einem nichtarchimedischen Körper mit einem infinitesimalen $r > 0$ anwenden. Ähnlich wie in Beispiel 5.3.7 bekommt man dann wieder eine Darstellung über $\mathbb{R}$. In beiden Fällen muss man überprüfen, dass das Ergebnis links nicht das Nullpolynom ergibt.
Man bekommt dann eine Darstellung

$$y = \sigma_0 + \sigma_1(y - x^3) + \sigma_2 y + \sigma_3(1 - y) + \sigma_4(x + 1) \tag{6.4}$$

mit $\sigma_i(0, 0) = 0$ für alle $i \neq 1$. Nun setzt man hier $y = 0$ und erhält

$$0 = \sigma_0(x, 0) + \sigma_1(x, 0)(-x^3) + \sigma_3(x, 0) + \sigma_4(x, 0)(x + 1).$$

Da $-x^3$ und $x + 1$ auf dem ganzen Intervall $[-1, 0]$ nichtnegativ sind, folgt insbesondere $\sigma_1(x, 0) = 0$, d.h. y^2 teilt σ_1. Setzt man nun in (6.4) $x = 0$, erhält man

$$y = \sigma_0(0, y) + \sigma_1(0, y)y + \sigma_2(0, y)y + \sigma_3(0, y)(1 - y) + \sigma_4(0, y).$$

Da $\sigma_i(0, 0) = 0$ für alle $i \neq 1$ gilt, ist y^2 ein Teiler aller $\sigma_i(0, y)$ für $i \neq 1$. Es ist aber y^2 ein Teiler von σ_1, und also von $\sigma_1(0, y)$. Also teilt y^2 die ganze rechte Seite, also auch y, ein Widerspruch.
Der Widerspruch entsteht wirklich durch die Annahme einer simultanen Gradschranke, die es uns erst erlaubt, den Grenzübergang $a \to (0, 0)$ zu machen. Man kann zeigen, dass alle ℓ_a wirklich zum quadratischen Modul $M(y - x^3, y, 1 - y, 1 + x)$ gehören. Nur gehen die Grade der Quadratsummen gegen unendlich, wenn a gegen den Ursprung läuft.
Man kann dieses Beispiel zu einem allgemeineren Satz ausbauen. Sobald eine konvexe und basisch abgeschlossene Menge W eine *nichtexponierte Seite* hat, kann es Gradschranken für die Darstellung linearer Polynome nicht geben. In unserem Beispiel ist der Ursprung eine nichtexponierte Seite (sogar ein Extrempunkt). Für solche Mengen kann man mit der Methode von Lasserre also nicht (direkt) zeigen, dass sie spektraedrische Schatten sind. △

Kapitel 7

Das Momentenproblem

Das *Momentenproblem* ist eine klassische Frage der Funktionalanalysis. Man möchte wissen, welche linearen Funktionale auf einem Raum von Funktionen durch Integration anhand eines Maßes gegeben sind. Für Polynomringe stellt der Satz von Haviland dabei einen Bezug zu positiven Polynomen her. Kann man dann positive Polynome durch Elemente einer endlich erzeugten Präordnung ersetzen, erhält man eine Charakterisierung von Funktionalen mit Maßdarstellung, die relativ einfach, und zum Beispiel mit semidefiniter Optimierung testbar ist. Wir wollen dazu die Grundbegriffe im nächsten Abschnitt einführen.

7.1 Das Momentenproblem und der Satz von Haviland

Sei $\varphi\colon \mathbb{R}[x] \to \mathbb{R}$ eine lineare Abbildung (auch *Funktional* genannt). Wir interessieren uns für die Frage, ob es ein (Borel)-Maß μ auf dem $\mathbb{R}^n$ gibt, so dass

$$\varphi(p) = \int_{\mathbb{R}^n} p \, d\mu$$

für alle $p \in \mathbb{R}[x]$ gilt. Man beachte, dass dabei nur spezielle Maße überhaupt in Frage kommen, die Integrale auf der rechten Seite müssen ja immer endlich sein. Insbesondere muss zum Beispiel $\mu(\mathbb{R}^n) = \int 1 d\mu = \varphi(1) < \infty$ gelten.
Die Frage nach einer Klassifizierung solcher Funktionale mit Maßdarstellung nennt man das *Momentenproblem*. Klassische Ergebnisse sind dazu beispielsweise die folgenden Aussagen, die wir noch genauer herleiten werden.

Satz 7.1.1 (Hamburgers Momentenproblem)**.** *Für ein Funktional $\varphi\colon \mathbb{R}[t] \to \mathbb{R}$ gibt es genau dann ein darstellendes Maß μ auf $\mathbb{R}$, wenn $\varphi(p^2) \geq 0$ für alle $p \in \mathbb{R}[t]$.*

Man beachte, dass die Bedingung $\varphi(p^2) \geq 0$ für alle $p \in \mathbb{R}[t]$ offensichtlich notwendig für die Existenz einer Maßdarstellung ist. Integriert man Quadrate, sogar allgemeiner nichtnegative Polynome, ist das Ergebnis immer nichtnegativ. Man beachte weiter, dass die Bedingung $\varphi(p^2)$ für alle $p \in \mathbb{R}[t]$ auch formuliert werden kann als

$$\varphi \in M_d(1)^\vee$$

für alle d, mit der Notation aus Beispiel 6.3.4. Die Bedingung lässt sich also als Folge von Zugehörigkeitsproblemen zu einem Spektraeder formulieren, und ist damit der semidefiniten Optimierung zugänglich. Weitere klassische Ergebnisse sind:

Satz 7.1.2 (Stieltjes Momentenproblem)**.** *Für ein Funktional $\varphi\colon \mathbb{R}[t] \to \mathbb{R}$ gibt es genau dann ein darstellendes Maß μ auf $[0, \infty)$ (d.h. $\mu((-\infty, 0)) = 0$), wenn*

$$\varphi(p^2) \geq 0 \text{ und } \varphi(p^2 \cdot t) \geq 0$$

für alle $p \in \mathbb{R}[t]$ (d.h. $\varphi \in M_d(t)^\vee$ für alle d).

Satz 7.1.3 (Hausdorffs Momentenproblem)**.** *Für ein Funktional $\varphi\colon \mathbb{R}[t] \to \mathbb{R}$ gibt es genau dann ein darstellendes Maß μ auf $[0, 1]$, wenn*

$$\varphi(p^2) \geq 0 \text{ und } \varphi(p^2 \cdot t) \geq 0 \text{ und } \varphi(p^2(1-t)) \geq 0$$

für alle $p \in \mathbb{R}[t]$ (d.h. $\varphi \in M_d(t, 1-t)^\vee$ für alle d).

Eine ganz allgemeine Klassifikation der Funktionale mit Maßdarstellung liefert der Satz von Haviland:

Satz 7.1.4 (Satz von Haviland)**.** *Sei $W \subseteq \mathbb{R}^n$ eine (beliebige!) abgeschlossene Menge, und $\varphi\colon \mathbb{R}[x] \to \mathbb{R}$ ein lineares Funktional. Dann sind äquivalent:*

(i) Es gibt ein Maß μ auf W mit $\varphi(p) = \int_W p \, d\mu$ für alle $p \in \mathbb{R}[x]$.

(ii) Es gilt $\varphi(p) \geq 0$ für alle $p \in \mathbb{R}[x]$ mit $p \geq 0$ auf W.

Bemerkung 7.1.5. (i) Die Richtung $(i) \Rightarrow (ii)$ im Satz von Haviland ist offensichtlich: das Integral einer nichtnegativen Funktion ist nichtnegativ. Die andere Richtung ist also die interessante. Man kann den Satz von Haviland zurückführen auf den Satz von Riesz, der dieselbe Aussage im Fall des Rings der stetigen Funktionen mit kompakten Träger auf einem lokalkompakten Hausdorffraum

liefert. Da wir den Satz von Riesz hier nicht beweisen wollen, verzichten wir auch auf diese Reduktion (man kann sie aber im Buch [4] von Marshall nachlesen).
(ii) Die Bedingung $\varphi(p) \geq 0$ für alle nichtnegativen Polynome p ist im Allgemeinen nicht einfacher zu überprüfen als die ursprüngliche Frage nach einer Maßdarstellung. Genau hier kommen nun die Ergebnisse der reellen algebraischen Geometrie ins Spiel. Wir fragen uns, wann man die Menge der auf W nichtnegativen Polynome durch eine geeignete endlich erzeugte Präordnung oder einen quadratischen Modul ersetzen kann. Dann wird die Bedingung nämlich einfacher und zum Beispiel der semidefiniten Optimierung zugänglich. Das kann funktionieren, selbst wenn nicht jedes auf W nichtnegative Polynome zur so einer Präordnung gehört.
(iii) Jedes global nichtnegative Polynom $p \in \mathbb{R}[t]$ in einer Variablen ist eine Quadratsumme (Satz 2.2.1). Der Satz von Hamburger folgt damit unmittelbar aus Havilands Satz.
(iv) Jedes auf $[0, \infty)$ nichtnegative Polynom liegt im quadratischen Modul $M(t)$ (Übungsaufgabe 31). Der Satz von Stieltjes folgt also ebenfalls unmittelbar aus dem Satz von Haviland.
(v) Jedes auf $[0, 1]$ nichtnegative Polynom liegt im quadratischen Modul $M(t, 1 - t)$ (vergleiche Übungsaufgabe 31). Der Satz von Hausdorff folgt also ebenso aus dem Satz von Haviland. △

Die bisherigen Betrachtungen rechtfertigen die folgenden Definitionen. Seien dazu immer $p_1, \ldots, p_r \in \mathbb{R}[x] = \mathbb{R}[x_1, \ldots, x_n]$. Den endlich erzeugten quadratischen Modul

$$M = M(p_1, \ldots, p_r) = \left\{\sigma_0 + \sigma_1 p_1 + \cdots + \sigma_r p_r \mid \sigma_i \in \Sigma\mathbb{R}[x]^2\right\}$$

kennen wir schon. Man beachte nochmals, dass der Fall einer endliche erzeugten Präordnung damit ebenfalls abgedeckt ist; die Präordnung ist der quadratische Modul, der von den Produkten der p_i erzeugt wird.
Wir betrachten nun den *dualen Kegel* im algebraischen Dualraum $\mathbb{R}[x]^*$

$$M^\vee = M(p_1, \ldots, p_r)^\vee = \{\varphi \colon \mathbb{R}[x] \to \mathbb{R} \text{ linear} \mid \varphi \geq 0 \text{ auf } M\}$$

und schließlich das *Doppel-Dual*

$$M^{\vee\vee} = M(p_1, \ldots, p_r)^{\vee\vee} = \{p \in \mathbb{R}[x] \mid \varphi(p) \geq 0 \,\forall \varphi \in M^\vee\}\,,$$

das wir allerdings nicht in $(\mathbb{R}[x]^*)^*$ sondern nur in $\mathbb{R}[x]$ betrachten. Die basisch abgeschlossene Menge

$$W = W_{\mathbb{R}}(p_1, \ldots, p_r) = \{a \in \mathbb{R}^n \mid p_1(a) \geq 0, \ldots, p_r(a) \geq 0\}$$

kennen wir auch bereits. Mit der Menge der auf W nichtnegativen Polynome haben wir auch schon gearbeitet. Wir nennen sie ab jetzt die *Saturierung* von M:

$$M^{\mathrm{sat}} = M(p_1, \ldots, p_r)^{\mathrm{sat}} = \{p \in \mathbb{R}[x] \mid p \geq 0 \text{ auf } W(p_1, \ldots, p_r)\}\,.$$

Ab jetzt lassen wir die Bedingung $-1 \notin M$ für quadratische Moduln weg. Dann müssen wir nicht immer eine zusätzliche Fallunterscheidung machen. Im Fall $-1 \in M$, d.h. $M = \mathbb{R}[x]$, ist ohnehin alles trivial.

Satz 7.1.6. *$M^{\vee\vee}$ ist ein quadratischer Modul, und sogar eine Präordnung, wenn M eine war. Es ist M^{sat} immer eine Präordnung. Es gilt die Inklusionskette*

$$M \subseteq M^{\vee\vee} \subseteq M^{\mathrm{sat}}.$$

Beweis. $M \subseteq M^{\vee\vee}$ ist klar. Wir zeigen nun, dass $M^{\vee\vee}$ ein quadratischer Modul ist. Für $p, q \in M^{\vee\vee}$ und $\varphi \in M^{\vee}$ ist $\varphi(p+q) = \varphi(p) + \varphi(q) \geq 0$. Also ist $M^{\vee\vee}$ abgeschlossen unter $+$. Sei nun $f \in \mathbb{R}[x]$ beliebig. Wir müssen zeigen dass auch $\varphi(f^2 p) \geq 0$ gilt. Dazu definieren wir ein neues Funktional

$$\psi\colon \mathbb{R}[x] \to \mathbb{R}; \quad g \mapsto \varphi(f^2 g).$$

Für $m \in M$ ist $\psi(m) = \varphi(f^2 m) \geq 0$, da $f^2 m \in M$. Also ist $\psi \in M^{\vee}$, und somit

$$0 \leq \psi(p) = \varphi(f^2 p).$$

Damit ist also $M^{\vee\vee}$ ein quadratischer Modul. Ganz analog zeigt man die Präordnungseigenschaft von $M^{\vee\vee}$, vorausgesetzt M hat sie. Die Menge M^{sat} aller auf W nichtnegativen Polynom ist offensichtlich eine Präordnung, siehe Bemerkung 3.1.2 (iii) (dort nannten wir sie T_W). Es bleibt noch $M^{\vee\vee} \subseteq M^{\mathrm{sat}}$ zu zeigen. Für jeden Punkt $a \in W$ ist der Auswerungshomomorphismus

$$\delta_a\colon \mathbb{R}[x] \to \mathbb{R}; \quad g \mapsto g(a)$$

aber ein lineares Funktional, das offensichtlich in $M^{\vee}$ liegt (Polynome aus M sind nichtnegativ auf W!). Damit ist für $p \in M^{\vee\vee}$

$$0 \leq \delta_a(p) = p(a),$$

d.h. $p \in M^{\mathrm{sat}}$. □

Bemerkung 7.1.7. Man kann zeigen, dass $M^{\vee\vee}$ ein topologischer Abschluss von M ist, und zwar in der feinsten lokalkonvexen Topologie auf dem Vektorraum $\mathbb{R}[x]$. In dieser Topologie ist eine Menge genau dann abgeschlossen, wenn ihr Schnitt mit jedem endlich-dimensionalen Teilraum von $\mathbb{R}[x]$ abgeschlossen ist (in der kanonischen Topologie, die man auf solchen endlich-dimensionalen Teilräumen hat). Deshalb nennen wir $M^{\vee\vee}$ auch den *Abschluss* von M. △

Definition 7.1.8. (i) Wir nennen M *abgeschlossen*, wenn $M = M^{\vee\vee}$ gilt.
(ii) Wir sagen dass M die *starke Momenteneigenschaft* (SMP, *strong moment property*) hat, falls $M^{\vee\vee} = M^{\mathrm{sat}}$ gilt.
(iii) Wir nennen M *saturiert*, wenn $M = M^{\mathrm{sat}}$ gilt. △

Der folgende Satz fasst nochmal zusammen, warum die starke Momenteneigenschaft wichtig ist:

Satz 7.1.9. *Für einen endlich erzeugen quadratischen Modul $M = M(p_1, \ldots, p_r)$ und $W = W_{\mathbb{R}}(p_1, \ldots, p_r)$ sind äquivalent:*

(i) M hat (SMP).

(ii) Die linearen Funktionale auf $\mathbb{R}[x]$ mit Maßdarstellung auf W sind genau die Elemente von $M^{\vee}$.

Beweis. $(i) \Rightarrow (ii)$: Jedes Funktional mit Maßdarstellung auf W ist immer in $M^{\vee}$. Sei umgekehrt $\varphi \in M^{\vee}$. Dann ist φ nichtnegativ auf $M^{\vee\vee} = M^{\mathrm{sat}}$, und also hat φ nach dem Satz von Haviland eine Maßdarstellung auf W.
$(ii) \Rightarrow (i)$: Sei $p \in M^{\mathrm{sat}}$ und $\varphi \in M^{\vee}$. Sei μ ein darstellendes Maß für φ auf W. Es ist also

$$\varphi(p) = \int_W p \, d\mu \geq 0,$$

und also $p \in M^{\vee\vee}$. Wir haben damit $M^{\vee\vee} = M^{\mathrm{sat}}$ gezeigt. □

Beispiel 7.1.10. In den Momentenproblemen von Hamburger, Stieltjes und Hausdorff sind die quadratischen Moduln $M(1), M(t)$ und $M(t, 1-t)$ saturiert (siehe Bemerkung 7.1.5 (iii)-(v)). Insbesondere haben sie (SMP), und bereits daraus folgen die Sätze von Hamburger, Stieltjes und Hausdorff. △

Wir erhalten aus den archimedischen Positivstellensätzen sofort das folgende starke Ergebnis. Man beachte nochmal, dass die Archimedizität eines quadratischen Moduls M die Kompaktheit der Menge $W_{\mathbb{R}}$ impliziert, und im Falle einer Präordnung sogar äquivalent dazu ist. Es handelt sich also bei folgendem Satz um eine Lösung des Momentenproblems im kompakten Fall.

Satz 7.1.11. *Seien $p_1, \dots, p_r \in \mathbb{R}[x]$ so, dass $M = M(p_1, \dots, p_r)$ archimedisch ist. Dann hat M die starke Momenteneigenschaft.*

Beweis. Sei $p \in M^{\mathrm{sat}}$. Dann ist $p+\epsilon > 0$ auf $W_{\mathbb{R}}$, für alle $\epsilon > 0$, und also $p+\epsilon \in M$, nach Satz 5.4.3. Für jedes $\varphi \in M^\vee$ ist also

$$0 \leq \varphi(p + \epsilon) = \varphi(p) + \epsilon\varphi(1),$$

und also $\varphi(p) \geq 0$. Das zeigt $p \in M^{\vee\vee}$. □

Beispiel 7.1.12. Wenn wir wissen wollen, ob $\varphi\colon \mathbb{R}[x] \to \mathbb{R}$ von einem Maß auf der Einheitskugel im $\mathbb{R}^n$ kommt, müssen wir

$$\varphi(p^2) \geq 0, \qquad \varphi(p^2 \cdot (1 - x_1^2 - \cdots - x_n^2)) \geq 0$$

für alle $p \in \mathbb{R}[x]$ überprüfen. Der quadratische Modul $M(1 - x_1^2 - \cdots - x_n^2)$ ist ja archimedisch. △

Im Fall einer nichtkompakten Menge gibt es einen weiteren Satz von Schmüdgen, der es erlaubt, bei der Frage nach (SMP) die Dimension zu veringern. Wir illustrieren die Methode nur an einem Beispiel:

Beispiel 7.1.13. Sei $p_1 = 1 - x^2 \in \mathbb{R}[x, y]$. Dann ist $W = W_{\mathbb{R}}(p_1)$ der vertikale Streifen über dem Intervall $[-1, 1]$. Bisher wissen wir nicht, ob $M = M(p_1)$ die Eigenschaft (SMP) hat oder nicht. Der Nichtnegativstellensatz liefert zwar einen Zusammenhang zwischen M^{sat} und M; dabei tauchen aber Nenner auf.
Es gibt nun ein nichtkonstantes Polynom q, das auf W beschränkt bleibt. Man kann zum Beispiel $q = x$ nehmen. Der *Fasersatz* von Schmüdgen besagt nun, dass wir auf die Fasern eines beschränkten Polynoms einschränken können, um (SMP) zu testen. Eine Faser ist dabei das Urbild einer Zahl unter q. Das Urbild von $r \in [-1, 1]$ unter $q = x$ ist die vertikale Gerade $\{x = r\}$. Darauf einzuschränken bedeutet, für x überall r einzusetzen. Das liefert also den quadratischen Modul $M(1 - r^2) = \Sigma\mathbb{R}[y]^2 \subseteq \mathbb{R}[y]$. Dieser quadratische Modul hat aber (SMP), wie wir durch Hamburgers Momentenproblem wissen. Da das für alle r, und damit für alle Fasern des beschränkten Polynoms x stimmt, hat auch M selbst die Momenteneigenschaft. Ein Funktional $\varphi\colon \mathbb{R}[x, y] \to \mathbb{R}$ besitzt also genau dann eine Maßdarstellung im senkrechten Streifen, wenn $\varphi(p^2) \geq 0$ und $\varphi(p^2 \cdot (1 - x^2)) \geq 0$ für alle $p \in \mathbb{R}[x, y]$ gilt. △

Nachdem wir nun einige positive Resultate zum Momentenproblem erhalten haben, beschäftigen wir uns im nächsten Abschnitt mit negativen Resultaten.

7.2 Stabilität

Definition 7.2.1. Für $p_1, \ldots, p_r \in \mathbb{R}[x]$ heißt der quadratische Modul $M(p_1, \ldots, p_r)$ *stabil,* falls für jedes $d \in \mathbb{N}$ ein $d' \in \mathbb{N}$ existiert mit

$$M(p_1, \ldots, p_r) \cap \mathbb{R}[x]_d \subseteq M_{d'}(p_1, \ldots, p_r). \qquad \triangle$$

Bemerkung 7.2.2. (i) Die Stabilität besagt also gerade, dass jedes Polynom $p \in M(p_1, \ldots, p_r)$ eine Darstellung besitzt, in der die Grade der Quadratsummen beschränkt werden können, und zwar nur in Abhängigkeit vom Grad von p.
(ii) Im Satz von Schmüdgen kann es solche Gradschranken nicht geben, wie wir uns in Beispiel 4.2.4 (v) schon überlegt haben. Quadratische Moduln sind also nicht immer stabil.
(iii) Stabilität hängt strenggenommen per Definition nicht vom quadratischen Modul ab, sondern von seinen Erzeugern $p_1, \ldots, p_r$. Man kann aber relativ leicht zeigen, dass sie eben doch unabhängig von den Erzeugern ist (Übungsaufgabe 52).
(iv) Der nächste Satz illustriert, dass Stabilität häufig gerade dann auftritt, wenn die semialgebraische Menge $W_{\mathbb{R}}(p_1, \ldots, p_r)$ sehr unkompakt ist. $\triangle$

Satz 7.2.3. *Wenn $W_{\mathbb{R}}(p_1, \ldots, p_r)$ einen voll-dimensionalen konvexen Kegel enthält, ist $M(p_1, \ldots, p_r)$ stabil.*

Beweis. Sei $B \subseteq \mathbb{R}^n$ eine Kugel mit nichtleerem Inneren, und

$$\mathrm{cc}(B) = \{\lambda a \mid a \in B, \lambda \geq 0\} \subseteq W_{\mathbb{R}}(p_1, \ldots, p_r) =: W.$$

Falls ein Polynom q auf W nichtnegativ ist, gilt also $0 \leq q(\lambda a)$ für alle $a \in B$ und $\lambda \geq 0$. Der homogene Summand höchsten Grades von q muss dann auf B nichtnegativ sein. Schreibt man nämlich $q = q_0 + \cdots + q_d$ als Summe homogener Summanden, ist

$$0 \leq q(\lambda a) = q_0 + \lambda q_1(a) + \cdots + \lambda^d q_d(a),$$

und ein Polynom mit negativem Leitkoeffizient wäre für große λ irgendwann negativ.
Sind zwei Polynome auf einer Menge B mit nichtleeren Inneren nichtnegativ, kann ihre Summe nicht Null sein. Es gibt ja einen Punkt, an dem beide positiv sind.

Also kann beim Addieren von zwei auf W nichtnegativen Polynomen der Grad nicht sinken, weil sich ihre Terme höchsten Grades nicht aufheben können. Wendet man das auf die Terme in einem Ausdruck

$$\sigma_0 + \sigma_1 p_1 + \cdots + \sigma_r p_r$$

an, erhält man die Stabilität sogar mit $d' = d$. □

Bemerkung 7.2.4. In Satz 7.2.3 erhält man nicht nur die starke Aussage $d' = d$, sondern sogar, dass in *jeder* Darstellung eines Elements aus $M(p_1, \ldots, p_r)$ diese Gradschranken gelten. Für die Stabilität wäre das gar nicht unbedingt nötig, man bräucht nur *eine* solche Darstellung mit Gradschranken für jedes Element. △

Beispiel 7.2.5. (i) Der quadratische Modul $M(1) = \Sigma\mathbb{R}[x]^2$ ist stabil. Das ist per Definition allerdings trivial. Da $W_{\mathbb{R}}(1) = \mathbb{R}^n$ einen volldimensionalen Kegel enthält, bekommen wir mit dem Argument aus dem letzen Satz allerdings sogar mehr, nämlich genau die Aussage von Lemma 2.2.14.
(ii) Der quadratische Modul $M(t) \subseteq \mathbb{R}[t]$ ist stabil, denn $W_{\mathbb{R}}(t) = [0, \infty) \subseteq \mathbb{R}$ enthält einen volldimensionalen Kegel. Wenn

$$p = \sigma_0 + \sigma_1 t$$

mit Quadratsummen σ_i gilt, folgt $\deg(\sigma_0), \deg(\sigma_1 t) \leq \deg(p)$.
(iii) Der quadratische Modul $M(x, y) \subseteq \mathbb{R}[x, y]$ ist stabil, denn $W_{\mathbb{R}}(x, y)$ ist der erste Orthant in $\mathbb{R}^2$, und der enthält auch einen volldimensionalen Kegel. △

Die Stabilität erlaubt es uns, in der Kette

$$M \subseteq M^{\vee\vee} \subseteq M^{\mathrm{sat}}$$

eine Gleichheit zu zeigen:

Satz 7.2.6. *Falls $M = M(p_1, \ldots, p_r)$ stabil ist und $W_{\mathbb{R}}(p_1, \ldots, p_r)$ nichtleeres Inneres in $\mathbb{R}^n$ hat, ist M abgeschlossen, d.h. es gilt*

$$M = M^{\vee\vee}.$$

Beweis. Wir zeigen, dass $M \cap U$ in U abgeschlossen ist, für jeden endlich-dimensionalen Teilraum U von $\mathbb{R}[x]$. Das ist gerade die Abgeschlossenheit in der feinsten lokalkonvexen Topologie, und das ist gleichbedeutend mit $M = M^{\vee\vee}$.

Da jeder endlich-dimensionale Teilraum U in einem $\mathbb{R}[x]_d$ enthalten ist, reicht es zu zeigen, dass $M \cap \mathbb{R}[x]_d$ abgeschlossen in $\mathbb{R}[x]_d$ ist. Aufgrund der Stabilität gilt

$$M \cap \mathbb{R}[x]_d = M_{d'} \cap \mathbb{R}[x]_d,$$

und also genügt es, die Abgeschlossenheit jedes $M_{d'}$ zu zeigen.
Das geht aber wie in Beispiel 6.4.7. Da in einer Folge von Elementen aus $M(p_1, \ldots, p_r)_{d'}$ die Grade der Quadratsummen beschränkt bleiben, kann man einen Grenzübergang machen. Man verwendet zum Beispiel nach Skalierung den Satz von Bolzano-Weierstraß koeffizientenweise. Man benötigt dass $W_{\mathbb{R}}(p_1, \ldots, p_r)$ nichtleeres Inneres hat, um zu sehen dass die Skalierung eigentlich nicht nötig ist (Übungsaufgabe 53). □

Beispiel 7.2.7. (i) Die Quadratsummen $\Sigma\mathbb{R}[x]^2$ sind stabil, und damit stets abgeschlossen.
(ii) Der quadratische Modul $M(x, y) \subseteq \mathbb{R}[x, y]$ ist abgeschlossen, d.h. es gilt

$$M(x, y) = M(x, y)^{\vee\vee} \subseteq M(x, y)^{\mathrm{sat}}. \qquad \triangle$$

Ursprünglich interessieren wir uns in diesem Kapitel ja für die starke Momenteneigenschaft (SMP). Der folgende Satz stellt den Zusammenhang zur Stabilität her:

Satz 7.2.8. *Sei $n \geq 2$ und $W_{\mathbb{R}}(p_1, \ldots, p_r)$ habe nichtleeres Inneres in $\mathbb{R}^n$. Dann schließen sich Stabilität und (SMP) für $M(p_1, \ldots, p_r)$ gegenseitig aus.*

Beweis. Wir nehmen o.B.d.A. $p_i \neq 0$ für alle i an. Es muss nun einen Punkt in $W = W_{\mathbb{R}}(p_1, \ldots, p_r)$ geben, an dem alle p_i strikt positiv sind. Sonst wäre $p_1 \cdots p_r = 0$ auf W, und damit $p_1 \cdots p_r = 0$ in $\mathbb{R}[x]$, da W nichtleeres Inneres hat. Wir nehmen nun o.B.d.A an, dass der Ursprung so ein Punkt ist, d.h. alle p_i haben einen strikt positiven konstanten Term.
Angenommen $M = M(p_1, \ldots, p_r)$ ist stabil und hat (SMP). Nach Satz 7.2.6 gilt dann sogar $M = M^{\mathrm{sat}}$, d.h. M ist saturiert.
Sei nun $p \in \mathbb{R}[x]$ ein global nichtnegatives Polynom, das keine Quadratsumme ist. Für $n \geq 2$ gibt es das immer, man kann zum Beispiel das Motzkinpolynom nehmen. Für $\lambda > 0$ setzen wir

$$p_\lambda := p(\lambda x) = p(\lambda x_1, \ldots, \lambda x_n) \in \mathbb{R}[x].$$

Dann ist jedes p_λ global nichtnegativ, liegt also insbesondere in M^{sat} und damit in M. Da der Grad von p_λ sich mit λ nicht ändert, liegen aufgrund der Stabilität alle p_λ aber in einem festen M_d. Wir ersetzen in allen diesen Darstellungen wieder x durch $\frac{1}{\lambda}x$ und erhalten so Darstellungen

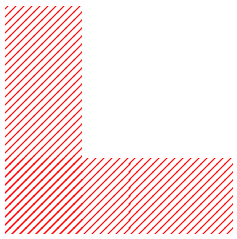
$$p = \sigma_0^{(\lambda)} + \sigma_1^{(\lambda)} p_1(\lambda^{-1}x) + \cdots + \sigma_r^{(\lambda)} p_r(\lambda^{-1}x)$$

mit Quadratsummen $\sigma_i^{(\lambda)}$ von beschränktem Grad. Wie in Beispiel 6.4.7 und Satz 7.2.6 bilden wir den Grenzwert für $\lambda \to \infty$, nach Skalierung um alle auftretenden Koeffizienten beschränkt zu halten, und mit dem Satz von Bolzano-Weierstraß. Die $p_i(\lambda^{-1}x)$ konvergieren dabei gegen $p_i(0) > 0$. Falls eine Skalierung also wirklich notwendig war, entsteht links 0 und rechts eine nichttriviale Quadratsumme. Das ist ein Widerspruch. Falls keine Skalierung nötig war, entsteht rechts eine Quadratsumme, und links steht immer noch p. Auch das ist ein Widerspruch. □

Beispiel 7.2.9. (i) Die Quadratsummen $\Sigma\mathbb{R}[x]^2$ sind stabil, abgeschlossen, und haben ab $n \geq 2$ nicht die Momenteneigenschaft. Es gibt ab $n \geq 2$ also Funktionale $\varphi\colon \mathbb{R}[x] \to \mathbb{R}$, die auf den Quadratsummen nichtnegativ sind, aber keine Maßdarstellung haben.
(ii) Im Fall $n = 1$ ist $\Sigma\mathbb{R}[t]^2$ stabil, und hat (SMP). Wir haben ja schon gesehen, dass jedes global nichtnegative Polynom in einer Variablen eine Quadratsumme ist. $\Sigma\mathbb{R}[t]^2$ ist also sogar saturiert. Die Bedingung $n \geq 2$ kann also in Satz 7.2.8 nicht weggelassen werden.
(iii) $M(x, y) \subseteq \mathbb{R}[x, y]$ ist stabil, also abgeschlossen und hat nicht (SMP).
(iv) Wir sehen hier nochmal, dass archimedische quadratische Moduln ab Dimension 2 niemals stabil sind (vergleiche Bemerkung 7.2.2 (ii)). Sie haben nach Satz 7.1.11 ja die starke Momenteneigenschaft. △

Wir haben das folgende Phänomen beobachtet, zumindest ab Dimension 2: ist die Menge W relativ klein (z.B. kompakt), so hat M tendenziell (SMP) und ist nicht stabil. Ist W dagegen groß (z.B. wenn es einen voll-dimensionalen Kegel enthält), so ist M tendenziell stabil und hat nicht (SMP). Eine gute und exakte Definition von *groß* ist beispielsweise, dass es auf W keine nichtkonstanten Polynome gibt, die als Funktion beschränkt bleiben. Falls W zum Beispiel einen voll-dimensionalen Kegel enthält, ist das erfüllt. Falls W kompakt ist, ist es offensichtlich nicht erfüllt. Allerdings kann man die Aussage über Stabilität für solche großen Mengen bisher nicht in voller Allgemeinheit beweisen.

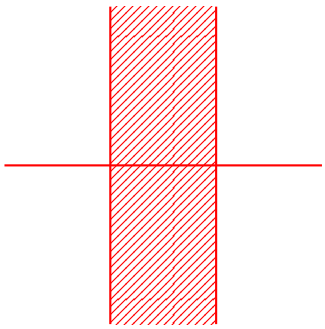
Beispiel 7.2.10. (i) Enthalte $W \subseteq \mathbb{R}^2$ zum Beispiel voll-dimensionale Streifen in Richtung beider Koordinatenachsen:



Auch dann gibt es auf W keine nichttrivialen beschränkten Polynome. Aufgrund des senkrechten Streifens kann ein beschränktes Polynom nämlich die eine Variable nicht enthalten, und aufgrund des waagerechten Streifens nicht die andere. Nun kann man sich überlegen, dass beim Addieren zweier auf W nichtnegativer Polynome der Grad zwar sinken kann (anders als in Satz 7.2.3), allerdings nur um maximal die Hälfte (Übungsaufgabe 54). Wendet man das auf die Summanden von Elementen aus $M(p_1, \ldots, p_r)$ an, erhält man

$$M \cap \mathbb{R}[x]_d \subseteq M_{2d},$$

und damit ist M stabil.
(ii) Man betrachte den quadratischen Modul $M((1-x^2)y^2) \subseteq \mathbb{R}[x, y]$. Die Menge $W_{\mathbb{R}}((1-x^2)y^2)$ ist ein vertikaler Streifen, zusammen mit der x-Achse:



Wieder gibt es keine nichttrivialen beschränkten Polynome. Der senkrechte Streifen impliziert nämlich, dass y in einem beschränkten Polynom nicht auftritt. Ein Polynom in x bleibt aber auf der x-Achse nicht beschränkt, es sei denn es ist konstant.
Wäre nun $M((1-x^2)y^2)$ stabil, so auch $M(1-x^2)$ (Übungsaufgabe 55). Dieser quadratische Modul hat aber (SMP), wie wir in Beispiel 7.1.13 gesehen haben, und

kann damit nicht stabil sein. Hier haben wir also ein Beispiel, das etwas aus der Reihe fällt. Allerdings ist die Menge nicht überall voll-dimensional, was oft zu seltsamem Verhalten führt. △

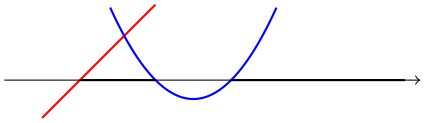
7.3 Saturiertheit

In diesem Abschnitt wollen wir uns noch genauer mit der Saturiertheit eines quadratischen Moduls befassen. Ein Modul $M = M(p_1, \dots, p_r) \subseteq \mathbb{R}[x]$ ist saturiert, wenn $M = M^{\text{sat}}$ gilt, also jedes auf $W = W_{\mathbb{R}}(p_1, \dots, p_r)$ nichtnegative Polynom zu M gehört. Saturiertheit bedeutet also, dass M abgeschlossen ist und gleichzeitig (SMP) hat, also die betrachtete Inklusionskette eine Identitätskette ist:

$$M = M^{\vee\vee} = M^{\text{sat}}.$$

Diese Frage haben wir außer für Beispiele in Dimension 1 bisher noch überhaupt nicht behandelt. Die archimedischen Positivstellensätze liefern ja immer nur Aussage über *strikt positive* Polynome.
Wir beschäftigen uns zunächst nochmal mit Dimension 1. Hier klappt alles sehr schön, siehe Satz 7.3.2. Dann zeigen wir, dass ab Dimension 3 *keine* saturierten quadratischen Moduln mehr exisitieren (Satz 7.3.4). Mit Dimension 2 beschäftigen wir uns dann erst im nächsten Kapitel. Man beachte dass wir die Abgeschlossenheit $M = M^{\vee\vee}$ bisher nur über die Stabilität zeigen können, mit Satz 7.2.6. Ab Dimension 2 schließt sich das mit (SMP) aus, und wir können nicht erwarten, so die Saturiertheit zeigen zu können.
Sei zunächst $n = 1$, d.h. wir betrachten Polynome in einer Variablen t und semialgebraische Teilmengen von $\mathbb{R}$. Die basisch-abgeschlossenen Teilmengen sind dabei nach Satz 1.5.5 gerade endliche Vereinigungen von abgeschlossenen Intervallen $(-\infty, a], [a, b], [b, \infty)$. Jede solche basisch-abgeschlossene Menge W hat sogenannte *natürliche Erzeuger*. Das sind die Polynome, die einem offensichtlich als erstes einfallen, um die Menge zu definieren. Definert sind sie folgendermaßen. Falls W ein kleinstes Element a besitzt, sei $t - a$ einer der natürlichen Erzeuger. Falls W ein größtes Element b besitzt, gehöre $b - t$ zu den natürlichen Erzeugern. Falls $a, b \in W, a < b$ und $W \cap (a, b) = \emptyset$, so sei weiter $(t-a)(t-b)$ ein natürlicher Erzeuger. Auf diese Weise erhalten wir Polynome $p_1, \dots, p_r \in \mathbb{R}[t]$, mit $W = W_{\mathbb{R}}(p_1, \dots, p_r)$.

Beispiel 7.3.1. Die Menge $W = [-1, 0] \cup [1, \infty)$ hat die natürlichen Erzeuger $t+1$ und $t(t-1)$:



△

Satz 7.3.2. *Seien $p_1, \dots, p_r \in \mathbb{R}[t]$ die natürlichen Erzeuger für $W \subseteq \mathbb{R}$. Dann ist die Präordnung $T = T(p_1, \dots, p_r)$ saturiert und stabil.*

Beweis. Sei $p \geq 0$ auf W. Wir faktorisieren p in irreduzible Faktoren wie in Satz 1.2.11. Wir zeigen, dass jeweils die einzelnen Faktoren oder Produkte aus zwei von ihnen zu T gehören. Da T als Präordnung multiplikativ abgeschlossen ist, zeigt das die Saturiertheit.
Faktoren vom Grad 2 sind global strikt positive Quadratsummen und gehören also zu T. Ohne diese Faktoren ist das Polynom immer noch nichtnegativ auf W. Falls in der Faktorisierung ein Linearfaktor $t-c$ mit einem $c \in \operatorname{int}(W)$ auftaucht, muss er in gerade Vielfachheit auftreten. Sonst wäre p auf W irgendwo negativ. Auch diese Faktoren können wir also vernachlässigen.
Falls W ein kleinstes Element a besitzt, und in p ein Faktor $t - a'$ für ein $a' \leq a$ auftritt, liegt dieser in T, da $t - a' = (t - a) + (a - a')$ und $t - a$ als natürlicher Erzeuger in T liegt. Auch solche Faktoren können wir also vernachlässigen. Den Fall von einem größten Element erledigt man genauso.
Es kann nun noch passieren, dass p eine Nullstelle c in einem Intervall $[a, b]$ hat, mit $a, b \in W, a < b$ und $W \cap (a, b) = \emptyset$. Dann muss es aber im selben Intervall noch eine Nullstelle d geben, sonst wäre p nicht nichtnegativ auf W. Also enthält p einen Faktor $(t - c)(t - d)$ mit $a \leq c \leq d \leq b$. Wir zeigen

$$(t - c)(t - d) \in T((t - a)(t - b))$$

und sind dann fertig, da $(t-a)(t-b) \in T$. Man überlegt sich nun, dass man ein $\lambda \in (0, 1]$ finden kann, so dass

$$(t - c)(t - d) - \lambda(t - a)(t - b)$$

global nichtnegativ ist (Übungsaufgabe 57). Damit ist es eine Quadratsumme, und das zeigt die Behauptung.
Die Stabilität haben wir gleich mitbewiesen. In jedem Schritt hatten die jeweiligen Faktoren Darstellungen mit den bestmöglichen Gradschranken an die Quadratsummen. Somit gilt das auch für das Produkt und daraus folgt

$$T \cap \mathbb{R}[t]_d \subseteq T_d.$$

□

Bemerkung 7.3.3. Ohne die natürlichen Erzeuger ist die Saturiertheit im Allgemeinen nicht gegeben. Es ist zum Beispiel $t \notin T(t^3)$. △

Wir kommen nun wie angekündigt zur Dimension ≥ 3. Der folgende Satz zeigt, dass $M(p_1, \ldots, p_r)$ dort niemals saturiert ist, bzw. dass M^{sat} niemals endlich erzeugt als quadratischer Modul sein kann.

Satz 7.3.4. *Sei $n \geq 3$ und $p_1, \ldots, p_r \in \mathbb{R}[x]$ so, dass $W_{\mathbb{R}}(p_1, \ldots, p_r)$ in $\mathbb{R}^n$ nichtleeres Inneres hat. Dann ist $M(p_1, \ldots, p_r)$ nicht saturiert.*

Beweis. Es muss wieder einen Punkt in $W_{\mathbb{R}}(p_1, \ldots, p_r)$ geben, an dem alle p_i strikt positiv sind; o.B.d.A. sei das der Ursprung, d.h. alle p_i haben einen strikt positiven konstanten Term.
Sei nun $h \in \mathbb{R}[x]$ in *homogenes* und global nichtnegatives Polynom, das keine Quadratsumme ist. Für $n \geq 3$ gibt es das immer, man kann zum Beispiel das homogenisierte Motzkinpolynom nehmen. Offensichtlich gilt $h \in M^{\mathrm{sat}}$. Angenommen $h \in M(p_1, \ldots, p_r)$, d.h. es gibt eine Darstellung

$$h = \sigma_0 + \sigma_1 p_1 + \cdots + \sigma_r p_r \tag{7.1}$$

mit Quadratsummen σ_i. In jedem der Terme $\sigma_i p_i$ ist der homogene Summand von niedrigstem Grad eine Quadratsumme. Dabei verwenden wir, dass alle p_i strikt positiven konstanten Term haben. Der homogene Summand von niedrigstem Grad auf der rechten Seite in (7.1) ist also eine nichttriviale Quadratsumme. Da h homogen ist, muss er mit h übereinstimmen, und das kann nicht sein. □

Bemerkung 7.3.5. Im Beweis von Satz 7.3.4 sieht es so aus, als wäre immer dasselbe Polynom h nicht in $M(p_1, \ldots, p_r)$. Dabei könnte man es ja einfach zu den Erzeugern dazunehmen. Wir haben h aber homogen in Bezug auf den Ursprung gewählt, und angenommen, dass alle p_i dort positiv sind. Man sieht aber: es gibt ein festes Polynom h (zum Beispiel das homogenisiert Motzkinpolynom), von dem immer eine geeignete Verschiebung nicht in M liegt. △

Nur in Dimension 2 können also noch interessante Fälle von Saturiertheit auftreten. Allzu einfach kann es dort aber nicht sein sie zu zeigen, da wir nicht gleichzeitig (SMP) und Stabilität erwarten können. Das Lokal-Global-Prinzip von Scheiderer erlaubt es uns, trotzdem positive Ergebnisse zu bekommen. Wir wollen uns im nächsten Kapitel ausführlicher damit beschäftigen.

Kapitel 8

Ein Lokal-Global-Prinzip von Scheiderer

In diesem Kapitel wollen wir einen Satz beweisen, der uns für gewisse semialgebraische Mengen in der Ebene erlaubt, die Saturiertheit einer dazugehörigen Präordnung zu zeigen. Dabei zeigt man die Saturiertheit zunächst lokal, und erhält sie dann für die ursprüngliche Präordnung.

8.1 Zwei Lemmas

Wie immer sei A ein kommutativer Ring mit $\mathbb{Q} \subseteq A$.

Lemma 8.1.1. *Sei X ein kompakter Hausdorffraum und*

$$\Phi\colon A \to C(X, \mathbb{R})$$

ein Ringhomomorphismus, für den $\Phi(A)$ die Punkte von X trennt (d.h. für $x \neq y$ aus X gibt es ein $a \in A$ mit $\Phi(a)(x) \neq \Phi(a)(y)$). Seien $a, b \in A$ mit $\Phi(a), \Phi(b) \geq 0$ auf X. Falls es eine Gleichung

$$1 = ra + sb$$

in A gibt, gibt es auch eine solche Gleichung mit $\Phi(r), \Phi(s) > 0$ auf X.

Beweis. Zur Vereinfachung der Notation lassen wir Φ im Beweis einfach weg, schreiben also $a \geq 0$ auf X statt $\Phi(a) \geq 0$ etc... Nach dem Satz von Stone-Weierstrass liegt A dicht in $C(X, \mathbb{R})$ bezüglich der Supremumsnorm (hierfür

verwenden wir dass A die Punkte von X trennt). Sei nun zunächst eine Gleichung $1 = fa + gb$ mit $f, g \in A$ beliebig gegeben.
1. Behauptung: Es gibt $N_1, N_2 \geq 1$ mit $N_1 b > -f$ und $N_2 a > -g$ auf X.
Sei $Y := \{x \in X \mid f(x) \leq 0\}$. Dann ist Y eine abgeschlossene und damit kompakte Teilmenge von X. Auf Y gilt $gb = 1 - fa \geq 1$, und also muss b auf Y strikt positiv sein. Wähle ein $n \in \mathbb{N}$ mit $b > \frac{1}{n}$ und $-f < n$ auf Y. Dann gilt $n^2 b > -f$ auf Y und $n^2 b \geq 0 > -f$ auf $X \setminus Y$. Da also $n^2 b > -f$ überall gilt, kann man $N_1 = n^2$ nehmen. Die Existenz von N_2 folgt aus Symmetriegründen.
Wir definieren nun zwei Funktionen $\varphi, \psi \colon X \to \mathbb{R}$ durch die Vorschrift

$$\varphi := \max\left\{-N_1, -\frac{g}{a}\right\}, \quad \psi := \min\left\{N_2, \frac{f}{b}\right\}.$$

Dabei treffen wir die Konvention $\frac{r}{0} = \infty$ für alle $r \in \mathbb{R}$.
2. Behauptung: Die Funktionen φ und ψ sind stetig auf X.
Wir müssen die Behauptung nur für φ zeigen, für ψ folgt sie aus Symmetriegründen. Falls $a(x) \neq 0$ an einem Punkt gilt, ist φ in einer Umgebung von x das Maximum zweier stetiger Funktionen, und also selbst stetig an x. Sei also $a(x) = 0$. Wir zeigen, dass dann $\varphi = -N_1$ in einer Umgebung von x gilt, und damit ist φ dann auch stetig an x. Sei dazu $k \in \mathbb{N}$ mit $k > b$ und $k > f$ auf X. Wähle dann y so nah bei x, dass $a(y) < \frac{1}{k(N_1+1)}$ gilt. Wir können zusätzlich annehmen, dass $a(y) \neq 0$ ist, sonst wären wir ja schon fertig. Also ist $a(y) > 0$. Wir erhalten

$$g(y)b(y) = 1 - f(y)a(y) > 1 - \frac{k}{k(N_1+1)} > 0,$$

also ist $g(y), b(y) > 0$. Es ist nun

$$\frac{g(y)}{a(y)} > \frac{g(y)b(y)}{ka(y)} = \frac{1 - f(y)a(y)}{ka(y)} > \frac{1 - ka(y)}{ka(y)} = \frac{1}{ka(y)} - 1 > N_1.$$

Somit gilt in einer Umgebung von x schon $\varphi = -N_1$.
3. Behauptung: Es gilt $\varphi < \psi$ auf X.
Auf ganz X gelten die folgenden Ungleichungen:

$$-N_1 < N_2, \; -N_1 < \frac{f}{b}, \; -\frac{g}{a} < N_2.$$

Die erste Ungleichung ist klar, die zweite und die dritte folgen aus der ersten Behauptung. Es bleibt noch zu zeigen, dass $-\frac{g}{a} < \frac{f}{b}$ gilt an jedem Punkt x mit $a(x), b(x) \neq 0$ gilt. Aufgrund der Positivität von a und b folgt das aus

$$f(x)a(x) + g(x)b(x) = 1 > 0.$$

Wir verwenden nun den Satz von Stone-Weierstrass und wählen ein $c \in A$ mit

$$\varphi < c < \psi$$

auf X. Damit definieren wir

$$r = f - cb,\ s = g + ca.$$

Damit gilt $ra + sb = (f - cb)a + (g + ca)b = fa + gb = 1$ wie gewünscht. Die Aussage folgt also aus der
4. Behauptung: $r, s > 0$ auf X.
Ist $b(x) = 0$, so gilt $1 = f(x)a(x)$, also $r(x) = f(x) > 0$. Andernfalls ist

$$r(x) = f(x) - c(x)b(x) > f(x) - \psi(x)b(x) \geq f(x) - \frac{f(x)}{b(x)}b(x) = 0.$$

Der Beweis für s geht analog. □

Lemma 8.1.2. *Seien $p_1, \dots p_r \in \mathbb{R}[x]$ mit $M = M(p_1, \dots, p_r)$ archimedisch. Sei $p \in \mathbb{R}[x]$ und $p \geq 0$ auf $W = W_{\mathbb{R}}(p_1, \dots, p_r)$. Dann gilt*

$$p \in M \Leftrightarrow p \in M + (p^2).$$

Beweis. Man beachte dass $M + (p^2)$ selbst wieder ein endlich erzeuger quadratischer Modul ist. Mit dem üblichen Trick aus Bemerkung 1.1.11 (ii) sieht man nämlich $M + (p^2) = M(p_1, \dots, p_r, -p^2)$. Die dazugehörige semialgebraische Menge ist also gerade $W \cap V_{\mathbb{R}}(p)$.
"⇒" ist trivial. Für "⇐" habe p eine Darstellung $p = m - \sigma p^2$ mit $m \in M, \sigma \in \Sigma\mathbb{R}[x]^2$. Also gilt $p(1 + \sigma p) \in M$. Wegen $p, 1 + \sigma p \geq 0$ auf W und $1 = -\sigma \cdot p + 1 \cdot (1 + \sigma p)$ können wir Lemma 8.1.1 an wenden. Dabei fassen wir Polynome als stetige Funktionen auf $X = W$ auf. Wir erhalten $r, s \in \mathbb{R}[x]$ mit $r, s > 0$ auf W und

$$1 = rp + s(1 + \sigma p). \tag{8.1}$$

Nach Satz 5.4.3 gilt $r, s, rs \in M$. Multiplizieren wir (8.1) mit sp ergibt sich

$$sp = rsp^2 + s^2p(1 + \sigma p) \in M,$$

und bei Multiplikation mit p damit

$$p = rp^2 + sp(1 + \sigma p) = rp^2 + sp + s\sigma p^2 \in M.$$ □

8.2 Komplettierungen von Ringen

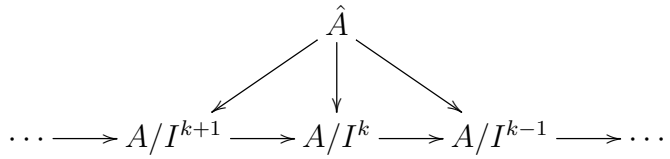
Wir wiederholen einige Fakten aus der kommutativen Algebra. Die Beweise der Aussagen sind dabei Übungsaufgabe 58. Sei wieder A ein kommutativer Ring mit 1 und $I \subsetneq A$ ein echtes Ideal. Die *Komplettierung von A bezüglich I* ist wieder ein Ring, der wie folgt definiert ist. Sei

$$I^k = \left\{ \sum_{\text{endl.}} i_1 \cdots i_k \mid i_j \in I \right\}$$

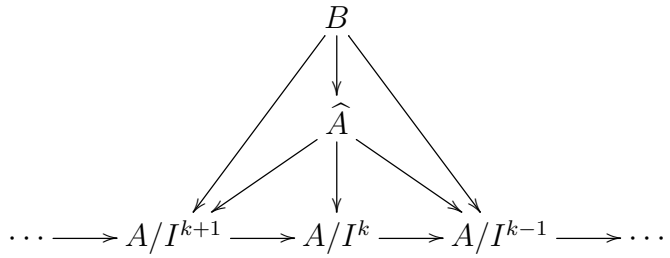
die k-te Potenz von I. Die I^k bilden eine absteigende Kette von Idealen, und deshalb gibt es kanonische surjektive Homomorphismen

$$\cdots \to A/I^{k+1} \to A/I^k \to A/I^{k-1} \to \cdots \to A/I.$$

Die Komplettierung $\widehat{A}$ ist nun der sogenannte *inverse Limes* dieser Kette. Mithilfe einer universellen Eigenschaft formuliert man das so. Es ist $\widehat{A}$ ein Ring, zusammen mit Homomorphismen $\widehat{A} \to A/I^k$ für alle k, so dass die Dreiecke



alle kommutieren, und der universell bezüglich dieser Eigenschaft ist. Das heißt, falls für einen Ring B ebenfalls solche Homomorphismen $B \to A/I^k$ existieren, gibt es einen eindeutig bestimmten Homomorphismus $B \to \widehat{A}$, so dass wieder alles kommutiert, was kommutieren kann:



Durch diese Eigenschaft ist der inverse Limes bis auf eindeutigen Isomorphismus eindeutig bestimmt. Da für A selbst die kanonischen Homomorphismen

$A \to A/I^k$ existieren, gibt es also auch einen eindeutigen Homomorphismus $A \to \widehat{A}$ wie oben. Um zu sehen ob solch ein inverser Limes überhaupt existiert, kann man ihn explizit konstruieren. Dazu betrachtet man zunächst das direkte Produkt von Ringen

$$\prod_{k\geq 1} A/I^k$$

und darin den Unterring

$$\widehat{A} = \left\{(a_k + I^k)_k \mid a_{k+1} \equiv a_k \mod I^k \, \forall k \geq 1\right\}.$$

Die Abbildungen $\widehat{A} \to A/I^k$ sind dann einfach die Projektionen auf die k-ten Einträge der Tupel. Diese Konstruktion erfüllt die universelle Eigenschaft des inversen Limes. Der Homomorphismus $A \to \widehat{A}$ ist gegeben durch

$$a \mapsto (a + I, a + I^2, a + I^3, \ldots).$$

Falls $\cap I^k = (0)$ gilt, ist es eine Einbettung.
Wir verwenden nun diese Konstruktion für den Ring $A = \mathbb{R}[x]$ und das maximale Ideal

$$I = \mathfrak{m}_a = (x_1 - a_1, \ldots, x_n - a_n) = \{p \in \mathbb{R}[x] \mid p(a) = 0\}$$

für $a \in \mathbb{R}^n$. Wir bezeichnen die Komplettierung dann auch mit $\widehat{\mathbb{R}[x]}_a$. Es gibt dabei eine kanonische Identifikation von $\widehat{\mathbb{R}[x]}_a$ mit dem Potenzreihenring $\mathbb{R}[[x_1 - a_1, \ldots, x_n - a_n]]$.

8.3 Das Lokal-Global-Prinzip

Der folgende Satz kann als Verallgemeinerung des Satzes von Schmüdgen aufgefasst werden, denn er erlaubt Nullstellen von Polynomen. Wenn ein Polynom an allen seinen Nullstellen "lokal" zur Präordnung gehört, gehört es auch global dazu. Lokal ist die Bedingung oft einfacher zu überprüfen, wie wir dann im nächsten Abschnitt sehen werden.

Satz 8.3.1 (Lokal-Global-Prinzip von Scheiderer). *Seien $p_1, \ldots, p_r \in \mathbb{R}[x]$ mit $W = W_{\mathbb{R}}(p_1, \ldots, p_r)$ beschränkt. Sei $p \in \mathbb{R}[x]$ mit $p \geq 0$ auf W, und p habe nur endlich viele Nullstellen in W. Dann sind äquivalent:*

(i) $p \in T = T(p_1, \ldots, p_r)$

(ii) *Für jede Nullstelle a von p in W liegt p in der von $p_1, \ldots, p_r$ erzeugten Präordnung in $\widehat{\mathbb{R}[x]}_a$.*

Beweis. "$(i) \Rightarrow (ii)$" ist klar. Für "$(ii) \Rightarrow (i)$" betrachten wir zusätzlich die Präordnung $T' := T + (p^2)$ und das Ideal $J = \operatorname{supp}(T') = T' \cap -T'$. Die zu T' gehörige semialgebraische Menge ist die endliche Menge $W' = W \cap V_{\mathbb{R}}(p)$.
1. Behauptung: Jedes Primideal von $\mathbb{R}[x]$, welches J enthält, ist von der Gestalt $\mathfrak{m}_a$ für ein $a \in W'$.
Sei dazu $\mathfrak{p}$ ein solches Primideal. Da W' endlich ist, können wir das Ideal $\mathfrak{q} := \prod_{a \in W'} \mathfrak{m}_a$ betrachten. Für jedes $q \in \mathfrak{q}$ gilt dann $q = 0$ auf W', und aus dem konkreten Nichtnegativstellensatz 3.3.3 folgt $q^{2m} \in -T'$ für ein $m \geq 0$ (Übungsaufgabe 60). Somit ist $q^{2m} \in T' \cap -T' = J \subseteq \mathfrak{p}$, und aus der Primidealeigenschaft folgt $q \in \mathfrak{p}$. Wir haben also $\mathfrak{q} \subseteq \mathfrak{p}$ gezeigt, und aus der Primidealeigenschaft folgt dann $\mathfrak{m}_a \subseteq \mathfrak{p}$ für ein $a \in W'$. Aus der Maximalität von $\mathfrak{m}_a$ folgt Gleichheit.
Aus der Behauptung folgt, dass $\mathbb{R}[x]/J$ ein 0-dimensionaler (und noetherscher) Ring ist, und damit artinsch. Daraus wiederum folgt, dass der Schnitt über eine Potenz seiner maximalen Ideale das Nullideal ergibt. Für $\mathbb{R}[x]$ folgt daraus

$$\bigcap_{a \in W'} \mathfrak{m}_a^k \subseteq J$$

für ein festes k.
Wir setzen nun voraus, dass p für jedes $a \in W'$ in der von den p_i in $\widehat{\mathbb{R}[x]}_a$ erzeugten Präordnung liegt. Wenn wir den kanonischen Homomorphismus $\widehat{\mathbb{R}[x]}_a \to \mathbb{R}[x]/{\mathfrak{m}_a}^k$ anwenden, erhalten wir eine Darstellung

$$p \equiv \sum_{e \in \{0,1\}^r} \sigma_e^{(a)} p_1^{e_1} \cdots p_r^{e_r} \mod \mathfrak{m}_a^k$$

mit Quadratsummen $\sigma_e^{(a)}$, für alle $a \in W'$. Da mit den maximalen Idealen $\mathfrak{m}_a$ auch ihre Potenzen $\mathfrak{m}_a^k$ paarweise teilferfremd sind, gibt es nach dem Chinesischen Restsatz Elemente ω_a in $\mathbb{R}[x]$ mit $\omega_a \equiv 1 \mod \mathfrak{m}_a^k$ und $\omega_a \equiv 0 \mod \mathfrak{m}_b^k$ für $b \neq a$. Wir setzen

$$t := \sum_{b \in W'} \sum_e \omega_b^2 \sigma_e^{(b)} p_1^{e_1} \cdots p_r^{e_r} \in T.$$

Modulo jedem $\mathfrak{m}_a^k$ gilt nun

$$t \equiv \sum_e \sigma_e^{(a)} p_1^{e_1} \cdots p_r^{e_r} \equiv p,$$

also

$$p - t \in \bigcap_{a \in W'} \mathfrak{m}_a^k \subseteq J,$$

also

$$p \in T + J \subseteq T' = T + (p^2).$$

Nach Satz 4.2.2 folgt aus der Beschränktheit von W die Archimedizität von T. Also können wir Lemma 8.1.2 verwenden, und erhalten $p \in T$. □

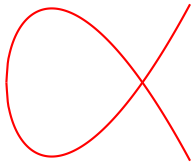
8.4 Anwendungen in Dimension 2

Wir wollen das Lokal-Global-Prinzip nun in Dimension 2 anwenden. Dafür zeigen wir zunächst, dass die Annahme nur *endlich vieler Nullstellen von* p keine wirkliche Einschränkung darstellt.
Sei dazu $p \in \mathbb{R}[x, y]$ und $V_{\mathbb{R}}(p) \subseteq \mathbb{R}^2$ die von p definierte Varietät (auch Kurve genannt). Ein Punkt $a \in V_{\mathbb{R}}(p)$ heißt *singuläre* Nullstelle von p, falls für beide partielle Ableitungen gilt

$$(\partial_x p)(a) = 0 \text{ und } (\partial_y p)(a) = 0.$$

Ansonsten heißt a *reguläre* oder *glatte* Nullstelle. Der Gradientenvektor $((\partial_x p)(a), (\partial_y p)(a))$ definiert dann die *Normalenrichtung* von $V_{\mathbb{R}}(p)$ am Punkt a, und ein dazu orthogonaler Vektor definiert die *Tangentenrichtung*.

Beispiel 8.4.1. (*i*) Sei $p = 1 - x^2 - y^2$. Dann ist $V_{\mathbb{R}}(p)$ der Einheitskreis. Für $a \in V_{\mathbb{R}}(p)$ ist $a_1^2 + a_2^2 = 1$, und also entweder $a_1 \neq 0$ oder $a_2 \neq 0$. Wegen $\partial_x p = -2x$ und $\partial_y p = -2y$ ist jede Nullstelle glatt. Die Normalenrichtung am Punkt $a = (1, 0)$ ist $(-2, 0)$ und die Tangentenrichtung $(0, 1)$.
(*ii*) Sei $p = y^2 - x^2(x + 1)$. Dann ist $V_{\mathbb{R}}(p)$ eine Schleifenkurve:



Es ist $\partial_x p = -3x^2 - 2x$ und $\partial_y p = 2y$. Der Punkt $a = (0,0)$ ist also eine singuläre Nullstelle. △

Lemma 8.4.2. *(i) Seien $p, q \in \mathbb{R}[x,y]$, p irreduzibel und $p \nmid q$. Dann haben p und q nur endlich viele gemeinsame Nullstellen.*
(ii) Jedes irreduzible $p \in \mathbb{R}[x,y]$ hat nur endlich viele singuläre Nullstellen.
(iii) Ist $a \in \mathbb{R}^2$ eine reguläre Nullstelle von $p \in \mathbb{R}[x,y]$, so nimmt p in jeder Umgebung von a sowohl positive als auch negative Werte an.

Beweis. (i) Wir betrachten $p \in \mathbb{R}(x)[y]$ als Polynom in y über dem Körper $\mathbb{R}(x)$. Dann ist p entweder invertierbar (wenn es Grad 0 in y hat), oder es ist nach dem Lemma von Gauß auch irreduzibel in $\mathbb{R}(x)[y]$, und teilt q auch dort nicht. In beiden Fällen ist $(p,q) = 1$ in $\mathbb{R}(x)[y]$, denn nichttriviale Primideale sind maximal in Polynomringen über Körpern. Es gibt also eine Gleichung

$$1 = fp + gq$$

mit $f, g \in \mathbb{R}(x)[y]$. Nach Multiplikation mit dem Hauptnenner ergibt sich

$$b = cp + dq$$

mit $0 \neq b \in \mathbb{R}[x], c, d \in \mathbb{R}[x,y]$. Für jede gemeinsame Nullstelle $a = (a_1, a_2)$ von p und q ist also a_1 eine Nullstelle von b, und dafür gibt es nur endlich viele Möglichkeiten. Aus Symmetriegründen gilt dasselbe für a_2.
(ii) Es ist o.B.d.A. $\partial_x p \neq 0$, also $p \nmid \partial_x p$ aus Gradgründen. Eine singuläre Nullstelle ist eine gemeinsame Nullstelle von $p, \partial_x p$. Die Aussage folgt also aus (i).
(iii) Ist o.B.d.A. $a = (0,0)$ eine reguläre Nullstelle von p, so hat p keinen konstanten Term, und der lineare Term von p ist $p_1 = (\partial_x p)(a) \cdot x + (\partial_y p)(a) \cdot y$. In einer kleinen Umgebung der Null bestimmt aber der lineare Teil von p das Verhalten, und es ist

$$p(\lambda \cdot (\partial_x p(a), \partial_y p(a)) = \lambda \cdot (\underbrace{\partial_x p(a)^2 + \partial_y p(a)^2}_{>0}) + \text{ Terme höherer Ordnung.}$$

In Normalenrichtung ist p also nahe a positiv, in Gegenrichtung negativ. □

Ab jetzt stellen wir an unsere semialgebraischen Mengen eine gewisse (nicht sehr starke) Regularitätsbedingung. Wir betrachten nur noch irreduzible Polynome $p_1, \ldots, p_r \in \mathbb{R}[x,y]$ für die $W = W_{\mathbb{R}}(p_1, \ldots, p_r)$ die Bedingung

$$W = \overline{\mathrm{int(W)}}$$

erfüllt. W muss der Abschluss seines Inneren sein, darf also keine eindimensionalen Stücke enthalten. Im nächsten Lemma verwenden wir, dass $\mathbb{R}[x,y]$ ein faktorieller Ring ist, also eindeutige Primfaktorzerlegungen erlaubt.

Lemma 8.4.3. *Schreibe $0 \neq p \in \mathbb{R}[x,y]$ also Produkt*

$$p = q \cdot \bar{p} \cdot h^2,$$

wobei $\bar{p}$ ein Produkt der p_i und q quadratfrei ist, mit $p_i \nmid q$ für alle i. Dann gilt

$$p \geq 0 \text{ auf } W \Leftrightarrow q \geq 0 \text{ auf } W.$$

Beweis. "$\Leftarrow$" ist klar. Für "$\Rightarrow$" sei angenommen q nicht nichtnegativ auf W. Dann gibt es eine offene Kreisscheibe $U \subseteq W$ mit $q < 0$ auf U. Hier verwenden wir $W = \overline{\mathrm{int}(W)}$. Es gibt nun einen Punkt $a \in U$ mit $(\bar{p}h^2)(a) \neq 0$. An diesem Punkt ist p dann negativ, ein Widerspruch. □

Offensichtlich folgt aus $q \in T(p_1, \ldots, p_r)$ schon $p \in T(p_1, \ldots, p_r)$. Im folgenden können wir uns also auf quadratfreie Polynome beschränken, die von keinem der p_i geteilt werden.

Lemma 8.4.4. *Sei $0 \neq p \in \mathbb{R}[x,y]$ quadratfrei, und $p_i \nmid p$ für alle i. Falls $p \geq 0$ auf W ist, hat p in W nur endlich viele Nullstellen.*

Beweis. Angenommen p hat unendlich viele Nullstellen in W. Dann gibt es einen irreduziblen Faktor q von p, der auch unendlich viele Nullstellen in W hat (und q taucht aufgrund der Quadratfreiheit nur einfach in p auf). Insbesondere muss q eine reguläre Nullstelle $a \in W$ haben, die gleichzeitig keine Nullstelle eines der p_i und eines anderen irreduziblen Faktors von p ist. Das folgt aus Lemma 8.4.2 (i) und (ii). Inbesondere gilt $p_i(a) > 0$ für alle i, und a liegt also im Inneren von W. Da q nach Lemma 8.4.2 (iii) in jeder Umgebung von a das Vorzeichen wechselt, die anderen irreduziblen Faktoren von p das aber nicht tun, wechselt auch p sein Vorzeichen in W. Das ist ein Widerspruch. □

Wir beschäftigen uns als nächstes mit Potenzreihen in zwei Variablen.

Satz 8.4.5. *Seien $p, p_1, p_2 \in \mathbb{R}[x,y]$ so dass p_1 und p_2 am Ursprung mit linear unabhängigen Tangentenrichtungen verschwinden. Sei weiter $U \subseteq \mathbb{R}^2$ eine offene Umgebung des Ursprungs.*
(i) Ist $p \geq 0$ auf U, so ist p eine Quadratsumme in $\mathbb{R}[[x,y]]$.

(ii) Ist $p \geq 0$ auf $U \cap \{p_1 \geq 0\}$, so liegt p in der von p_1 erzeugten Präordnung in $\mathbb{R}[[x,y]]$.
(iii) Ist $p \geq 0$ auf $U \cap \{p_1 \geq 0, p_2 \geq 0\}$, so liegt p in der von p_1 und p_2 erzeugten Präordnung von $\mathbb{R}[[x,y]]$.

Beweis. In allen drei Fällen impliziert die geometrische Nichtnegativität, dass p in jeder Anordnung von $\mathbb{R}[[x,y]]$ liegt, die die entsprechenden Gleichungen enthält; also in Fall (i) in jeder Anordnung, in Fall (ii) in jeder Anordnung die p_1 enthält, und in Fall (iii) in jeder Anordnung, die p_1 und p_2 enthält. Das folgt aus dem Transferprinzip (Übungsaufgabe 61).
Ist beispielsweise $(\partial_x p_1)\,(0,0) \neq 0$, so ist $\mathbb{R}[[x,y]] = \mathbb{R}[[p_1,y]]$. Ebenso ist $\mathbb{R}[[x,y]] = \mathbb{R}[[p_1,p_2]]$ (Übungsaufgabe 62). Wir können also $p_1 = x$ und $p_2 = y$ annehmen.
Man kann nun zeigen, dass der Durchschnitt aller Anordnungen von $\mathbb{R}[[x,y]]$ die Quadratsummen ergibt. Genauer reicht es für ein Element in $\mathbb{R}[[x,y]]$ bereits aus, in jeder Anordnung des Quotientenkörpers $\mathrm{Quot}(\mathbb{R}[[x,y]]) = \mathbb{R}((x,y))$ zu liegen, um in $\mathbb{R}[[x,y]]$ eine Quadratsumme zu sein. Wir führen diesen relativ komplizierten Beweis hier nicht, man kann ihn aber in [4] nachlesen.
Aussage (i) ist damit bewiesen. Die anderen beiden Aussagen kann man darauf reduzieren. Dazu betrachten wir den Körper $K := \mathbb{R}((x,y))[\sqrt{x}]$. Jede Anordnung auf K induziert eine Anordnung auf $\mathbb{R}((x,y))$, die x enthält. Somit ist in (ii) das Polynom p in jeder Anordnung von K enthalten. K kann aber mit $\mathbb{R}((\sqrt{x},y))$ identifiziert werden, und also ist p eine Quadratsumme in $\mathbb{R}[[\sqrt{x},y]] = \mathbb{R}[[x,y]][\sqrt{x}]$.
Schreibt man $p = \sum_i q_i^2$ mit $q_i = q_{i1} + q_{i2}\sqrt{x}$ mit $q_{i1}, q_{i2} \in \mathbb{R}[[x,y]]$, sieht man, dass

$$p = \sum_i q_{i1}^2 + q_{i2}^2 x$$

gelten muss, und also liegt p in der von x erzeugten Präordnung. Der Fall (iii) geht analog. □

Nun können wir den Hauptsatz dieses Kapitels formulieren:

Satz 8.4.6. *Seien $p_1, \ldots, p_r \in \mathbb{R}[x,y]$ irreduzibel mit $W = W_{\mathbb{R}}(p_1, \ldots, p_r)$ beschränkt. Für jeden Randpunkt a von W gelte eine der beiden Aussagen:*

(i) a ist reguläre Nullstelle eines p_i, und W ist in einer Umgebung von a schon allein durch die Gleichung $p_i \geq 0$ definiert.

(*ii*) *a ist reguläre Nullstelle von p_i und p_j, die an a linear unabhängige Tangentenrichtungen haben, und W ist in einer Umgebung von a schon allein durch die Gleichungen $p_i \geq 0, p_j \geq 0$ definiert.*

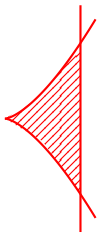
Dann ist die Präordnung $T = T(p_1, \ldots, p_r)$ saturiert.

Beweis. Die Bedingungen implizieren $W = \overline{\mathrm{int(W)}}$. Wenn wir zeigen wollen, dass T jedes auf W nichtnegative Polynom p enthält, können wir also nach Lemma 8.4.3 und Lemma 8.4.4 annehmen, dass p nur endlich viele Nullstellen in W hat. Wir können also Satz 8.3.1 anwenden, und müssen nur zeigen, dass p für jede Nullstelle a in der von den p_i in $\widehat{\mathbb{R}[x,y]}_a$ erzeugten Präordnung liegt. Wir nehmen dabei o.B.d.A $a = (0,0)$ an. Dann ist $\widehat{\mathbb{R}[x,y]}_a = \mathbb{R}[[x,y]]$. Die Aussage folgt damit unmittelbar aus Satz 8.4.5. □

Beispiel 8.4.7. (*i*) Sei $p_1 = 1 - x^2 - y^2$. Dann ist $W_{\mathbb{R}}(p_1)$ die Einheitskreisscheibe, und jeder Randpunkt erfüllt Bedingung (i) aus Satz 8.4.6. Also ist die Präordnung $T(p_1)$ saturiert, d.h. jedes auf der Kreisscheibe nichtnegative Polynom p ist von der Gestalt

$$p = \sigma_0 + \sigma_1(1 - x^2 - y^2)$$

(vergleiche Beispiel 4.2.4(*i*)). Die gleiche Aussage stimmt offensichtlich, wenn W beschränkt und durch irgendeine irreduzible Gleichung $p_1 \geq 0$ definiert ist, und jeder Randpunkt von W eine reguläre Nullstelle von p_1 ist.
(*ii*) Sei das Quadrat $[0,1]^2 \subseteq \mathbb{R}^2$ definiert durch die Polynome $x, 1-x, y, 1-y$. An jeder Ecke gilt dann (2) aus Satz 8.4.6, an den Seiten gilt (1). Also ist $T(x, 1-x, y, 1-y)$ saturiert. Das gleiche stimmt offensichtlich für beliebige Polytope in der Ebene.
(*iii*) Die Bedingungen (*i*) und (*ii*) können im Allgemeinen nicht abgeschwächt werden. Man betrachte etwa $p_1 = x^3 - y^2$ und $p_2 = 1 - x$.



Es gilt $x \geq 0$ auf $W_{\mathbb{R}}(p_1, p_2)$, aber $x \notin T(p_1, p_2)$. Der Ursprung ist hier kein glatter Punkt von p_1. △

Satz 8.4.6 setzt die Beschränktheit der Menge W voraus. Es gibt in der Ebene im Prinzip nur ein einziges bekanntes Beispiel mit nicht-beschränkter Menge und saturierter Präordnung.

Beispiel 8.4.8. Sei $p_1 = 1 - x^2 \in \mathbb{R}[x, y]$ wie in Beispiel 7.1.13. Die Menge W ist ein senkrechter Streifen. Bisher wissen wir nur, dass $T(p_1)$ die Momenteneigenschaft (SMP) hat. Marshall hat aber gezeigt, dass $T(p_1)$ sogar saturiert ist. Der Beweis ist sehr technisch und völlig an dieses eine Beispiel angepasst. Es ist unklar, ob es noch signifikant mehr nicht-kompakte saturierte Beispiele in der Ebene gibt. △

Kapitel 9

Nicht-kommutative reelle algebraische Geometrie

In diesem Kapitel beschäftigen wir uns mit einem Teilgebiet der reellen algebraischen Geometrie, das noch ziemlich neu und erst relativ wenig entwickelt ist: der nicht-kommutativen reellen algebraischen Geometrie. Dabei betrachtet man nicht-kommutative Ringe und Algebren und versucht dort Positivstellensätze zu beweisen. Man lässt sich dabei von den wichtigsten Beispielen solcher Algebren leiten: Matrixalgebren oder allgemeiner Algebren von Operatoren auf Hilberträumen.

Dabei treten neue Schwierigkeiten auf. Zum Beispiel muss man sich überlegen, was man unter Positivität überhaupt verstehen will. Nicht-kommutative Algebren bestehen eben nicht aus Polynomen oder reellwertigen Funktionen. Ein weiteres Problem ist, dass Produkte von positiven Elementen nicht unbedingt wieder positiv sind, man also über Präordnungen gar nicht sprechen kann. Einige solcher Fragen und Ergebnisse wollen wir hier besprechen.

9.1 Beispiele: Matrizen und Operatoren auf Hilberträumen

Zunächst betrachten wir zwei Klassen von Beispielen, von denen wir uns im weiteren leiten lassen.

Matrixalgebren

Sei $d \in \mathbb{N}$ und $\mathrm{M}_d(\mathbb{C})$ die Menge der $d \times d$-Matrizen über $\mathbb{C}$. Die Menge $\mathrm{M}_d(\mathbb{C})$ trägt die Struktur einer $\mathbb{C}$-Algebra mit multiplikativ neutralem Element I_d, die für $d \geq 2$ nicht mehr kommutativ ist. Im Kontext von Positivität haben wir bisher meistens $\mathbb{R}$ als Konstantenkörper gewählt. Da wir im nicht-kommutativen immer eine *Involution* verwenden, können wir hier aber zu $\mathbb{C}$ übergehen, was manche Schritte einfacher macht. Die Involution ist hier das Konjugieren und Transponieren einer Matrix, d.h. für $M = (m_{ij})_{i,j=1,\dots,d}$ setzen wir

$$M^* = (\overline{m}_{ji})_{i,j=1,\dots,d}\,.$$

Die Involution ist ein konjugiert-linearer selbstinverser Antiautomorphismus, d.h. sie erfüllt

$$(\lambda M + \gamma N)^* = \overline{\lambda} M^* + \overline{\gamma} N^*, \quad (M^*)^* = M \text{ und } (MN)^* = N^* M^*$$

für alle Matrizen $M, N \in \mathrm{M}_d(\mathbb{C})$ und $\lambda, \gamma \in \mathbb{C}$. Einen Fixpunkt der Involution, also eine Matrix mit $M^* = M$, nennt man *hermitesche Matrix*, und die Menge dieser Elemente bildet einen $\mathbb{R}$-Untervektorraum von $\mathrm{M}_d(\mathbb{C})$, *nicht* jedoch einen $\mathbb{C}$-Unterraum:

$$\mathrm{Her}_d(\mathbb{C}) = \mathrm{M}_d(\mathbb{C})_h = \{M \in \mathrm{M}_d(\mathbb{C}) \mid M^* = M\}\,.$$

Es ist $\mathrm{Her}_d(\mathbb{C})$ für $d \geq 2$ auch *keine* Unteralgebra, also nicht abgeschlossen unter Multiplikation.
Wir kennen bereits einen Begriff von *Positivität* (siehe Definition 2.2.4 und Lemma 2.2.3), der auch für komplexe Matrizen sinnvoll ist. Eine Matrix $M \in \mathrm{Her}_d(\mathbb{C})$ heißt *positiv semidefinit*, falls

$$v^* M v \geq 0 \quad \text{für alle } v \in \mathbb{C}^d$$

gilt. Dabei ist v^* der konjugiert-transponierte Vektor zu $v \in \mathbb{C}^d$. Man beachte, dass wir Positivität nur für hermitesche Matrizen definieren (sowie wir es früher auch nur für symmetrischen reelle Matrizen getan haben). Das ist auch sinnvoll, da die Zahl $v^* M v$ sonst im Allgemeinen gar nicht reell ist, und von Positivität dann gar nicht gesprochen werden kann. *Positivität spielt sich also im Raum der hermiteschen Elemente der Algebra ab.* Es kann hier nun passieren, dass ein Produkt von positiven Elementen gar nicht wieder hermitesch und nicht wieder positiv ist:

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

Abgesehen davon, dass das Produkt nicht hermitesch ist, ist beispielsweise

$$\begin{pmatrix} -\frac{3}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} -\frac{3}{2} \\ 1 \end{pmatrix} < 0.$$

Der Begriff einer Präordnung ist deshalb im nicht-kommutativen nicht ohne weiteres definierbar.
Die Aussage von Lemma 2.2.3, angepasst an das hermitesche Setup, stimmt aber genauso. Die wichtigste Aussage ist, dass M genau dann positiv semidefinit ist, wenn es eine Darstellung

$$M = S^*S$$

für eine (sogar hermitesche und positiv semidefinite) Matrix $S \in \mathrm{M}_d(\mathbb{C})$ gibt. Man diagonalisiert dafür einfach $U^*MU = D$ mit einer unitären Matrix U und setzt $S = U\sqrt{D}U^*$. Dieses S wird auch mit $\sqrt{M}$ bezeichnet. Man erhält S auch als Limes $p_n(M)$ für $n \to \infty$ von *polynomialen Ausdrücken* in M. Dabei müssen die $p_n(t) \in \mathbb{R}[t]$ so gewählt sein, dass sie die Wurzelfunktion auf dem Spektrum von M approximieren. Wählt man die p_n so, dass sie die Wurzelfunktion auf immer größeren Intervallen immer besser approximieren, funktioniert diese Folge also für alle M simultan.
Insgesamt haben wir hier einen sehr schönen Positivstellensatz: M ist genau dann positiv semidefinit, wenn M in $\mathrm{M}_d(\mathbb{C})$ ein Quadrat ist, genauer ein hermitesches Quadrat. *Im nicht-kommutativen Setup werden Quadrate und Quadratsummen immer mithilfe der Involution gebildet, wir betrachten also immer hermitesche Quadrate und Quadratsummen*. Normale Quadrate der Gestalt $Q \cdot Q$ liefern im Allgemeinen nämlich keine positiven Matrizen, wie man bereits im 1×1-Fall sieht: $i \cdot i = -1$.
Wir beenden dieses Beispiel mit zwei interessanten Sätzen über Matrixalgebren:

Satz 9.1.1. $\mathrm{M}_d(\mathbb{C})$ *ist eine zentral-einfache Algebra, d.h. sie besitzt keine nichttrivialen beidseitigen Ideale, und das Zentrum besteht nur aus* $\mathbb{C} \cdot I_d$.

Beweis. Sei $0 \neq M \in \mathrm{M}_d(\mathbb{C})$. Wir zeigen dass für das von M erzeugte beidseitige Ideal gilt

$$\langle M \rangle = \mathrm{M}_d(\mathbb{C}).$$

Sei E_{ij} die Matrix mit einer 1 an der (i,j)-Position, und Nullen überall sonst. Die E_{ij} spannen $\mathrm{M}_d(\mathbb{C})$ als $\mathbb{C}$-Vektorraum auf, also genügt es

$$E_{ij} \in \langle M \rangle$$

für alle i, j zu zeigen. Aufgrund der Gleichung

$$E_{ki}E_{ij}E_{jl} = E_{kl}$$

genügt es aber, das für *eine* Matrix E_{ij} zu zeigen. Falls der Eintrag von M an einer Stelle (i, j) gerade $m \neq 0$ ist, so gilt

$$E_{ij} = m^{-1} \cdot E_{ii}ME_{jj} \in \langle M \rangle.$$

Falls nun M im Zentrum von $\mathrm{M}_d(\mathbb{C})$ liegt, also mit allen Matrizen kommutiert, sieht man an den Gleichungen

$$E_{ij}M = ME_{ij}$$

bereits leicht, dass M diagonal mit identischen Diagonaleinträgen ist. □

Im nächsten Satz untersuchen wir Unteralgebren von Matrixalgebren. Sei dazu $\mathcal{A} \subseteq \mathrm{M}_d(\mathbb{C})$ eine $*$-Unteralgebra, d.h. mit M gehöre auch M^* zu $\mathcal{A}$. Falls $\mathcal{A}$ einen echten invarianten Unterraum besitzt, d.h. ein

$$\{0\} \subsetneq V \subsetneq \mathbb{C}^d$$

mit $MV \subseteq V$ für alle $M \in \mathcal{A}$, so ist auch $V^\perp$ ein solcher invarianter Unterraum. Dafür verwendet man, dass $\mathcal{A}$ abgeschlossen unter $*$ ist. Nach einem unitären Basiswechsel haben alle Matrizen in $\mathcal{A}$ Blockgestalt, d.h. $\mathcal{A}$ ist eine Unteralgebra von einer Algebra $\mathrm{M}_{d_1}(\mathbb{C}) \oplus \mathrm{M}_{d_2}(\mathbb{C})$ mit $1 \leq d_1, d_2$ und $d_1 + d_2 = d$, und wir sind in einer einfacheren Situation. Falls es keinen invarianten Unterraum gibt, gilt $\mathcal{A} = \mathrm{M}_d(\mathbb{C})$:

Satz 9.1.2 (Satz von Burnside)**.** *Falls die Unteralgebra $\mathcal{A} \subseteq \mathrm{M}_d(\mathbb{C})$ keinen nichttrivialen invarianten Unterraum besitzt, gilt $\mathcal{A} = \mathrm{M}_d(\mathbb{C})$.*

Beweis. $\mathcal{A}$ operiert transitiv auf $\mathbb{C}^d$: für jedes $0 \neq v \in \mathbb{C}^d$ ist ja

$$\{0\} \subsetneq \{Mv \mid M \in \mathcal{A}\}$$

ein invarianter Unterraum, stimmt also mit $\mathbb{C}^d$ überein. Wir zeigen zunächst, dass $\mathcal{A}$ eine Matrix von Rang 1 enthält. Sei dazu $0 \neq P \in \mathcal{A}$. Falls $\mathrm{rang}(P) \geq 2$ ist, wähle $v_1, v_2 \in \mathbb{C}^d$ mit Pv_1, Pv_2 linear unabhängig. Wähle dann $M \in \mathcal{A}$ mit $MPv_1 = v_2$. Dann sind also $PMPv_1$ und Pv_1 linear unabhängig und $PMP - \lambda P \neq 0$ gilt also für alle $\lambda \in \mathbb{C}$. Es gibt aber ein λ_0 für welches $PM - \lambda_0 I_d$ auf

dem Raum $P\mathbb{C}^d$ nicht invertierbar ist, denn $\mathbb{C}$ ist algebraisch abgeschlossen, und jede lineare Abbildung besitzt also einen Eigenwert (hier sehen wir warum wir lieber über $\mathbb{C}$ als über $\mathbb{R}$ arbeiten). Also hat

$$(PM - \lambda_0 I_d)P$$

einen echt kleineren Rang als P, ist aber nicht Null. Iterativ erhalten wir also eine Matrix Q vom Rang 1 in $\mathcal{A}$.
Jede andere Matrix vom Rang 1 liegt dann ebenso in $\mathcal{A}$. Dafür benötigt man nochmal die Transitivität von $\mathcal{A}$ auf $\mathbb{C}^d$ (Übungsaufgabe 63). Da jede Matrix eine Summe von Matrizen von Rang 1 ist, folgt $\mathcal{A} = \mathrm{M}_d(\mathbb{C})$. □

Beispiel 9.1.3. Sei $\mathcal{A} \subseteq \mathrm{M}_d(\mathbb{C})$ eine *kommutative* $*$-Unteralgebra. Durch iteratives Anwenden von Satz 9.1.2 und der davor beschriebenen Reduktionsmethode erreichen wir

$$\mathcal{A} \subseteq \mathrm{M}_{d_1}(\mathbb{C}) \oplus \cdots \oplus \mathrm{M}_{d_r}(\mathbb{C})$$

mit $d_1 + \cdots + d_r = d$, und die Projektion auf jeden d_i-Block ist auf $\mathcal{A}$ bereits surjektiv. Andererseits kommutieren die Elemente von $\mathcal{A}$ miteinander, und mit Satz 9.1.1 folgt daraus $d_i = 1$ für alle i. *Die Algebra $\mathcal{A}$ besteht also nur aus Diagonalmatrizen* (nach unitärer Konjugation). △

Operatoren auf Hilberträumen

Das zweite wichtige Beispiel sind Algebren von Operatoren auf Hilberträumen, also eine Verallgemeinerung von Matrixalgebren. Sei $\mathcal{H}$ ein Hilbertraum über $\mathbb{C}$. Eine lineare Abbildung

$$T\colon \mathcal{H} \to \mathcal{H}$$

(auch *Operator* genannt) ist genau dann stetig wenn sie *beschränkt* ist, d.h. wenn

$$\|Tv\| \leq C\|v\|$$

für ein $C \geq 0$ und alle $v \in \mathcal{H}$ gilt. Das kleinste solche C heißt *Operatornorm* von T, und wird mit $\|T\|_{\mathrm{op}}$ bezeichnet. Die Menge aller stetigen Operatoren auf $\mathcal{H}$ wird mit

$$\mathcal{B}(\mathcal{H})$$

bezeichnet. Die Menge $\mathcal{B}(\mathcal{H})$ trägt einerseits die Struktur eines Banachraums (bezüglich der Operatornorm), ist andererseits aber auch eine Algebra mit Involution. Die Multiplikation ist die Hintereinanderausführung von Operatoren,

die Involution besteht aus dem Bilden des *adjungierten Operators*, der eindeutig definiert ist durch die Bedingung

$$\langle Tv, w\rangle = \langle v, T^*w\rangle$$

für alle $v, w \in \mathcal{H}$, und der für jedes T in $\mathcal{B}(\mathcal{H})$ existiert. Ein Fixpunkt unter der Involution heißt *selbstadjungierter* Operator, und wieder ist die Menge

$$\mathcal{B}(\mathcal{H})_h := \{T \in \mathcal{B}(\mathcal{H}) \mid T^* = T\}$$

aller selbstadjungierten Operatoren ein $\mathbb{R}$-Unterraum von $\mathcal{B}(\mathcal{H})$.

Beispiel 9.1.4. Es sei

$$\mathcal{H} = \ell^2(\mathbb{Z}) = \left\{ (a_i)_{i\in\mathbb{Z}} \mid a_i \in \mathbb{C}, \sum_i |a_i|^2 < \infty \right\}$$

der Hilbertraum der quadratsummierbaren Folgen in $\mathbb{C}$, mit dem Skalarprodukt

$$\langle (a_i)_i, (b_i)_i\rangle = \sum_i a_i \overline{b}_i.$$

(i) Sei $m = (m_i)_{i\in\mathbb{Z}} \in \ell^\infty(\mathbb{Z})$ eine *beschränkte* Folge, d.h. $|m_i| \leq C$ für ein $C \geq 0$ und alle i. Dann definiert die folgende Vorschrift einen stetigen *Multiplikationsoperator* M_m auf $\mathcal{H}$, mit $\|M_m\|_{\mathrm{op}} \leq C$:

$$M_m \colon (a_i)_i \mapsto (m_i a_i)_i.$$

Der adjungierte Operator M_m^* ist gerade die Multiplikation mit der komplex konjugierten Folge $\overline{m} = (\overline{m}_i)_i$, und somit ist M_m genau dann selbstadjungiert, wenn alle m_i reell sind.
(ii) Der *Shiftoperator* S ist definiert durch

$$S \colon (a_i)_i \mapsto (a_{i-1})_i.$$

Er ist normerhaltend, d.h. $\|Sv\| = \|v\|$ für alle $v \in \mathcal{H}$, insbesondere $\|S\|_{\mathrm{op}} = 1$ und $S \in \mathcal{B}(\mathcal{H})$. Sein adjungierter Operator ist gerade der Shift in die andere Richtung

$$S^* \colon (a_i)_i \mapsto (a_{i+1})_i.$$

Also ist S unitär, aber nicht selbstadjungiert. △

Ein selbstadjungierter Operator $T \in \mathcal{B}(\mathcal{H})_h$ heißt *positiv semidefinit,* wenn

$$\langle Tv, v\rangle \geq 0 \quad \forall v \in \mathcal{H}$$

gilt. Der Begriff ist also eine direkte Verallgemeinerung des Positivitätsbegriffs für Matrizen. Auch hier benötigen wir die Selbstadjungiertheit, damit der Wert $\langle Tv, v\rangle$ überhaupt reell ist:

$$\overline{\langle Tv, v\rangle} = \langle v, Tv\rangle = \langle T^*v, v\rangle = \langle Tv, v\rangle.$$

Die Spektraltheorie und das Funktionalkalkül für selbstadjungierte Operatoren in Hilberträumen liefert nun ein genaues Analog zum Positivstellensatz für Matrizen: Ein selbstadjungierter Operator $T \in \mathcal{B}(\mathcal{H})_h$ ist genau dann positiv semidefinit, wenn

$$T = S^*S \text{ für ein } S \in \mathcal{B}(\mathcal{H})$$

gilt.

Beispiel 9.1.5. Der Multiplikationsoperator M_m aus Beispiel 9.1.4 (i) mit einer reellen Folge $m = (m_i)_i$ ist selbstadjungiert, und genau dann positiv semidefinit, wenn alle $m_i \geq 0$ sind. In diesem Fall liefert der Multiplikationsoperator $M_{\sqrt{m}}$ mit $\sqrt{m} = (\sqrt{m_i})_i$ eine Quadratzerlegung von M_m. △

Vereinzelt ist es auch nötig, *unbeschränkte,* d.h. nicht-stetige Operatoren zu betrachten. Dazu muss der zugrundeliegende Raum auch nicht unbedingt ein Hilbertraum sein. Sei also $\mathcal{D}$ ein $\mathbb{C}$-Vektorraum mit Skalarprodukt (nicht unbedingt vollständig). Dann bezeichnen wir mit

$$\mathcal{L}(\mathcal{D})$$

die Menge der linearen Abbildungen $T\colon \mathcal{D} \to \mathcal{D}$, diesmal ohne Stetigkeitsbedingung. Sie bilden wieder eine Algebra. Ist $T \in \mathcal{L}(\mathcal{D})$ stetig auf $\mathcal{D}$, so kann man T eindeutig linear auf die Vervollständigung $\mathcal{H}$ von $\mathcal{D}$ fortsetzen, d.h. man kann T als Element von $\mathcal{B}(\mathcal{H})$ auffassen (siehe Aufgabe 68 für eine allgemeinere Aussage).
Da die Definition des adjungierten Operators in $\mathcal{L}(\mathcal{D})$ nicht ohne weiteres möglich ist, halten wir nur folgendes fest: Ein Operator $T \in \mathcal{L}(\mathcal{D})$ heißt *selbstadjungiert* falls

$$\langle Tv, w\rangle = \langle v, Tw\rangle$$

für alle $v, w \in \mathcal{D}$ gilt, und ein selbstadjungierter Operator heißt *positiv semidefinit,* falls

$$\langle Tv, v\rangle \geq 0$$

für alle $v \in \mathcal{D}$ gilt.

9.2 $*$-Algebren und Darstellungen

Wir betrachten nun allgemeine $*$-Algebren und deren Darstellungen. Nach wie vor besitze jede Algebra ein multiplikativ neutrales Element 1, und Algebrahomomorphismen bilden 1 auf 1 ab.

Definition 9.2.1. (i) Eine *$*$-Algebra* ist eine (nicht notwendigerweise kommutative) $\mathbb{C}$-Algebra $\mathcal{A}$, auf der eine Involution $*$ existiert, d.h. es gelte

$$(\lambda a + \gamma b)^* = \overline{\lambda} a^* + \overline{\gamma} b^*, \quad (a^*)^* = a, \quad (ab)^* = b^* a^*$$

für alle $a, b \in \mathcal{A}, \lambda, \gamma \in \mathbb{C}$.
(ii) Für eine $*$-Algebra $\mathcal{A}$ heißt der $\mathbb{R}$-Unterraum

$$\mathcal{A}_h := \{a \in \mathcal{A} \mid a^* = a\}$$

Unterraum der *hermiteschen Elemente*.
(iii) Mit

$$\sum \mathcal{A}^2 = \left\{ \sum_{i=1}^{m} a_i^* a_i \mid m \in \mathbb{N}, a_i \in \mathcal{A} \right\}$$

wird die Menge der *hermiteschen Quadratsummen* bezeichnet. Es ist $\sum \mathcal{A}^2$ ein konvexer Kegel in $\mathcal{A}_h$, mit

$$a^* \cdot \sum \mathcal{A}^2 \cdot a \subseteq \sum \mathcal{A}^2$$

für alle $a \in \mathcal{A}$.
(iv) Ein *$*$-Algebrahomomorphismus* ist ein Algebrahomomorphismus

$$\pi \colon \mathcal{A} \to \mathcal{B}$$

zwischen $*$-Algebren, der $\pi(a^*) = \pi(a)^*$ für alle $a \in \mathcal{A}$ erfüllt.
(v) Eine *(beschränkte) $*$-Darstellung* von $\mathcal{A}$ ist ein $*$-Algebrahomomorphismus

$$\pi \colon \mathcal{A} \to \mathcal{B}(\mathcal{H})$$

für einen Hilbertraum $\mathcal{H}$. Eine Darstellung heißt *endlich-dimensional,* falls $\mathcal{H}$ endlichdimensional ist. Nach Wahl einer Basis erhält man dabei also

$$\pi \colon \mathcal{A} \to \mathrm{M}_d(\mathbb{C}).$$

(vi) Eine *unbeschränkte $*$-Darstellung* ist ein Algebrahomomorphismus

$$\pi \colon \mathcal{A} \to \mathcal{L}(\mathcal{D})$$

für einen Vektorraum $\mathcal{D}$ mit Skalarprodukt, der zusätzlich

$$\langle \pi(a)v, w\rangle = \langle v, \pi(a^*)w\rangle$$

für alle $v, w \in \mathcal{D}$ und alle $a \in \mathcal{A}$ erfüllt.
(vii) Ein *Zustand* auf $\mathcal{A}$ ist ein $\mathbb{C}$-lineares Funktional $\varphi\colon \mathcal{A} \to \mathbb{C}$, das

$$\varphi(1) = 1 \text{ und } \varphi(a^*a) \geq 0$$

für alle $a \in \mathcal{A}$ erfüllt. Oft wird noch zusätzlich $\varphi(a^*) = \overline{\varphi(a)}$ gefordert, das folgt aber bereits aus der Positivität auf den Quadraten. (Übungsaufgabe 64). Insbesondere ist $\varphi\colon \mathcal{A}_h \to \mathbb{R}$ ein $\mathbb{R}$-lineares Funktional. △

Beispiel 9.2.2. (i) Sei $\mathcal{A}$ eine kommutative $*$-Algebra und $\pi\colon \mathcal{A} \to \mathrm{M}_d(\mathbb{C})$ eine endlich-dimensionale $*$-Darstellung. Laut Beispiel 9.1.3 können wir nach unitärer Konjugation annehmen, dass $\pi\colon \mathcal{A} \to \mathrm{M}_1(\mathbb{C}) \oplus \cdots \oplus \mathrm{M}_1(\mathbb{C})$ gilt, d.h. π besteht aus einem d-Tupel von $*$-Algebrahomomorphismen $\varphi_i\colon \mathcal{A} \to \mathbb{C}$.
(ii) Sei $\mathcal{A} = \mathbb{C}[x_1, \ldots, x_n]$ mit der Involution, die die Koeffizienten komplex konjugiert und die Variablen invariant lässt. Es gilt

$$\mathbb{C}[x_1, \ldots, x_n]_h = \mathbb{R}[x_1, \ldots, x_n] \text{ und } \sum \mathbb{C}[x_1, \ldots, x_n]^2 = \sum \mathbb{R}[x_1, \ldots, x_n]^2.$$

Die endlich-dimensionalen $*$-Darstellungen von $\mathbb{C}[x_1, \ldots, x_n]$ sind (nach Basiswechsel) direkte Summen von Einsetzungen von Punkten aus $\mathbb{R}^n$. △

Bemerkung 9.2.3. (i) Sei $a \in \sum \mathcal{A}^2$ und sei π eine (beschränkte oder unbeschränkte) $*$-Darstellung von $\mathcal{A}$. Dann ist $\pi(a)$ positiv semidefinit. Falls nämlich $a = \sum_i a_i^* a_i$ ist, gilt

$$\langle \pi(a)v, v\rangle = \sum_i \langle \pi(a_i^*)\pi(a_i)v, v\rangle = \sum_i \langle \pi(a_i)v, \pi(a_i)v\rangle = \sum_i \|\pi(a_i)v\|^2 \geq 0.$$

(ii) Jede (beschränkte oder unbeschränkte) $*$-Darstellung π von $\mathcal{A}$ liefert viele Zustände φ auf $\mathcal{A}$. Für $v \in \mathcal{H}$ (bzw. $v \in \mathcal{D}$) mit $\|v\| = 1$ setzt man

$$\varphi(a) := \langle \pi(a)v, v\rangle$$

und rechnet die Eigenschaften direkt nach. △

Die wichtige Konstruktion von *Gelfand, Neumark* und *Segal*, auch *GNS-Konstruktion* genannt, liefert eine Umkehrung zur letzten Bemerkung. Man startet mit einem Zustand $\varphi\colon \mathcal{A} \to \mathbb{C}$ und konstruiert dazu eine Darstellung

$$\pi_\varphi\colon \mathcal{A} \to \mathcal{L}(\mathcal{D})$$

und einen Vektor $v \in \mathcal{D}$, mit $\varphi(a) = \langle \pi_\varphi(a)v, v\rangle$ für alle $a \in \mathcal{A}$. Wir beweisen dafür zunächst zwei Lemmas.

Lemma 9.2.4 (Cauchy-Schwarz-Ungleichung)**.** *Sei $\varphi\colon \mathcal{A} \to \mathbb{C}$ ein Zustand. Dann gilt für alle $a, b \in \mathcal{A}$:*

$$|\varphi(b^*a)|^2 \leq \varphi(b^*b)\varphi(a^*a).$$

Beweis. Die hermitesche Matrix

$$M := \begin{pmatrix} \varphi(a^*a) & \varphi(a^*b) \\ \varphi(b^*a) & \varphi(b^*b) \end{pmatrix} \in \mathrm{Her}_2(\mathbb{C})$$

ist positiv semidefinit. Mit $v = (v_1, v_2)^t \in \mathbb{C}^2$ gilt nämlich

$$v^*Mv = \varphi\left((v_1a + v_2b)^*(v_1a + v_2b)\right) \geq 0.$$

Damit hat M eine nichtnegative Determinante, und die ist gerade

$$\det(M) = \varphi(a^*a)\varphi(b^*b) - \varphi(a^*b)\varphi(b^*a) = \varphi(a^*a)\varphi(b^*b) - |\varphi(b^*a)|^2.$$

Daraus folgt die Behauptung. □

Lemma 9.2.5. *Sei $\varphi\colon \mathcal{A} \to \mathbb{C}$ ein Zustand. Dann ist*

$$N_\varphi := \{a \in \mathcal{A} \mid \varphi(a^*a) = 0\}$$

ein (echtes) Linksideal in $\mathcal{A}$.

Beweis. Übungsaufgabe 65. □

Die GNS-Konstruktion funktioniert nun wie folgt. Dabei sind alle nicht bewiesenen Aussagen Übungsaufgabe 66. Sei φ ein Zustand auf $\mathcal{A}$. Zunächst versieht man den $\mathbb{C}$-Vektorraum $\mathcal{A}$ mit einer Sesquilinearform

$$\langle a, b\rangle_\varphi := \varphi(b^*a),$$

die offensichtlich positiv semidefinit ist, i.e. $\langle a, a\rangle_\varphi \geq 0$ erfüllt. Um $\langle \cdot, \cdot\rangle_\varphi$ positiv definit, d.h. zu einem Skalarprodukt zu machen, muss man noch N_φ ausdividieren: auf dem $\mathbb{C}$-Vektorraum

$$\mathcal{D} := \mathcal{A}/N_\varphi$$

ist $\langle \cdot, \cdot\rangle_\varphi$ ein wohldefiniertes Skalarprodukt. Da N_φ ein Linksideal ist, ist auf $\mathcal{D}$ die Multiplikation von links mit Elementen von $\mathcal{A}$ wohldefiniert., d.h. jedes $a \in \mathcal{A}$ definiert einen linearen Operator

$$m_a \colon \mathcal{D} \to \mathcal{D};\ d \mapsto ad.$$

Auf diese Weise erhält man eine $*$-Darstellung

$$\begin{aligned} \pi_\varphi \colon \mathcal{A} &\to \mathcal{L}(\mathcal{D}) \\ a &\mapsto m_a \end{aligned}$$

von $\mathcal{A}$. Wählt man als $v \in \mathcal{D}$ gerade die Restklasse der 1, so gilt $\|v\|_\varphi = 1$ und

$$\langle \pi_\varphi(a)v, v\rangle_\varphi = \varphi(1^* \cdot a \cdot 1) = \varphi(a)$$

für alle $a \in \mathcal{A}$. Das ist genau die oben aufgestellte Behauptung.

Um Positivstellensätze für $*$-Algebren beweisen zu können, müssen wir noch Positivität definieren. Das macht man gewöhnlich über Darstellungen:

Definition 9.2.6. Sei $\mathcal{A}$ eine $*$-Algebra und $\mathcal{F}$ eine Klasse von $*$-Darstellungen von $\mathcal{A}$. Ein Element $a \in \mathcal{A}_h$ heißt *$\mathcal{F}$-nichtnegativ*, falls $\pi(a)$ ein positiv semidefiniter Operator für jedes $\pi \in \mathcal{F}$ ist. △

Bemerkung 9.2.7. Eine hermitesche Quadratsumme aus $\mathcal{A}$ ist nach Bemerkung 9.2.3 (i) $\mathcal{F}$-nichtnegativ für jede Klasse $\mathcal{F}$ von Darstellungen. △

Beispiel 9.2.8. Sei $\mathcal{A} = \mathbb{C}[x_1, \ldots, x_n]$ mit der bekannten Involution. Sei $\mathcal{F}$ die Menge der endlich-dimensionalen $*$-Darstellungen von $\mathcal{A}$. Nach Beispiel 9.2.2 (ii) ist ein Polynom $p \in \mathbb{R}[x_1, \ldots, x_n]$ genau dann $\mathcal{F}$-nichtnegativ, wenn es als Polynom global nichtnegativ auf $\mathbb{R}^n$ ist. Dafür genügt es bereits, die eindimensionalen $*$-Darstellungen zu betrachten. △

Wir definieren schließlich

$$\left(\sum \mathcal{A}^2\right)^{\vee\vee} = \{a \in \mathcal{A}_h \mid \varphi(a) \geq 0 \text{ für alle Zustände } \varphi \text{ auf } \mathcal{A}\}$$

genauso wie früher für Quadratsummen in $\mathbb{R}[x]$ (die Zusatzbedingung $\varphi(1) = 1$ für Zustände ändert die Definition nicht, siehe Übungsaufgabe 67). Wiederum ist das gerade der Abschluss von $\sum \mathcal{A}^2$ in der feinsten lokalkonvexen Topologie auf dem $\mathbb{R}$-Vektorraum $\mathcal{A}_h$. Der folgende Satz ist ein allgemeiner Positivstellensatz für $*$-Algebren.

Satz 9.2.9. *Sei $\mathcal{A}$ eine $*$-Algebra und $\mathcal{F}$ die Klasse aller $*$-Darstellungen (beschränkte und unbeschränkte) von $\mathcal{A}$. Dann gilt für $a \in \mathcal{A}_h$*

$$a \text{ ist } \mathcal{F}\text{-nichtnegativ} \quad \Leftrightarrow \quad a \in \left(\sum \mathcal{A}^2\right)^{\vee\vee}$$

Beweis. "$\Leftarrow$": Sei $\pi\colon \mathcal{A} \to \mathcal{L}(\mathcal{D})$ eine $*$-Darstellung und $v \in \mathcal{D}$ mit $\|v\| = 1$. Dann ist

$$b \mapsto \langle \pi(b)v, v\rangle$$

ein Zustand auf $\mathcal{A}$. Nach Voraussetzung gilt also $\langle \pi(a)v, v\rangle \geq 0$, und also ist a $\mathcal{F}$-nichtnegativ.
Für "$\Rightarrow$" sei $\varphi\colon \mathcal{A} \to \mathbb{C}$ ein Zustand. Mit der GNS-Konstruktion erhalten wir eine $*$-Darstellung

$$\pi_\varphi\colon \mathcal{A} \to \mathcal{L}(\mathcal{D})$$

und ein $v \in \mathcal{D}$ mit $\varphi(a) = \langle \pi_\varphi(a)v, v\rangle_\varphi$. Wegen $\pi_\varphi \in \mathcal{F}$ gilt also $\varphi(a) \geq 0$. □

Bemerkung 9.2.10. Im Beweis von Satz 9.2.9 sieht man, dass man $\mathcal{F}$ als die Menge aller $*$-Darstellungen wählen kann, in denen

$$\dim(\mathcal{D}) \leq \dim(\mathcal{A})$$

gilt. Die Darstellung π_φ liefert ja als $\mathcal{D}$ gerade einen Quotienten von $\mathcal{A}$. △

Beispiel 9.2.11. Für $\mathbb{C}[x]$ haben wir in Beispiel 7.2.7 (i) schon

$$\left(\sum \mathbb{C}[x]^2\right)^{\vee\vee} = \left(\sum \mathbb{R}[x]^2\right)^{\vee\vee} = \sum \mathbb{R}[x]^2$$

gesehen. Also ist $p \in \mathbb{R}[x]$ genau dann eine Quadratsumme, wenn p bei jeder (beschränkten und unbeschränkten) $*$-Darstellung von $\mathbb{C}[x]$ positiv semidefinit ist. Die Positivität in allen endlich-dimensionalen Darstellungen reicht dabei im Allgemeinen nicht aus, denn sie ist nach Beispiel 9.2.8 nur äquivalent zur Positivität von p auf dem $\mathbb{R}^n$. △

Definition 9.2.12. Wir nennen $\sum \mathcal{A}^2$ *archimedisch*, wenn für jedes $a \in \mathcal{A}_h$ ein $r > 0$ existiert mit

$$r - a \in \sum \mathcal{A}^2.$$

Nach Teilung durch r ist das offensichtlich äquivalent zu

$$1 - \varepsilon a \in \sum \mathcal{A}^2,$$

d.h. dazu dass 1 ein *algebraisch innerer Punkt* von $\sum \mathcal{A}^2$ im $\mathbb{R}$-Vektorraum $\mathcal{A}_h$ ist, man also von 1 ein Stück in jede Richtung laufen kann, ohne die Quadratsummen zu verlassen. △

Der folgende Satz ist ein archimedischer Positivstellensatz für $*$-Algebren. Man kann dabei die Voraussetzungen von Satz 9.2.9 abschwächen und erhält zusätzlich ein stärkeres Ergebnis:

Satz 9.2.13. *Sei $\mathcal{A}$ eine $*$-Algebra in der $\sum \mathcal{A}^2$ archimedisch ist, und sei $\mathcal{F}$ die Klasse der beschränkten $*$-Darstellungen von $\mathcal{A}$. Dann gilt für $a \in \mathcal{A}_h$*

$$a \text{ ist } \mathcal{F}\text{-nichtnegativ} \quad \Leftrightarrow \quad a + \varepsilon \in \sum \mathcal{A}^2 \quad \forall \varepsilon > 0.$$

Beweis. "⇐" folgt aus Satz 9.2.9, bzw. wird genauso bewiesen. Für "⇒" reicht es zu zeigen, dass $a \in (\sum \mathcal{A}^2)^{\vee\vee}$ gilt. Da 1 ein innerer Punkt von $\sum \mathcal{A}^2$ ist folgt daraus dann nämlich $a + \varepsilon \in \sum \mathcal{A}^2$ für alle $\varepsilon > 0$, nach einer bekannten Version des Satzes von Hahn-Banach.
Das wiederum geht wie im Beweis von Satz 9.2.9. Wir müssen uns nur überlegen dass Zustände auf $\mathcal{A}$ in der GNS-Konstruktion hier *beschränkte* Darstellungen liefern. Sei also φ ein Zustand auf $\mathcal{A}$ und π_φ die dazu konstruierte GNS-Darstellung auf $\mathcal{D} = \mathcal{A}/N_\varphi$. Für $a \in \mathcal{A}$ gibt es ein $r > 0$ mit $r - a^*a \in \sum \mathcal{A}^2$, aufgrund der Archimedizität. Jeder Vektor $v \in \mathcal{D}$ ist nun die Restklasse eines $b \in \mathcal{A}$, und es gilt

$$b^*(r - a^*a)b = rb^*b - b^*a^*ab \in \sum \mathcal{A}^2.$$

Damit erhält man

$$\|\pi_\varphi(a)v\|_\varphi^2 = \langle \pi_\varphi(a)v, \pi_\varphi(a)v\rangle_\varphi = \varphi(b^*a^*ab) \leq \varphi(rb^*b) = r\langle b, b\rangle_\varphi = r\|v\|_\varphi^2,$$

wobei die Ungleichung daraus folgt, dass φ auf $\sum \mathcal{A}^2$ nichtnegativ ist. Also ist $\pi_\varphi(a)$ ein beschränkter Operator auf $\mathcal{D}$ (mit Operatornorm $\leq \sqrt{r}$, die nicht mal

von φ abhängt), und setzt sich eindeutig linear und stetig fort auf die Vervollständigung $\mathcal{H}$. Es kann also

$$\pi_\varphi \colon \mathcal{A} \to \mathcal{B}(\mathcal{H})$$

als beschränkte $*$-Darstellung aufgefasst werden (siehe Aufgabe 68). □

Bemerkung 9.2.14. Auch in Satz 9.2.13 kann man sich auf die Klasse der stetigen Darstellungen auf Hilberträumen $\mathcal{H}$ beschränken, die einen dichten Unterraum $\mathcal{D} \subseteq \mathcal{H}$ mit $\dim(\mathcal{D}) \leq \dim(\mathcal{A})$ enthalten. Im abzählbar-dimensionalen Fall sind das gerade die *separablen Hilberträume*. △

In den folgenden Abschnitten betrachten wir spezielle Beispiele von Algebren, für die wir teilweise noch deutlich stärkere Positivstellensätze beweisen können.

9.3 Nicht-kommutative Polynome

In diesem Abschnitt betrachten wir die Algebra $\mathcal{A} = \mathbb{C}\langle z\rangle = \mathbb{C}\langle z_1, \ldots, z_n\rangle$ der Polynome in *nicht-kommutierenden Variablen*. $\mathcal{A}$ wird auch *freie Algebra* genannt. Ein *Wort* (oder *Monom*) ω in den Variablen $z_1, \ldots, z_n$ ist ein endliches formales Produkt

$$\omega = z_{i_1} z_{i_2} \cdots z_{i_k}$$

der Variablen. Die Zahl k nennt man dabei *Wortlänge* (oder *Grad*) von ω. Da die Variablen hier nicht kommutieren, sind beispielsweise $z_1 z_2$ und $z_2 z_1$ unterschiedliche Wörter. Zwei Wörter werden durch das formale Aneinanderhängen miteinander multipliziert, z.B. ist

$$z_1 z_3 \cdot z_2 z_1 = z_1 z_3 z_2 z_1.$$

Beim Multiplizieren addieren sich die Wortlängen.
Als Vektorraum besitzt $\mathbb{C}\langle z\rangle = \mathbb{C}\langle z_1, \ldots, z_n\rangle$ die Menge aller Wörter als Basis:

$$\mathbb{C}\langle z\rangle = \left\{ \sum_\omega c_\omega \cdot \omega \mid c_\omega \in \mathbb{C} \right\},$$

wobei die Summen immer endlich sind. Mit $\mathbb{C}\langle z\rangle_d$ bezeichnen wir den endlich-dimensionalen Unterraum der von allen Wörtern mit Länge $\leq d$ aufgespannt wird.

Die Multiplikation von Wörtern setzt sich eindeutig zu einer distributiven und assoziativen Multiplikation auf $\mathbb{C}\langle z\rangle$ fort, und macht es so zu einer Algebra, die für $n \geq 2$ nicht kommutativ ist. Schließlich betrachten wir noch eine Involution, die auf $\mathbb{C}\langle z\rangle$ eindeutig festgelegt ist durch die Bedingungen

$$z_i^* = z_i \text{ und } \lambda^* = \overline{\lambda}$$

für $\lambda \in \mathbb{C}$. Man beachte, dass Involutionen immer die Produktreihenfolge vertauschen, also hier nicht einfach nur auf die Koeffizienten eines Polynoms angewandt werden. Es ist zum Beispiel

$$(iz_1z_2)^* = -iz_2z_1.$$

Daraus folgt auch, dass $\mathbb{C}\langle z\rangle_h$ im Gegensatz zum kommutativen Fall hier *nicht* mit $\mathbb{R}\langle z\rangle$ übereinstimmt. Beispielsweise z_1z_2 nicht hermitesch, $iz_1z_2 - iz_2z_1$ hingegen schon. Man beachte allerdings dass

$$\sum \mathbb{C}\langle z\rangle^2 \cap \mathbb{R}\langle z\rangle = \sum \mathbb{R}\langle z\rangle^2$$

trotzdem gilt.

Proposition 9.3.1. *Im reellen Vektorraum $\mathbb{C}\langle z\rangle_h$ gilt*

$$\sum \mathbb{C}\langle z\rangle^2 = \left(\sum \mathbb{C}\langle z\rangle^2\right)^{\vee\vee}.$$

Beweis. Das beweist man im wesentlich gleich wie im kommutativen Fall in Abschnitt 7.2 (Übungsaufgabe 69). □

Bemerkung 9.3.2. Da die Variablen z_i keine Relationen erfüllen (außer $z_i^* = z_i$), sind die $*$-Darstellungen von $\mathbb{C}\langle z\rangle$ einfach zu beschreiben. Für jedes n-Tupel von selbstadjungierten Operatoren $T_1, \ldots, T_n \in \mathcal{L}(\mathcal{D})$ (bzw. $\mathcal{B}(\mathcal{H})$) erhält man eine $*$-Darstellung

$$\begin{aligned} \pi\colon \mathbb{C}\langle z_1, \ldots, z_n\rangle &\to \mathcal{L}(\mathcal{D}) \quad (\text{bzw. } \mathcal{B}(\mathcal{H})) \\ p &\mapsto p(T_1, \ldots, T_n), \end{aligned}$$

und jede $*$-Darstellung ist von dieser Gestalt. △

Obwohl die Quadratsummen in $\mathbb{C}\langle z\rangle$ nicht archimedisch sind, gibt es überraschenderweise den folgenden starken Positivstellensatz, der im kommutativen Fall gerade nicht stimmt (vergleiche Beispiel 9.2.11):

Satz 9.3.3 (Satz von Helton). *Sei $\mathcal{F}$ die Menge der endlich-dimensionalen $*$-Darstellungen von $\mathbb{C}\langle z_1, \dots, z_n\rangle$. Dann gilt für $p \in \mathbb{C}\langle z_1, \dots, z_n\rangle_h$*

$$p \text{ ist } \mathcal{F}\text{-nichtnegativ} \quad \Leftrightarrow \quad p \in \sum \mathbb{C}\langle z_1, \dots, z_n\rangle^2.$$

Beweis. Nach Satz 9.2.9 und Proposition 9.3.1 reicht es zu zeigen, dass aus der Nichtnegativität von p in den endlich-dimensionalen $*$-Darstellungen bereits die Nichtnegativität in allen $*$-Darstellungen folgt.
Sei also $\mathcal{D}$ ein Vektorraum mit Skalarprodukt und seien $T_1, \dots, T_n \in \mathcal{L}(\mathcal{D})$ selbstadjungierte Operatoren. Wir müssen für jeden Vektor $v \in \mathcal{D}$ zeigen

$$\langle p(T_1, \dots, T_n)v, v\rangle \geq 0.$$

Wähle $d \in \mathbb{N}$ mit $p \in \mathbb{C}\langle z\rangle_d$, d.h. in p kommen höchstens Wörter der Länge d vor. Nun betrachten wir

$$\mathcal{H} := \{q(T_1, \dots, T_n)v \mid q \in \mathbb{C}\langle z\rangle_d\}.$$

Offensichtlich ist $\mathcal{H}$ ein endlich-dimensionaler Untervektorraum von $\mathcal{D}$ mit $v \in \mathcal{H}$. Sei $P \in \mathcal{L}(\mathcal{D})$ die orthogononale Projektion auf $\mathcal{H}$, d.h. wenn $u_1, \dots, u_r$ eine Orthonormalbasis von $\mathcal{H}$ ist, ist

$$P(w) = \sum_{j=1}^{r} \langle w, u_j\rangle \cdot u_j$$

für alle $w \in \mathcal{D}$. Setze nun

$$M_i := P \circ T_i \in \mathcal{L}(\mathcal{H}) = \mathcal{B}(\mathcal{H}) \quad \text{für } i = 1, \dots, n.$$

Als endlich-dimensionaler Raum ist $\mathcal{H}$ nämlich automatisch ein Hilbertraum und jede lineare Abbildung ist stetig.
Man überlegt sich nun leicht dass die M_i auf $\mathcal{H}$ wieder selbstadjungiert sind (man verwendet dabei dass auch P selbstadjungiert ist). Auf diese Weise erhalten wir eine neue, und diesmal endlich-dimensionale $*$-Darstellung

$$\begin{aligned} \pi \colon \mathbb{C}\langle z\rangle &\to \mathcal{B}(\mathcal{H}) \\ q &\mapsto q(M_1, \dots, M_n). \end{aligned}$$

Aufgrund der Definition von $\mathcal{H}$ gilt für ein Produkt $M_{i_1} \cdots M_{i_k}$ mit $k \leq d$ nun

$$(M_{i_1} \circ \cdots \circ M_{i_k})v = (T_{i_1} \circ \cdots \circ T_{i_k})v.$$

Solange man nämlich höchstens d viele der T_i nacheinander auf v anwendet bleibt man in $\mathcal{H}$, und muss deshalb P gar nicht anwenden. Insbesondere gilt

$$p(M_1, \dots, M_n)v = p(T_1, \dots, T_n)v$$

und damit

$$\langle p(T_1, \dots, T_n)v, v\rangle = \langle p(M_1, \dots, M_n)v, v\rangle \geq 0,$$

wobei wir die Nichtnegativität von p in allen endlich-dimensionalen $*$-Dar-stellungen verwendet haben. □

Bemerkung 9.3.4. (i) In Satz 9.3.3 bedeutet die Bedingung $\mathcal{F}$-nichtnegativ zu sein gerade

$$p(M_1, \dots, M_n) \succeq 0$$

für alle n-Tupel von hermiteschen Matrizen $M_1, \dots, M_n \in \mathrm{Her}_d(\mathbb{C})$ (und alle $d \geq 1$).
(ii) In Satz 9.3.3 kann man auch die Dimensionen der Darstellungen beschränken. Im Beweis hat der Raum $\mathcal{H}$ Dimension höchstens

$$\dim_{\mathbb{C}}\langle z\rangle_d = \frac{n^{d+1}-1}{n-1}$$

falls $p \in \mathbb{C}\langle z\rangle_d$. Man muss die Nichtnegativität von p also nur auf Matrizen dieser Größe testen.
(iii) Sei $m \in \mathbb{C}\langle z_1, z_2\rangle_h$ eine nicht-kommutative Variante des Motzkinpolynoms (d.h. macht man die Variablen kommutativ entsteht das bekannte Motzkinpolynom). Man kann zum Beispiel

$$m = \frac{1}{2}z_1^4z_2^2 + \frac{1}{2}z_2^2z_1^4 + \frac{1}{2}z_1^2z_2^4 + \frac{1}{2}z_2^4z_1^2 - \frac{3}{2}z_1^2z_2^2 - \frac{3}{2}z_2^2z_1^2 + 1$$

nehmen. Dann gibt es ein $2 \leq d \leq 127$ und $M_1, M_2 \in \mathrm{Her}_d(\mathbb{C})$ so dass

$$m(M_1, M_2) \in \mathrm{Her}_d(\mathbb{C})$$

nicht positiv semidefinit ist. Andernfalls wäre m nach Satz 9.3.3 nämlich eine hermitesche Quadratsumme, und nach Kommutativmachung der Variablen wäre das kommutative Motzkinpolynom ebenfalls eine Quadratsumme. △

9.4 Gruppenalgebren

Sei Γ eine (multiplikativ geschriebene) Gruppe mit neutralem Element e. Die (komplexe) Gruppenalgebra $\mathbb{C}\Gamma$ besitzt als Vektorraum gerade die Elemente von Γ als Basis:

$$\mathbb{C}\Gamma = \left\{ \sum_{g\in\Gamma} c_g \cdot g \mid c_g \in \mathbb{C}, \text{ nur endlich viele } c_g \neq 0 \right\}.$$

Die Multiplikation auf Γ liefert eine Multiplikation auf den Basisvektoren von $\mathbb{C}\Gamma$, und damit eine distributive und assoziative Multiplikation auf $\mathbb{C}\Gamma$:

$$\left(\sum_g c_g \cdot g\right) \cdot \left(\sum_g c'_g \cdot g\right) = \sum_g \left(\sum_{f\cdot h=g} c_f c'_h\right) \cdot g.$$

Auf diese Weise wird $\mathbb{C}\Gamma$ zu einer Algebra, die genau dann kommutativ ist, wenn Γ kommutativ war. Das neutrale Element bezüglich der Multiplikation ist $1 \cdot e$. Wir versehen $\mathbb{C}\Gamma$ mit der folgenden Involution:

$$\left(\sum_g c_g \cdot g\right)^* = \sum_g \overline{c_g} \cdot g^{-1} = \sum_g \overline{c_{g^{-1}}} \cdot g.$$

Ein Element $\sum_g c_g \cdot g$ ist also genau dann hermitesch, wenn

$$\overline{c_g} = c_{g^{-1}}$$

für alle $g \in \Gamma$ gilt.

Beispiel 9.4.1. (i) Sei $\Gamma = \mathbb{Z}$ mit der Addition als Gruppenverknüpfung. Dann ist $\mathbb{C}\mathbb{Z}$ kommutativ und es gibt einen Algebrenisomorphismus

$$\mathbb{C}\mathbb{Z} \cong \mathbb{C}[t, t^{-1}]$$

mit der Algebra der Laurentpolynome in einer Variablen. Dabei entspricht dem Basisvektor $i \in \mathbb{Z}$ von $\mathbb{C}\mathbb{Z}$ auf der rechten Seite gerade t^i. Die Involution erfüllt dann $(t^i)^* = t^{-i}$. Man kann $\mathbb{C}[t, t^{-1}]$ auch weiter mit

$$\mathbb{C}[x, y]/(xy - 1)$$

identifizieren, der Algebra der polynomialen Funktionen auf der Varietät $\{xy = 1\}$. Dabei entsprich t gerade x und t^{-1} entspricht y. Damit die hermiteschen Elemente gerade die Polynome mit reellen Koeffizienten sind, würde man hier allerdings die Involution $x^* = x$ und $y^* = y$ wählen wollen, die dann nicht mit der ursprünglichen Involution auf $\mathbb{C}\mathbb{Z}$ übereinstimmt. Man kann aber $\mathbb{C}[t, t^{-1}]$ auch mit den hermiteschen Elementen

$$a = \frac{t + t^{-1}}{2} \text{ und } b = \frac{t - t^{-1}}{2i}$$

erzeugen. Sie erfüllen die Relationen

$$a^2 + b^2 = 1,$$

und also gibt es einen $*$-Algebrenisomorphismus

$$\mathbb{C}\mathbb{Z} \cong \mathbb{C}[x, y]/(x^2 + y^2 - 1) = \mathbb{C}[S^1]$$

mit der Algebra der polynomialen Funktionen auf dem Einheitskreis, diesmal mit der üblichen Involution $x^* = x, y^* = y$. Ein hermitesches Element ist hier dann ein Polynom mit reellen Koeffizienten. Allgemeiner ist $\mathbb{C}\mathbb{Z}^n$ isomorph zur Algebra

$$\mathbb{C}[\underbrace{S^1 \times \cdots \times S^1}_{n}]$$

der polynomialen Funktionen auf dem n-dimensionalen Torus, mit der kanonischen Involution.
(ii) Ist $\Gamma = S_3$ die Permutationsgruppe von 3 Elementen, erhält man eine 6-dimensionale Algebra $\mathbb{C}S_3$ die nicht kommutativ ist.
(iii) Mit $\Gamma = F_n$ bezeichnen wir die *freie Gruppe* mit n Erzeugern (gewöhnlich $z_1, \ldots, z_n$). Ein Element in F_n ist also gerade ein Wort in den Buchstaben

$$z_1, \ldots, z_n \text{ und } z_1^{-1}, \ldots, z_n^{-1}.$$

Die einzigen Relationen dabei sind

$$z_i^{-1} z_i = z_i z_i^{-1} = e$$

für alle i, wobei e das leere Wort bezeichne. Die Gruppenverknüpfung ist das Aneinanderhängen von Worten. Die Gruppenalgebra $\mathbb{C}F_n$ enthält die freie Algebra $\mathbb{C}\langle z_1, \ldots, z_n\rangle$ als Unteralgebra, *nicht* jedoch als $*$-Unteralgebra. In $\mathbb{C}\langle z\rangle$ gilt ja $z_i^* = z_i$, in der Gruppenalgebra $\mathbb{C}F_n$ gilt $z_i^* = z_i^{-1}$. △

Bemerkung 9.4.2. Wir beschreiben wieder die $*$-Darstellungen von $\mathbb{C}\Gamma$. Jede $*$-Darstellung

$$\pi\colon \mathbb{C}\Gamma \to \mathcal{L}(\mathcal{D})$$

liefert einen Gruppenhomomorphismus

$$\begin{aligned}\pi\colon \Gamma &\to \mathcal{U}(\mathcal{D})\\ g &\mapsto \pi(g).\end{aligned}$$

in die Gruppe der unitären Operatoren auf $\mathcal{D}$. Dabei heißt $T \in \mathcal{L}(\mathcal{D})$ *unitär,* falls ein $S \in \mathcal{L}(\mathcal{D})$ existiert mit

$$\langle Tv, w\rangle = \langle v, Sw\rangle \ \forall v, w \in \mathcal{D} \text{ und } TS = ST = \mathrm{id}_{\mathcal{D}}.$$

Dafür verwenden wir dass $g^* = g^{-1}$ in $\mathbb{C}\Gamma$ gilt. Daraus folgt auch, dass jede Darstellung von $\mathbb{C}\Gamma$ eigentlich eine beschränkte Darstellung ist, denn unitäre Operatoren sind beschränkt von Norm 1. Umgekehrt definiert jeder Gruppenhomomorphismus

$$\pi\colon \Gamma \to \mathcal{U}(\mathcal{D})$$

eine $*$-Darstellung von $\mathbb{C}\Gamma$ durch die Vorschrift

$$\sum_g c_g \cdot g \mapsto \sum_g c_g \cdot \pi(g). \qquad \triangle$$

Beispiel 9.4.3. Die Gruppenhomomorphismen der freie Gruppe F_n in eine beliebige Gruppe G erhält man gerade durch die Vorgabe eines beliebigen Bildes $g_i \in G$ für jeden der Erzeuger z_i von F_n. Ein Wort wie etwa $z_1 z_2^{-1} z_1 z_3$ wird dann auf $g_1 g_2^{-1} g_1 g_3$ abgebildet.
Die $*$-Darstellungen von $\mathbb{C}F_n$ entstehen also gerade durch die Vorgabe von n unitären Operatoren $U_i \in \mathcal{U}(\mathcal{H})$ auf einem Hilbertraum. $\triangle$

Proposition 9.4.4. *Für jede Gruppe Γ ist $\sum \mathbb{C}\Gamma^2$ archimedisch.*

Beweis. Für $a = \sum_g c_g \cdot g \in \mathbb{C}\Gamma$ setzen wir

$$\|a\|_1 = \sum_g |c_g|.$$

Dann rechnet man die folgende Formel direkt nach:

$$\|a\|_1^2 - a^*a = \frac{1}{2}\sum_{g,h\in\Gamma} |c_g c_h|\left(1 - \frac{c_g\overline{c}_h}{|c_g c_h|}h^{-1}g\right)^*\left(1 - \frac{c_g\overline{c}_h}{|c_g c_h|}h^{-1}g\right).$$

Also ist $\|a\|_1^2 - a^*a \in \sum \mathbb{C}\Gamma^2$. Für ein hermitesches Element $a \in \mathbb{C}\Gamma_h$ und $r = \|a\|_1$ gilt damit

$$r - a = \frac{1}{2r}\left((r-a)^*(r-a) + (r^2 - a^*a)\right) \in \sum \mathbb{C}\Gamma^2. \qquad \square$$

Der nächste Satz folgt aus dem Satz von Schmüdgen (Satz 4.2.3):

Satz 9.4.5. *Sei Γ eine endlich erzeugte abelsche Gruppe und sei $\mathcal{F}$ die Menge der endlich-dimensionalen $*$-Darstellungen von $\mathbb{C}\Gamma$. Dann gilt für $a \in \mathbb{C}\Gamma_h$*

$$a \text{ ist } \mathcal{F}\text{-nichtnegativ} \;\Leftrightarrow\; a + \varepsilon \in \sum \mathbb{C}\Gamma^2 \quad \forall \varepsilon > 0.$$

Beweis. Übungsaufgabe 70, man lasse sich von Beispiel 9.4.1 (i) anleiten. $\square$

Für die Gruppenalgebra der freien Gruppe können wir den gleichen Positivstellensatz beweisen:

Satz 9.4.6. *Sei $\Gamma = F_n$ die freie Gruppe und $\mathcal{F}$ die Menge der endlich-dimensionalen $*$-Darstellungen von $\mathbb{C}\Gamma$. Dann gilt für $a \in \mathbb{C}\Gamma_h$*

$$a \text{ ist } \mathcal{F}\text{-nichtnegativ} \;\Leftrightarrow\; a + \varepsilon \in \sum \mathbb{C}\Gamma^2 \quad \forall \varepsilon > 0.$$

Beweis. Nach Satz 9.2.13 und Proposition 9.4.4 reicht es zu zeigen, dass die $\mathcal{F}$-Nichtnegativität von a die Nichtnegativität in jeder (beschränkten) $*$-Darstellung impliziert. Sei also

$$\pi \colon \mathbb{C}\Gamma \to \mathcal{B}(\mathcal{H})$$

eine $*$-Darstellung und $v \in \mathcal{H}$ fest gewählt. Wir müssen

$$\langle \pi(a)v, v\rangle \geq 0$$

zeigen. Sei $d \in \mathbb{N}$ so, dass $a \in \mathbb{C}\Gamma_d$, d.h. eine Linearkombination von Worten in den z_i und z_i^{-1} der Länge höchstens d ist. Wir betrachten den endlich-dimensionalen Untervektorraum

$$\mathcal{H}' = \{\pi(b)v \mid b \in \mathbb{C}\Gamma_d\}$$

von $\mathcal{H}$ und die orthogonale Projektion $P \colon \mathcal{H} \to \mathcal{H}$ auf $\mathcal{H}'$. Mit

$$T_i := \pi(z_i) \in \mathcal{U}(\mathcal{H})$$

setzen wir

$$M_i := P \circ T_i \in \mathcal{B}(\mathcal{H}').$$

Dann gilt $M_i^* = P \circ T_i^* = P \circ T_i^{-1}$ auf $\mathcal{H}'$. Für ein Produkt $M_{i_1}^{\delta_1} \circ \cdots \circ M_{i_k}^{\delta_k}$ mit $\delta_i \in \{1, *\}$ und $k \leq d$ gilt also wie im Beweis von Satz 9.3.3

$$(M_{i_1}^{\delta_1} \circ \cdots \circ M_{ik}^{\delta_k})v = (T_{i_1}^{\delta_1} \circ \ldots \circ T_{i_k}^{\delta_k})v.$$

Nun sind die M_i im Allgemeinen nicht unitär und definieren deshalb keine neue und endlich-dimensionale $*$-Darstellung von $\mathbb{C}\Gamma$. Allerdings sind die M_i *Kontraktionen*, d.h.

$$\mathrm{id}_{\mathcal{H}'} - M_i^* M_i \text{ und } \mathrm{id}_{\mathcal{H}'} - M_i M_i^*$$

sind positiv semidefinit auf $\mathcal{H}'$. Das rechnet man direkt nach, unter Verwendung dass die T_i unitär und also normerhaltend sind. Die folgende Konstruktion ist bekannt als *Choi's Matrixtrick*. Man setzt

$$\widetilde{M_i} := \begin{pmatrix} M_i & \sqrt{\mathrm{id}_{\mathcal{H}'} - M_i M_i^*} \\ \sqrt{\mathrm{id}_{\mathcal{H}'} - M_i^* M_i} & -M_i^* \end{pmatrix} \in \mathcal{B}(\mathcal{H}' \oplus \mathcal{H}').$$

Die $\widetilde{M_i}$ sind nun unitär (Übungsaufgabe 71), und für ein Produkt $\widetilde{M}_{i_1}^{\delta_1} \circ \cdots \circ \widetilde{M}_{i_k}^{\delta_k}$ wie oben und $\tilde{v} := (v, 0) \in \mathcal{H}' \oplus \mathcal{H}'$ gilt

$$\langle (\widetilde{M}_{i_1}^{\delta_1} \circ \cdots \circ \widetilde{M}_{i_k}^{\delta_k})\tilde{v}, \tilde{v} \rangle = \langle (M_{i_1}^{\delta_1} \circ \cdots \circ M_{i_k}^{\delta_k})v, v \rangle = \langle (T_{i_1}^{\delta_1} \circ \cdots \circ T_{i_k}^{\delta_k})v, v \rangle$$

(Übungsaufgabe 71). Da die $\widetilde{M_i}$ als unitäre Operatoren eine endlich-dimensionale $*$-Darstellung $\tilde{\pi}$ von $\mathbb{C}\Gamma$ liefern, und $a \in \mathbb{C}\Gamma_d$ gilt, folgt

$$0 \leq \langle \tilde{\pi}(a)\tilde{v}, \tilde{v} \rangle = \langle \pi(a)v, v \rangle. \qquad \square$$

Bemerkung 9.4.7. (i) Man kann in Satz 9.4.6 sogar noch die Aussage

$$a + \varepsilon \in \sum \mathbb{C}\Gamma^2 \quad \forall \varepsilon > 0$$

verstärken zu

$$a \in \sum \mathbb{C}\Gamma^2.$$

Dann wird der Beweis aber nochmal deutlich technischer.
(ii) Für $\Gamma = \mathbb{Z}^n$ mit $n \geq 3$ erhält man in Satz 9.4.5 diese stärkere Bedingung $a \in \sum \mathbb{C}\Gamma^2$ *nicht*! Im Kommutativen gibt es ab Dimension drei nach Satz 7.3.4 ja keine saturierten endlich erzeugten Präordnungen mehr (wir haben das allerdings nur für Mengen mit nichtleerem Inneren im affinen Raum $\mathbb{R}^n$ bewiesen; es stimmt aber auch für Varietäten). $\triangle$

9.5 Matrixpolynome

Sei $\mathcal{A} = \mathrm{M}_d(\mathbb{C}[x_1, \ldots, x_n])$ die Algebra der $d \times d$-Matrizen mit *polynomialen* Einträgen (allerdings *kommutative* Polynome). Addition und Multiplikation sind kanonisch definiert, und die Involution wie folgt:

$$(p_{ij})^* := \left(p_{ji}^*\right)$$

wobei $(\sum_\alpha c_\alpha x^\alpha)^* = \sum_\alpha \overline{c}_\alpha x^\alpha$ die kanonische Involution auf dem Polynomring $\mathbb{C}[x] = \mathbb{C}[x_1, \ldots, x_n]$ ist. Elemente von $\mathcal{A}$ heißen auch *Matrixpolynome* (oder *polynomiale Matrizen*). Jeder Punkt $a \in \mathbb{R}^n$ liefert eine endlich-dimensionale $*$-Darstellung

$$\begin{aligned} \pi_a \colon \mathrm{M}_d(\mathbb{C}[x]) &\to \mathrm{M}_d(\mathbb{C}) \\ M = (p_{ij}) &\mapsto M(a) = (p_{ij}(a)). \end{aligned}$$

Sei $\mathcal{F} = \{\pi_a \mid a \in \mathbb{R}^n\}$. Ein hermitesches Matrixpolynom ist genau dann $\mathcal{F}$-nichtnegativ, wenn es punktweise auf dem $\mathbb{R}^n$ positiv semidefinit ist.
Der folgende Satz sieht zwar aus wie ein Satz der nicht-kommutativen reellen algebraischen Geometrie, ist allerding doch eher ein Satz der kommutativen Theorie, wie man am Beweis sieht. Hilberts 17. Problem entspricht dem Fall $d = 1$.

Satz 9.5.1 (Gondard & Ribenboim)**.** *Es sei* $\mathcal{A} = \mathrm{M}_d(\mathbb{C}[x_1, \ldots, x_n])$ *und* $\mathcal{F} = \{\pi_a \mid a \in \mathbb{R}^n\}$. *Dann gilt für* $M \in \mathcal{A}_h$

$$M \text{ ist } \mathcal{F}\text{-nichtnegativ} \;\Leftrightarrow\; p^*p \cdot M \in \sum \mathcal{A}^2 \text{ für ein } 0 \neq p \in \mathbb{C}[x_1, \ldots, x_n].$$

Beweis. "$\Leftarrow$": Für $a \in \mathbb{R}^n$ ist

$$0 \preceq \pi_a(p^*p \cdot M) = |p(a)|^2 \cdot M(a).$$

Für $p(a) \neq 0$ folgt daraus schon $M(a) \succeq 0$, und aus Dichtheitsgründen gilt dasselbe für alle $a \in \mathbb{R}^n$.
"$\Rightarrow$" Wir nehmen zunächst $M \in \mathrm{M}_d(\mathbb{R}[x])$ an, d.h. M ist symmetrisch. Über dem Körper $\mathbb{R}(x)$ lässt sich M dann wie folgt diagonalisieren:

$$M = P^t \cdot \mathrm{diag}(f_1, \ldots, f_d) \cdot P \tag{9.1}$$

für ein $P \in \mathrm{Gl}_d(\mathbb{R}(x))$. Aus der $\mathcal{F}$-Nichtnegativität von M folgt dass die $f_i \in \mathbb{R}(x)$ überall dort nichtnegativ sind, wo sie definiert sind. Aus Satz 2.1.1 folgt

$f_i \in \sum \mathbb{R}(x)^2$, wie man sich leicht überlegt. Durch Multiplizieren von Gleichung (9.1) mit einem geeigneten p^2 kann man alle auftretenden Nenner auf der rechten Seite aufheben, und das liefert die Behauptung.
Der allgemeine Fall $M \in \mathrm{M}_d(\mathbb{C}[x])_h$ wird auf den reellen Fall zurückgeführt. Schreibe $M = M_1 + iM_2$ mit $M_1, M_2 \in \mathrm{M}_d(\mathbb{R}[x])$. Dann ist

$$\widetilde{M} = \begin{pmatrix} M_1 & M_2 \\ -M_2 & M_1 \end{pmatrix} \in \mathrm{M}_{2d}(\mathbb{R}[x])_h$$

symmetrisch und wieder punktweise positiv semidefinit, und man erhält das Ergebnis für M aus der Aussage für $\widetilde{M}$ (Übungsaufgabe 72). □

9.6 Connes' Einbettungsvermutung

Im letzten Abschnitt beschreiben wir noch eine bekannte offene Vermutung, die sich in der Sprache der nicht-kommutativen reellen algebraischen Geometrie formulieren lässt. Die sogenannte *Einbettungsvermutung von Alain Connes* aus dem Jahr 1976 ist eine Frage aus der Theorie der Operatoralgebren. In ihrer ursprünglichen Formulierung ist sie nicht leicht zu verstehen, aber es gibt eine äquivalente Formulieren von Klep und Schweighofer, die wir hier erklären können.
Sei dazu $\mathcal{A} = \mathbb{R}\langle z\rangle = \mathbb{R}\langle z_1, \ldots, z_n\rangle$ der nicht-kommutative Polynomring über $\mathbb{R}$, mit der Involution $z_i^* = z_i$. Anstelle der Quadratsummen betrachten wir nun einen endlich erzeugten quadratischen Modul:

$$M = M(1-z_1^*z_1, \ldots, 1-z_n^*z_n) = \left\{ \sum_{i=0}^{n} \sum_{j=1}^{m} a_{ij}^*(1 - z_i^*z_i)a_{ij} \mid m \in \mathbb{N}, a_i \in \mathcal{A} \right\}.$$

Dabei setzen wir $z_0 := 0$, um $\sum \mathcal{A}^2 \subseteq M$ zu ereichen. M ist gerade die kleinste Menge die 1 und alle $1 - z_i^*z_i$ enthält, und $M + M \subseteq M$ sowie $a^*Ma \subseteq M$ für alle $a \in \mathcal{A}$ erfüllt. Analog wie in Proposition 4.2.1 kann man sich überlegen, dass M archimedisch ist, bzw. 1 als algebraisch inneren Punkt enthält.
Sei weiter C der von den Kommutatoren in $\mathcal{A}$ aufgespannte Unterraum, d.h.

$$C = \left\{ \sum_{i=1}^{m} (a_ib_i - b_ia_i) \mid m \in \mathbb{N}, a_i, b_i \in \mathcal{A} \right\}.$$

C ist ein unter $*$ abgeschlossener Untervektorraum von $\mathcal{A}$.

Eine *selbstadjungierte Kontraktion* ist eine symmetrische Matrix $Q \in \mathrm{Sym}_d(\mathbb{R})$ mit $I_d - Q^*Q \succeq 0$. Mit tr bezeichen wir die Spur einer quadratischen Matrix, d.h. die Summe ihrer Diagonaleinträge bzw. ihrer Eigenwerte.

Vermutung 9.6.1 (Connes' Einbettungsvermutung, Version von Klep und Schweighofer)**.** *Sei $p \in \mathbb{R}\langle z\rangle_h$ mit*

$$\mathrm{tr}(p(Q_1, \ldots, Q_n)) \geq 0$$

für alle n-Tupel selbstadjungierter Kontraktionen Q_i (von jeder Größe d). Dann gilt

$$p + \epsilon \in M + C$$

für alle $\varepsilon > 0$.

Bemerkung 9.6.2. (i) Elemente aus $M + C$ liefern bei Einsetzung von selbstadjungierten Kontraktionen immer eine nichtnegative Spur. Dafür verwendet man dass Kommutatoren immer Spur Null haben (da $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ gilt), und positiv semidefinite Matrizen nichtnegative Spur haben. Aus $p + \epsilon \in M + C$ für alle $\epsilon > 0$ folgt also offensichtlich $\mathrm{tr}(p(Q_1, \ldots, Q_n)) \geq 0$. Die Einbettungsvermutung fragt gerade nach der Umkehrung.
(ii) Die Einbettungsvermutung ist eine nicht-kommutative Spurversion des archimedischen Positivstellensatzes 5.4.3, und eine archimedische Spurversion des Satzes von Helton (Satz 9.3.3).
(iii) Die entsprechende Version der Vermutung *ohne Spur* und deshalb *ohne Kommutatoren* stimmt: ist

$$p(Q_1, \ldots, Q_n) \succeq 0$$

für alle Tupel von selbstadjungierten Kontraktionen, ist $p + \varepsilon \in M$ für alle $\varepsilon > 0$. Das ist eine Variante von Satz 9.3.3, die man ganz analog beweist (Übungsaufgabe 73). △

The End

(man bearbeite nun Übungsaufgabe 74)

Literaturverzeichnis

[1] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in real algebraic geometry*, vol. 10 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, second edn., 2006.

[2] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*, vol. 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3)*. Springer-Verlag, Berlin, 1998.

[3] M. Knebusch and C. Scheiderer. *Einführung in die reelle Algebra*, vol. 63 of *Vieweg Studium: Aufbaukurs Mathematik [Vieweg Studies: Mathematics Course]*. Friedr. Vieweg & Sohn, Braunschweig, 1989.

[4] M. Marshall. *Positive polynomials and sums of squares*, vol. 146 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2008.

[5] A. Prestel and C. N. Delzell. *Positive polynomials*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2001.

[6] C. Scheiderer. Positivity and sums of squares: a guide to recent results. In *Emerging applications of algebraic geometry*, vol. 149 of *IMA Vol. Math. Appl.*, pp. 271–324. Springer, New York, 2009.

[7] K. Schmüdgen. Noncommutative real algebraic geometry—some basic concepts and first ideas. In *Emerging applications of algebraic geometry*, vol. 149 of *IMA Vol. Math. Appl.*, pp. 325–350. Springer, New York, 2009.

Übungsaufgaben

Aufgabe 1. (i) Zeigen Sie, dass sich für $a \in \mathbb{R}$ die Anordnungen $\leq_{a_-}$ und $\leq_{a_+}$ auf $\mathbb{Q}(t)$ aus Beispiel 1.1.4 genau dann unterscheiden, wenn a algebraisch über $\mathbb{Q}$ ist. (ii) Zeigen Sie, dass die Anordnungen $\leq_{a_-}$ bzw. $\leq_{a_+}$ auf $\mathbb{Q}(t)$ genau dann archimedisch sind, wenn a transzendent über $\mathbb{Q}$ ist.

Aufgabe 2. Zeigen Sie, dass die in Satz 1.1.8 definierte Abbildung wirklich ein ordnungstreuer Ringhomomorphismus ist.

Aufgabe 3. Führen Sie den Beweis von Satz 1.1.12 aus.

Aufgabe 4. Zeigen Sie, dass die beiden Mengen

$$P_1 = \left\{ f/g \mid fg = \sum_{i=k}^{d} a_i t^i, a_d > 0 \right\} \cup \{0\}$$

$$P_2 = \left\{ f/g \mid fg = \sum_{i=k}^{d} a_i t^i, a_k > 0 \right\} \cup \{0\}$$

aus Beispiel 1.1.13 auf $\mathbb{R}(t)$ Anordnungen sind. Welchen der bereits bestimmten Anordnungen entsprechen sie?

Aufgabe 5. Bestimmen Sie alle Anordnungen von $\mathbb{Q}(\sqrt{2})$.

Aufgabe 6. Beweisen Sie Korollar 1.2.10.

Aufgabe 7. Beweisen Sie den Satz von Rolle für Polynome über reell abgeschlossenen Körpern (Satz 1.2.12).

Aufgabe 8. Sei K ein Körper und $v_1, \dots, v_d$ linear unabhängige (Spalten)-Vektoren aus K^n. Zeigen Sie:
(i) Die Matrix $M = \sum_{i=1}^{d} v_i v_i^t$ hat Rang d.
(ii] Die Matrix $N = \sum_{i=1}^{s} v_i v_i^t - \sum_{i=s+1}^{d} v_i v_i^t$ hat Signatur $s-(d-s)$, bezüglich jeder Anordnung von K.

Aufgabe 9. Sei $(K, \leq) \subseteq (L, \leq)$ eine Erweiterung angeordneter Körper. Sie heißt *archimedisch,* falls für alle $b \in L$ ein $a \in K$ existiert mit $b \leq a$. Zeigen Sie dass die Erweiterung immer archimedisch ist, wenn L/K algebraisch ist.

Aufgabe 10. Sei $(K, \leq)$ ein angeordneter Körper und $A \subseteq K$ ein Teilring. Zeigen Sie: Die Menge

$$\mathcal{O} := \{b \in K \mid \pm b \leq a \text{ für ein } a \in A\}$$

ist ebenfalls ein Teilring von K, mit den folgenden Eigenschaften:

- $A \subseteq \mathcal{O}$
- $b_1 \leq c \leq b_2$ mit $b_1, b_2 \in \mathcal{O} \Rightarrow c \in \mathcal{O}$ (Ordnungskonvexität)
- $c \in K \setminus \mathcal{O} \Rightarrow c^{-1} \in \mathcal{O}$ ($\mathcal{O}$ ist ein *Bewertungsring* von K).

Aufgabe 11. Sei $\mathcal{O} \subseteq K$ ein Bewertungsring des Körpers K (d.h. $a \in K \setminus \mathcal{O} \Rightarrow a^{-1} \in \mathcal{O}$). Zeigen Sie dass $\mathcal{O}$ genau ein maximales Ideal besitzt, und zwar

$$\mathfrak{m} = \left\{a \in \mathcal{O} \mid a^{-1} \notin \mathcal{O}\right\} = \mathcal{O} \setminus \mathcal{O}^{\times}.$$

Aufgabe 12. Sei $(K, \leq) = (K, P)$ ein angeordneter Körper und A ein Teilring. Sei $\mathcal{O}$ der konvexe Bewertungsring aus Aufgabe 10, $\mathfrak{m}$ sein maximales Ideal und $\pi \colon \mathcal{O} \to \mathcal{O}/\mathfrak{m}$ der Restklassenhomomorphismus. Zeigen Sie, dass

$$\pi(P \cap \mathcal{O})$$

eine Anordnung des Restklassenkörpers $\mathcal{O}/\mathfrak{m}$ ist. Falls $A = \mathbb{Z}$, so ist sie archimedisch.

Aufgabe 13. Bestimmen Sie die Hermite Matrix (siehe Definition 1.3.3) des allgemeinen Polynoms $p = t^3 + at^2 + bt + c \in \mathbb{R}[t]$. Geben Sie eine Bedingung an die Koeffizienten a, b, c an, die besagt, dass p mindestens zwei reelle Nullstellen hat.

Aufgabe 14. Sei $(K, \leq)$ ein angeordneter Körper und $0 \neq p \in K[t]$ ein Polynom, dessen Nullstellen alle in K liegen. Dann sind diese Nullstellen genau dann alle ≥ 0, wenn die Koeffizienten von p alternierend sind, d.h. $p = \sum_{i=0}^{d} a_i t^i$ mit $(-1)^i a_i \geq 0$.

Aufgabe 15. Geben Sie explizit eine Anordnung $\leq$ auf $\mathbb{R}(x, y)$ an, so dass für Polynome $p \in \mathbb{R}[x, y]$ gilt

$$0 < p(0, 0) \Rightarrow 0 \leq p.$$

Aufgabe 16. Zeigen Sie, dass die in Beispiel 1.5.4 angegebenen Mengen nicht semialgebraisch sind.

Aufgabe 17. (i) Formulieren Sie den Zwischenwertsatz für Polynome vom festen Grad d als $\mathbb{Z}$-Aussage.
(ii) Formulieren sie die Aussage, dass jedes Polynom (von festem Grad d) auf jedem Intervall $[a, b]$ ein Maximum annimmt, als $\mathbb{Z}$-Aussage.
(iii) Kann man die Archimedizität der Anordnung $\leq$ auf $\mathbb{R}$ als $\mathbb{R}$-Aussage formulieren?

Aufgabe 18. (i) Zeigen Sie Korollar 1.5.11, d.h. das A-polynomiale Bild einer A-semialgebraischen Menge ist wieder A-semialgebraisch.
(ii) Zeigen Sie dass der Abschluss und das Innere einer semialgebraischen Teilmenge von $\mathbb{R}^n$ wieder semialgebraisch ist.

Aufgabe 19. Sei R reell abgeschlossen. Eine Funktion $f\colon R \to R$ heiß *definierbar*, wenn ihr Graph

$$\Gamma(f) = \{(\alpha, f(\alpha)) \mid \alpha \in R\} \subseteq R^2$$

semialgebraisch ist.
(i) Zeigen Sie dass es ein nichttriviales Polynom $p \in R[x, y]$ gibt mit $p = 0$ auf $\Gamma(f)$.
(ii) Zeigen Sie dass es ein $q \in R[t]$ gibt mit $|f(\alpha)| \leq q(\alpha)$ für alle $\alpha \in R$ groß genug.
(iii) Gilt (ii) auch auf ganz R?

Aufgabe 20. Sei $S = \delta(R) = \delta'(R) \subseteq R^n$ eine semialgebraische Menge, definiert durch die Formeln δ und δ' in den freien Variablen $x_1, \ldots, x_n$.
(i) Zeigen Sie, dass für jeden reell abgeschlossenen Oberkörper R_1 von R gilt

$$\delta(R_1) = \delta'(R_1).$$

Man bezeichnet diese Menge mit S_{R_1}.
(ii) Ist $S \subseteq R^2$ der Graph einer Funktion, so ist auch $S_{R_1} \subseteq R_1^2$ der Graph einer Funktion.

Aufgabe 21. Sei $f\colon R \to R$ eine definierbare Abbildung. Nach Aufgabe 20 (ii) gibt es für jeden reell abgeschlossenen Oberkörper S von R eine kanonische Fortsetzung $f_S\colon S \to S$. Zeigen Sie dass sich Injektivität/Surjektivität von f auf f_S überträgt.

Aufgabe 22. Zeigen Sie, dass für Polynome $p, q \in K[x_1, \ldots, x_n]$ (über einem beliebigen Körper K) stets gilt

$$\mathcal{N}(pq) = \mathcal{N}(p) + \mathcal{N}(q).$$

Aufgabe 23. Sei $A = \mathbb{R}[[t]]$ der Ring der formalen Potenzreihen in einer Variablen.
(i) Zeigen Sie dass A ein Integritätsring ist.
(ii) Zeigen Sie dass $p \in A$ genau dann invertierbar ist, wenn $p = \sum_{i=0}^{\infty} p_i t^i$ mit $p_0 \neq 0$. Schließen Sie daraus, dass A genau ein maximales Ideal besitzt (d.h. A ist ein *lokaler Ring*).
(iii) Zeigen Sie dass $p \in A$ genau dann ein Quadrat in A ist, wenn $p = \sum_{i=k}^{\infty} p_i t^i$ mit k gerade und $p_k > 0$.
(iv) Bestimmen Sie alle Ringanordnungen von A.

Aufgabe 24. Beweisen Sie die Aussage aus Beispiel 3.1.5 (iv).

Aufgabe 25. Beweisen Sie Proposition 3.1.12.

Aufgabe 26. Sei X ein kompakter Hausdorffraum. Bestimmen Sie die maximalen Anordnungen des Rings $C(X, \mathbb{R})$ aller stetigen reellwertigen Funktionen auf X.

Aufgabe 27. Vervollständigen Sie den Beweis von Satz 3.1.22.

Aufgabe 28. Beweisen Sie Korollar 3.1.24.

Aufgabe 29. Beweisen Sie den abstrakten Nichtnegativstellensatz (Satz 3.2.3) und den abstrakten reellen Nullstellensatz (Satz 3.2.4).

Aufgabe 30. Zeigen Sie dass das Ideal $(1 - x^2 - y^2) \subseteq R[x, y]$ reell ist.

Aufgabe 31. Sei $T(p_1, \ldots, p_m) \subseteq R[t]$ die von den Polynomen $p_1, \ldots, p_m$ erzeugte Präordnung. Zeigen Sie:
(i) $t \notin T(t^3)$.
(ii) Falls $p \geq 0$ auf $[0, \infty)$, so $p \in T(t)$.
(iii) Falls $p \geq 0$ auf $[-1, 1]$, so $p \in T(1 - t, 1 + t)$.
(iv) $T(1 - t, 1 + t) = T(1 - t^2)$.

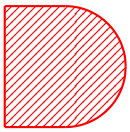
Aufgabe 32. Sei $R[[x_1, \ldots, x_n]]$ der Potenzreihenring in n Variablen. Zeigen Sie, dass für Polynome $p \in R[x_1, \ldots, x_n]$ gilt

$$p \geq 0 \text{ in einer Umgebung des Ursprungs} \Rightarrow \hat{p} \geq 0 \text{ auf } \mathrm{Sper}(R[[x_1, \ldots, x_n]]).$$

Aufgabe 33. Bezeichne B die Einheitskreisscheibe in $\mathbb{R}^2$. Zeigen Sie, dass die Menge

$$([-1, 0] \times [-1, 1]) \cup B \subseteq \mathbb{R}^2$$

nicht basisch abgeschlossen semialgebraisch ist, also nicht von der Gestalt $W_{\mathbb{R}}(p_1, \ldots, p_r)$ für gewisse $p_i \in \mathbb{R}[x, y]$.



Aufgabe 34. Sei I eine nichtleere Menge. Ein *Filter* auf I ist eine Teilmenge $\mathcal{F} \subseteq \mathcal{P}(I)$ der Potenzmenge von I, mit den folgenden drei Eigenschaften:

$$\emptyset \notin \mathcal{F} \qquad A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F} \qquad A \in \mathcal{F}, A \subseteq B \Rightarrow B \in \mathcal{F}$$

Ein *Ultrafilter* ist ein Filter $\mathcal{F}$ mit der zusätzlichen Eigenschaft

$$A \notin \mathcal{F} \Rightarrow I \setminus A \in \mathcal{F}.$$

Zeigen Sie:
(i) Für $i \in I$ ist $\mathcal{F}_i := \{A \subseteq I \mid i \in A\}$ ein Ultrafilter (solche Ultrafilter heißen *Haupt-Ultrafilter*).
(ii) Falls I unendlich ist, ist $\{A \subseteq I \mid I \setminus A \text{ ist endlich}\}$ ein Filter (genannt der *Filter der koendlichen Mengen*).
(iii) Jeder Filter ist ein einem Ultrafilter enthalten.
(iv) Ein Ultrafilter ist genau dann kein Haupt-Ultrafilter, wenn er den Filter der koendlichen Mengen enthält.

Aufgabe 35. Sei I eine unendliche Menge, und $\mathcal{F}$ ein Ultrafilter auf I. Sei $(K_i, \leq_i)$ ein angeordneter Körper, für jedes $i \in I$. Wir betrachten den kommutativen Ring

$$R = \prod_{i \in I} K_i = \{(a_i)_{i \in I} \mid a_i \in K_i \text{ für alle } i \in I\}.$$

Zeigen Sie:
(i) Die Menge

$$\mathfrak{m} = \{(a_i)_i \mid \{i \mid a_i = 0\} \in \mathcal{F}\}$$

ist ein maximales Ideal in R.
(ii) Die Relation

$$(a_i)_i \leq (b_i)_i :\Leftrightarrow \{i \mid a_i \leq_i b_i\} \in \mathcal{F}$$

induziert eine wohldefinierte Körperanordnung auf $R/\mathfrak{m}$.
(iii) Falls $\mathcal{F}$ kein Haupt-Ultrafilter und I abzählbar ist, ist der angeordnete Körper aus (2) nicht archimedisch.

Aufgabe 36. Sei I eine Indexmenge und $\mathcal{F}$ ein Ultrafilter auf I. Sei $(K_i, \leq_i)$ ein angeordneter Körper für jedes $i \in I$, und $(K, \leq)$ der angeordnete Körper, den wir in Aufgabe 35 (ii) konstruiert haben. Sei φ eine $\mathbb{Z}$-Formel in den freien Variablen $x_1, \ldots, x_n$, und seien $a_1 = \overline{(a_{1i})}_{i\in I}, \ldots, a_n = \overline{(a_{ni})}_{i\in I}$ Elemente aus $K = \left(\prod_{i\in I} K_i\right)/\mathfrak{m}$. Beweisen Sie den *Satz von Łos*:

In $(K, \leq)$ gilt $\varphi(a_1, \ldots, a_n) \Leftrightarrow \{i \in I \mid \varphi(a_{1i}, \ldots, a_{ni})$ gilt in $(K_i, \leq_i)\} \in \mathcal{F}$.

(Hinweis: Gehen Sie induktiv über den Aufbau der Formel vor; verwenden Sie statt $\forall$ lieber $\exists$.)

Aufgabe 37. Zeigen Sie, dass sich die Eigenschaft *archimedisch* eines angeordneten Körpers nicht als $\mathbb{Z}$-Aussage formulieren lässt (d.h. es gibt keine $\mathbb{Z}$-Aussage, die in einem angeordneten Körper genau dann gilt, wenn er archimedisch ist).

(Hinweis: Verwenden Sie Aufgabe 35 und 36.)

Aufgabe 38. Beweisen Sie die Behauptungen aus Beispiel 5.1.4.

Aufgabe 39. Finden Sie Polynome $p, p_1, \ldots, p_m \in \mathbb{R}[x, y]$, so dass

$$p > 0 \text{ auf } W_{\mathbb{R}}(p_1, \ldots, p_m)$$

gilt, aber nicht

$$\hat{p} > 0 \text{ auf } \widetilde{W}(p_1, \ldots, p_m) \subseteq \operatorname{Semisper} \mathbb{R}[x, y].$$

Aufgabe 40. Sei $\mathcal{F}$ ein Nicht-Hauptultrafilter auf $\mathbb{N}$, sei für alle $i \in \mathbb{N}$ jeweils $R_i = R$ derselbe reell abgeschlossene Körper, und sei schließlich

$$\widetilde{R} := \left(\prod_i R\right)/\mathfrak{m}$$

der reell abgeschlossene Erweiterungskörper von R, den wir in den Aufgaben 35 und 36 konstruiert und betrachtet haben (warum ist $\widetilde{R}$ reell abgeschlossen? warum enthält er R?). Zeigen Sie die sogenannte $\aleph_1$-*Saturiertheit* von $\widetilde{R}$, d.h.:

Wenn eine semialgebraische Menge in $\widetilde{R}^n$ durch abzählbar viele semialgebraische Mengen überdeckt wird, so reichen schon endlich viele davon aus.

(Hinweis: Falls keine endliche Überdeckung existiert, gibt es eine Folge von Elementen a_i, so dass a_i in der ursprünglichen Menge, aber nicht in den ersten i der Überdeckungsmengen liegt. Konstruieren Sie mit einem Diagonalargument ein neues Element a, das in der ursprünglichen Menge, aber in keiner der Überdeckungsmengen liegt.)

Aufgabe 41. Finden Sie $p_1, \ldots, p_m \in \mathbb{R}[x_1, \ldots, x_n]$, so dass $W_{\mathbb{R}}(p_1, \ldots, p_m) = \emptyset$ gilt, aber

$$-1 \notin M(p_1, \ldots, p_m).$$

Aufgabe 42. Zeigen Sie, dass man bei Diagonalmatrizen $M, M_1, \ldots, M_m$ die Bedingung der strikten Zulässigkeit im Dualitätssatz 6.1.4 weglassen kann (es handelt sich dabei dann um den Dualitätssatz der *linearen Optimierung*).

Aufgabe 43. Beweisen Sie das *Lemma von Farkas* (siehe auch Beispiel 6.4.5):

Seien $\ell_1, \ldots, \ell_m \in \mathbb{R}[x] = \mathbb{R}[x_1, \ldots, x_n]$ Polynome vom Grad ≤ 1, und

$$P = \{a \in \mathbb{R}^n \mid \ell_1(a) \geq 0, \ldots, \ell_m(a) \geq 0\},$$

der davon definierte Polyeder, sei nicht leer. Sei $\ell \in \mathbb{R}[x]$ ein weiteres Polynom vom Grad ≤ 1. Dann gilt

$$\ell \geq 0 \text{ auf } P \;\Leftrightarrow\; \ell = r_0 + r_1\ell_1 + \cdots + r_m\ell_m \text{ mit gewissen } r_i \in \mathbb{R}_{\geq 0}.$$

(Hinweis: Sie können zum Beispiel den Dualitätssatz der semidefiniten Optimierung mit geeigneten Diagonalmatrizen verwenden.)

Aufgabe 44. Beweisen Sie die *Existenz von Gradschranken in Hilberts 17. Problem*: Für $n, d \in \mathbb{N}$ gibt es eine Zahl $d' \in \mathbb{N}$, so dass für jeden reell abgeschlossenen Körper R und jedes global nichtnegative Polynom $p \in R[x_1, \ldots, x_n]$ vom Grad d ein Darstellung

$$q^2 p = p_1^2 + \cdots + p_m^2$$

existiert, mit Polynomen $q, p_1, \ldots, p_m \in \mathbb{R}[x_1, \ldots, x_n]$ vom Grad höchstens d'.

(Hinweis: Verwenden Sie die Lösung von Hilberts 17. Problem zunächst für einen $\aleph_1$-saturierten Körper R; vergleiche Aufgabe 40.)

Aufgabe 45. Finden Sie ein Beispiel für ein semidefinites Optimierungsproblem, in dem die starke Dualität $d^* = p^*$ nicht gilt.

Aufgabe 46. Finden Sie ein Beispiel für ein semidefinites Optimierungsproblem, in dem sowohl das primale als auch das duale Problem zulässige Punkte besitzen, aber in einem der beiden Probleme der Optimalwert nicht angenommen wird.

Aufgabe 47. Bestimmen Sie mit einer Methode Ihrer Wahl (z.B. der Lagrange-Methode) das Minimum und das Maximum des Polynoms $x^2 + y + 1$ auf der Einheitskreisscheibe $W(1 - x^2 - y^2)$ in der Ebene.

Aufgabe 48. Sei $S \subseteq \mathbb{R}^n$ ein spektraedrischer Kegel mit nichtleerem Inneren. Zeigen Sie dass es symmetrische Matrizen $M_1, \ldots, M_n$ gibt mit $S = \mathcal{S}(M_1, \ldots, M_n)$, so dass für das Innere von S gilt

$$\operatorname{int}(S) = \{a \in \mathbb{R}^n \mid a_1M_1 + \cdots + a_nM_n \succ 0\}.$$

Aufgabe 49. Sei $h \in \mathbb{R}[x_1, \ldots, x_n]$ hyperbolisch in Richtung $e \in \mathbb{R}^n$. Zeigen Sie
(i) Der Hyperbolizitätskegel $\Lambda_e(h)$ ist ein abgeschlossener konvexer Kegel.
(ii) Für $e' \in \operatorname{int}(\Lambda_e(h))$ ist h auch hyperbolisch in Richtung e'.
(iii) Für e' wie in (ii) gilt $\Lambda_e(h) = \Lambda_{e'}(h)$.
(Hinweis: Sie dürfen den Satz von Helton & Vinnikov 6.3.12 verwenden. Warum?)

Aufgabe 50. Zeigen Sie:
(i) Das Innere eines spektraedrischen Kegels ist ein spektraedrischer Schatten.
(ii) Das Innere eines spektraedrischen Schattens ist ein spektraedrischer Schatten.

Aufgabe 51. Zeigen Sie:
(i) Der duale Kegel eines spektraedrischen Kegels S ist ein spektraedrischer Schatten.
(Hinweis: o.B.d.A. hat S nichtleeres Inneres. Schauen Sie sich nun den Beweis des Dualitätssatzes 6.1.4 noch einmal an).
(ii) Der duale Kegel eines spektraedrischen Schattens ist ein spektradrischer Schatten.

Aufgabe 52. Zeigen Sie, dass Stabilität eines quadratischen Moduls nicht von den Erzeugern abhängt, d.h. falls $M(p_1, \ldots, p_r) = M(q_1, \ldots, q_s)$ gilt, und für alle $d \in \mathbb{N}$ ein $d' \in \mathbb{N}$ existiert mit

$$M(p_1, \ldots, p_r) \cap \mathbb{R}[x]_d \subseteq M_{d'}(p_1, \ldots, p_r),$$

so gilt dasselbe auf für $M(q_1, \ldots, q_s)$, eventuell mit einem anderen d'.

Aufgabe 53. Führen Sie den Grenzübergang aus Satz 7.2.6 exakt durch.

Aufgabe 54. Beweisen Sie die Aussage aus Beispiel 7.2.10 (i): Enthält die Menge $W_{\mathbb{R}}(p_1, \ldots, p_r)$ einen senkrechten und einen waagerechten Streifen, so ist $M(p_1, \ldots, p_r)$ stabil.

Aufgabe 55. Beweisen Sie die Aussage aus Beispiel 7.2.10 (ii): $M((1 - x^2)y^2)$ ist nicht stabil.

Aufgabe 56. Ein endlich erzeugter quadratischer Modul $M = M(q_1, \ldots, q_r)$ in $\mathbb{R}[x]$ heißt *entscheidbar*, falls die Menge $M \cap \mathbb{R}[x]_d$ für alle $d \in \mathbb{N}$ eine semialgebraische Teilmenge des endlich-dimensionalen Raums $\mathbb{R}[x]_d$ ist. Zeigen Sie:
(i) Falls M stabil ist, so ist M entscheidbar.
(ii) Falls M saturiert ist, so ist M entscheidbar.
(iii) Finden Sie ein Beispiel für einen unentscheidbaren quadratischen Modul?

Aufgabe 57. Seien $a, b, c, d \in \mathbb{R}$ mit $a \leq c \leq d \leq b$. Zeigen Sie dass es ein $\lambda \in [0, 1]$ gibt, so dass das Polynom

$$(t - c)(t - d) - \lambda(t - a)(t - b)$$

global nichtnegativ ist.

Aufgabe 58. Beweisen Sie alle Aussagen aus Abschnitt 8.2, die Ihnen nicht völlig klar sind.

Aufgabe 59. Finden Sie eine möglichst explizite Beschreibung der Komplettierungen der folgenden Ringe A bezüglich der jeweiligen Ideale I:
(*i*) $A = \mathbb{R}[x, y],\ I = (x)$
(*ii*) $A = \mathbb{Z}$, $I = (p)$ mit $p \in \mathbb{Z}$ prim.

Aufgabe 60. Seien $p, p_1, \ldots, p_r \in \mathbb{R}[x]$ und $p = 0$ auf $W_{\mathbb{R}}(p_1, \ldots, p_r)$. Zeigen Sie, dass

$$-p^{2m} \in T(p_1, \ldots, p_r)$$

für ein $m \geq 1$ gilt.

Aufgabe 61. Zeigen Sie die Aussage vom Anfang des Beweises von Satz 8.4.5, also dass die geometrische Nichtnegativität von Polynomen die entsprechende Zugehörigkeit zu den Anordnungen in $\mathbb{R}[[x, y]]$ impliziert.

Aufgabe 62. Seien $p_1, p_2 \in \mathbb{R}[x, y]$ Polynome, die am Ursprung mit linear unabhängigen Tangentenrichtungen verschwinden. Zeigen Sie

$$\mathbb{R}[[x, y]] = \mathbb{R}[[p_1, p_2]]$$

(und überlegen Sie sich, inwiefern diese Aussage überhaupt sinnvoll ist).

Aufgabe 63. Vervollständigen Sie den Beweis des Satzes von Burnside (Satz 9.1.2): Falls $\mathcal{A} \subseteq \mathrm{M}_d(\mathbb{C})$ eine Unteralgebra ist, die transitiv auf $\mathbb{C}^d$ operiert und eine Matrix von Rang 1 enthält, gilt $\mathcal{A} = \mathrm{M}_d(\mathbb{C})$.

Aufgabe 64. Sei $\varphi\colon \mathcal{A} \to \mathbb{C}$ ein Zustand. Zeigen Sie dass $\varphi(a^*) = \overline{\varphi(a)}$ für alle $a \in \mathcal{A}$ gilt.

Aufgabe 65. Beweisen Sie Lemma 9.2.5.

Aufgabe 66. Rechnen Sie alle Details der GNS-Konstruktion aus Abschnitt 9.2 nach.

Aufgabe 67. Sei $\varphi\colon \mathcal{A} \to \mathbb{C}$ ein Zustand mit $\varphi(1) = 0$. Zeigen Sie $\varphi \equiv 0$.

Aufgabe 68. Sei $\mathcal{D}$ ein Vektorraum mit Skalarprodukt und $\pi\colon \mathcal{A} \to \mathcal{L}(\mathcal{D})$ eine $*$-Darstellung, so dass $\pi(a)$ ein beschränkter Operator auf $\mathcal{D}$ ist, für alle $a \in \mathcal{A}$. Zeigen Sie, dass man

$$\pi\colon \mathcal{A} \to \mathcal{B}(\mathcal{H})$$

als beschränkte $*$-Darstellung auf der Vervollständigung $\mathcal{H}$ von $\mathcal{D}$ auffassen kann.

Aufgabe 69. Beweisen Sie Proposition 9.3.1.

Aufgabe 70. Beweisen Sie Satz 9.4.5.

Aufgabe 71. Überprüfen Sie Aussagen von Choi's Matrixtrick im Beweis von Satz 9.4.6.

Aufgabe 72. Führen Sie die Reduktion vom komplexen auf den reellen Fall im Beweis von Satz 9.5.1 durch.

Aufgabe 73. Beweisen Sie die Aussage aus Bemerkung 9.6.2 (iii).

Aufgabe 74. Lehnen Sie sich zufrieden zurück.