Trace-positive polynomials, sums of hermitian squares and the tracial moment problem

Dissertation zur Erlangung des akademischen
Grades eines Doktors der Naturwissenschaften

vorgelegt an der
Universität Konstanz, Fachbereich Mathematik und Statistik
und an der
Université de Rennes 1, IRMAR, U.F.R. de Mathématique

Sabine Burgdorf

Tag der mündlichen Prüfung: 3. Februar 2011

1. Referent: Prof. Dr. A. Prestel (Universität Konstanz)
2. Referent: Prof. Dr. K. Dykema (Texas A&M University)

Sabine Burgdorf
Universität Konstanz,
Fachbereich Mathematik und Statistik,
78457 Konstanz,
Germany
sabine.burgdorf@uni-konstanz.de

This thesis is part of a binational PhD project at Universität Konstanz and Université de Rennes 1.

Acknowledgements

It is a pleasure to thank those people whose support has been of great value for me during the work on my thesis.

First of all I would like appreciate my scientific advisers Igor Klep, Markus Schweighofer and Alexander Prestel for supporting me in many ways. Particularly, I thank Alexander Prestel and Markus Schweighofer for the opportunity to return to their research area after a one year break where I finished my studies in physics. I also want to thank explicitly Igor Klep for his continuous mathematical support and the great collaboration during the last two years. I enjoyed two stays in Slovenia with many intensive and fruitful discussions in his research group. Many results in this work were developed or improved during these two visits. I also thank Ken Dykema for having kindly accepted to serve as a reviewer of my thesis.

I am also indebted to many colleagues in Konstanz and Rennes for the friendly atmosphere and the support they offered to me. Special thanks go to Andrew Dolphin, Mathilde Herblot and Clément Dunand for their effort on improving the English and the French parts of my thesis.

Financial support for my research from the Landesgrauiertenförderung Baden-Württemberg, the German Academic Exchange Service DAAD and the Zukunftskolleg Konstanz are gratefully acknowledged.

Konstanz im Oktober 2010. *Sabine Burgdorf*

Abstract

A polynomial f in non-commuting variables is *trace-positive* if the trace of $f(\underline{A})$ is positive for all tuples $\underline{A}$ of symmetric matrices of the same size. The investigation of trace-positive polynomials and of the question of when they can be written as a sum of hermitian squares and commutators of polynomials are motivated by their connection to two famous conjectures: The BMV conjecture from statistical quantum mechanics and the embedding conjecture of Alain Connes concerning von Neumann algebras.

First, results on the question of when a trace-positive polynomial in two non-commuting variables can be written as a sum of hermitian squares and commutators are presented. For instance, any bivariate trace-positive polynomial of degree at most four has such a representation, whereas this is false in general if the degree is at least six. This is in perfect analogy to Hilbert's results from the commutative context. Further, a partial answer to the Lieb-Seiringer formulation of the BMV conjecture is given by presenting some concrete representations of the polynomials $S_{m,4}(X^2, Y^2)$ as a sum of hermitian squares and commutators.

The second part of this work deals with the *tracial moment problem*. That is, how can one describe sequences of real numbers that are given by tracial moments of a probability measure on symmetric matrices of a fixed size. The truncated tracial moment problem, where one considers only finite sequences, as well as the tracial analog of the K-moment problem are also investigated. Several results from the classical moment problem in Functional Analysis can be transferred to this context. For instance, a tracial analog of Haviland's theorem holds: A tracial linear functional L is given by the tracial moments of a positive Borel measure on symmetric matrices of a fixed size s if and only if L takes only positive values on all polynomials which are trace-positive on all tuples of symmetric $s \times s$-matrices. This result uses tracial versions of the results of Fialkow and Nie on positive extensions of truncated sequences. Further, tracial analogs of results of Stochel and of Bayer and Teichmann are given. Defining a tracial Hankel matrix in analogy to the Hankel matrix in the classical moment problem, the results of Curto and Fialkow concerning sequences with Hankel matrices of finite rank or Hankel matrices of finite size which admit a flat extension also hold true in the tracial context.

Finally, a relaxation for trace-minimization of polynomials using sums of hermitian squares and commutators is proposed. While this relaxation is not always exact, the tracial analogs of the results of Curto and Fialkow give a sufficient condition for the exactness of this relaxation.

Résumé

Un polynôme f en plusieurs variables non commutatives à coefficients réels a une *trace positive* si la trace de $f(\underline{A})$ est positive pour tout vecteur $\underline{A}$ de matrices symétriques de même taille. La recherche des polynômes à trace positive et la question de déterminer quand ils peuvent être écrits comme une somme de carrés hermitiens et de commutateurs sont liées à deux conjectures bien connues : la conjecture de Bessis, Moussa et Villani en mécanique quantique statistique, et la conjecture de plongement d'Alain Connes dans le domaine des algèbres de von Neumann.

La première partie présente des résultats sur ce sujet pour les polynômes en deux variables non commutatives. Tous ces polynômes de degré quatre à trace positive peuvent être écrits sous la forme de sommes de carrés hermitiens et de commutateurs. Mais, en général, ce n'est pas le cas si le degré est supérieur ou égal à six. Ceci est en parfaite analogie avec les résultats de Hilbert dans le contexte commutatif. En outre, on donnera des représentations concrètes des polynômes $S_{m,4}(X^2, Y^2)$, qui sont intimement liés à la conjecture de Bessis, Moussa et Villani, comme sommes de carrés hermitiens et de commutateurs.

La deuxième partie traite *le problème des moments traciaux*. C'est-à-dire, comment caractériser des suites de nombres réels, qui sont données par des moments traciaux d'une mesure de probabilité sur des matrices symétriques de taille fixée. On étudie également le problème tronqué des moments traciaux ainsi que le problème des K-moments traciaux. Certains résultats concernant le problème classique des moments peuvent être reformulés dans le contexte tracial. Par exemple, le théorème de Haviland a un analogue tracial, qui repose sur les résultats de Fialkow et Nie concernant des extensions positives des suites traciales. De plus, il existe des versions traciales de théorèmes de Stochel et de Bayer et Teichmann. Utilisant des matrices traciales de Hankel comme analogue des matrices de Hankel dans le problème classique des moments, les résultats de Curto et Fialkow, concernant des suites avec une matrice de Hankel de rang fini ou avec une matrice de Hankel admettant une extension plate, sont également vrais dans le contexte tracial.

Enfin, nous présentons une version plus faible du probléme de minimisation de la trace d'un polynôme utilisant des sommes de carrés hermitiens et de commutateurs. Bien que cet affaiblissement ne soit pas toujours exact, on considère le problème dual. Les versions traciales des résultats de Curto et Fialkow fournissent une condition suffisante d'exactitude de cette version faible.

Zusammenfassung

Ein Polynom f in nicht-kommutierenden Variablen mit reellen Koeffizienten heißt *spurpositiv*, falls die Spur von $f(\underline{A})$ für alle Tupel $\underline{A}$ von symmetrischen Matrizen gleicher Größe stets positiv ist. Die Untersuchung spurpositiver Polynome sowie die Frage, wann man diese als Summe hermitescher Quadrate und Kommutatoren von Polynomen schreiben kann, ist motiviert durch deren Verbindung zu zwei bekannten Vermutungen: Die BMV-Vermutung aus der statistischen Quantenmechanik und die Einbettungsvermutung von Alain Connes über Von-Neumann-Algebren.

Es werden zunächst Ergebnisse präsentiert, wann sich ein spurpositives Polynom in zwei nicht-kommutierenden Variablen als Summe hermitescher Quadrate und Kommutatoren schreiben lässt. Beispielsweise besitzt jedes bivariate spurpositive Polynom vom Grad höchstens vier eine solche Darstellung, wohingegen dieses für ein spurpositives Polynom vom Grad mindestens sechs im Allgemeinen nicht zutrifft. Außerdem wird eine partielle Antwort zur Lieb-Seiringer-Formulierung der BMV-Vermutung gegeben, indem verschiedene Darstellungen für die Polynome $S_{m,4}(X^2, Y^2)$ als Summe hermitescher Quadrate und Kommutatoren bewiesen werden.

Ein weiteres Thema dieser Arbeit ist das *spurige Momentenproblem*. Dieses ist die Frage, wodurch reelle Folgen charakterisiert sind, die durch spurige Momente eines Wahrscheinlichkeitsmaßes auf symmetrischen Matrizen fester Größe gegeben sind. Darüber hinaus wird das entsprechende trunkierte spurige Momentenproblem sowie das spurige Analogon des K-Momentenproblems behandelt. Verschiedene Ergebnisse hinsichtlich des klassischen Momentenproblems können auf diesen Kontext übertragen werden. Beispielsweise gilt ein spurige Analogon des Satzes von Haviland, welcher auf der spurigen Version der Ergebnisse von Fialkow und Nie über positive Erweiterungen trunkierter Folgen beruht. Des Weiteren gelten spurige Versionen der Resultate von Stochel sowie von Bayer und Teichmann. Definiert man eine spurige Hankelmatrix in Analogie zur Hankelmatrix im kommutativen Kontext, so gelten die Resultat von Curto und Fialkow über Folgen mit positiv semidefiniter Hankelmatrix von endlichem Rang und über Folgen, deren Hankelmatrix eine flache Erweiterung besitzt, entsprechend im spurigen Kontext.

Abschließend wird eine Relaxierung für die numerische Bestimmung des Spurinfimums eines Polynoms mit Hilfe von Summen hermitescher Quadrate und Kommutatoren vorgestellt. Obgleich diese Relaxierung im Allgemeinen nicht exakt ist, liefern die spurigen Versionen der Sätze von Curto und Fialkow eine hinreichende Bedingung für die Exaktheit dieser Relaxierung.

This thesis includes results from the following articles:

- *Sums of Hermitian squares as an approach to the BMV conjecture*,
  Linear and Multilinear Algebra, 59(1), 2011 , 1–9
  see Section 3.4

- *Trace-positive polynomials and the quartic tracial moment problem*,
  Comptes Rendus Mathématiques, 348(13-14), 2010, 721–726
  joint work with Igor Klep;
  see Section 3.3

- *The truncated tracial moment problem*,
  to appear in J. Operator Theory; `http://arxiv.org/abs/1001.3679`
  joint work with Igor Klep;
  see Sections 4.1.1, 4.1.2, 4.2, 4.3.4, 4.3.5 and Section 5.3

- *Semidefinite programming certificates for tracial matrix inequalities*,
  preprint, available from
  `http://www.optimization-online.org/DB_HTML/2010/04/2595.html`
  joint work with Kristijan Cafuta, Igor Klep, and Janez Povh;
  see Sections 3.2, 4.3.1, 4.3.2 and Chapter 6

Contents

Introduction

A real polynomial in non-commuting variables is called *trace-positive* if all its evaluations by symmetric matrices have positive trace.

The theory of trace-positive polynomials is intimately connected to deep open problems from e.g. operator algebras and mathematical physics. In fact, Connes' embedding conjecture on type II_1 von Neumann algebras is equivalent to a problem of describing polynomials which are trace-positive on tuples of matrices of norm at most 1. Further, the Bessis-Moussa-Villani conjecture in an algebraic reformulation of Lieb and Seiringer states that for all $m \in \mathbb{N}_0$ and all positive semidefinite matrices A, B, the polynomial

$$p(t) := \mathrm{Tr}((A + tB)^m) \in \mathbb{R}[t]$$

has only positive coefficients. In other words, the polynomial $S_{m,k}(X^2, Y^2)$, which describes the coefficient of t^k in $(X^2 + tY^2)^m$, is trace-positive. These connections are the main motivation for the present work and will be discussed in more detail in Chapter 2.

Another aim in investigating trace-positive polynomials is to find trace-inequalities involving symmetric matrices. That is, they propose a dimension-free approach to attain trace-inequalities, i.e. they provide certificates holding irrespective of the matrix-size.

To verify trace-inequalities, we use the fact that a matrix has positive trace if and only if it is a sum of a positive semidefinite matrix (i.e. a hermitian square of matrices) and a trace zero matrix (i.e. a commutator of matrices). The main idea in systematizing the verification of trace-inequalities is to look for certificates involving sums of hermitian squares and commutators at the level of polynomials. Let $\mathbb{R}\langle \underline{X} \rangle$ denote the ring of polynomials in the non-commuting variables $\underline{X} = (X_1, \ldots, X_n)$. A hermitian square is a polynomial in $\mathbb{R}\langle \underline{X} \rangle$ of the form g^*g for some $g \in \mathbb{R}\langle \underline{X} \rangle$, where the involution * reverses the order of variables in each monomial of g and models the conjugate transpose of matrices. We are interested in polynomials which can be written as a sum of hermitian squares and commutators of polynomials, i.e. for which $f \in \mathbb{R}\langle \underline{X} \rangle$ does there exist polynomials $g_i, p_j, q_j \in \mathbb{R}\langle \underline{X} \rangle$ such that

$$f = \sum_i {g_i}^* g_i + \sum_j (p_j q_j - q_j p_j)?$$

Let Θ^2 be the set of these polynomials. Obviously, any f in Θ^2 is trace-positive, hence gives rise to a trace-inequality. Let us explain this with a simple example.

For symmetric matrices A, B of the same size we have

$$\mathrm{Tr}(A^2B^2 - ABAB) \geq 0.$$

In fact, consider the polynomial $f = X^2Y^2 - XYXY \in \mathbb{R}\langle X, Y \rangle$. Since f can be written as

$$\begin{aligned} f &= \frac{1}{2}\big(XY^2X + YX^2Y + XYXY + YXYX\big) \\ &\quad + \frac{1}{2}\big(XYX \cdot Y - Y \cdot XYX + X \cdot XY^2 - XY^2 \cdot X + X^2Y \cdot Y - Y \cdot X^2Y\big) \\ &= \frac{1}{2}(XY - YX)^*(XY - YX) + (\text{sum of commutators}), \end{aligned}$$

$f(A,B)$ is a sum of hermitian squares and commutators for all symmetric matrices A, B of the same size. Hence $f(A,B)$ has positive trace.

Trace-positive polynomials lie between two well-investigated classes of polynomials. On the one side, there are polynomials in commuting variables that are positive on a semialgebraic set of $\mathbb{R}^n$. On the other side, there are polynomials in non-commuting variables with only positive semidefinite matrix evaluations. Therefore the natural question that arises is: Which results for these two classes of polynomials do also hold for trace-positive polynomials?

The polynomials whose evaluations by symmetric matrices are all positive semidefinite are exactly the sums of hermitian squares (without commutators). On the other hand, not all trace-positive polynomials are a sum of hermitian squares and commutators. For example, the following version of the Motzkin polynomial

$$M = X^2Y^4 + X^4Y^2 - 3X^2Y^2 + 1$$

in non-commuting variables is trace-positive, but it cannot be written as sum of hermitian squares and commutators. This is in analogy to the commutative case: Not all positive polynomials in commuting variables are sums of squares. Therefore, we investigate analogies for trace-positive polynomials of classical results in Real Algebra for positive polynomials in commuting variables. For polynomials of low degree we establish a tracial analog of the classical result of Hilbert on positive bivariate quartics.

Theorem. *Let $f \in \mathbb{R}\langle X, Y\rangle$ be of degree 4. The following statements are equivalent:*

(i) *f is trace-positive;*

(ii) $\operatorname{Tr}(f(A,B)) \geq 0$ *for all symmetric 2×2-matrices A, B;*

(iii) *f is a sum of four hermitian squares and some commutators;*

(iv) $f \in \Theta^2$.

Moreover, this implies that any trace-inequality of degree four in two symmetric matrices that holds for all symmetric 2×2-matrices holds also for any pair of symmetric $s \times s$-matrices for arbitrary $s \in \mathbb{N}$. This will be handled in Chapter 3. Further, we present there representations of the polynomials $S_{m,4}(X^2, Y^2)$ as sum of hermitian squares and commutators, which imply that, independent of the positive semidefinite matrices A, B, the coefficients of t^k in $p(t) = \operatorname{Tr}((A+tB)^m)$ for $k \leq 4$ are positive. In particular, we derive that the coefficients in p of t^4 are positive for any choice of symmetric matrices A, B of the same size, if the power m is of the form $m = 4r + 2$.

By duality one derives the tracial moment problem, another main topic of this thesis. The moment problem is a classical question in Functional Analysis, which is well studied because of its importance and the variety of its applications. A simple example is the (univariate) Hamburger moment problem: Which linear functionals L on univariate real polynomials are integration with respect to some positive Borel measure μ? By Haviland's theorem this is the case if and only if L is positive on all polynomials that are positive on $\mathbb{R}$. Thus Haviland's theorem relates the moment problem to positive polynomials. It holds in several variables and also if we restrict the support of μ to some appropriate set K. The duality between the moment problem and positive polynomials has been used, for example, in Schmüdgen's celebrated solution of the moment problem on compact basic closed semialgebraic sets, which then implies Schmüdgen's Positivstellensatz.

In Chapter 4 we define the tracial moment problem including tracial Riesz functionals and tracial Hankel matrices, which correspond to the given linear functional in the same way as in the classical case. The truncated tracial moment problem, where one considers only finite sequences, as well as the tracial analog of the K-moment problem are also investigated. We establish several analogies which hold between the classical moment problem and its tracial version. For instance, a tracial analog of Haviland's theorem holds.

Theorem. *Let L be a tracial linear functional on $\mathbb{R}\langle \underline{X} \rangle$. Then there is a positive Borel measure μ on symmetric $s \times s$-matrices such that for all monomials w,*

$$L(w) = \int \mathrm{Tr}(w)\, d\mu,$$

if and only if L takes only positive values on polynomials that are trace-positive on all tuples of symmetric $s \times s$-matrices.

In more detail, a sequence of real numbers which is labelled by monomials in non-commuting variables and with values invariant under cyclic permutations of the indices is called a tracial sequence. The *tracial moment problem* asks for a characterization of tracial sequences y for which there exists an integer $s \in \mathbb{N}$ and a probability measure μ on symmetric $s \times s$ matrices such that any value y_w of y can be written as

$$y_w = \int \mathrm{Tr}(w)\, d\mu. \tag{R}$$

These sequences are called tracial moment sequences. We present several results on the general structure of tracial sequences with a representation (R). For instance, we emphasize the truncated version is more general than the full tracial moment problem.

Theorem. *Let y be a tracial sequence. If there is an $s \in \mathbb{N}$ such that for all $k \in \mathbb{N}$ there exists a measure μ_k on symmetric $s \times s$-matrices with $y_w = \int \mathrm{Tr}(w)\, d\mu_k$ for all $w \in \langle \underline{X} \rangle$ of degree at most k, then y is a tracial moment sequence.*

Further, the tracial analog of the theorem of Bayer and Teichmann from the classical context holds. That is, the representation of a truncated tracial sequence y using integrals with a positive Borel measure can be replaced by a representation using a finitely atomic measure. Tracial moment sequences satisfy some necessary conditions, which are similar to the ones in the classical case. For instance, the tracial Hankel matrix of a tracial moment sequence is positive semidefinite. These necessary conditions are in general not sufficient therefore we also present some conditions for (R) to hold. We present tracial analogs of the classical results of Curto and Fialkow on Hankel matrices. For the full tracial moment problem we have the following:

Theorem. *Let y be a tracial sequence. Then y is a tracial moment sequence if its corresponding tracial Hankel matrix is positive semidefinite and of finite rank.*

For the truncated moment problem flatness governs the existence of a representation (R) for truncated tracial sequences, resembling the situation in the classical moment problem. Furthermore, the tracial Riesz functionals can be used, as in the commutative case, to obtain sufficient conditions of a tracial sequence y to have a representation (R). Indeed, if the Riesz functional admits a positive extension, then the corresponding truncated tracial sequence has such a representation. Finally, in analogy to results of Fialkow and Nie, we prove that if the tracial Riesz functional admits only strictly positive values on all polynomials trace-positive on symmetric matrices of a fixed size, then the corresponding sequence y is a tracial moment sequence. In some cases we can even restrict the support of the representing measure μ.

In a different vein, the classical theorem of Hilbert was used by Fialkow and Nie to solve the bivariate quartic truncated moment problem to some extent. The duality between positive polynomials and the moment problem extends to the tracial non-commutative setting. This is handled in Chapter 5 which also summarizes the previous results in terms of convex cones and shows the duality of sums of hermitian squares and commutators and the tracial moment problem in notions of conic duality.

In the last chapter we combine several results from the previous chapters to give an application of our theory. The question of when a given polynomial can be written as a sum of hermitian squares and commutators can be answered numerically by an algorithm using semidefinite programming. This is based on the tracial analog of the Gram matrix method, which is explained in Section 3.2. We apply this method and its dual in Chapter 6. Indeed, the optimization problem that looks for the trace-infimum of a given polynomial over all tuples of symmetric matrices can be relaxed by the optimization problem

$$f_{\text{sos}} := \sup\{a \in \mathbb{R} \mid f - a \in \Theta^2\}.$$

While this relaxation is not always exact, it is easy to compute and gives convenient bounds on the optima. To test for exactness the solution of the dual semidefinite program is investigated. If it satisfies a certain condition, which is directly connected to the tracial moment problem, then the relaxation is exact. In this case it is shown how to extract global trace-optimizers with a procedure based on the methods from Chapter 4.

1 Preliminaries

In this chapter we introduce the basic terminology to set the stage for this work. We recall some basic notions from Real Algebra and present their non-commutative analogs in our setting. Further, we list some well-known facts on von Neumann algebras and say a few words on terminology and results in measure theory needed in the sequel.

We set $\mathbb{N} = \{1, 2, \ldots\}$ and $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$. If we use the word *positive*, we mean nonnegative, i.e. we allow that the value zero might be taken, and we will say *strictly positive* otherwise.

1.1 Polynomials

The ring of polynomials in n commuting variables $\underline{x} = (x_1, \ldots, x_n)$ is well known and will be denoted by $\mathbb{R}[\underline{x}]$. In this section we fix the notation for polynomials in non-commuting variables. For better distinction between commuting and non-commuting variables, if needed, we use x_i for commuting variables and capital X_i for the non-commuting ones.

We denote by $\langle \underline{X} \rangle$ the monoid which is freely generated by the n non-commuting letters $\underline{X} = (X_1, \ldots, X_n)$. Its elements are called *words*, including the empty word denoted by 1. Let $\mathbb{R}\langle \underline{X} \rangle$ denote the monoid ring of $\langle \underline{X} \rangle$ over $\mathbb{R}$. That is, $\mathbb{R}\langle \underline{X} \rangle$ is the unital associative algebra freely generated by $X_1, \ldots, X_n$. The elements f of $\mathbb{R}\langle \underline{X} \rangle$ are thus polynomials in the non-commuting variables $X_1, \ldots, X_n$ with coefficients in $\mathbb{R}$. If we deal with polynomials in two variables we replace the variables X_1, X_2 by X, Y.

An element of the form aw, where $0 \neq a \in \mathbb{R}$ and $w \in \langle \underline{X} \rangle$, is called a *monomial* and a its *coefficient*. Thus words are monomials with coefficient 1. Instead of the multi-index $\alpha \in \mathbb{N}_0^n$, often used to abbreviate multivariate polynomials as $\sum_\alpha a_\alpha \underline{x}^\alpha$ for some $a_\alpha \in \mathbb{R}$, we use the words $w \in \langle \underline{X} \rangle$ itself as index, i.e. we write polynomials $f \in \mathbb{R}\langle \underline{X} \rangle$ as finite sums

$$f = \sum_{w \in \langle \underline{X} \rangle} f_w w \in \mathbb{R}\langle \underline{X} \rangle$$

with $f_w \in \mathbb{R}$. Let $\check{} : \mathbb{R}\langle \underline{X} \rangle \to \mathbb{R}[\underline{x}]$ be the algebra homomorphism mapping each X_i to the commuting variable x_i. The image $\check{f} \in \mathbb{R}[\underline{x}]$ of a given polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is called the *commutative collapse* of f.

The (total) *degree* of a polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is the length of the longest word appearing in f and is denoted by $\deg(f)$. The set of all monomials of degree $\leq d$ for a given degree bound $d \in \mathbb{N}_0$ will be denoted by $\langle \underline{X} \rangle_d$. The polynomials of degree $\leq d$ are denoted in the same way by $\mathbb{R}\langle \underline{X} \rangle_d$.

As in the commutative case, one can identify $\mathbb{R}\langle \underline{X} \rangle_d$ with a finite dimensional vector space, namely $\mathbb{R}^t$, where $t = t(d) = \sum_{\ell=0}^{d} n^\ell = \dim \mathbb{R}\langle \underline{X} \rangle_d < \infty$. In fact, the map

$$\begin{aligned} \varphi : \mathbb{R}\langle \underline{X} \rangle_d &\to \mathbb{R}^t \\ f &\mapsto \vec{f} \end{aligned}$$

which sends a polynomial $f = \sum_w f_w w \in \mathbb{R}\langle \underline{X} \rangle_d$ onto its (column) vector $\vec{f} \in \mathbb{R}^t$, which is given by the coefficients f_w (with $\deg w \leq d$) in a fixed order, is an isomorphism. In the same way, we can also identify $\mathbb{R}\langle \underline{X} \rangle$ with the vector space of column vectors $\vec{f} = [f_w]_w$ in the product space $\mathbb{R}^{\langle \underline{X} \rangle}$ with only finitely many entries $f_w \neq 0$.

Instead of evaluating a polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ in tuples of real numbers resulting in a real number we substitute $\underline{X}$ by tuples $\underline{A} = (A_1, \ldots, A_n)$ of symmetric matrices $A_1, \ldots A_n \in \mathbb{R}^{s\times s}$ for some $s \in \mathbb{N}$. This works as follows. The empty word goes to the identity matrix $\mathbf{1}_s$ of size s and a word $w = X_{i_1} \ldots X_{i_r}$ becomes $w(\underline{A}) := A_{i_1} \cdots A_{i_r}$. Thus $f(\underline{A})$ is an $s \times s$-matrix. To model the symmetry $A_i^T = A_i$ of the matrices we plug in, where ^{T} denotes the matrix transpose, we endow $\mathbb{R}\langle \underline{X} \rangle$ with the *involution* $^* : \mathbb{R}\langle \underline{X} \rangle \to \mathbb{R}\langle \underline{X} \rangle$, $p \mapsto p^*$ that fixes $\mathbb{R} \cup \{\underline{X}\}$ pointwise. As any involution, * has the properties $(f+g)^* = f^* + g^*$, $(fg)^* = g^* f^*$ and $f^{**} = f$ for all $f, g \in \mathbb{R}\langle \underline{X} \rangle$. Thus for each word $w \in \langle \underline{X} \rangle$, w^* is its reverse. As an example, we have

$$(X_1 X_2^2 - X_2 X_1)^* = X_2^2 X_1 - X_1 X_2.$$

This involution is compatible with the matrix transpose, i.e. $f^*(\underline{A}) = f(\underline{A})^T$ for all tuples $\underline{A}$ of symmetric matrices of the same size. In fact, taking $w = X_i = w^*$ we have $A_i = w^*(\underline{A}) = w(\underline{A})^T = A_i^T$.

We regard $\mathbb{R}\langle \underline{X} \rangle$ as $*$-algebra and equip it with the finest locally convex topology, which makes all seminorms continuous. Every finite dimensional vector space of $\mathbb{R}\langle \underline{X} \rangle$ then inherits the euclidean topology.

A polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is *symmetric* if $f^* = f$. Let $\mathcal{S}\mathbb{R}\langle \underline{X} \rangle$ denote the set of symmetric elements in $\mathbb{R}\langle \underline{X} \rangle$, i.e. $\mathcal{S}\mathbb{R}\langle \underline{X} \rangle = \{f \in \mathbb{R}\langle \underline{X} \rangle \mid f^* = f\}$. In the same way, we have $\mathcal{S}\mathbb{R}\langle \underline{X} \rangle_d := \mathcal{S}\mathbb{R}\langle \underline{X} \rangle \cap \mathbb{R}\langle \underline{X} \rangle_d$ for the set of symmetric polynomials of degree at most d. Further we write $\mathcal{S}\mathbb{R}^{s\times s}$ to denote the set of all symmetric matrices of size s. The set of all n-tuples $\underline{A}$ consisting of symmetric matrices $A_1, \ldots, A_n$ of the *same* (arbitrary) size are denoted by $\mathcal{S}^n$, i.e.

$$\mathcal{S}^n := \bigcup_{s\in\mathbb{N}} (\mathcal{S}\mathbb{R}^{s\times s})^n.$$

A symmetric matrix $A \in \mathcal{S}\mathbb{R}^{s\times s}$ is *positive-semidefinite*, denoted as $A \succeq 0$, if $\vec{z}^T A \vec{z} \geq 0$ for all $\vec{z} \in \mathbb{R}^s$. Equivalently, all its eigenvalues are positive, or it arises as the Gram matrix of some set of vectors $\vec{v_1}, \ldots, \vec{v_s} \in \mathbb{R}^s$, i.e. $A_{ij} = \langle \vec{v_i}, \vec{v_j} \rangle$ for all $i, j = 1 \ldots, s$.

1.2 Positivity

In the sequel we will distinguish three different kinds of positivity. Namely, positivity of polynomials in commuting variables, matrix-positivity and trace-positivity.

First, we define what we mean by a positive polynomial in commuting variables. This is one of the main definitions from Real Algebra and the other notions of positivity below will be extensions to polynomials in non-commuting variables.

1.1 Definition. A polynomial $f \in \mathbb{R}[\underline{x}]$ is *positive* (denoted as $f \geq 0$) if

$$f(\underline{a}) \geq 0 \text{ for all } \underline{a} \in \mathbb{R}^n.$$

If $f(\underline{a}) \geq 0$ for all $\underline{a} \in K$ where $K \subseteq \mathbb{R}^n$, we call f positive on K.

In the free non-commutative setting one evaluates polynomials in symmetric matrices and considers matrix-positivity.

1.2 Definition. A polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is *matrix-positive* (for short $f \succeq 0$) if

$$f(\underline{A}) \text{ is positive semidefinite for all } \underline{A} \in \mathcal{S}^n.$$

Such a polynomial f is necessarily symmetric. For a given set $K \subseteq \mathcal{S}^n$, we call f matrix-positive on K if $f(\underline{A})$ is positive semidefinite for all $\underline{A} \in K$.

The most important concept of positivity in this work will be trace-positivity of polynomials in non-commuting variables. Since we are interested in a dimension-free approach, which is independent of the size of matrices we plug in, we consider the *normalized* trace Tr instead of the canonical matricial trace tr, i.e.

$$\operatorname{Tr}(A) = \frac{1}{s}\operatorname{tr}(A) = \frac{1}{s}\sum_{i=1}^{s} A_{ii} \quad \text{for } A \in \mathbb{R}^{s\times s}.$$

1.3 Definition. A polynomial $f \in \mathbb{R}\langle \underline{X}\rangle$ is *trace-positive* if

$$\operatorname{Tr}(f(\underline{A})) \geq 0 \text{ for all tuples } \underline{A} \in \mathcal{S}^n.$$

If $\operatorname{Tr}(f(\underline{A})) \geq 0$ for all $\underline{A}$ of a given set $K \subseteq \mathcal{S}^n$ of symmetric matrices, we call f trace-positive on K.

These three notions of positivity are connected but they describe different sets of polynomials. Since positive semidefinite matrices have positive trace a matrix-positive polynomial is also trace-positive. Moreover, if f is trace-positive, then $\check{f}$ is positive. However the converse implications do not hold in general as shown in the following example.

1.4 Example.

(a) The commutative collapse $\check{f} = x^2y^2 \in \mathbb{R}[x,y]$ of the polynomial $f = XYXY \in \mathbb{R}\langle X,Y\rangle$ is positive on $\mathbb{R}^2$, but f is not trace-positive. For instance, taking

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

we obtain

$$\operatorname{Tr}\left(f(A,B)\right) = \operatorname{Tr}\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -1.$$

(b) The polynomial $g = X^2Y^2 + XYXY \in \mathbb{R}\langle X,Y\rangle$ as well as its symmetrized version $f = \frac{1}{2}(X^2Y^2 + Y^2X^2 + XYXY + YXYX)$ is trace-positive but not matrix-positive. In fact, for arbitrary $A, B \in \mathcal{S}\mathbb{R}^{s\times s}$ we have

$$\begin{aligned}\operatorname{Tr}(g(A,B)) &= \operatorname{Tr}(f(A,B)) \\ &= \frac{1}{2}\operatorname{Tr}(AB^2A + ABAB + BABA + BA^2B) \\ &= \frac{1}{2}\operatorname{Tr}((AB+BA)^T(AB+BA)) \geq 0.\end{aligned}$$

Since g is not symmetric, it can not be matrix-positive. To show that f is not matrix-positive take for instance

$$A = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

which gives

$$f(A,B) = \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix}.$$

1.3 Cyclic equivalence

We are interested in the class of trace-positive polynomials. Therefore we endow the free algebra $\mathbb{R}\langle \underline{X} \rangle$ with an equivalence relation to model the invariance of the trace under cyclic permutations. This motivates the following definition of cyclic equivalence [KS1, p. 1817].

1.5 Definition. An element of the form $[p, q] = pq - qp$ for $p, q \in \mathbb{R}\langle \underline{X} \rangle$ is called a *commutator*. Two polynomials $f, g \in \mathbb{R}\langle \underline{X} \rangle$ are *cyclically equivalent* ($f \overset{\text{cyc}}{\sim} g$) if $f - g$ is a sum of commutators:

$$f - g = \sum_{i=1}^{k} (p_i q_i - q_i p_i) \text{ for some } k \in \mathbb{N} \text{ and } p_i, q_i \in \mathbb{R}\langle \underline{X} \rangle.$$

1.6 Example. The polynomials $f = 2XY^2X - Y^2X^2 + 2YXY$ and $g = X^2Y^2 + 2XY^2$ are cyclically equivalent since we can write $f - g$ as

$$f - g = [XY^2, X] + [X, Y^2X] + 2[Y, XY].$$

On the other hand X^2Y^2 and $XYXY$ are not cyclically equivalent. This can be seen by the following remarks, which show that it can easily be checked whether two polynomials are cyclically equivalent and motivate its name.

1.7 Remark.

1. Two words $v, w \in \langle \underline{X} \rangle$ are cyclically equivalent if and only if w is a cyclic permutation of v, i.e. there exist $u_1, u_2 \in \langle \underline{X} \rangle$ such that $v = u_1 u_2$ and $w = u_2 u_1$.
2. Two polynomials $f = \sum_{w \in \langle \underline{X} \rangle} f_w w$ and $g = \sum_{w \in \langle \underline{X} \rangle} g_w w$ are cyclically equivalent if and only if for each $v \in \langle \underline{X} \rangle$,
$$\sum_{\substack{w \in \langle \underline{X} \rangle \\ w \overset{\text{cyc}}{\sim} v}} f_w = \sum_{\substack{w \in \langle \underline{X} \rangle \\ w \overset{\text{cyc}}{\sim} v}} g_w. \tag{1.1}$$
3. If $f \overset{\text{cyc}}{\sim} g$ then $\operatorname{Tr}(f(\underline{A})) = \operatorname{Tr}(g(\underline{A}))$ for all $\underline{A} \in \mathcal{S}^n$. Less obvious is the following partial converse: If $f \in \mathcal{S}\mathbb{R}\langle \underline{X} \rangle$ and $\operatorname{Tr}(f(\underline{A})) = 0$ for all $\underline{A} \in \mathcal{S}^n$, then $f \overset{\text{cyc}}{\sim} 0$ [KS1, Theorem 2.1].
4. Although $f \overset{\text{cyc}}{\nsim} f^*$ in general, by evaluating f in real matrices we still have $\operatorname{Tr}(f(\underline{A})) = \operatorname{Tr}(f^*(\underline{A}))$ for all $f \in \mathbb{R}\langle \underline{X} \rangle$ and $\underline{A} \in \mathcal{S}^n$. Therefore f is often assumed to be a symmetric polynomial.

Each polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ has a canonical representative $[f]$ with respect to $\overset{\text{cyc}}{\sim}$ which represents the equivalence class of f in $\mathbb{R}\langle \underline{X} \rangle / _{\overset{\text{cyc}}{\sim}}$.

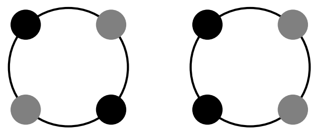
1.8 Definition. Let $w \in \langle \underline{X} \rangle$. The canonical representative $[w]$ of w is the smallest word (with respect to a fixed order) among all words cyclically equivalent to w. We define the *canonical representative* $[f]$ of a polynomial $f = \sum_w f_w w \in \mathbb{R}\langle \underline{X} \rangle$ as $[f] := \sum_{[w]} f_{[w]} [w]$. That is, $[f]$ contains only canonical representatives of words from f with coefficients $f_{[w]} := \sum_{u \overset{\text{cyc}}{\sim} w} f_u$.

As an example, for $f = 2Y^2X^2 - XY^2X + XY - YX$ we have $[f] = X^2Y^2$ if we take the lexicographic order. By Remark 1.7 2., two polynomials $f, g \in \mathbb{R}\langle \underline{X} \rangle$ are cyclically equivalent if and only if they have the same canonical representative:

$$f \overset{\text{cyc}}{\sim} g \Leftrightarrow [f] = [g].$$

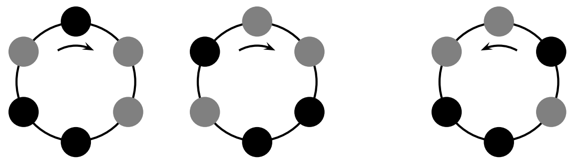
Hence any polynomial of an equivalence class in $\mathbb{R}\langle \underline{X} \rangle / _{\overset{\text{cyc}}{\sim}}$ has the same canonical representative. Therefore $[f]$ denotes the polynomial $[f] \in \mathbb{R}\langle \underline{X} \rangle$ as well as its equivalence class $[f] \in \mathbb{R}\langle \underline{X} \rangle / _{\overset{\text{cyc}}{\sim}}$.

1.9 Remark. The equivalence classes in $\langle \underline{X} \rangle$ with respect to the equivalence relation $\overset{\text{cyc}}{\sim}$ are called *necklaces* in Combinatorics. Intuitively, a necklace connects $\deg(w)$ beads of up to n colours on a circle. To obtain a representative of a necklace one “reads" the beads clockwise starting from an arbitrary bead. If for example grey beads represent the letter X and black beads the letter Y, the following two necklaces represent the equivalence classes of $XYXY$ and X^2Y^2 respectively.



One easily sees that these two graphical representations of necklaces are not congruent for any rotation. Hence X^2Y^2 and $XYXY$ are not cyclically equivalent.

The reversal of strings is respected, that is, necklaces represent circular collections of beads in which the necklace may not be turned over. For example, the following two necklaces on the left hand side are equal since their graphical representations are congruent if we turn the first one step to the left or equivalently five steps to the right.



The graphical representation of the third necklace would be congruent to one of the others after a rotation if we allow a turn-over, i.e. the corresponding word is obtained by reading the beads counter-clockwise. This corresponds to a reversal of the corresponding word. In fact, in our example the two necklaces on the left hand side represent $w = XYX^2Y^2$ whereas the necklace on the right represents $w^* = Y^2X^2YX$. If reversal (or equivalently, a turn-over) is allowed one calls the corresponding equivalence class a *bracelet*.

In general a polynomial in two non-commuting variables is already different in the behaviour from polynomials in commuting variables. However the following class of cyclically sorted polynomials, introduced in [KS1], will turn out to be quite similar to the commutative case.

1.10 Definition. A polynomial $f \in \mathbb{R}\langle X, Y \rangle$ is called *cyclically sorted* if f is cyclically equivalent to $\sum_{i,j} a_{ij} X^i Y^j$ for some $i, j \in \mathbb{N}_0$, $a_{ij} \in \mathbb{R}$.

In particular, the canonical representative with respect to the lexicographic order of a cyclically sorted polynomial is of the form $\sum_{i,j} a_{ij} X^i Y^j$ for some $i, j \in \mathbb{N}_0$, $a_{ij} \in \mathbb{R}$.

1.4 Real Algebra

The following notions are – in the commutative setting – standard knowledge from Real Algebra, see e.g. [BCR, Mar, PD].

Classical Real Algebra involves the investigation of the (convex) cone of positive polynomials. Since this is hard in general, one tries to find simple algebraic certificates that make the positive character evident. A good candidate for global positivity is the cone $\sum \mathbb{R}[\underline{x}]^2$ of sums of squares of polynomials, i.e. elements of the form $\sum_i g_i^2$ for $g_i \in \mathbb{R}[\underline{x}]$, which are obviously positive on $\mathbb{R}^n$. More generally one considers quadratic modules of $\mathbb{R}[\underline{x}]$. A *quadratic module* of $\mathbb{R}[\underline{x}]$ is a subset M of $\mathbb{R}[\underline{x}]$ such that $M + M \subseteq M$, $\sum \mathbb{R}[\underline{x}]^2 \cdot M \subseteq M$ and $1 \in M$. For $g_1, \ldots, g_r \in \mathbb{R}[\underline{x}]$ the smallest quadratic module containing $g_1, \ldots, g_r$ consists of all elements of the form

$$\sigma_0 + \sigma_1 g_1 + \cdots + \sigma_r g_r,$$

where $\sigma_i \in \sum \mathbb{R}[\underline{x}]^2$. It is called the quadratic module generated by $g_1, \dots, g_r$ and will be denoted by $\mathrm{QM}(g_1, \dots, g_r)$. In particular, $\sum \mathbb{R}[\underline{x}]^2$ can be considered as the quadratic module generated by $g = 1$ in $\mathbb{R}[\underline{x}]$.

Sometimes one is interested in positivity on a given (semialgebraic) set $K \subseteq \mathbb{R}^n$, where K is defined by $g_1, \dots, g_r \in \mathbb{R}[\underline{x}]$ in the sense that

$$K = \{\underline{a} \in \mathbb{R}^n \mid g_1(\underline{a}) \geq 0, \dots, g_r(\underline{a}) \geq 0\}. \tag{1.2}$$

A distinguished set of polynomials being positive on K is then given by the quadratic module generated by $g_1, \dots, g_r$. A classical examples of such a set $K \subseteq \mathbb{R}^n$ is the *hypercube* $[-1, 1]^n$, which is defined by $g_i = 1 - x_i^2$ for $i = 1, \dots, n$.

Two famous results concerning positivity of polynomials are the following results of Hilbert and of Putinar respectively. A *quartic* is a polynomial f of degree four. If f is in addition homogeneous, we call f a quartic form. Hilbert's theorem deals with bivariate quartics and sums of squares whereas Putinar's theorem gives a certificate for polynomials being strictly positive on the hypercube.

1.11 Theorem (Hilbert)**.** *Let $f \in \mathbb{R}[x, y]_4$. Then $f \geq 0$ on $\mathbb{R}^2$ if and only if $f \in \sum \mathbb{R}[x, y]^2$. In particular, f can be written as a sum of three squares.*

Theorem 1.11 has originally been formulated for ternary quartic forms, which can easily be derived from Theorem 1.11 by homogenization. A modern treatment of Hilbert's proof is given in [PRSS], a more elementary proof, which does not give the sharp bound on the number of squares needed, is given in [CL].

1.12 Remark. This result arose in Hilbert's classification of the cases when the cone of positive polynomials in n variables of degree d is equal to the cone $\sum \mathbb{R}[\underline{x}]^2_{d/2}$ of sums of squares of polynomials of degree at most $d/2$ [Hilb]. This is trivially true for univariate polynomials ($n = 1$) of any degree and also for quadratic polynomials ($d = 2$) in arbitrary many variables. Theorem 1.11 shows that it also holds true for $n = 2, d = 4$. By homogenization, we also have equality in some additional cases if we only consider forms, i.e. homogeneous polynomials. From the univariate case we obtain that every positive form in two variables is a sum of squares of forms. Further, Theorem 1.11 implies that this holds true for forms with $n = 3, d = 4$. Hilbert also showed by abstract arguments that these cases are the only cases where equality holds. Several years later, Motzkin [Mot] presented the Motzkin polynomial

$$f_{\mathrm{Motz}} = x^2y^4 + x^4y^2 - 3x^2y^2 + 1 \in \mathbb{R}[x, y]_6,$$

the first concrete example of a positive polynomial which is not a sum of squares. See also [Rez] for a modern survey on this topic.

1.13 Theorem (Putinar)**.** *Let $f \in \mathbb{R}[\underline{x}]$. Then $f \geq 0$ on $[-1, 1]^n$ if and only if for all $\varepsilon \in \mathbb{R}_{>0}$, $f + \varepsilon$ lies in the quadratic module generated by $1 - x_i^2$ for $i = 1, \dots, n$.*

1.14 Remark. Putinar showed a more general statement for archimedean quadratic modules, i.e. quadratic modules M satisfying that for all $p \in \mathbb{R}[\underline{x}]$ there is an integer $N \in \mathbb{N}$ such that $N \pm p \in M$. He proved [Put] that for a given archimedean quadratic module $\mathrm{QM}(g_1, \dots, g_r)$ any polynomial $f \in \mathbb{R}[\underline{x}]$, that is strictly positive on the semialgebraic set K defined by $g_1 \dots, g_r$ as in (1.2), lies in $\mathrm{QM}(g_1, \dots, g_r)$. Since the quadratic module $\mathrm{QM}(1 - x_1^2, \dots, 1 - x_n^2)$ is archimedean [PD, Cor. 5.1.14], Theorem 1.13 follows from his original statement.

We now present the tracial analogs of this terminology concerning trace-positive polynomials. Basically, one derives these by adding commutators to the free non-commutative analogs. The tracial analog of Theorem 1.11 will be presented and proved in Section 3.3, the tracial version of Theorem 1.13 is connected to Connes' embedding conjecture presented in Section 2.2.

1.4.1 Sums of hermitian squares

A positive semidefinite matrix A has a square root $\sqrt{A}$, i.e. it can be written as $\sqrt{A}^T\sqrt{A}$. To model such a decomposition on the polynomial level one considers the set of sums of hermitian squares instead of sums of squares of polynomials. A polynomial of the form g^*g in $\mathbb{R}\langle \underline{X}\rangle$ is called a *hermitian square* and the set of all sums of hermitian squares will be denoted by Σ^2, i.e.

$$\Sigma^2 = \{f \in \mathbb{R}\langle \underline{X}\rangle \mid f = \sum_{i=1}^{r} g_i{}^* g_i \text{ for some } g_i \in \mathbb{R}\langle \underline{X}\rangle, r \in \mathbb{N}_0\}.$$

Clearly, $\Sigma^2 \subseteq \mathcal{S}\mathbb{R}\langle \underline{X}\rangle$ and any element $f \in \Sigma^2$ is matrix-positive and symmetric. Moreover, Σ^2 coincides with the cone of matrix-positive polynomials as proved by Helton [Hel, Theorem 1.1] and independently by McCullough [McC, Theorem 0.2], see also [MP] for a proof of Helton's theorem using a separation argument.

1.15 Theorem (Helton/McCullough). *Let $f \in \mathbb{R}\langle \underline{X}\rangle$. Then f is matrix-positive if and only if $f \in \Sigma^2$.*

Since the trace of a matrix does not change if we add commutators of matrices, any matrix-positive polynomials stays trace-positive if we add commutators of polynomials. Hence the tracial analog of sums of squares is formed by sums of hermitian squares and commutators.

1.16 Definition. Let

$$\Theta^2 := \{f \in \mathbb{R}\langle \underline{X}\rangle \mid f \overset{\text{cyc}}{\sim} g \text{ for some } g \in \Sigma^2\}.$$

denote the set of all polynomials cyclically equivalent to a sum of hermitian squares. By definition, the elements in Θ^2 are exactly the polynomials which can be written as a sum of hermitian squares and commutators.

Clearly, any $f \in \Theta^2$ is trace-positive. Further, Σ^2 is a proper subset of Θ^2 if $n \geq 2$, since $\Sigma^2 \subseteq \mathcal{S}\mathbb{R}\langle \underline{X}\rangle$ and Θ^2 contains polynomials which are not symmetric. Furthermore, we have $(\Theta^2 \cap \mathcal{S}\mathbb{R}\langle \underline{X}\rangle) \setminus \Sigma^2 \neq \varnothing$ as shown in the following example.

1.17 Example. We have

$$f = X^2Y^2 + Y^2X^2 + XYXY + YXYX \in (\Theta^2 \cap \mathcal{S}\mathbb{R}\langle \underline{X}\rangle) \setminus \Sigma^2.$$

In fact, $f \overset{\text{cyc}}{\sim} XY^2X + YX^2Y + XYXY + YXYX = (XY + YX)^*(XY + YX)$, and thus $f \in \Theta^2$. The polynomial f is not matrix-positive by the same argument as in Example 1.4(b), and therefore $f \notin \Sigma^2$.

We proceed by showing that Θ^2 is a closed convex cone in $\mathbb{R}\langle \underline{X}\rangle$ with respect to the finest locally convex topology. To do this we set

$$\Theta^2_{n,k} := \Theta^2 \cap \mathbb{R}\langle \underline{X}\rangle_{2k}.$$

The index n, which denotes the number of variables in $\mathbb{R}\langle \underline{X}\rangle$, will only be important in Chapter 5.

1.18 Remark. Since the highest degree terms do not cancel, one easily sees that

$$\Theta^2_{n,k} = \{f \in \mathbb{R}\langle \underline{X}\rangle_{2k} \mid f \overset{\text{cyc}}{\sim} \sum_{i=1}^{r} g_i{}^* g_i \text{ for some } g_i \in \mathbb{R}\langle \underline{X}\rangle_k, r \in \mathbb{N}_0\}.$$

Indeed, suppose $\deg g_i = t > k$ for some $i \in \{1, \dots, r\}$. Let h_i be the homogeneous part of degree t of the g_i and $r_i = g_i - h_i$. Then $\deg r_i < t$ and

$$\sum_i {g_i}^* g_i = \sum_i {h_i}^* h_i + \sum_i ({r_i}^* r_i + {h_i}^* r_i + {r_i}^* h_i).$$

Since each monomial in ${h_i}^* r_i$, ${r_i}^* h_i$ and ${r_i}^* r_i$ has degree $< 2t$, none of these can be cyclically equivalent to a monomial in ${h_i}^* h_i$, where each monomial is of degree $2t$. Thus we have $\sum_i {h_i}^* h_i \overset{\text{cyc}}{\sim} 0$ which implies by [KS2, Lemma 3.2] that $h_i = 0$ for all $i \in \{1, \dots, r\}$.

Further, we can also assume that the commutators are of degree at most $2k$. In fact, if $f \in \Theta^2_{n,k}$ and we have a representation $f = \sum_i {g_i}^* g_i + \sum_j [p_j, q_j]$ for some $g_i, p_j, q_j \in \mathbb{R}\langle \underline{X} \rangle$, we can split the sum of commutators $c := \sum_j [p_j, q_j]$ into a sum of commutators of monomials, i.e. each commutator is a difference of two monomials of the same degree. Since $c = f - \sum_i {g_i}^* g_i$ we get that $\deg c \le 2k$. Hence all homogeneous parts of c of degree greater than $2k$ are equal to zero and we can omit them. Hence we have a representation of f as sum of hermitian squares and commutators where the commutators have degree at most $2k$. Therefore we have

$$\Theta^2_{n,k} = \{f \mid f = \sum_i {g_i}^* g_i + \sum_j [p_j, q_j] \text{ for some } g_i \in \mathbb{R}\langle \underline{X} \rangle_k, [p_i, q_j] \in \mathbb{R}\langle \underline{X} \rangle_{2k}\}.$$

To show that Θ^2 is closed with respect to the finest locally convex topology on $\mathbb{R}\langle \underline{X} \rangle$ it suffices to prove that $\Theta^2_{n,k}$ is closed in $\mathbb{R}\langle \underline{X} \rangle_{2k}$ for all $k \in \mathbb{N}$ with respect to the norm topology. The proof can also be found in [BK1, Lemma 4.5].

1.19 Proposition. *$\Theta^2_{n,k}$ is a closed convex cone in $\mathbb{R}\langle \underline{X} \rangle_{2k}$.*

Proof. It is clear that $\Theta^2_{n,k}$ is a convex cone. To show the closedness, endow $\mathbb{R}\langle \underline{X} \rangle_{2k}$ with a norm $\|\llcorner\!\lrcorner\|$ and the quotient space $\mathbb{R}\langle \underline{X} \rangle_{2k}/_{\overset{\text{cyc}}{\sim}}$ with the quotient norm

$$\|\pi(f)\| := \inf\left\{\|f + h\| \mid h \overset{\text{cyc}}{\sim} 0\right\}, \tag{1.3}$$

for all $f \in \mathbb{R}\langle \underline{X} \rangle_{2k}$. Here $\pi : \mathbb{R}\langle \underline{X} \rangle_{2k} \to \mathbb{R}\langle \underline{X} \rangle_{2k}/_{\overset{\text{cyc}}{\sim}}$ denotes the quotient map. Note that the infimum on the right-hand side of (1.3) is attained since $\mathbb{R}\langle \underline{X} \rangle_{2k}$ is finite-dimensional. Since $\Theta^2_{n,k} = \pi^{-1}\big(\pi(\Theta^2_{n,k})\big)$, it suffices to show that $\pi(\Theta^2_{n,k})$ is closed. Let $t_k = \dim \mathbb{R}\langle \underline{X} \rangle_{2k}$. By Carathéodory's theorem [Bar, p. 10] each element $f \in \Theta^2_k$ can be written as a convex combination of t_k elements of Θ^2_k. Hence the image of

$$\varphi : (\mathbb{R}\langle \underline{X} \rangle_k)^{t_k} \to \mathbb{R}\langle \underline{X} \rangle_{2k}/_{\overset{\text{cyc}}{\sim}}$$
$$(g_i)_{i=1,\dots,t_k} \mapsto \pi\Big(\sum_{i=1}^{t_k} {g_i}^* g_i\Big)$$

equals $\pi(\Theta^2_{n,k})$. In $(\mathbb{R}\langle \underline{X} \rangle_k)^{t_k}$ let $S := \{g = (g_i) \mid \|g\| = 1\}$. Note that S is compact, thus $V := \varphi(S) \subseteq \pi(\Theta^2_{n,k})$ is compact as well. By [KS2, Lemma 3.2 (b)], a sum of hermitian squares which is cyclically equivalent to 0 is already equal to zero. Hence, since $0 \notin S$, we see that $0 \notin V$.

Let $(f_\ell)_\ell$ be a sequence in $\pi(\Theta^2_{n,k})$ which converges to $\pi(f)$ for some $f \in \mathbb{R}\langle \underline{X} \rangle_{2k}$. Write $f_\ell = \lambda_\ell v_\ell$ for $\lambda_\ell \in \mathbb{R}_{\ge 0}$ and $v_\ell \in V$. Since V is compact there exists a subsequence $(v_{\ell_j})_j$ of v_ℓ converging to $v \in V$. Then

$$\lambda_{\ell_j} = \frac{\|f_{\ell_j}\|}{\|v_{\ell_j}\|} \overset{j\to\infty}{\longrightarrow} \frac{\|\pi(f)\|}{\|v\|}.$$

Thus $f_\ell \overset{j\to\infty}{\longrightarrow} \pi(f) = \frac{\|\pi(f)\|}{\|v\|} v \in \pi(\Theta^2_{n,k})$. ∎

Since Proposition 1.19 holds true for all $k \in \mathbb{N}_0$, we get the closedness of Θ^2 with respect to the finest locally convex topology on $\mathbb{R}\langle \underline{X} \rangle$.

1.20 Corollary. *The cone* Θ^2 *is closed in* $\mathbb{R}\langle \underline{X} \rangle$.

We will need later the concept of a tracial state which is intimately connected to the cone Θ^2.

1.21 Definition. Let $\mathcal{A}$ be an $\mathbb{R}$-algebra with involution $*$. We call a linear map $L : \mathcal{A} \to \mathbb{R}$ a *state* if $L(1) = 1$, $L(a^*a) \geq 0$ and $L(a^*) = L(a)$ for all $a \in \mathcal{A}$. If all the commutators have value 0, i.e. if $L(ab) = L(ba)$ for all $a, b \in \mathcal{A}$, then L is called a *tracial state*.

A state on $\mathbb{R}\langle \underline{X} \rangle$ is therefore a linear map L (with $L(f^*) = L(f)$) satisfying $L(\Sigma^2) \subseteq [0, \infty)$. If $L(\Theta^2) \subseteq [0, \infty)$, we have in particular $L(pq - qp) = 0$, hence L is a tracial state.

1.4.2 Tracial quadratic module

A quadratic module of the $\mathbb{R}$-algebra $\mathbb{R}\langle \underline{X} \rangle$ is a subset M of $\mathcal{S}\mathbb{R}\langle \underline{X} \rangle$ such that

$$M + M \subseteq M, p^*Mp \subseteq M \text{ for all } p \in \mathbb{R}\langle \underline{X} \rangle \text{ and } 1 \in M.$$

As for the cone of sums of hermitian squares we add commutators to the elements $f \in M$ to obtain its tracial analog.

1.22 Definition. The *tracial quadratic module* trM of a given quadratic module M in $\mathbb{R}\langle \underline{X} \rangle$ is defined as

$$\mathrm{trM} := \{f \in \mathcal{S}\mathbb{R}\langle \underline{X} \rangle \mid f \overset{\text{cyc}}{\sim} h \text{ for some } h \in M\}.$$

For $\underline{g} = (g_1, \ldots, g_r)$ with $g_1, \ldots, g_r \in \mathbb{R}\langle \underline{X} \rangle$ the quadratic module $\mathrm{QM}(\underline{g})$ generated by $\underline{g}$ in $\mathbb{R}\langle \underline{X} \rangle$ is the smallest quadratic module in $\mathbb{R}\langle \underline{X} \rangle$ containing $g_1, \ldots, g_r$. It consists of all sums of elements of the form p^*g_ip for $i = 0, \ldots, r$, where $p \in \mathbb{R}\langle \underline{X} \rangle$ and $g_0 := 1$. Again, by adding commutators, we obtain its tracial analog.

Let $\underline{g} = (g_1, \ldots, g_r) \in (\mathbb{R}\langle \underline{X} \rangle)^r$ be given and let $g_0 := 1$. The tracial quadratic module $\mathrm{trQM}(\underline{g})$ generated by $\underline{g}$ in $\mathbb{R}\langle \underline{X} \rangle$ is the set of all symmetric polynomials cyclically equivalent to an element in $\mathrm{QM}(\underline{g})$, i.e.

$$\begin{aligned} \mathrm{trQM}(\underline{g}) : &= \{f \in \mathcal{S}\mathbb{R}\langle \underline{X} \rangle \mid f \overset{\text{cyc}}{\sim} h \text{ for some } h \in \mathrm{QM}(\underline{g})\} \\ &= \{f \in \mathcal{S}\mathbb{R}\langle \underline{X} \rangle \mid f \overset{\text{cyc}}{\sim} \sum_{j=1}^{N} \sum_{i=0}^{r} p_{ij}{}^* g_i p_{ij} \text{ for some } p_{ij} \in \mathbb{R}\langle \underline{X} \rangle, N \in \mathbb{N}\}. \end{aligned}$$

For example, Θ^2 is the tracial quadratic module generated by $g_0 = 1$ in $\mathbb{R}\langle \underline{X} \rangle$. Any element $f \in \mathrm{trQM}(\underline{g})$ is trace-positive on the set

$$K(\underline{g}) := \{\underline{A} \in \mathcal{S}^n \mid g_i(\underline{A}) \succeq 0 \text{ for all } i = 1, \ldots, r\},$$

which replaces the semialgebraic set $K \subseteq \mathbb{R}^n$ defined by $g_1, \ldots, g_r$ in $\mathbb{R}[\underline{x}]$.

Be aware that in general $f \in \mathrm{trQM}(\underline{g})$ does not imply that f is trace-positive on

$$K_{\mathrm{Tr}}(\underline{g}) := \{\underline{A} \in \mathcal{S}^n \mid \mathrm{Tr}(g_i(\underline{A})) \geq 0 \text{ for all } i = 1 \ldots, r\},$$

since the product p^*g_ip, which is cyclically equivalent to the product pp^*g_i of the trace-positive polynomial g_i and the matrix-positive polynomial pp^*, might not be trace-positive.

1.23 Example. The non-commutative hypercube is defined by $g_i = 1 - X_i^2$ for $i = 1, \dots, n$, and will be denoted by

$$K_{\text{hc}} := \{\underline{A} \in \mathcal{S}^n \mid \mathbf{1} - A_i^2 \succeq 0 \text{ for all } i = 1, \dots, n\}.$$

One easily sees that $K_{\text{hc}} = \{\underline{A} \in \mathcal{S}^n \mid \|A_i\| \leq 1 \text{ for all } i = 1, \dots, n\}$, hence the non-commutative hypercube consists of all n-tuples of symmetric contractions. The corresponding tracial quadratic module will be denoted by trQM_{hc}.

1.5 Von Neumann algebras

To present an exact formulation of Connes' embedding conjecture (see Section 2.2), we need some preliminaries on von Neumann algebras. We will present the basic notions and some related results needed. The main reference for this introduction is [Tak].

A *von Neumann algebra* N is a unital $*$-subalgebra of the $*$-algebra $L(H)$ of bounded operators on a Hilbert space H that is closed in the weak operator topology. This can also be described in an algebraic way by the double commutant theorem of von Neumann [vN1]. Let N be a unital $*$-subalgebra of $L(H)$. The *commutant* of N is then defined as

$$N' = \{x \in L(H) \mid xa = ax \text{ for every } a \in N\}.$$

The double commutant theorem states that the following are equivalent:

(i) N is closed in the weak operator topology,

(ii) N is closed in the strong operator topology,

(iii) $N'' := (N')' = N$.

Thus a unital $*$-algebra N of bounded operators is a von Neumann algebra if and only if it is equal to its bicommutant N''.

A *factor* $\mathcal{F}$ is a von Neumann algebra with trivial center, i.e. a center which consists only of scalar multiples of the identity operator. Von Neumann proved that every von Neumann algebra on a separable Hilbert space is isomorphic to a direct integral of factors [vN2]. Thus one considers only separable factors instead of arbitrary von Neumann algebras acting on a separable Hilbert space. A factor $\mathcal{F}$ is *separable* if it can be represented faithfully into $L(H)$ where H is a separable Hilbert space. Equivalently, its predual $\mathcal{F}_*$, which is the unique Banach space X such that the Banach space dual $X^\vee$ is equal to $\mathcal{F}$, is norm-separable.

A factor $\mathcal{F}$ is *finite* if it possesses a normal, faithful, tracial state $\tau : \mathcal{F} \to \mathbb{C}$. This tracial state τ, called the canonical center valued trace, is unique and gives rise to the Hilbert-Schmidt norm on $\mathcal{F}$ given by $\|a\|_2^2 := \tau(a^*a)$ for $a \in \mathcal{F}$. This norm induces on $\mathcal{F}$ a topology which coincides on bounded sets with the strong operator topology.

Factors $\mathcal{F}$ can be classified into types by the behaviour of projections in $\mathcal{F}$. This was an early achievement of Murray and von Neumann [MvN1]. A *projection* $p \in N$ in a von Neumann algebra N is an operator satisfying $p = p^* = p^2$. Two projections are equivalent if there is an $a \in N$ such that $p = a^*a$ and $q = aa^*$. A given projection p is finite, if there is no $q \in N$, equivalent to p but $q \neq p$, such that $p - q = a^*a$ for some $a \in N$.

We are only interested in finite factors equipped with a canonical (center valued) trace τ. If the range of τ over all projection $p \in \mathcal{F}$ is discrete, then $\mathcal{F}$ is of type I. The classification of these algebras is complete as they are isomorphic to $L(H)$ for some finite-dimensional Hilbert space H. Hence any finite type I factor $\mathcal{F}$ is isomorphic to a matrix algebra over $\mathbb{C}$.

The key objects for studying finite von Neumann algebras are thus II_1 factors, i.e. factors where τ maps projections (surjectively) onto $[0, 1]$. An important question is to which extent II_1 factors are close to matrix algebras. Murray and von Neumann showed that there is a unique II_1 factor $\mathcal{R}$ which is generated as a von Neumann algebra by a union of an increasing sequence of finite-dimensional von Neumann subalgebras [MvN2]. This factor $\mathcal{R}$ is called the *hyperfinite* II_1 *factor*. Let τ_0 be its trace. They are several constructions of $\mathcal{R}$, e.g., as group von Neumann algebra of a discrete countable, amenable group with the infinite conjugacy class (i.c.c.) property or as the infinite tensor product $\overline{\bigotimes}_{n\in\mathbb{N}}(\mathbb{C}^{2\times 2})$ of the von Neumann algebras $\mathbb{C}^{2\times 2}$, which is the weak closure of the algebraic tensor product $\bigotimes_{n\in\mathbb{N}}(\mathbb{C}^{2\times 2})$.

Finally, we need the ultrapower $\mathcal{R}^\omega$ of the hyperfinite II_1 factor. Let $(a_k)_{k\in\mathbb{N}}$ be a sequence in a Hausdorff space X and ω be an ultrafilter on $\mathbb{N}$. Then $\lim_{k\to\omega} a_k = a$ means that for every neighbourhood U of a we have $\{k \in \mathbb{N} \mid a_k \in U\} \in \omega$. This limit is unique and exists for compact X. Consider the C^*-algebra

$$\ell^\infty(\mathcal{R}) := \{(a_k)_{k\in\mathbb{N}} \in \mathcal{R}^{\mathbb{N}} \mid \sup_{k\in\mathbb{N}} \|a_k\| < \infty\}$$

endowed with the supremum norm. Every ultrafilter ω on $\mathbb{N}$ defines a closed ideal

$$I_\omega := \{(a_k)_{k\in\mathbb{N}} \in \ell^\infty(\mathcal{R}) \mid \lim_{k\to\omega} \|a_k\|_2 = 0\}$$

in $\ell^\infty(\mathcal{R})$. The quotient C^* algebra $\mathcal{R}^\omega := \ell^\infty(\mathcal{R})/I_\omega$ is called the *ultrapower* of $\mathcal{R}$ (with respect to ω) and is a II_1 factor with trace $\tau_{0,\omega} : (a_k)_{k\in\mathbb{N}} + I_\omega \mapsto \lim_{k\to\omega} \tau_0(a_k)$.

1.6 Measure Theory

In this section we present the basic terminology concerning Borel measures and an auxiliary proposition on sequences of Borel measures which will be needed in Chapter 4. The main reference for the following is [Rud].

Let X be a locally compact Hausdorff space. Being locally compact means that for each $x \in X$ there is an open set U containing x, whose closure $\overline{U}$ is compact. We will later set $X = \mathbb{R}^n$ or $X = (\mathcal{S}\mathbb{R}^{s\times s})^n$ for some $s \in \mathbb{N}$. We consider X as topological space and write $\mathcal{C}_c(X)$ for the set of all continuous real valued functions $f : X \to \mathbb{R}$ with compact support. By $\mathcal{C}_0(X)$ we denote the set of continuous real valued functions f on X that vanish at infinity, i.e. for all $\varepsilon > 0$ the set $\{A \in X \mid |f(A)| \geq \varepsilon\}$ is compact. The space $\mathcal{C}_0(X)$ equipped with the supremum norm $\|f\|_\infty = \sup_{A\in X} |f(A)|$ is a Banach space. Further $\mathcal{C}_0(X)$ is the completion of $\mathcal{C}_c(X)$ relative to the supremum norm, hence $\mathcal{C}_c(X)$ is dense in $\mathcal{C}_0(X)$; and for compact X equality holds, i.e. $\mathcal{C}_0(X) = \mathcal{C}_c(X)$.

Let $\mathcal{B}(X)$ denote the Borel σ-algebra of X, i.e. the smallest collection of subsets of X containing all open sets and being closed under set differences, countable unions and intersections. A Borel measure μ on X is a function $\mu : \mathcal{B}(X) \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ satisfying $\mu(\varnothing) = 0$ and $\mu(\bigcup_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} \mu(A_i)$ for any pairwise disjoint Borel sets $A_i \in \mathcal{B}(X)$. By a *measure* we will always mean a Borel measure. In particular, any measure μ is assumed to be positive. If $\mu(A_i) < \infty$ for all $A_i \in \mathcal{B}(X)$ then the measure μ is finite. A probability measure is a finite measure with $\mu(X) = 1$.

Given a measure μ on $(\mathcal{S}\mathbb{R}^{s\times s})^n$ its support $\operatorname{supp}\mu$ is the smallest closed set $S \subseteq (\mathcal{S}\mathbb{R}^{s\times s})^n$ for which $\mu((\mathcal{S}\mathbb{R}^{s\times s})^n \setminus S) = 0$. If $\operatorname{supp}\mu \subseteq K$ for some set $K \subseteq (\mathcal{S}\mathbb{R}^{s\times s})^n$, we say that μ is supported in K.

Let $\mathcal{C}_0(X)'$ denote the dual Banach space of $\mathcal{C}_0(X)$ consisting of all continuous linear maps $L : \mathcal{C}_0(X) \to \mathbb{R}$. We call L *positive* if $L(f) \geq 0$ for all $f \in \mathcal{C}_0(X)$ that only takes positive values and denote the subspace of positive functionals in $\mathcal{C}_0(X)'$ by $\mathcal{C}_0(X)'_{\geq 0}$. The Riesz representation theorem states then that $\mathcal{C}_0(X)'_{\geq 0}$ is isometrically isomorphic to the space of finite regular measures on X equipped with the norm $\|\mu\| = \mu(X)$. That is, to each finite regular Borel measure μ on X corresponds a functional $\hat{\mu} \in \mathcal{C}_0(X)'$ defined by

$$\hat{\mu}(f) = \int_X f \, d\mu \text{ for all } f \in \mathcal{C}_0(X)$$

and vice versa. In general, the Riesz representation theorem states that the whole space $\mathcal{C}_0(X)'$ is isometrically isomorphic to the space of finite regular Borel measures on X. However, since we suppose that a measure is positive we only have an isomorphism on the subspace $\mathcal{C}_0(X)'_{\geq 0}$.

The following proposition is a consequence of the Lebesgue monotone convergence theorem. A similar statement for $K \subseteq \mathbb{R}^n$, i.e. the case $s = 1$, has been shown by Stochel [Sto, Prop. 1].

1.24 Proposition. *Les K be a closed subset of $(\mathcal{S}\mathbb{R}^{s\times s})^n$ for some $s \in \mathbb{N}$ and let $\varrho : K \to \mathbb{R}$ be a positive continuous function. Further, let $(\mu_k)_{k\in\mathbb{N}}$ be a sequence of finite measures supported in K and μ be a finite measure supported in K with $\lim_{k\to\infty} \int f \, d\mu_k = \int f \, d\mu$ for all $f \in \mathcal{C}_0(K)$ and $\sup_k \int \varrho \, d\mu_k < \infty$. Then $\int \varrho \, d\mu < \infty$ and*

$$\lim_{k\to\infty} \int f \varrho \, d\mu_k = \int f \varrho \, d\mu \text{ for all } f \in \mathcal{C}_0(K).$$

The proof works with the same line of reasoning as [Sto, Prop. 1].

Proof. If K is compact, the supremum $\|\varrho\|_\infty = \sup \varrho$ is attained. Hence

$$\int \varrho d\mu \leq \|\varrho\|_\infty \int d\mu < \infty.$$

If K is not compact, let $(U_\ell)_{\ell\in\mathbb{N}}$ be an increasing sequence of compact subsets of K satisfying $\bigcup_{\ell=1}^\infty U_\ell = K$. By Urysohn's lemma [Rud, 2.12] there is for each $\ell \geq 1$ a continuous function $\tau_\ell \in \mathcal{C}_c(K)$ such that $0 \leq \tau_\ell \leq 1$ and $\tau_\ell = 1$ on U_ℓ. Then, by the Lebesgue monotone convergence theorem, we obtain

$$\begin{aligned} \int \varrho d\mu = \lim_{\ell\to\infty} \int_{U_\ell} \varrho \, d\mu &\leq \limsup_{\ell\to\infty} \int_{U_\ell} \varrho \, d\mu \\ &\leq \limsup_{\ell\to\infty} \lim_{k\to\infty} \int \tau_\ell \varrho \, d\mu_k \leq \limsup_{k\to\infty} \int \varrho \, d\mu_k < \infty. \end{aligned}$$

Hence the Borel measure ν, given by $d\nu(\underline{A}) = \varrho(\underline{A}) \, d\mu(\underline{A})$, is finite. We define the Borel measures ν_k on K analogously by $d\nu_k(\underline{A}) = \varrho(\underline{A}) \, d\mu_k(\underline{A})$. Then, by assumption, the sequence $(\widehat{\nu}_k)_{k\in\mathbb{N}}$ of linear functionals is uniformly bounded and converges pointwise to $\widehat{\nu}$ on $\mathcal{C}_c(K)$. Since $\widehat{\nu} \in \mathcal{C}_0(K)'$ and $\mathcal{C}_c(K)$ is dense in $\mathcal{C}_0(K)$, we get that $(\widehat{\nu}_k)_k$ converges to $\widehat{\nu}$ in the $\sigma(\mathcal{C}_0'(K), \mathcal{C}_0(K))$-topology which is exactly what we wanted to show. ∎

2 Conjectures

In this chapter we give an overview of mostly algebraic results on two famous conjectures concerning trace-positivity: The Bessis-Moussa-Villani (BMV) conjecture from quantum statistical physics and Connes' embedding conjecture from operator algebra. These conjectures are the main motivation for the investigation of trace-positive polynomials and the tracial moment problem.

2.1 The BMV conjecture

In their attempt to extend variational and perturbation methods to quantum statistical mechanics Bessis, Moussa and Villani conjectured 1975 [BMV] that for any hermitian $s \times s$ matrices A and B with B positive semidefinite, the function

$$\begin{aligned} \varphi_{A,B} : \mathbb{R} &\to \mathbb{R} \\ t &\mapsto \operatorname{tr}(e^{A-tB}) \end{aligned}$$

is a Laplace transform of a positive measure $\mu_{A,B}$ supported in $\mathbb{R}_{\geq 0}$.

2.1 Conjecture (Bessis, Moussa, Villani)**.** *For all hermitian $A, B \in \mathbb{C}^{s \times s}$, $s \in \mathbb{N}$, where $B \succeq 0$, there exists a positive measure $\mu_{A,B}$ supported in $\mathbb{R}_{\geq 0}$ such that*

$$\varphi_{A,B}(t) = \int e^{-tx}\, d\mu_{A,B}(x).$$

If this conjecture was true it would imply that a series of Padé approximations of partition functions in quantum statistical mechanics converges. This would lead to explicit lower and upper bounds of energy levels in multiple particle systems by a generalization of the well-know Rayleigh-Ritz variational principle [BMV]. Bessis, Moussa and Villani showed in the same paper [BMV] that Conjecture 2.1 holds for some physical examples, for instance for an N particle bosonic system in a local interaction without external magnetic field, whose partition function can be represented by Wiener integrals. Further, Conjecture 2.1 holds true for commuting matrices and for symmetric matrices A, B where A is diagonal and B has only positive entries. From this one deduces easily that the BMV conjecture holds for symmetric 2×2-matrices.

Since its introduction about 35 years ago many other partial results mostly obtained by analytic methods have been given, see e.g. [Mou] for a review up to the year 1998. Conjecture 2.1 is true in an average sense. That is, the expectation of $\varphi_{A,B}$ over independently distributed Gaussian random matrices A, B is positive if A and B are sufficiently large. More specific, the BMV conjecture holds for semicircular self-adjoint operators in a type II_1 von Neumann algebra with faithful normal tracial state τ, which are in a free relation; see [FP] for details, or [Boz] for an extension to generalized random matrices. Drmota, Schachermayer and Teichmann [DST] showed that Conjecture 2.1 is true for 3×3-matrices of a specific structure. In their approach they use hypergeometric identities to reduce the BMV conjecture in their case to a summation problem in the theory of hypergeometric series. The advantage of their approach is, that they can explicitly construct a positive measure $\mu_{A,B}$.

By Bernstein's theorem, Conjecture 2.1 is equivalent to the question whether the function $\varphi_{A,B}$ is completely monotone, i.e.

$$(-1)^\ell \frac{d^\ell}{dt^\ell} \varphi_{A,B}(t) \geq 0 \text{ for all } \ell \in \mathbb{N}_0, t \in \mathbb{R}_{\geq 0}. \tag{2.1}$$

In 2004 Lieb and Seiringer [LiS] opened the way to more algebraic attempts on the BMV conjecture. They restated Conjecture 2.1 in the following purely algebraic form, which follows from (2.1) with help of the Taylor expansion series of the exponential function [LiS, Theorem 1]. We present a slightly different formulation which is equivalent to the original one.

2.2 Conjecture. *For all positive semidefinite matrices $A, B \in \mathcal{S}\mathbb{R}^{s\times s}$, $s \in \mathbb{N}$, and all $m \in \mathbb{N}_0$, the polynomial $p(t) := \operatorname{Tr}((A+tB)^m) \in \mathbb{R}[t]$ has only positive coefficients.*

This reformulation allows for numerical experiments which have been extensively carried out. So far all these experiments agree with the BMV conjecture and lead to partial results confirming it.

Let $S_{m,k}(X,Y) \in \mathbb{R}\langle \underline{X}\rangle$ be the polynomial obtained by adding all words of total degree m where Y appears exactly k times. For example,

$$S_{4,2}(X,Y) = X^2Y^2 + XYXY + XY^2X + YXYX + Y^2X^2 + YX^2Y.$$

Then the coefficient of t^k in $p(t) = \operatorname{Tr}((A+tB)^m)$ for a given m is equal to the trace of $S_{m,k}(A,B)$. To model the positive semidefiniteness of A and B we consider $S_{m,k}(X^2,Y^2)$ as a polynomial in X^2 and Y^2 if necessary. Thus the Lieb-Seiringer reformulation asks if for all (m,k) the polynomial $S_{m,k}(X^2,Y^2)$ is trace-positive. We call these polynomials $S_{m,k}(X^2,Y^2)$ the BMV polynomials since they are intimately connected to the BMV conjecture. In literature, the expression $S_{m,k}(A,B)$ is often called the k-th Hurwitz product of A and B.

It is easy to see that the BMV polynomials are in general not matrix-positive. For instance, with

$$A = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{3/8} \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 4 \\ 4 & 32 \end{bmatrix}$$

we have

$$S_{2,1}(A^2,B^2) = A^2B^2 + B^2A^2 = \begin{bmatrix} 34 & 363/2 \\ 363/2 & 780 \end{bmatrix},$$

which is not positive semidefinite. Furthermore, not all words appearing in a BMV polynomial are trace-positive itself. For example, the word $X^2Y^2X^4Y^4$ appearing in $S_{6,3}(X^2,Y^2)$ can have negative trace [HJ1, p. 919]. In fact, $\operatorname{Tr}(ABA^2B^2) = -1582$ for the positive semidefinite matrices

$$A = \begin{bmatrix} 1 & 20 & 210 \\ 20 & 402 & 4240 \\ 210 & 4240 & 44903 \end{bmatrix} \text{ and } B = \begin{bmatrix} 36501 & -3820 & 190 \\ -3820 & 401 & -20 \\ 190 & -20 & 1 \end{bmatrix}.$$

This implies that $X^2Y^2X^4Y^4$ has negative trace if one replaces X^2, Y^2 by A, B. Alternatively, one can replace X and Y by the square roots $\sqrt{A}, \sqrt{B}$ to obtain the same result. These events are extremely rare and lie in a narrow range [HJ1]. Therefore there seems to be no hope of finding a counter example for Conjecture 2.2 by trying randomly generated pairs of matrices.

There have been several approaches to prove Conjecture 2.2. Using a variational approach Hillar introduced a fundamental pair of matrix equations, called Euler-Lagrange equations, which are satisfied by matrices A and B that minimize for some fixed pair (m,k) the quantity $\operatorname{Tr}(S_{m,k}(X,Y))$ over all positive semidefinite matrices of norm 1 [Hil, Theorem 1.3], see also [HJ2, Theorem 2.4]. Using this equations he reduced Conjecture 2.2 to the case of singular matrices A, B if one wants to show it by induction over $m \in \mathbb{N}_0$ [Hil, Theorem 1.13]. Further, he deduced that if Conjecture 2.2 is true for some fixed $M \in \mathbb{N}_0$ then it is also true for all $m < M$. More precisely, he showed the following descent theorem [Hil, Theorem 1.10].

2.3 Theorem (Hillar)**.** *Let $M, K \in \mathbb{N}_0$ with $M \geq K$. If the polynomial $S_{M,K}(X^2, Y^2)$ is trace-positive, then $S_{m,k}(X^2, Y^2)$ is trace-positive for all pairs $(m, k) \in \mathbb{N}_0^2$ satisfying $m \leq M, k \leq K$ and $m - k \leq M - K$.*

Thus one only needs the asymptotic behaviour of Conjecture 2.2 for large $m \in \mathbb{N}_0$. Following an analytic or alternatively a combinatorial approach, one can show that for fixed positive semidefinite matrices A, B and fixed $k \in \mathbb{N}_0$ the trace of $S_{m,k}(A, B)$ is positive whenever m is large enough [FF, Theorem 1.1]. Since this lower bound on m is dependent of A and B, this result does not imply the BMV conjecture via Theorem 2.3.

In the last few years there has been much activity around the strategy based on the work of Hägele [Häg] to identify the pairs (m, k) for which $S_{m,k}(X^2, Y^2) \in \Theta^2$ holds. An affirmative answer for all (m, k) would imply the BMV conjecture. In fact, not all pairs (m, k) are needed due to Theorem 2.3. This approach has been investigated thoroughly. On the one hand, there are positive results showing that $S_{m,k}(X^2, Y^2) \in \Theta^2$ holds for some specific (m, k) as well as for some infinite classes only depending on m. On the other hand, this approach does not provide a proof for Conjecture 2.2. In the sequel we will give an overview of these results.

2.1.1 Positive results of the Θ^2-approach

By exchanging the variables X and Y it is clear that $S_{m,k}(X^2, Y^2) \in \Theta^2$ holds if and only if $S_{m,m-k}(X^2, Y^2) \in \Theta^2$ holds. Although in general not every word in $S_{m,k}(X^2, Y^2)$ is trace-positive, as shown above, this holds true for $0 \leq k \leq 2$ (or $m - 2 \leq k \leq m$). In fact, each word in $S_{m,k}(X^2, Y^2)$ where $0 \leq k \leq 2$ is cyclically equivalent to a hermitian square of a word, and hence lies in Θ^2 and has positive trace. Therefore Conjecture 2.2 is true for $m \leq 5$. This was first mentioned explicitly by Hillar and Johnson [HJ1, Corollary 5]. They also verified the first non-trivial case, namely $m = 6, k = 3$, for positive semidefinite 3×3 matrices with help of a computer algebra system [HJ2, Theorem 4.1].

Hägele [Häg] laid the foundation to the attempt of writing $S_{m,k}(X^2, Y^2)$ as sum of hermitian squares and commutators. He verified the case $m = 7$. Theorem 2.3 then implies that the BMV conjecture holds true for all $m \leq 7$. By exploiting Hägele's approach Klep and Schweighofer derived that Conjecture 2.2 is true for $m \leq 13$ [KS2, Theorem 1.2]. Using the tracial Gram matrix method, which will be explained in Section 3.2, and semidefinite programming, they found exact representations of $S_{14,4}(X^2, Y^2)$ and $S_{14,6}(X^2, Y^2)$ as sum of hermitian squares and commutators with rational coefficients. By Theorem 2.3, this implies Conjecture 2.2 for $m \leq 13$.

These results give concrete representations of $S_{m,k}(X^2, Y^2)$ as a sum of hermitian squares and commutators for fixed m and fixed k. We found by a combination of numerical experiments and combinatorial arguments a representation of $S_{m,k}(X^2, Y^2)$ as a sum of hermitian squares and commutators for $k = 4$ (or $k = m - 4$) and *arbitrary* $m \geq 4$. A proof of this result can be found in Section 3.4 of this work or in [Bur]. Independently, Landweber and Speer [LS, Theorem 2] showed the same result for *odd* $m \in \mathbb{N}$ without results in the even case.

To summarize, the BMV polynomials $S_{m,k}(X^2, Y^2)$ lie in Θ^2 if $k = 0, 1, 2, 4$ and $m \geq k$, or if $(m, k) \in \{(14, 6), (14, 8), (7, 3), (11, 3), (11, 8)\}$.

2.1.2 Negative results of the Θ^2-approach

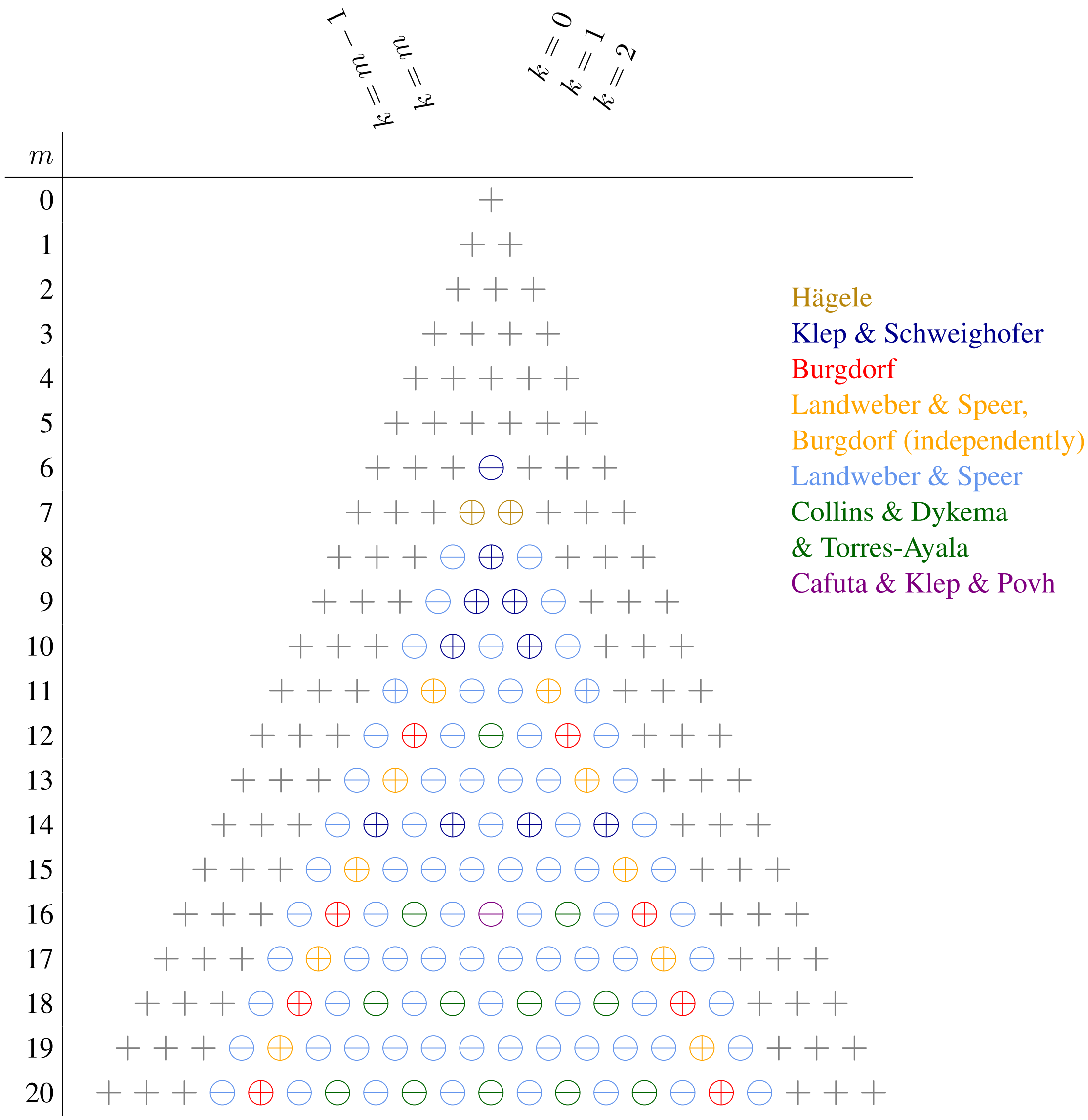
However, the above mentioned cases are the only cases where the BMV polynomial $S_{m,k}(X^2, Y^2)$ admits a Θ^2-certificate. First, Hägele showed [Häg, p. 1169f] that, in contrast to $S_{7,4}(X^2, Y^2)$, which is a sum of hermitian squares and commutators, the polynomial

$$6X^6Y^6 + 6X^4Y^2X^2Y^4 + 6X^4Y^4X^2Y^2 + 2X^2Y^2X^2Y^2X^2Y^2,$$

which is cyclically equivalent to $S_{6,3}(X^2, Y^2)$, is not a sum of hermitian squares $\sum_i {g_i}^* g_i$ where the g_i are of the form $a_i X^3 Y^3 + b_i XY^2 X^2 Y$ for some $a_i, b_i \in \mathbb{C}$. However, he speculated that a more general g_i may give a sum of hermitian squares and commutators. Klep and Schweighofer finally showed that $S_{6,3}(X^2, Y^2) \notin \Theta^2$ [KS2, Example 3.5]. They found a reduction of words needed in a representation as a sum of hermitian squares and commutators. In fact, if there is no representation with this reduced set of words, then there is no such representation at all. This implies that Hägele's argument is sufficient to obtain $S_{6,3}(X^2, Y^2) \notin \Theta^2$.

Exploring the methods of Hägele and of Klep and Schweighofer, Landweber and Speer [LS] proved that the BMV polynomials $S_{m,k}(X^2, Y^2)$ do not lie in Θ^2 if $k = 3$ (or $k = m - 3$) and $m \geq 6$, $m \neq 7, 11$; or if $5 \leq k \leq m - 5$, $m \geq 10$ and m or k odd. Collins, Dykema and Torres-Ayala [CDT] extended this result to all remaining cases, except the case $(m, k) = (16, 8)$, where they only got numerical evidence but no exact proof for the fact that $S_{16,8}(X^2, Y^2) \notin \Theta^2$. This last case has recently been solved by Cafuta, Klep and Povh [CKP2, Corollary 2.6]. They show that $S_{16,8}(X^2, Y^2) \notin \Theta^2$ by giving an exact hyperplane with rational parameters that separates $S_{16,8}(X^2, Y^2)$ from Θ^2.

A graphical overview of the positive and negative results concerning the question whether $S_{m,k}(X^2, Y^2) \in \Theta^2$ holds for fixed (m, k), is given by the following tree:



The cases $k = 0, 1, 2$ are trivial, the other cases (m, k) have been proved ($\oplus$) or disproved ($\ominus$) to satisfy the Θ^2-certificate. The colouring indicates the persons who were involved in the specific case. The tree continues as in the lines 17–20.

2.2 Connes' embedding conjecture

Another important conjecture which motivates the study of trace-positive polynomials is the embedding conjecture of Alain Connes, one of the most important open problems in Operator Algebras. This conjecture would imply a fundamental approximation property for finite II_1 von Neumann algebras. Furthermore, it is related to Lance's WEP conjecture [Kir] concerning C^*-algebras, and to the question whether all i.c.c. groups are hyperfinite [R2]. Alain Connes stated 1976 in his ingenious paper on the classification of injective factors [Con, Section V] the following conjecture.

2.4 Conjecture (Connes)**.** *If ω is a free ultrafilter on $\mathbb{N}$ and $\mathcal{F}$ is a II_1 factor with separable predual, then $\mathcal{F}$ can be embedded into the ultrapower $\mathcal{R}^\omega$.*

This conjecture is not only interesting in of itself. Kirchberg [Kir] has shown that Connes' embedding conjecture has several other equivalent reformulations in operator algebras and Banach space theory which are related to the QWEP conjecture for separable von Neumann algebras. For example, there is the statement that there exists a unique C^*-norm on the tensor product of the universal C^*-algebra of a free group with itself. Moreover, Voiculescu [Voi] introduced the notion of free entropy in free probability theory following the classical concept of Boltzmann, whose behaviour is intimately connected with Connes' embedding conjecture. More concretely, Conjecture 2.4 is equivalent to the question of whether every generating set for $\mathcal{F}$ has matricial microstates [Con, Voi].

2.5 Definition. Let N be a von Neumann algebra and $\mathcal{X} = \{A_1, \dots, A_n\}$ be a finite subset of $\mathcal{S}N := \{A \in N \mid A^* = A\}$. We say that $\mathcal{X}$ has *matricial microstates* if for every $k \in \mathbb{N}$, every $\varepsilon \in \mathbb{R}_{>0}$ there exists an $s \in \mathbb{N}$ and self-adjoint $s \times s$-matrices $B_1, \dots, B_n$ such that for every $w \in \langle \underline{X} \rangle_k$:

$$|\tau(w(A_1, \dots, A_n)) - \mathrm{Tr}(w(B_1, \dots, B_n))| < \varepsilon.$$

Voiculescu showed that without loss of generality one can choose the B_i to be of operator norm at most $\|A_i\|$ [Voi, Remark 2.5].

It is well-known that many II_1 factors do embed into an ultrapower of the hyperfinite II_1 factor, see for example [R2]. However Conjecture 2.4 still remains open and is the subject of deep ongoing research. Rădulescu established a relationship between Connes' embedding conjecture and some analog of Hilbert's $17th$ problem on positive polynomials [R3, Corollary 1.2]. Further, Hadwin [Had] studied a non-commutative moment problem concerning C^*-algebras. Both work with *weak limits of sums of hermitian squares*. The main idea of these approaches is to show (mostly via a Gelfand-Naimark-Segal construction) a specific statement involving the elements of an arbitrary II_1 von Neumann algebra with faithful trace τ which would also hold for matrices if Conjecture 2.4 is true.

In 2009, the same concept was used by Klep and Schweighofer to obtain a purely algebraic statement which is equivalent to Conjecture 2.4. Namely, the following statement on specific representations of polynomials which are trace-positive on the non-commutative hypercube is equivalent to Connes' embedding conjecture [KS1, Theorem 1.6]. Recall that the non-commutative hypercube

$$K_{\mathrm{hc}} = \{\underline{A} \in \mathcal{S}^n \mid \mathbf{1} - A_i^2 \succeq 0 \text{ for all } i = 1, \dots, n\}$$

is equal to the set of all tuples of symmetric contractions, see Example 1.23. Further, the corresponding tracial quadratic module $\mathrm{trQM}_{\mathrm{hc}}$ is the tracial quadratic module which is generated by $1 - X_i^2$ for $i = 1, \dots, n$.

2.6 Conjecture. *Suppose $f \in \mathcal{S}\mathbb{R}\langle \underline{X} \rangle$. Then the following are equivalent:*

(i) *f is trace-positive on K_{hc};*

(ii) *For all $\varepsilon \in \mathbb{R}_{>0}$, $f + \varepsilon$ lies in $\mathrm{trQM}_{\mathrm{hc}}$.*

The implication (ii) $\Longrightarrow$ (i) is obvious as elements in $\mathrm{trQM}_{\mathrm{hc}}$ are trace-positive on K_{hc}. Hence for any $\underline{A} \in K_{\mathrm{hc}}$ we have $\mathrm{Tr}(f(\underline{A})) \geq -\varepsilon$ for all $\varepsilon > 0$ which implies $\mathrm{Tr}(f(\underline{A})) \geq 0$. The implication (i) $\Longrightarrow$ (ii) is by [KS1, Theorem 1.5] equivalent to Conjecture 2.4. To show this, one uses the following theorem concerning positivity of polynomials on II_1 von Neumann algebras [KS1, Theorem 3.12].

2.7 Proposition (Klep, Schweighofer)**.** *For $f \in \mathcal{S}\mathbb{R}\langle \underline{X} \rangle$ the following statements are equivalent:*

(i) *$\tau(f(A_1, \dots, A_n)) \geq 0$ for every II_1 factor $\mathcal{F}$ with separable predual and faithful trace τ and all self-adjoint contractions $A_1, \dots, A_n \in \mathcal{F}$;*

(ii) *For every $\varepsilon \in \mathbb{R}_{>0}$, $f + \varepsilon \in \mathrm{trQM}_{\mathrm{hc}}$.*

From the proof it follows easily that it suffices to consider in statement (i) only II_1 factors $\mathcal{F}$ which are generated by n self-adjoint elements.

If f is trace-positive and Conjecture 2.4 holds, then $\tau(f(A_1, \dots, A_n)) \geq 0$ for all II_1 factors $\mathcal{F}$ with separable predual and faithful trace τ and all self-adjoint contractions $A_1, \dots, A_n \in \mathcal{F}$. Hence Proposition 2.7 gives $f + \varepsilon \in \mathrm{trQM}_{\mathrm{hc}}$. For the other direction one shows that Conjecture 2.6 implies that $\mathcal{F}$ has matricial microstates.

2.8 Remark. One can easily replace the non-commutative hypercube K_{hc} by the non-commutative unit ball $K(1 - X_1^2 - \cdots - X_n^2)$ and the corresponding tracial quadratic module generated by $1 - X_1^2 - \cdots - X_n^2$. The proof works similarly as the proof for the non-commutative hypercube, see [KS1], and will not be presented in this work.

2.2.1 Comparison with Positivstellensätze

Connes' embedding conjecture is intimately connected to the investigation of Positivstellensätze concerning trace-positivity. As seen above, Conjecture 2.6 is connected to a theorem on trace-positivity of polynomials on II_1 von Neumann algebras. It also reveals the analogies to Positivstellensätze in the commutative and the free non-commutative setting.

In the ring $\mathbb{R}[\underline{x}]$ of polynomials in commuting variables two words are cyclically equivalent if they are equal. Hence we can replace the cyclic equivalence by equality. Further, the involution becomes the identity. In this context, we have to evaluate the polynomial f in pairwise commuting symmetric matrices A_i. Thus they can be simultaneously diagonalized and it suffices to consider 1×1 matrices of norm ≤ 1, that is, the tuples $\underline{A}$ will be replaced by tuples $\underline{a}$ lying in the commutative hypercube $[-1, 1]^n$. One therefore obtains naturally the specific case of Putinar's Positivstellensatz, see also Theorem 1.13.

2.9 Theorem (Putinar)**.** *Let $f \in \mathbb{R}[\underline{x}]$. Then $f \geq 0$ on $[-1, 1]^n$ if and only if for all $\varepsilon \in \mathbb{R}_{>0}$, $f + \varepsilon$ lies in the quadratic module generated by $1 - x_i^2$ for $i = 1, \dots, n$.*

If we ask for a similar statement for polynomials $f \in \mathbb{R}\langle \underline{X} \rangle$ which are matrix-positive on K_{hc} instead of trace-positive, we omit adding commutators to f, that is, we consider the quadratic module $\mathrm{QM}_{\mathrm{hc}} := \mathrm{QM}(1 - X_1^2, \dots, 1 - X_n^2)$ in $\mathbb{R}\langle \underline{X} \rangle$. Thus the natural counterpart to Conjecture 2.6 in this setting is a particular case of [HM, Theorem 1.2].

2.10 Theorem (Helton, McCullough). *Let $f \in S\mathbb{R}\langle \underline{X} \rangle$. Then f is matrix-positive on K_{hc} if and only if for all $\varepsilon \in \mathbb{R}_{>0}$, $f + \varepsilon$ lies in* $\mathrm{QM}_{\mathrm{hc}}$.

Conjecture 2.6 holds true in the easiest non-commutative case: the case of cyclically sorted polynomials [KS1, Prop. 4.2], which have been introduced in Definition 1.10.

2.11 Proposition (Klep, Schweighofer). *Let $f \in S\mathbb{R}\langle X, Y \rangle$ be cyclically sorted. Then f is trace-positive on K_{hc} if and only if for all $\varepsilon \in \mathbb{R}_{>0}$, $f + \varepsilon$ lies in* $\mathrm{trQM}(1 - X^2, 1 - Y^2)$.

2.2.2 Reduction of parameters

The statement of Conjecture 2.6 allows an additional reduction. We will show that it suffices to show a seemingly easier statement where one reduces the degree of polynomials $f \in \mathbb{R}\langle \underline{X} \rangle$ in Conjecture 2.6.

We recall that Connes' embedding conjecture is equivalent to a question on matricial microstates. In an earlier work, Rădulescu found that to show Conjecture 2.4 it suffices to approximate the moments of elements of separable II_1 factors $\mathcal{F}$ by moments of matrices only for certain words, which have degree at most four. Namely, he proved the following proposition, which remarks in [R1].

2.12 Proposition (Rădulescu). *Let $\mathcal{F}$ be a II_1 factor with separable predual and τ its normal, faithful trace. If for every $n \in \mathbb{N}$, $\varepsilon \in \mathbb{R}_{>0}$ and every $A_1, \ldots, A_n \in \mathcal{F}$ there is an $s \in \mathbb{N}$ and matrices $B_1, \ldots, B_n \in \mathbb{C}^{s \times s}$ such that for all $i, j, k \in \{1, \ldots, n\}$ we have*

$$|\tau(A_i A_j) - \mathrm{Tr}(B_i B_j)| < \varepsilon, \quad |\tau(A_i A_j A_k) - \mathrm{Tr}(B_i B_j B_k)| < \varepsilon \text{ and}$$

$$|\tau(A_i^2 A_j^2) - \mathrm{Tr}(B_i^2 B_j^2)| < \varepsilon,$$

then for every $n, k \in \mathbb{N}$, $\varepsilon \in \mathbb{R}_{>0}$ and every $A_1, \ldots, A_n \in \mathcal{F}$ one can find an $s \in \mathbb{N}$ and matrices $B_1, \ldots, B_n \in \mathbb{C}^{s \times s}$ such that for all $w \in \langle \underline{X} \rangle_k$:

$$|\tau(w(A_1, \ldots, A_n)) - \mathrm{Tr}(w(B_1, \ldots, B_n))| < \varepsilon.$$

In particular, $\mathcal{F}$ is embeddable in $\mathcal{R}^\omega$ for all free ultrafilter ω on $\mathbb{N}$.

The following corollary is an immediate consequence of Proposition 2.12.

2.13 Corollary. *Let $\mathcal{F}$ be a separable II_1 factor with faithful trace τ. Then the following statements are equivalent:*

(i) *For every free ultrafilter ω on $\mathbb{N}$, $\mathcal{F}$ is embeddable in $\mathcal{R}^\omega$;*

(ii) *There exists a free ultrafilter ω on $\mathbb{N}$ such that $\mathcal{F}$ is embeddable in $\mathcal{R}^\omega$;*

(iii) *For all $\varepsilon \in \mathbb{R}_{>0}$, $n \in \mathbb{N}$ and self-adjoint contractions $A_1, \ldots, A_n \in \mathcal{F}$, there is an $s \in \mathbb{N}$ and self-adjoint contractions $B_1, \ldots, B_n \in \mathbb{C}^{s \times s}$ such that for all $w \in \langle \underline{X} \rangle_4$:*

$$|\tau(w(A_1, \ldots, A_n)) - \mathrm{Tr}(w(B_1, \ldots, B_n))| < \varepsilon;$$

(iv) *For all $\varepsilon \in \mathbb{R}_{>0}$, $n \in \mathbb{N}$ and all $A_1, \ldots, A_n \in \mathcal{F}$, there is an $s \in \mathbb{N}$ and matrices $B_1, \ldots, B_n \in \mathbb{C}^{s \times s}$ such that for all $w \in \langle \underline{X} \rangle_4$:*

$$|\tau(w(A_1, \ldots, A_n)) - \mathrm{Tr}(w(B_1, \ldots, B_n))| < \varepsilon.$$

Proof. The implication (i) $\implies$ (ii) is trivial. Implication (ii) $\implies$ (iii) is well known using the fact that $\mathcal{R}$ is generated as a von Neumann algebra by a union of an increasing sequence of finite-dimensional von Neumann subalgebras, see e.g. [Con, Voi] or [CD, Proposition 3.3]. In fact, (ii) implies that $\mathcal{F}$ has matricial microstates which implies (iii) by taking $k = 4$ in Definition 2.5. Statement (iii) implies (iv) since any element $A \in \mathcal{F}$ can be written as $\mathbb{C}$-linear combination of self-adjoint contractions in $\mathcal{F}$. If (iv) holds, we can apply Proposition 2.12 which implies (i). ∎

Using Corollary 2.13 one can reduce the statement of Conjecture 2.6 to a statement on polynomials $f \in \mathcal{S}\mathbb{R}\langle \underline{X} \rangle$ of degree four to obtain Conjecture 2.4. The proof is an adaptation of the proof of [KS1, Prop.3.17] with some modifications.

2.14 Proposition. *The following statements are equivalent:*

(i) *Connes embedding conjecture holds;*

(ii) *For each $n, k \in \mathbb{N}$ and $f \in \mathcal{S}\mathbb{R}\langle \underline{X} \rangle_k$, f trace-positive on K_{hc}, implies that for all $\varepsilon \in \mathbb{R}_{>0}$, $f + \varepsilon$ lies in $\mathrm{trQM}_{\mathrm{hc}}$;*

(iii) *For each $n \in \mathbb{N}$ and $f \in \mathcal{S}\mathbb{R}\langle \underline{X} \rangle_4$, f trace-positive on K_{hc}, implies that for all $\varepsilon \in \mathbb{R}_{>0}$, $f + \varepsilon$ lies in $\mathrm{trQM}_{\mathrm{hc}}$;*

Proof. To show (i) $\implies$ (ii) let $f \in \mathcal{S}\mathbb{R}\langle \underline{X} \rangle$ be trace-positive on K_{hc}. We will show that $\tau(f(\underline{A})) \geq 0$ for every II_1 factor $\mathcal{F}$ with separable predual and faithful trace τ and all self-adjoint contractions $A_1, \dots, A_n \in \mathcal{F}$. Then (ii) follows by Proposition 2.7.

Since $\mathcal{R}$ is generated as a von Neumann algebra by a union of an increasing sequence of finite-dimensional von Neumann subalgebras, we have $\tau_0(f(A_1, \dots, A_n)) \geq 0$ for all self-adjoint contractions $A_1, \dots, A_n \in \mathcal{R}$. Let ω be a free ultrafilter on $\mathbb{N}$. By construction of the ultrapower $\mathcal{R}^\omega$ we have

$$\tau_{0,\omega}(f(A_1, \dots, A_n)) \geq 0 \text{ for all self-adjoint contractions } A_1, \dots, A_n \in \mathcal{R}^\omega. \tag{2.2}$$

In fact, by continuity, we may assume that there exists $\varepsilon \in \mathbb{R}_{>0}$ such that $\|A_i\| \leq 1 - \varepsilon$ for $i = 1, \dots, n$. Then each A_i has a representative $(A_i^{(j)} + B_i^{(j)})_{j \in \mathbb{N}}$ where each $A_i^{(j)}$ is a self-adjoint contraction in $\mathcal{R}$ and $(B_i^{(j)})_{j \in \mathbb{N}} \in I_\omega$, which implies

$$\begin{aligned}\tau_{0,\omega}(f(A_1, \dots, A_n)) &= \lim_{j \to \omega} \tau_0(f(A_1^{(j)} + B_1^{(j)}, \dots, A_n^{(j)} + B_n^{(j)})) \\ &= \lim_{j \to \omega} \tau_0(f(A_1^{(j)}, \dots, A_n^{(j)})) \geq 0.\end{aligned}$$

Let $\mathcal{F}$ be a II_1 factor with separable predual and faithful trace τ. Since $\mathcal{F}$ is embeddable in $\mathcal{R}^\omega$ by (i), statement (2.2) implies that $\tau(f(\underline{A})) \geq 0$ for all self-adjoint contractions $A_1, \dots, A_n \in \mathcal{F}$.

The implication (ii) $\implies$ (iii) is obvious.

Instead of (iii) $\implies$ (i) we show that (iii) implies statement (iii) in Corollary 2.13 for every II_1 factor $\mathcal{F}$ with separable predual and faithful trace τ. Let such a factor $\mathcal{F}$ be fixed. Further fix $n \in \mathbb{N}$. To show the desired statement, consider the finite-dimensional $\mathbb{C}$-vector space $\mathbb{C}\langle \underline{X} \rangle_4$ and its dual space $\mathbb{C}\langle \underline{X} \rangle_4^*$. Let an n-tuple $\underline{A}$ of self-adjoint contractions in $\mathcal{F}$ be given and consider the linear form $L \in \mathbb{C}\langle \underline{X} \rangle_4^*$ given by

$$L(p) = \tau(p(\underline{A}))$$

for $p \in \mathbb{C}\langle \underline{X} \rangle_4$. Since L is defined via τ the linear form is tracial, that is, $L(pq) = L(qp)$ for all $p, q \in \mathbb{C}\langle \underline{X} \rangle$ with $\deg(pq) \leq 4$. Let C denote the convex hull of all linear forms on $\mathbb{C}\langle \underline{X} \rangle_4$ given by

$$p \mapsto \mathrm{Tr}(p(\underline{B})) \tag{2.3}$$

for some $s \in \mathbb{N}$ and $\underline{B} \in (\mathbb{C}^{s\times s})^n$ a tuple of self-adjoint contractions. Let $\overline{C} \subseteq \mathbb{C}\langle \underline{X} \rangle_4^*$ be its closure. We claim that $L \in \overline{C}$. Then every ε-neighbourhood of L in $\mathbb{C}\langle \underline{X} \rangle_4^*$ contains an element of C, i.e. a convex combination of linear forms as in (2.3). Since $\mathbb{Q}$ is dense in $\mathbb{R}$, every such neighbourhood also contains such a convex combination with rational coefficients. This can be built to a linear form as in (2.3) by taking matrices in block diagonal form, which implies (iii) of Corollary 2.13 and hence (i) since $\mathcal{F}$ was arbitrary.

To show the claim assume $L \notin \overline{C}$. Then by the complex Hahn-Banach separation theorem there is a polynomial $f \in \mathbb{C}\langle \underline{X} \rangle_4 \cong \mathbb{C}\langle \underline{X} \rangle_4^{**}$ such that $\operatorname{Re}(L(f)) < 0 < \operatorname{Re}(L'(f))$ for all $L' \in \overline{C}$. Let $g := f + f^* \in \mathcal{S}\mathbb{R}\langle \underline{X} \rangle_4$. The linear forms $L' \in \overline{C}$ are tracial and we have $L'(g) = L(f + f^*) = 2\operatorname{Re}(L'(f)) > 0$ for all $L' \in \overline{C}$. In particular, g is trace-positive on K_{hc} and thus by (iii) we have for all $\varepsilon \in \mathbb{R}_{>0}$ that $g + \varepsilon \in \mathrm{trQM}_{\mathrm{hc}}$. This implies $\tau(g(\underline{A})) \geq 0$ in contradiction to $L(g) = L(f + f^*) < 0$. ■

3 Sums of hermitian squares and commutators

The cone of matrix-positive polynomials coincides with the cone Σ^2 of sums of hermitian squares, see Theorem 2.10. On the other hand, we will see that not all trace-positive polynomials are sums of hermitian squares and commutators. This is in analogy to the commutative context: Not every positive polynomial is a sum of squares. Thus further investigation is needed.

In this chapter we present results concerning the cone Θ^2 of sums of hermitian squares and commutators. After a short introduction, presenting some classes of polynomials for which we can apply results from the polynomial ring $\mathbb{R}[\underline{x}]$, and giving some examples of polynomials which are trace-positive but not a member of Θ^2, we present in Section 3.2 the tracial analog of the Gram matrix method. This is a method to calculate possible sum of hermitian squares and commutators decompositions of polynomials. In the last two Sections 3.3 and 3.4 we present results on the question which trace-positive polynomials $f \in \mathbb{R}\langle \underline{X} \rangle$ are elements of Θ^2. Namely, we prove the tracial analog of Hilbert's result on bivariate quartics, see also Theorem 1.11; and we present our results concerning the BMV polynomials and representations as a sum of hermitian squares and commutators.

3.1 Introduction

This introduction summarizes easy results and examples that were known before.

3.1.1 Reduction to sums of squares

For any polynomial $f \in \Theta^2$ the commutative collapse $\check{f} \in \mathbb{R}[\underline{x}]$ is a sum of squares in $\mathbb{R}[\underline{x}]$ since the commutative collapse of a hermitian square in $\mathbb{R}\langle \underline{X} \rangle$ is a square in $\mathbb{R}[\underline{x}]$. In some easy cases the converse also holds true. This will be the starting point of the following investigation.

Univariate polynomials

By trivial reasons, any polynomial in one variable is a sum of hermitian squares if and only if it is a sum of squares of polynomials. In fact, for $g_i \in \mathbb{R}\langle X \rangle$ the sum of hermitian squares $\sum_i g_i{}^* g_i$ is equal to the sum of squares $\sum_i g_i^2$. In particular, using that univariate positive polynomials in $\mathbb{R}[x]$ are sums of squares, see e.g. [Mar, Prop. 1.2.1], the following proposition holds.

3.1 Proposition. *Let $f \in \mathbb{R}\langle X \rangle$. Then $f \in \Theta^2$ if and only if $f \geq 0$ on $\mathbb{R}$.*

Quadratic polynomials

Let $f \in \mathbb{R}\langle \underline{X} \rangle$ be a quadratic polynomial. Each word of f is of the form $X_i X_j$, with $X_0 := 1$. Since $X_i X_j \overset{\mathrm{cyc}}{\sim} X_j X_i$ for any i, j, the polynomial f modulo $\overset{\mathrm{cyc}}{\sim}$ behaves like a polynomial in commuting variables. That is, two quadratic polynomials $f, g \in \mathbb{R}\langle \underline{X} \rangle_2$ are cyclically equivalent if and only if $\check{f} = \check{g} \in \mathbb{R}[\underline{x}]_2$. Thus $f \in \Theta^2$ if and only if $\check{f} \in \sum \mathbb{R}[\underline{x}]^2$. Since any positive quadratic polynomial is a sum of squares, we obtain the following proposition.

3.2 Proposition. *Let $f \in \mathbb{R}\langle \underline{X} \rangle_2$. Then $f \in \Theta^2$ if and only if $f \geq 0$ on $\mathbb{R}^n$.*

Cyclically sorted polynomials

A polynomial $f \in \mathbb{R}\langle X, Y\rangle$ is cyclically sorted if $f \overset{\text{cyc}}{\sim} \sum_{i,j} a_{ij} X^i Y^j$ for some $a_{ij} \in \mathbb{R}$, $i, j \in \mathbb{N}_0$, see also Definition 1.10. If two cyclically sorted polynomials $f, g \in \mathbb{R}\langle X, Y\rangle$ have the same commutative collapse, i.e. $\check{f} = \check{g}$, then $f \overset{\text{cyc}}{\sim} g$ holds. Cyclically sorted polynomials behave in positivity questions like bivariate polynomials in commuting variables in the sense of the following two propositions.

3.3 Proposition. *Let $f \in \mathbb{R}\langle X, Y\rangle$ be cyclically sorted. Then f is trace-positive if and only if $\check{f} \in \mathbb{R}[x, y]$ is positive.*

Proof. If f is trace-positive then $\check{f}$ is obviously positive. The other implication follows with the spectral theorem. Let $A, B \in \mathcal{S}\mathbb{R}^{s\times s}$ for some $s \in \mathbb{N}$ be fixed and let

$$A = \sum_\ell \lambda_\ell p_\ell, \; B = \sum_k \mu_k q_k$$

be its spectral decompositions with $\lambda_\ell, \mu_k \in \mathbb{R}$ and pairwise orthogonal projections p_ℓ for A and q_k for B respectively. Let $f = \sum_t c_t X^{i_t} Y^{j_t}$, then

$$\operatorname{Tr}(f(A, B)) = \sum_t c_t \operatorname{Tr}(A^{i_t} B^{j_t}) = \sum_t c_t \operatorname{Tr}\Big(\big(\sum_\ell p_\ell A\big)^{i_t}\big(\sum_k B q_k\big)^{j_t}\Big).$$

Since the projections p_ℓ, respectively q_k, are pairwise orthogonal we have $(\sum_\ell p_\ell A)^{i_t} = \sum_\ell p_\ell \lambda_\ell^{i_t}$ and $(\sum_k B q_k)^{j_t} = \sum_k \mu_k^{j_t} q_k$ respectively. Therefore

$$\begin{aligned}
\operatorname{Tr}(f(A, B)) &= \sum_t \sum_{\ell,k} c_t \operatorname{Tr}(p_\ell \lambda_\ell^{i_t} \mu_k^{j_t} q_k) \\
&= \sum_{\ell,k} \sum_t c_t \lambda_\ell^{i_t} \mu_k^{j_t} \operatorname{Tr}(p_\ell q_k) \\
&= \sum_{\ell,k} \check{f}(\lambda_\ell, \mu_k) \operatorname{Tr}(p_\ell q_k) \geq 0
\end{aligned}$$

since $\check{f}$ is positive on $\mathbb{R}^2$ and $\operatorname{Tr}(p_\ell q_k) \geq 0$ for all ℓ, k. ∎

This proof fails in general for polynomials which are not cyclically sorted, since we then obtain mixed products of the projections p_ℓ and q_k which might not have positive trace.

3.4 Proposition. *Let $f \in \mathbb{R}\langle X, Y\rangle$ be cyclically sorted. Then $f \in \Theta^2$ if and only if $\check{f} \in \mathbb{R}[x, y]$ is a sum of squares.*

The proof follows the same line of reasoning as the proof of [KS1, Prop. 4.2].

Proof. If $f \in \Theta^2$ holds, then $\check{f}$ is obviously a sum of squares. For the converse direction let a representation $\check{f} = \sum_i g_i^2$ with $g_i \in \mathbb{R}[x, y]$ be given. For any g_i there exists a linear combination $h_i \in \mathbb{R}\langle X, Y\rangle$ of words of the form $X^i Y^j$ such that $\check{h}_i = g_i$ and $h_i^* h_i$ is cyclically sorted. Hence $h := \sum_i h_i{}^* h_i$ is cyclically sorted and $\check{f} = \check{h}$ which implies $f \overset{\text{cyc}}{\sim} h = \sum_i h_i{}^* h_i$. ∎

In particular, the above propositions imply that a trace-positive cyclically sorted bivariate quartic lies in Θ^2 since the same statement is true in the commutative context, see Theorem 1.11. We will return to this topic in Section 3.3.

3.1.2 Counter examples

On the other hand, in contrast to Theorem 1.15 on matrix-positive polynomials, not every trace-positive polynomial can be written as a sum of hermitian squares and commutators. One example are the following versions of the Motzkin polynomial.

3.5 Example (Motzkin polynomial)**.** We present several non-commutative versions of the well-known Motzkin polynomial. They all do not lie in Θ^2 as their commutative collapse is not a sum of squares [Mar, p. 5].

(a) Let

$$M_{\mathrm{nc}} := X^2Y^4 + X^4Y^2 - 3X^2Y^2 + 1 \in \mathbb{R}\langle X, Y\rangle.$$

The fact that $M_{\mathrm{nc}}(A, B)$ has positive trace for all symmetric matrices A, B can be shown in several ways. It follows by replacing the non-commutative hypercube K_{hc} generated by $1 - X^2, 1 - Y^2$ in Proposition 2.11 by the hypercube generated by $C - X^2, C - Y^2$ for any scalar factor $C > 0$, and using the fact that $M_{\mathrm{nc}} + \varepsilon \in \mathrm{trQM}(C - X^2, C - Y^2)$ for all $\varepsilon \in \mathbb{R}_{>0}$, see [KS1, Example 3.4].

It also follows from Proposition 3.3 using that the Motzkin polynomial in $\mathbb{R}[x, y]$ is positive. Alternatively, the trace-positivity of M_{nc} is a consequence of the fact that $M_{\mathrm{nc}}(X^3, Y^3)$ lies in Θ^2. Indeed[1],

$$M_{\mathrm{nc}}(X^3, Y^3) \overset{\mathrm{cyc}}{\sim} g_1{}^*g_1 + g_2{}^*g_2 + g_3{}^*g_3 + \frac{3}{4}g_4{}^*g_4 + \frac{3}{4}g_5{}^*g_5 + \frac{3}{4}g_6{}^*g_6$$

with

$$\begin{aligned} g_1 &= X^2Y - \frac{1}{2}X^4Y^5 - \frac{1}{2}X^6Y^3, & g_2 &= XY^2 - \frac{1}{2}X^3Y^6 - \frac{1}{2}X^5Y^4, \\ g_3 &= 1 - \frac{1}{2}X^2Y^4 - \frac{1}{2}X^4Y^2, & g_4 &= X^2Y^4 - X^4Y^2, \\ g_5 &= X^3Y^6 - X^5Y^4 \text{ and} & g_6 &= X^4Y^5 - X^6Y^3. \end{aligned}$$

(b) Every polynomial which is cyclically equivalent to M_{nc} is also trace-positive but not a sum of hermitian squares and commutators. For instance, the symmetric polynomial

$$M_{\mathrm{sym}} := XY^4X + YX^4Y - 3XY^2X + 1 \in \mathcal{S}\mathbb{R}\langle X, Y\rangle$$

is trace-positive but does not lie in Θ^2.

(c) The following version of the Motzkin polynomial

$$M := X^2Y^4 + Y^2X^4 - 3XYXY + 1 \in \mathbb{R}\langle X, Y\rangle,$$

which is not cyclically sorted, is also trace-positive. Since

$$M - M_{\mathrm{nc}} = 3(X^2Y^2 - XYXY) \in \Theta^2$$

we have for all $A, B \in \mathcal{S}^2$ that $\mathrm{Tr}(M(A, B)) \geq \mathrm{Tr}(M_{\mathrm{nc}}(A, B)) \geq 0$.

3.6 Remark. Any example of a positive polynomial in the polynomial ring $\mathbb{R}[x, y]$ that can not be written as sum of squares in $\mathbb{R}[x, y]$ gives rise to a (cyclically sorted) trace-positive polynomial in $\mathbb{R}\langle X, Y\rangle$ which does not lie in Θ^2. The argumentation follows the same line of reasoning as for the Motzkin polynomial M_{nc}.

[1] calculated by Kristijan Cafuta

Another class of examples are given by several BMV polynomials.

3.7 Example (BMV polynomials)**.**

(a) The BMV polynomial $S_{6,3}(X^2, Y^2)$ is trace-positive as a consequence of Theorem 2.3 and the fact that $S_{7,4}(X^2, Y^2) \in \Theta^2$. However $S_{6,3}(X^2, Y^2)$ is not a sum of hermitian squares and commutators [KS2, Example 3.5]. There are several other BMV polynomials of this kind as mentioned in Section 2.1.2.

(b) The BMV polynomials $S_{m,3}(X^2, Y^2)$ with $m \geq 12$ are trace-positive, which will be shown in Section 3.4. However they cannot be written as sum of hermitian squares and commutators [LiS, Theorem 8].

3.2 The tracial Gram matrix method

Testing whether a given polynomial $f \in \mathbb{R}\langle \underline{X} \rangle$ is an element of Θ^2 can be done using semidefinite programming as first observed in [KS2, Section 3]. This is based on the *tracial Gram matrix method*, which is the topic of this section. In this work we only need the main idea of this method. A more detailed description can be found in [BCKP, Chapter 4].

First, we need some preliminary notation. The *minimal degree* $\operatorname{mindeg}(f)$ of f is the length of the shortest word appearing in f. We denote by $\mathbf{v}$ a column vector consisting of words $w \in \langle \underline{X} \rangle$. The operation * maps $\mathbf{v}$ to the row vector $\mathbf{v}^*$ where each word v_i in $\mathbf{v}$ is mapped to $v_i{}^*$.

The core of the tracial Gram matrix method is given by the following proposition, which can also be found in [BCKP, Prop. 2.8].

3.8 Proposition. *Suppose $f \in \mathbb{R}\langle \underline{X} \rangle$. Then $f \in \Theta^2$ if and only if there exists a positive semidefinite matrix G such that*

$$f \overset{\text{cyc}}{\sim} \mathbf{v}^* G \mathbf{v}, \tag{3.1}$$

where $\mathbf{v}$ is a vector consisting of all words $w \in \langle \underline{X} \rangle$ satisfying

$$\operatorname{mindeg}(f) \leq 2 \deg(w) \leq \deg(f). \tag{3.2}$$

Moreover, given such a positive semidefinite matrix G of rank r, one can construct polynomials $g_1, \ldots, g_r \in \mathbb{R}\langle \underline{X} \rangle$ such that

$$f \overset{\text{cyc}}{\sim} \sum_{i=1}^{r} g_i{}^* g_i. \tag{3.3}$$

The matrix G is called a tracial Gram matrix *for f with respect to* $\mathbf{v}$.

The tracial Gram matrix method is an extension of the Gram matrix method for sums of hermitian squares [Hel, Section 2.2], which is in turn a variant of the classical result for polynomials in commuting variables due to Choi, Lam and Reznick [CLR, Section 2]; see also [Par]. In the commutative case, the Gram matrix method can be refined by the Newton polytope method. The Newton polytope $N(f)$ is the convex hull of all exponents $\alpha \in \mathbb{N}^n$ arising in $f \in \mathbb{R}[\underline{x}]$. In general, it suffices to construct $\mathbf{v}$ by taking only the words $w = \underline{x}^\beta$ with $2\beta \in N(f)$. In most cases this reduces significantly the size of the vector $\mathbf{v}$. A modification for the non-commutative case concerning representations in Σ^2 is possible [KP, Chapter 3] as well as a tracial analog of the Newton polytope, which is explained in [BCKP, Section 4.1].

Proof. The proof is straightforward like in the commutative case. Let $g_i \in \mathbb{R}\langle \underline{X} \rangle$ be such that $f \overset{\mathrm{cyc}}{\sim} \sum_i g_i{}^* g_i$. Define $t := \max_i\{\deg g_i\}$. Then $2t \le \deg(f)$ since the highest degree terms do not cancel. Indeed, suppose $2t > \deg f$. Let h_i be the homogeneous part of degree t of the g_i and $r_i = g_i - h_i$. Then $\deg r_i < t$ and

$$\sum_i g_i{}^* g_i = \sum_i h_i{}^* h_i + \sum_i (r_i{}^* r_i + h_i{}^* r_i + r_i{}^* h_i).$$

Since each monomial w in $h_i{}^* r_i$, $r_i{}^* h_i$ and $r_i{}^* r_i$ has $\deg(w) < 2t$ none of these can be cyclically equivalent to a monomial in $h_i{}^* h_i$, where each monomial is of degree $2t$. Thus $\sum_i h_i{}^* h_i \overset{\mathrm{cyc}}{\sim} 0$ which implies by [KS2, Lemma 3.2] that $h_i = 0$ for all i contradicting $\deg h_i = t$. The second statement $\operatorname{mindeg} f \le 2 \deg g_i$ follows the same way.

Set

$$G := \sum_i \vec{g_i} \vec{g_i}^T,$$

where $\vec{g_i}$ is defined via $g_i = \vec{g_i}^T \mathbf{v}$. Then $\sum_i g_i{}^* g_i = \mathbf{v}^* G \mathbf{v}$ and G is obviously positive semidefinite. Conversely, given a positive semidefinite $G \in \mathbb{R}^{N \times N}$ of rank r satisfying (3.1), write $G = \sum_{i=1}^r \vec{g_i} \vec{g_i}^T$ for $\vec{g_i} \in \mathbb{R}^N$ and define $g_i := \vec{g_i}^T \mathbf{v}$ to obtain (3.3). ■

If $f = \mathbf{v}^* G \mathbf{v}$ for some vector $\mathbf{v}$ and some symmetric matrix G, then G is called an *exact* Gram matrix for f. If the exact Gram matrix is positive semidefinite, then f is a sum of hermitian squares, see [Hel, Section 2.2] or [KP, Theorem 3.1]. The following example will illustrate this method in more detail.

3.9 Example. Let

$$f = X^4 + 4Y^4 - YX^2Y + XYXY + YXYX \in \mathcal{S}\mathbb{R}\langle X, Y \rangle_4.$$

Then $f \in \Theta^2 \setminus \Sigma^2$ and hence f is trace-positive but not matrix-positive as we will show by the (tracial) Gram matrix method.

Since $f = \sum_w f_w w$ is homogeneous of degree 4 it suffices to consider words of degree 2 in X and Y by Condition (3.2). Thus let $\mathbf{v} = \begin{bmatrix} X^2 & XY & YX & Y^2 \end{bmatrix}^T$. The exact Gram matrix G_g of a homogeneous polynomial $g = \sum_w g_w w \in \mathbb{R}\langle X, Y \rangle_4$ is unique and of the form

$$G_g = \begin{bmatrix} g_{\mathsf{X}^4} & g_{\mathsf{X}^3\mathsf{Y}} & g_{\mathsf{X}^2\mathsf{YX}} & g_{\mathsf{X}^2\mathsf{Y}^2} \\ g_{\mathsf{YX}^3} & g_{\mathsf{YX}^2\mathsf{Y}} & g_{\mathsf{YXYX}} & g_{\mathsf{YXY}^2} \\ g_{\mathsf{XYX}^2} & g_{\mathsf{XYXY}} & g_{\mathsf{XY}^2\mathsf{X}} & g_{\mathsf{XY}^3} \\ g_{\mathsf{Y}^2\mathsf{X}^2} & g_{\mathsf{Y}^2\mathsf{XY}} & g_{\mathsf{Y}^3\mathsf{X}} & g_{\mathsf{Y}^4} \end{bmatrix}.$$

Hence the exact Gram matrix for f is given by

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

which is not positive semidefinite and thus $f \notin \Sigma^2$.

If we calculate $\mathbf{v}^* G \mathbf{v}$ for an arbitrary symmetric 4×4 matrix

$$G = \begin{bmatrix} G_{11} & G_{12} & G_{13} & G_{14} \\ G_{12} & G_{22} & G_{23} & G_{24} \\ G_{13} & G_{23} & G_{33} & G_{34} \\ G_{14} & G_{24} & G_{34} & G_{44} \end{bmatrix}$$

we obtain

$$\begin{aligned}\mathbf{v}^*G\mathbf{v} \stackrel{\text{cyc}}{\sim}\ & G_{11}X^4 + 2(G_{12}+G_{13})X^3Y + (G_{22}+G_{33}+2G_{14})X^2Y^2 \\ & + 2G_{23}XYXY + 2(G_{24}+G_{34})XY^3 + G_{44}Y^4.\end{aligned}$$

Thus the tracial Gram matrix $G_{[g]}$ of g with canonical representative $[g] = \sum_{[w]} g_{[w]}[w]$, where $g_{[w]} = \sum_{v \stackrel{\text{cyc}}{\sim} w} g_v$, is of the form

$$G_g = \begin{bmatrix} g_{\mathsf{X}^4} & r_1 & \frac{1}{2}g_{[\mathsf{X}^3\mathsf{Y}]} - r_1 & t_1 \\ r_1 & t_2 & \frac{1}{2}g_{[\mathsf{XYXY}]} & r_2 \\ \frac{1}{2}g_{[\mathsf{X}^3\mathsf{Y}]} - r_1 & \frac{1}{2}g_{[\mathsf{XYXY}]} & g_{[\mathsf{X}^2\mathsf{Y}^2]} - 2t_1 - t_2 & \frac{1}{2}g_{[\mathsf{XY}^3]} - r_2 \\ t_1 & r_2 & \frac{1}{2}g_{[\mathsf{XY}^3]} - r_2 & g_{\mathsf{Y}^4} \end{bmatrix}$$

for some $r_1, r_2, t_1, t_2 \in \mathbb{R}$. Taking $r_1 = r_2 = 0$ and $t_1 = -2, t_2 = 1$ we obtain the positive semidefinite matrix

$$G = \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ -2 & 0 & 0 & 4 \end{bmatrix}$$

which satisfies $f \stackrel{\text{cyc}}{\sim} \mathbf{v}^*G\mathbf{v}$. Hence G is a positive semidefinite tracial Gram matrix of f and $f \in \Theta^2$. Indeed, taking

$$\vec{g_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix} \qquad \vec{g_2} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \vec{g_3} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

leads to $G = \sum_{i=1}^3 \vec{g_i}\vec{g_i}^T$. That is,

$$\begin{aligned} f &\stackrel{\text{cyc}}{\sim} \sum_{i=1}^{3} {g_i}^* g_i \\ &= X^4 + 4Y^4 - 2X^2Y^2 - 2Y^2X^2 + 2XY^2X + YX^2Y + XYXY + YXYX \end{aligned}$$

with

$$\begin{aligned} g_1 &= X^2 - 2Y^2 \\ g_2 &= XY + YX \\ g_3 &= YX. \end{aligned}$$

A polynomial f does not in general have a unique tracial Gram matrix. For instance, in Example 3.9 we can also take

$$G' = \begin{bmatrix} 1 & 0 & 0 & -\frac{3}{2} \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ -\frac{3}{2} & 0 & 0 & 4 \end{bmatrix}.$$

Determining whether $f \in \Theta^2$ amounts to finding *a* positive semidefinite Gram matrix from the affine set of all tracial Gram matrices for f. Problems like this can be (in theory) solved *exactly* using real quantifier elimination. However, this only works for problems of small size, so a *numerical* approach is needed in practice, which can be done by semidefinite programming. This will be handled in Chapter 6.

3.3 Bivariate quartics

As we have seen in Section 3.1, univariate polynomials and quadratic polynomials are trace-positive if and only if they can be written as a sum of hermitian squares and commutators. On the other hand, trace-positive polynomials of degree at least 6, like the Motzkin polynomial, need not be in Θ^2. The degree gap for bivariate polynomials is bridged by the following theorem, see also [BK2, Theorem 3.1].

3.10 Theorem. *For $f \in \mathbb{R}\langle X, Y\rangle_4$, the following are equivalent:*

(i) *f is trace-positive;*

(ii) *f is trace-positive on $(\mathcal{S}\mathbb{R}^{2\times 2})^2$;*

(iii) *f is cyclically equivalent to a sum of four hermitian squares;*

(iv) *$f \in \Theta^2$.*

This theorem is the tracial analog of Theorem 1.11. In the commutative case we can write a positive bivariate quartic polynomial as a sum of three squares. This bound has to be increased to four in the tracial case, as we will see in Remark 3.14. In contrast to the commutative case, we cannot use the homogenization process to obtain the same result for ternary quartic forms. It is not known whether Theorem 3.10 holds true for homogeneous polynomials $f \in \mathbb{R}\langle X, Y, Z\rangle_4$. Theorem 3.10 implies that any trace-inequality of degree four which holds for all pairs of symmetric 2×2-matrices holds also for any pair of symmetric $s \times s$-matrices for arbitrary $s \in \mathbb{N}$.

The idea of the proof is as follows. The implications (i) $\Longrightarrow$ (ii), (iii) $\Longrightarrow$ (iv) and (iv) $\Longrightarrow$ (i) are obvious. To show (ii) $\Longrightarrow$ (iii), we apply a linear transformation of the variables X, Y to reduce the proof to the case that the coefficient belonging to X^2Y^2 of the canonical representative $[f]$ is at least as big as the one belonging to $XYXY$. Then the membership certificate $f \in \Theta^2$ can be explicitly constructed from a sum of squares certificate for the commutative collapse $\check{f}$ of f.

First, we present some auxiliary lemmas. Let $f = \sum_w a_w w \in \mathbb{R}\langle X, Y\rangle_4$. We recall that the coefficient of a word w in the canonical representative $[f]$ of f is given by $a_{[w]} = \sum_{v \overset{\mathrm{cyc}}{\sim} w} a_v$, confer Definition 1.10. If $a_w = a_{[w]}$ then we omit the brackets for simplicity in notation.

3.11 Lemma. *Let $f \in \mathbb{R}\langle X, Y\rangle_4$. If $\check{f} \geq 0$ on $\mathbb{R}^2$ and $a_{[\mathsf{X^2Y^2}]} \geq a_{[\mathsf{XYXY}]}$, then $f \in \Theta^2$. In fact, f is cyclically equivalent to a sum of four hermitian squares.*

Proof. Since $\check{f} \geq 0$ on $\mathbb{R}^2$, by Theorem 1.11 there exist $\check{g}_1, \check{g}_2, \check{g}_3 \in \mathbb{R}[x, y]$ such that $\check{f} = \sum_{i=1}^3 \check{g}_i^2$. Each $\check{g}_i$ can be chosen of degree at most 2, hence it can be lifted to

$$g_i := c_{\mathsf{1}}^{(i)} + c_{\mathsf{X}}^{(i)} X + c_{\mathsf{Y}}^{(i)} Y + c_{\mathsf{XY}}^{(i)} \frac{XY + XY}{2} + c_{\mathsf{X^2}}^{(i)} X^2 + c_{\mathsf{Y^2}}^{(i)} Y^2 \in \mathbb{R}\langle X, Y\rangle.$$

One easily verifies that

$$\sum_{i=1}^3 g_i{}^* g_i \overset{\mathrm{cyc}}{\sim} f - \left(\left(a_{[\mathsf{X^2Y^2}]} - \frac{a_{[\mathsf{X^2Y^2}]} + a_{[\mathsf{XYXY}]}}{2}\right) X^2Y^2 + \left(a_{[\mathsf{XYXY}]} - \frac{a_{[\mathsf{X^2Y^2}]} + a_{[\mathsf{XYXY}]}}{2}\right) XYXY \right).$$

With $g_4 := \frac{1}{\sqrt{2}} \sqrt{a_{[\mathsf{X^2Y^2}]} - a_{[\mathsf{XYXY}]}}(XY - YX) \in \mathbb{R}\langle X, Y\rangle$, we obtain $f \overset{\mathrm{cyc}}{\sim} \sum_{i=1}^4 g_i{}^* g_i \in \Theta^2$. ■

As a consequence of Lemma 3.11, we derive a criterion for biquadratic polynomials to be members of Θ^2. A polynomial $f \in \mathbb{R}\langle X, Y\rangle$ is called *biquadratic* if $\deg_X f \leq 2$ and $\deg_Y f \leq 2$.

3.12 Lemma. *Let $f \in \mathbb{R}\langle X, Y\rangle$ be biquadratic. Then $f \in \Theta^2$ if and only if f is trace-positive on $(\mathcal{S}\mathbb{R}^{2\times 2})^2$.*

Proof. Obviously, $f \in \Theta^2$ implies that f is trace-positive. Conversely, if f is trace-positive on $(\mathcal{S}\mathbb{R}^{2\times 2})^2$, then by considering

$$A_x := x \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } B_y := y \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

we obtain that the (commutative) polynomial

$$p(x, y) := \operatorname{Tr}(f(A_x, B_y)) = a_1 + a_{\mathsf{X}^2} x^2 + a_{\mathsf{Y}^2} y^2 + (a_{[\mathsf{X}^2\mathsf{Y}^2]} - a_{[\mathsf{XYXY}]}) x^2 y^2$$

is positive on $\mathbb{R}^2$. This implies $a_{[\mathsf{X}^2\mathsf{Y}^2]} - a_{[\mathsf{XYXY}]} \geq 0$. Since by assumption $\check{f} \geq 0$ on $\mathbb{R}^2$, Lemma 3.11 yields $f \in \Theta^2$. ∎

3.13 Lemma. *If $f \in \mathbb{R}\langle X, Y\rangle_4$ is trace-positive on $(\mathcal{S}\mathbb{R}^{2\times 2})^2$, then there exists a transformation matrix $T \in \mathrm{GL}_2(\mathbb{R})$, such that the coefficients $b_{[w]}$ of the canonical representative of $f(T[X \quad Y]^T)$ satisfy $b_{[\mathsf{X}^2\mathsf{Y}^2]} \geq b_{[\mathsf{XYXY}]}$.*

Proof. If $a_{[\mathsf{X}^2\mathsf{Y}^2]} \geq a_{[\mathsf{XYXY}]}$ we are done by taking $T = \mathbf{1}_2$. Now let $a_{[\mathsf{X}^2\mathsf{Y}^2]} < a_{[\mathsf{XYXY}]}$. Since $\check{f}$ is trace-positive on $(\mathcal{S}\mathbb{R}^{2\times 2})^2$, we obtain that at least one of the coefficients of X^4 and Y^4 is non zero. Indeed, assume that $a_{\mathsf{X}^4} = a_{\mathsf{Y}^4} = 0$. Then by positivity of $\check{f}$ we get that the words x^3y, xy^3, x^3 and y^3 cannot occur in $\check{f}$. Hence the same holds true for X^3Y, XY^3, X^3 and Y^3 and thus f is biquadratic. As in the proof of Lemma 3.12 then follows that $a_{[\mathsf{XYXY}]} \geq a_{[\mathsf{X}^2\mathsf{Y}^2]}$, a contradiction.

Without loss of generality, let f contain X^4. Then $a_{\mathsf{X}^4} > 0$ since $\check{f}$ is positive on $\mathbb{R}^2$. We set

$$T := \begin{bmatrix} 1 & c \\ 0 & -1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$$

for some

$$c \geq \frac{1}{2a_{\mathsf{X}^4}} \left(a_{[\mathsf{X}^2\mathsf{Y}]} + \sqrt{a_{[\mathsf{X}^2\mathsf{Y}]}^2 + 4a_{\mathsf{X}^4}(a_{[\mathsf{XYXY}]} - a_{[\mathsf{X}^2\mathsf{Y}^2]})} \right).$$

Then

$$b_{[\mathsf{X}^2\mathsf{Y}^2]} = 4c^2 a_{\mathsf{X}^4} - 2c\, a_{[\mathsf{X}^3\mathsf{Y}]} + a_{[\mathsf{X}^2\mathsf{Y}^2]} \geq 2c^2 a_{\mathsf{X}^4} - c\, a_{[\mathsf{X}^3\mathsf{Y}]} + a_{[\mathsf{XYXY}]} = b_{[\mathsf{XYXY}]},$$

as desired. ∎

Proof. (of Theorem 3.10) Since the implications (i) $\Longrightarrow$ (ii), (iii) $\Longrightarrow$ (iv) and (iv) $\Longrightarrow$ (i) are obvious, we are left to show (ii) $\Longrightarrow$ (iii). Suppose f is trace-positive on $(\mathcal{S}\mathbb{R}^{2\times 2})^2$. If we have $a_{[\mathsf{XYXY}]} > a_{[\mathsf{X}^2\mathsf{Y}^2]}$, then we apply Lemma 3.13 and obtain a trace-positive polynomial $g \in \mathbb{R}\langle X, Y\rangle$ that satisfies the assumptions of Lemma 3.11. Hence (iii) holds for g and thus also for f. If $a_{[\mathsf{XYXY}]} \leq a_{[\mathsf{X}^2\mathsf{Y}^2]}$, then (iii) holds by Lemma 3.11. ∎

3.14 Remark. In Theorem 3.10 we have shown that a trace-positive bivariate quartic polynomial is cyclically equivalent to a sum of four hermitian squares. This bound on the number of hermitian squares is sharp. For example, the polynomial

$$f = 1 + \frac{1}{2}X^2 + X^4 + Y^4 + 2XYXY$$

is cyclically equivalent to a sum of four but not three hermitian squares. By Proposition 3.8, f is cyclically equivalent to a sum of three hermitian squares if and only if there exists a positive semidefinite tracial Gram matrix G of rank 3. Consider $\mathbf{v} = \begin{bmatrix} 1 & X & Y & X^2 & XY & YX & Y^2 \end{bmatrix}^T$. Then any tracial Gram matrix G of f with respect to $\mathbf{v}$ is of the form

$$G = \begin{bmatrix} 1 & 0 & 0 & r_1 & -z_1 & -z_2 & r_2 \\ 0 & \frac{1}{2} - 2r_1 & z_1 + z_2 & 0 & -u_1 & -u_2 & v_1 + v_2 \\ 0 & z_1 + z_2 & -2r_2 & u_1 + u_2 & -v_1 & -v_2 & 0 \\ r_1 & 0 & u_1 + u_2 & 1 & s_1 & -s_1 & -t_1 \\ -z_1 & -u_1 & -v_1 & s_1 & t_2 & 1 & s_2 \\ -z_2 & -u_2 & -v_2 & -s_1 & 1 & 2t_1 - t_2 & -s_2 \\ r_2 & v_1 + v_2 & 0 & -t_1 & s_2 & -s_2 & 1 \end{bmatrix}$$

where $r_1, r_2, s_1, s_2, t_1, t_2, u_1, u_2, v_1, v_2, z_1, z_2 \in \mathbb{R}$. Positive semidefiniteness of G implies that the 4×4-submatrix

$$G_{4567} = \begin{bmatrix} 1 & s_1 & -s_1 & -t_1 \\ s_1 & t_2 & 1 & s_2 \\ -s_1 & 1 & 2t_1 - t_2 & -s_2 \\ -t_1 & s_2 & -s_2 & 1 \end{bmatrix}$$

is positive semidefinite. This is possible only if $s_1 = s_2 = 0$ and $t_1 = t_2 = 1$. Assuming that $\operatorname{rank} G = 3$ implies that all 4×4-minors of G are equal to 0. Taking for example the minor $\det G_{1567} = (z_1 - z_2)^2$ implies $z_1 = z_2$. Similarly, using the determinants of G_{2567}, G_{3567}, G_{3467}, G_{2457} and G_{1457} leads to the following tracial Gram matrix

$$G = \begin{bmatrix} 1 & 0 & 0 & r_1 & -z_1 & -z_1 & -r_1 \\ 0 & \frac{1}{2} - 2r_1 & 2z_1 & 0 & 0 & 0 & 0 \\ 0 & 2z_1 & 2r_1 & 0 & 0 & 0 & 0 \\ r_1 & 0 & 0 & 1 & 0 & 0 & -1 \\ -z_1 & 0 & 0 & 0 & 1 & 1 & 0 \\ -z_1 & 0 & 0 & 0 & 1 & 1 & 0 \\ -r_1 & 0 & 0 & -1 & 0 & 0 & 1 \end{bmatrix}.$$

Finally, $\det G_{1245} = \det G_{1345} = 0$ implies $z_1 = \sqrt{1 - r_1^2}$ which contradicts the positive semidefiniteness of G. Thus there is no positive semidefinite tracial Gram matrix of f of rank 3. Taking the positive semidefinite tracial Gram matrix

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{bmatrix}$$

of rank 4 proves that f is a sum of four hermitian squares.

3.4 BMV polynomials

The BMV conjecture introduced in Section 2.1 states that all BMV polynomials $S_{m,k}(X^2, Y^2)$ are trace-positive. In this section we will prove that the BMV polynomials $S_{m,4}(X^2, Y^2)$ are elements of Θ^2 for all $m \geq 4$. Hence $S_{m,k}(X^2, Y^2)$ is trace-positive for all $k \leq 4$ (or $k \geq m - 4$ respectively) and arbitrary m. Furthermore, we prove that $S_{4r+2,4}(X, Y) \in \Theta^2$ holds for all $r \in \mathbb{N}$.

3.15 Theorem. *For all $m \in \mathbb{N}$, $m \geq 4$, we have $S_{m,4}(X^2, Y^2) \in \Theta^2$.*

This theorem has already been proved in [Bur, Theorem 6]. We recall this proof with some additional examples in Section 3.4.1. A proof for the case of odd $m \in \mathbb{N}$ has independently been found by Landweber and Speer [LiS, Theorem 2].

The same result holds true by symmetry for $k = m - 4$. Hillar's descent Theorem 2.3 implies immediately the following corollary.

3.16 Corollary. *For all $m \in \mathbb{N}$, $k \leq 4$ the polynomials $S_{m,k}(X^2, Y^2)$ are trace-positive.*

While investigating Conjecture 2.2, Collins, Dykema and Torres-Ayala found for $m = 2r$, $k = 2$ and for $m = 4r + 2, k = 4$, where $r \in \mathbb{N}$ is arbitrary, that $S_{m,k}(X, Y) \in \Theta^2$ holds [CDT, Theorem 3.7]. Hence in these cases we have $\operatorname{Tr}(S_{m,k}(A, B)) \geq 0$ not only for positive semidefinite matrices, but also for all symmetric matrices $A, B \in \mathcal{S}^2$. The second statement has also been mentioned earlier by the author in [Bur] as a remark. Since this result seemed to be outside of the topic of the BMV conjecture a proof was omitted there but it will be presented in Section 3.4.2.

3.17 Theorem. *For all $r \in \mathbb{N}$, we have $S_{4r+2,4}(X, Y) \in \Theta^2$.*

By symmetry, the same statement holds true for $m = 4r + 2$ and $k = 4r - 2$ with $r \in \mathbb{N}$ arbitrary. Furthermore, the BMV polynomials $S_{4r+2,4}(X^2, Y^2)$ are cyclically equivalent to a sum of hermitian squares of polynomials in $\mathbb{R}\langle X^2, Y^2\rangle$.

The idea of the proofs of Theorem 3.15 and 3.17 is the following. We will present for the proof of Theorem 3.15 concrete constructions of a sum of hermitian squares depending on the parameter m. These constructions have been found with the aid of numerical experiments using NCAlgebra [HdOMS] in Mathematica as well as Yalmip [Lof] and SeDuMi [Stu] in Matlab. For the proof of Theorem 3.17 we present a positive semidefinite tracial Gram matrix depending on the parameter r, from which one can deduce a sum of hermitian squares decomposition. In a second step we show that the constructed sums of hermitian squares are in fact cyclically equivalent to $S_{m,4}(X^2, Y^2)$ (respectively $S_{4r+2,4}(X, Y)$).

3.18 Definition. The *order* $(\operatorname{ord} w)$ of a word $w = w_{(1)} \cdots w_{(m)}$ of length m, where $w_{(i)}$ denote the i-th letter of w, is the smallest positive integer k, such that $w_{(i+k)} = w_{(i)}$ for all $i = 1, \ldots, m$ where we identify $w_{(i+k)}$ with $w_{(i+k-m)}$ if $i + k > m$. Thus cyclically equivalent words have the same order. It can also be defined as the smallest integer $k \geq 1$ such that there exists a subword $v = v_{(1)} \cdots v_{(k)}$ of length k with $w = v \cdots v = v^{m/k}$.

3.4.1 Proof of $S_{m,4}(X^2, Y^2) \in \Theta^2$

In this section we present the proof of Theorem 3.15, which can also be found in [Bur]. For a better understanding we include some concrete sum of hermitian squares representations as examples.

The order of a word $w = v^{m/\operatorname{ord}(w)}$ in $S_{m,4}(X^2, Y^2)$ divides m. Further, since Y^2 appears the same number of times (in fact, either one, two or four times) in every subword v, we get that $\frac{m}{\operatorname{ord}(w)}$ divides 4. In particular,

$$\operatorname{ord}(w) \in \{m, \frac{m}{2}, \frac{m}{4}\} \cap \mathbb{N}.$$

For technical reasons, we distinguish between even and odd m.

Case m odd

Let m be fixed. Set $V := \{v \in \langle X, Y\rangle \mid \deg_X v = m-4, \deg_Y v = 4\}$ and define the subsets

$$V_0 = \{v \in \{X^2, Y^2\}^{\frac{m-1}{2}} X \mid v = X^k Y^2 X^\ell Y^2 X^{k'+1}, k \leq k'\} \cap V,$$
$$V_1 = \{v \in X\{X^2, Y^2\}^{\frac{m-1}{2}} \mid v = X^{k+1} Y^2 X^\ell Y^2 X^{k'}, k+1 \leq k'\} \cap V.$$

Let $\mathbf{v}$ be the vector given by all words in $V_0 \cup V_1$ in a fixed order. The goal is to give a construction of positive semidefinite tracial Gram matrices G_m for $S_{m,4}(X^2, Y^2)$, or equivalently, a decomposition as sum of hermitian squares obtained from G_m. This construction will be explained in the sequel.

We denote the exponents of X in a word $v_i \in V_0 \cup V_1$ by k_i, ℓ_i and k_i' as in the definition of V_0 (respectively V_1). Thus every $v_i \in V_0$ is of the form $v_i = X^{k_i} Y^2 X^{\ell_i} Y^2 X^{k_i'+1}$ where $k_i, \ell_i, k_i' \in 2\mathbb{N}$ satisfy the conditions $k_i + \ell_i + k_i' = m-5$ and $k_i \leq k_i'$. The exponent k_i (respectively $k_i + 1$ for $v_i \in V_1$) is bounded by d, the highest possible even (respectively odd) number which is less than or equal to $\frac{m-5}{2}$. Thus the maximum of these bounds is in any case $\frac{m-5}{2}$.

For a given $k \in \mathbb{N}$ let $k(2)$ denote the remainder of k modulo 2. Then we group the words $v_i \in V_0$ (respectively V_1) according to k_i. For every $k = 0, 1, 2, \ldots, \frac{m-5}{2}$ we add all words $v_i \in V_{k(2)}$ with $k_i + k(2) = k$ and obtain a polynomial f_k. By construction all words in ${f_k}^* f_k$ have even exponents in X and Y. Finally, we set

$$f := m \sum_{k=0}^{\frac{m-5}{2}} {f_k}^* f_k. \tag{3.4}$$

3.19 Example.

(a) $m = 7$: We have $V_0 = \{Y^2X^2Y^2X, Y^4X^3\}$ and $V_1 = \{XY^4X^2\}$ which leads by the proposed construction to

$$f_0 = Y^2X^2Y^2X + Y^4X^3 \quad \text{and} \quad f_1 = XY^4X^2$$

and finally

$$\begin{aligned} S_{7,4}(X^2, Y^2) &\overset{\text{cyc}}{\sim} 7\,({f_0}^* f_0 + {f_1}^* f_1) \\ &= 7\,(XY^2X^2Y^4X^2Y^2X + XY^2X^2Y^6X^3 + X^3Y^6X^2Y^2X + \\ &\qquad + X^3Y^8X^3 + XY^4X^4Y^4X). \end{aligned}$$

The decomposition $7\,({f_0}^* f_0 + {f_1}^* f_1)$ corresponds to the positive semidefinite tracial Gram matrix

$$G_7 := 7 \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where we chose the basis $\mathbf{v} = \begin{bmatrix} Y^2X^2Y^2 & Y^4X^3 & XY^4X^2 \end{bmatrix}^T$.

(b) $m = 9$: In this case we have the sets $V_0 = \{Y^2X^2Y^2X^3, Y^4X^5, X^2Y^4X^3, Y^2X^4Y^2X\}$ and $V_1 = \{XY^4X^4, XY^2X^2Y^2X^2\}$. One easily checks that $S_{9,4}(X^2, Y^2)$ decomposes as $9\,({f_0}^* f_0 + {f_1}^* f_1 + {f_2}^* f_2)$ with

$$\begin{aligned} f_0 &= Y^2X^2Y^2X^3 + Y^4X^5 + Y^2X^4Y^2X, \\ f_1 &= XY^2X^2Y^2X^2 + XY^4X^4 \quad \text{and} \\ f_2 &= X^2Y^4X^3. \end{aligned}$$

Let $\mathbf{v} = \begin{bmatrix} Y^2X^2Y^2X^3 & Y^2X^4Y^2X & Y^4X^5 & XY^2X^2Y^2X & XY^4X^4 & X^2Y^4X^3 \end{bmatrix}^T$ and G_9 be the positive semidefinite tracial Gram matrix

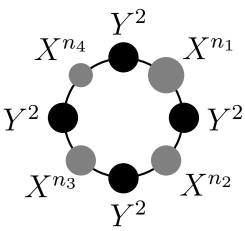
$$G_9 := 9 \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then we obtain the same decomposition $9\,({f_0}^*f_0 + {f_1}^*f_1 + {f_2}^*f_2)$ of $S_{9,4}(X^2, Y^2)$ as above.

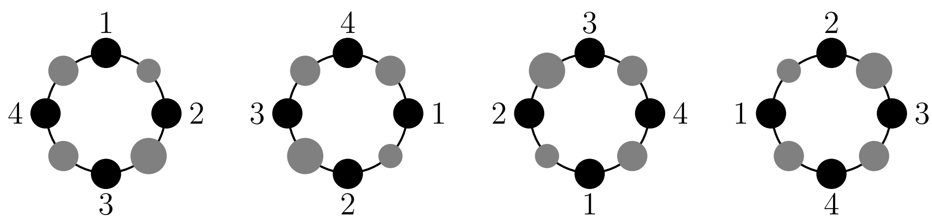
It is easy to generalize the construction of the tracial Gram matrices in Example 3.19 by extending the number and size of the block matrices with all entries equal to 1. The presented Construction (3.4) is based on this procedure.

To prove that f in (3.4) is cyclically equivalent to $S_{m,4}(X^2, Y^2)$ for all odd m, we first show that all words appearing in f are pairwise cyclically inequivalent. Since each word in f appears also in $S_{m,4}(X^2, Y^2)$ we finish by showing that the sums of coefficients in both polynomials are the same.

3.20 Remark. To compare two words appearing in f with respect to cyclic equivalence we use the following method. Since Y^2 appears exactly four times in each word w of f, we can write w as $w = X^{n_0}Y^2X^{n_1}Y^2X^{n_2}Y^2X^{n_3}Y^2X^{n_4'}$ for some $n_0, \dots, n_3, n_4'$. Further w is cyclically equivalent to $\tilde{w} = Y^2X^{n_1}Y^2X^{n_2}Y^2X^{n_3}Y^2X^{n_4}$ where $n_4 = n_4' + n_0$, i.e. $\tilde{w}$ consists of four groups $Y^2X^{n_i}$. Each such word can be represented by a necklace. For simplicity a black bead represents



a Y^2 and a grey bead represents a group X^{n_i} of $\tilde{w}$. The size of the grey beads is correlated to the exponent n_i. The word w is obtained in the necklace by shifting the starting point from the black bead on the top to X^{n_0} which is included implicitly in the upper left grey bead labelled by X^{n_4}. Let w' be a word with exponents m_i, i.e. $\tilde{w}' := Y^2X^{m_1}Y^2X^{m_2}Y^2X^{m_3}Y^2X^{m_4}$. Effectively, we will compare $\tilde{w}$ and $\tilde{w}'$. These words are cyclically equivalent if and only if $w \overset{\text{cyc}}{\sim} w'$ holds. Thus without loss of generality let w and w' start by Y^2. Then w and w' are equal or $n_i = m_{i-j}$ $(i - j \mod 4)$ for $i = 1, \dots, 4$ and $j = 1, 2, 3$, which can be obtained by "rotating" w' j times, i.e. for $j = 1$ one shifts the first group $Y^2X^{m_1}$ to the end, for $j = 2$ one shifts also the second group to the end and so on, thus m_i becomes m_{i-j}. This can also be represented via necklaces. We draw the necklace for w' and label the black beads according to Y^2 by $1, \dots, 4$. The groups X^{n_i} are given by grey beads. If we choose the upper black bead as starting point of the word which is represented by the necklace, then the words w and w' are cyclically equivalent if and only if one of the following four necklaces is equal to the necklace of w (without additional rotation).



In the following calculation we will consider these four necklaces and ask for conditions in each of these four cases to have equal necklaces. For simplicity we use the fact that rotating three times is the same as rotating once in the reverse direction, i.e. shifting $Y^2X^{m_4}$ to the beginning. Thus rotating w' three times is the same as fixing w' and rotating w once. Therefore we can omit $j = 3$ by symmetry.

3.21 Lemma. *All words appearing in f are pairwise cyclically inequivalent.*

Proof. By construction a word w in f is either a word in $\sum_{2k} f_{2k}{}^* f_{2k}$, and thus of the form $w = v_1{}^* v_2$ where $v_1, v_2 \in V_0$ and $k_1 = k_2$, i.e.

$$w = X^{k_2'+1}Y^2X^{\ell_1}Y^2X^{2k_1}Y^2X^{\ell_2}Y^2X^{k_2'+1} \overset{\text{cyc}}{\sim} Y^2X^{\ell_1}Y^2X^{2k_1}Y^2X^{\ell_2}Y^2X^{k_2'+k_2'+2},$$

or it is a word in $\sum_{2k} f_{2k+1}{}^* f_{2k+1}$, and thus of the form $w = v_1{}^* v_2$ where $v_1, v_2 \in V_1$ and $k_1 = k_2$. The same is true for any other word $w' = v_3{}^* v_4$. As is easily seen $\tilde{w} = \tilde{w}'$ is possible only if $v_1 = v_3$ and $v_2 = v_4$. We are left with the following cases.

If w and w' are words in $\sum_{2k} f_{2k}{}^* f_{2k}$ which are cyclically equivalent then we have to consider

(a) $\ell_1 = 2k_3, \quad 2k_1 = \ell_4, \quad \ell_2 = k_3' + k_4' + 2, \quad k_2' + k_2' + 2 = \ell_3$ or

(b) $\ell_1 = \ell_4, \quad 2k_1 = k_3' + k_4' + 2, \quad \ell_2 = \ell_3, \quad k_2' + k_2' + 2 = 2k_3.$

In (a) $2k_3 + k_1 + k_2' = \ell_1 + k_1 + k_2' = \ell_3 + k_3 + k_3' = k_2' + k_2' + 2 + k_3 + k_3'$ leads to $k_1 + k_3 = k_2' + k_3' + 2$ contradicting $k_1 + k_3 \le k_2' + k_3' < k_2' + k_3' + 2$. Subcase (b) leads to $2k_1 = k_3' + k_4' + 2 > 2k_3 = k_2' + k_2' + 2 > 2k_1$, which is not possible.

The case that w, w' are words in $\sum_{2k} f_{2k+1}{}^* f_{2k+1}$ works the same way.

If w is a word in $\sum_{2k} f_{2k}{}^* f_{2k}$ and w' a word in $\sum_{2k} f_{2k+1}{}^* f_{2k+1}$, then we have

(a) $\ell_1 = k_3' + k_4', \quad 2k_1 = \ell_3, \quad \ell_2 = 2k_3 + 2, \quad k_2' + k_2' + 2 = \ell_4$ or

(b) $\ell_1 = \ell_3, \quad 2k_1 = 2k_3' + 2, \quad \ell_2 = \ell_4, \quad k_2' + k_2' + 2 = k_3' + k_4'.$

In (a) $k_3' + k_4' + k_1 + k_2' = \ell_1 + k_1 + k_2' = \ell_3 + k_3 + k_3' = 2k_1 + k_3 + k_3'$ leads to $k_2' + k_4' = k_1 + k_3 = k_1 + k_4 < k_2' + k_4'$. Subcase (b) contradicts $k_1, k_3' \in 2\mathbb{N}$.

If w is a word in $\sum_{2k} f_{2k+1}{}^* f_{2k+1}$ and w' in $\sum_{2k} f_{2k}{}^* f_{2k}$, we exchange w and w'.

Summarizing, except in the trivial case where w and w' are constructed by the same subwords v_i, they cannot be cyclically equivalent. ∎

Thus every word in f has its order m as coefficient. Since up to cyclic equivalence this is the same in $S_{m,4}(X^2, Y^2)$, we are done by the following lemma.

3.22 Lemma. *The number of pairwise cyclically inequivalent words in $S_{m,4}(X^2, Y^2)$ is the same as in f.*

Proof. $S_{m,4}(X^2, Y^2)$ contains $\binom{m}{4}$ words. Since each word has order m, there are

$$\frac{1}{m}\binom{m}{4} = \frac{1}{6}(\frac{m-3}{2})(\frac{m-1}{2})(m-2)$$

pairwise cyclically inequivalent words in $S_{m,4}(X^2, Y^2)$.

Let $k \in \mathbb{N}$ be fixed. Then f_k consists of $\frac{m-3}{2} - k$ different words. For example, if k is even then there are $\frac{1}{2}(m - 5 - k_1) + 1$ possibilities for $k_1, \ell_1, k_2' \in 2\mathbb{N}$ with $\ell_1 + k_2' = m - 5 - k_1$ (namely $k_2' = m - 5 - k_1 - \ell_1, \ell_1 = 0, 2, \dots m - 5 - k_1$), the restriction $k_1 \le k_2'$ of V_0 excludes $\frac{k_1}{2}$ possibilities. Thus the number of words in f is given by

$$\sum_{k=0}^{\frac{m-5}{2}} (\frac{m-3}{2} - k)^2 = \sum_{k=0}^{\frac{m-3}{2}} k^2 = \frac{1}{6}(\frac{m-3}{2})(\frac{m-1}{2})(m-2).$$

∎

Case m even

Since words in $S_{m,4}(X^2,Y^2)$ have order m, $\frac{m}{2}$ or $\frac{m}{4}$ for even m, the constructed polynomial f in (3.4) will not be cyclically equivalent to $S_{m,4}(X^2,Y^2)$. Hence we modify the construction for even m using different weights on the summands.

Let m be fixed, $V := \{v \in \langle X,Y\rangle \mid \deg_X v = m-4, \deg_Y v = 4\}$ as above and

$$V_0 = \{v \in \{X^2,Y^2\}^{\frac{m}{2}} \mid v = X^k Y^2 X^\ell Y^2 X^{k'}, k \le k'\} \cap V,$$
$$V_1 = \{v \in X\{X^2,Y^2\}^{\frac{m-2}{2}} X \mid v = X^{k+1} Y^2 X^\ell Y^2 X^{k'+1}, k \le k'\} \cap V.$$

To distinguish even and odd exponents, we define $\hat{k_i} := k_i + 1$ and $\hat{k}_i' := k_i' + 1$. Then every $v_i \in V_0$ is of the form $v_i = X^{k_i} Y^2 X^{\ell_i} Y^2 X^{k_i'+1}$ and satisfies $k_i + \ell_i + k_i' = m-4$ with $\ell_i, k_i, k_i' \in 2\mathbb{N}$ and $k_i \le k_i'$, whereas every $v_i \in V_1$ satisfies $\hat{k_i} + \ell_i + \hat{k}_i' = m-4$. Thus the maximal possible exponent k_i respectively $\hat{k_i}$ (if m is not divisible by 4) is given by $\frac{m-4}{2}$.

Now we construct a sum of hermitian squares as follows. Let $k \in \mathbb{N}$ and let $k(2)$ denote the remainder of k modulo 2. For every $k = 0,1,2,\ldots \frac{m-4}{2}$ we add all words $v_i \in V_{k(2)}$ with $k_i + k(2) = k$ as in the case where m is odd, but we weight the words with $k_i < k_i'$ with coefficient 1 and the words with $k_i = k_i'$ with coefficient $\frac{1}{2}$. This leads to a polynomial f_k which contains exactly one word with coefficient $\frac{1}{2}$ whereas all other coefficients are 1. Finally we set

$$f := m \sum_{k=0}^{\frac{m-4}{2}} {f_k}^* f_k. \tag{3.5}$$

3.23 Example.

(a) $m = 8$: The two sets V_0 and V_1 are given by $V_0 = \{Y^2X^2Y^2X^2, Y^4X^4, X^2Y^4X^2, Y^2X^4Y^2\}$ and $V_1 = \{XY^4X^3, XY^2X^2Y^2X\}$. This leads to $S_{8,4}(X^2,Y^2) \overset{\text{cyc}}{\sim} 8({f_0}^*f_0 + {f_1}^*f_1 + {f_2}^*f_2)$ with

$$f_0 = Y^2X^2Y^2X^2 + Y^4X^4 + \frac{1}{2}Y^2X^4Y^2$$
$$f_1 = XY^4X^3 + \frac{1}{2}XY^2X^2Y^2X \quad \text{and}$$
$$f_2 = \frac{1}{2}X^2Y^4X^2.$$

Taking

$$\mathbf{v} = \begin{bmatrix} Y^2X^2Y^2X^2 & Y^4X^4 & Y^2X^4Y^2 & XY^4X^3 & XY^2X^2Y^2X & X^2Y^4X^2 \end{bmatrix}^T$$

and the positive semidefinite tracial Gram matrix

$$G_8 := 4\begin{bmatrix} 2 & 2 & 1 & 0 & 0 & 0 \\ 2 & 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

results in the same decomposition as sum of hermitian squares and commutators.

The different weights take care that the sum of coefficients of cyclically equivalent words does not exceed the corresponding order. For example, the coefficients of the two words $w = (Y^2X^2Y^2X^2)^*(Y^2X^4Y^2)$ and $w' = (XY^2X^2Y^2X)^*(XY^4X^3)$, which are cyclically equivalent, sum up to $\operatorname{ord}(w) = \operatorname{ord}(w') = 8$.

(b) $m = 10$: We have $S_{10,4}(X^2, Y^2) \overset{\text{cyc}}{\sim} 10(f_0{}^*f_0 + f_1{}^*f_1 + f_2{}^*f_2 + f_3{}^*f_3)$ where

$$\begin{aligned}
f_0 &= Y^2X^4Y^2X^2 + Y^2X^2Y^2X^4 + Y^4X^6 + \frac{1}{2}Y^2X^6Y^2, \\
f_1 &= XY^2X^2Y^2X^3 + XY^4X^5 + \frac{1}{2}XY^2X^4Y^2X, \\
f_2 &= X^2Y^4X^4 + \frac{1}{2}X^2Y^2X^2Y^2X^2 \quad \text{and} \\
f_3 &= \frac{1}{2}X^3Y^4X^3.
\end{aligned}$$

This is the same representation as sum of hermitian squares which will be obtained by the positive semidefinite tracial Gram matrix

$$G_{10} := 5\begin{bmatrix}
2 & 2 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 2 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 2 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2 & 2 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2 & 2 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & \frac{1}{2} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}
\end{bmatrix}$$

with respect to the vector $\mathbf{v}$ given by the words $Y^2X^4Y^2X^2, Y^2X^2Y^2X^4, Y^4X^6, Y^2X^6Y^2$, $XY^2X^2Y^2X^3$, $XY^4X^5, XY^2X^4Y^2X, X^2Y^4X^4, X^2Y^2X^2Y^2X^2, X^3Y^4X^3$.

Again, one can detect a block structure in the tracial Gram matrices G_m. It consists of $\frac{m}{2} - 1$ blocks of decreasing size starting by $\frac{m}{2} - 1$. This structure holds for every $m \in 2\mathbb{N}$ and the proposed construction of f is the resulting sum of hermitian squares representation coming from the Cholesky decomposition of these tracial Gram matrices.

The proof that f in (3.5) is cyclic equivalent to $S_{m,4}(X^2, Y^2)$ for even m will also be done in two steps. First, the sum of coefficients of cyclically equivalent words in f is less than or equal to their order. Second, since each word in f appears in $S_{m,4}(X^2, Y^2)$ we finish again by showing that the sums of coefficients are equal in both representations.

3.24 Lemma. *The sum of coefficients of cyclically equivalent words in f is equal or less than the order of the corresponding words.*

Proof. We will use the same method as explained in Remark 3.20 of the last section. Let w, w' be two different words appearing in f and $w \overset{\text{cyc}}{\sim} w'$.

If w and w' are words in $\sum f_{2k}{}^*f_{2k}$. Then either w and w' are equal or one of the following subcases holds:

(a) $\ell_1 = 2k_3, \quad 2k_1 = \ell_4, \quad \ell_2 = k_3' + k_4', \quad k_2' + k_2' = \ell_3$

(b) $\ell_1 = \ell_4, \quad 2k_1 = k_3' + k_4', \quad \ell_2 = \ell_3, \quad k_2' + k_2' = 2k_3$

In Subcase (a) we obtain $k_3+k_1 = k_2'+k_3'$ from $2k_3+k_1+k_2' = \ell_1+k_1+k_2' = \ell_3+k_3+k_3' = k_2'+k_2'+k_3+k_3'$, thus $k_1 = k_2'$ and $k_3 = k_3'$. Further we obtain from $\ell_1+k_1+k_2' = \ell_4+k_3+k_4'$ that $k_2'-k_1 = k_4'-k_3$. In (b) we obtain $2k_1 = k_3'+k_4' \geq 2k_3 = k_2'+k_2' \geq 2k_1$, thus equality holds, which leads to $w = w'$.

The other cases work in the same way by replacing k_i by $\hat{k}_i$ whenever w is a word in the sum $\sum f_{2k+1}{}^* f_{2k+1}$ and k_i' respectively if w' is a word in $\sum f_{2k+1}{}^* f_{2k+1}$. If w and w' are not in the same set $\sum f_{2k}{}^* f_{2k}$ or $\sum f_{2k+1}{}^* f_{2k+1}$, then they obviously cannot be equal.

Summarizing, we derive that when $w \overset{\text{cyc}}{\sim} w'$ but $w \neq w'$ then $k_1 = k_2',\ k_3 = k_3'$ and $k_2' - k_1 = k_4' - k_3$ or by symmetry (confer Remark 3.20) $k_3 = k_4',\ k_1 = k_2'$ and $k_3' - k_3 = k_2' - k_1$ holds, where the first set of equations describes the words which differ by one rotation, and the second set describes the case of three rotations.

Assuming, there are two different words w', w'' both cyclically equivalent to w. Then all three are pairwise cyclically equivalent and at least two of them (for example w', w'') are in $\sum_k f_k{}^* f_k$ (k even or odd). Thus each of them satisfies one set of equations, but then w' and w'' differ by two rotations, which leads to equality (Subcase (b)). Therefore there are at most two words in f which are pairwise cyclically equivalent.

To conclude the proof, if $w = v_1{}^* v_2$ with $k_1 = k_2' = k_2 = \frac{m-4}{4}'$ then $\ell_1 = m-4-2k_1 = \frac{m-4}{2} = \ell_2$, thus w has order $\frac{m}{4}$ which is equal to the coefficient of w in f. A cyclically equivalent word $w' = v_3{}^* v_4$ has to satisfy $k_3 = k_3' = k_4'$ and $2k_3 = \ell_1 = 2k_1$ which leads to $w = w'$. Therefore there is no other word $w' \overset{\text{cyc}}{\sim} w$ in f. In all other cases the coefficient of w is half of the order of w. Since there are at most two pairwise cyclically equivalent words the proof is finished. ∎

3.25 Lemma. *The sum of coefficients in f and $S_{m,4}(X^2, Y^2)$ is the same.*

Proof. The sum of coefficients in $S_{m,4}(X^2, Y^2)$ is $\binom{m}{4}$.

For every $k = 0, 1, 2 \ldots, \frac{m-4}{2}$ each polynomial f_k has one word with coefficient $\frac{1}{2}$ and $\frac{m-4}{2} - k$ times coefficient 1. Thus the sum of coefficients in f is

$$m \sum_{k=0}^{\frac{m-4}{2}} \Big(\frac{m-4}{2} - k + \frac{1}{2}\Big)^2 = \frac{m(m-2)}{8} + m \sum_{k=0}^{\frac{m-4}{2}} (k^2 + k)$$
$$= \frac{m}{24}\big(3(m-2) + (m-4)(m-2)m\big) = \binom{m}{4}.$$

∎

This finishes the proof of Theorem 3.15.

3.4.2 Proof of $S_{4r+2,4}(X,Y) \in \Theta^2$

In this section we establish Theorem 3.17 which implies that $\operatorname{Tr}(S_{4r+2,4}(A,B)) \geq 0$ for all matrices $A, B \in \mathcal{S}\mathbb{R}^{s\times s}$, $s \in \mathbb{N}$. We will construct a positive semidefinite tracial Gram matrix $G(r)$ for $S_{4r+2,4}(X,Y)$ depending on the parameter r, from which one can obtain a representation as a sum of hermitian squares and commutators as shown in Proposition 3.8. Th procedure works similarly to the proof of Theorem 3.15. First we prove that the constructed matrix is a tracial Gram matrix for $S_{4r+2}(X,Y)$ by comparing the coefficients of cyclically equivalent words. Then we show that this tracial Gram matrix is in fact positive semidefinite.

We consider the set $V := \{v \in \langle X, Y\rangle \mid \deg_X v = 2r-1, \deg_Y v = 2\}$, restricting the words appearing in $\mathbf{v}$ to the correct degree, and

$$V_0 := \{v \in \{X,Y\}^{2r-1} \mid v = X^k Y X^\ell Y X^{k'}, k \le k'\} \cap V.$$

In particular, we have $k \le r-1$. Let $\mathbf{v}$ contain all words in V_0 and let the order of these words in $\mathbf{v}$ be fixed. We denote then the i-th word of $\mathbf{v}$ by v_i and label the tracial Gram matrix G by $i, j \in \mathbb{N}$. The exponents of X in $v_i \in V_0$ may be k_i, ℓ_i and k_i' as before. Using this convention we define the tracial Gram matrix $G := G(r)$ by the following rules:

- $G_{i,j} = m$ if v_i, v_j satisfy $k_i = k_j$ and $k_i < k_i'$, $k_j < k_j'$;
- $G_{i,j} = \frac{m}{2}$ if $k_i = k_j$ and ($k_i = k_i'$ or $k_j = k_j'$);
- $G_{i,j} = \frac{m}{2}$ if ($k_i - k_j = 1$ and $k_j < k_j'$) or ($k_j - k_i = 1$ and $k_i < k_i'$);
- $G_{i,j} = 0$ otherwise.

One easily sees that G is symmetric by construction. We will illustrate this construction by two examples.

3.26 Example.

(a) $m = 10$: We consider the vector

$$\mathbf{v} = \begin{bmatrix} YX^2YX & YXYX^2 & Y^2X^3 & YX^3Y & XY^2X^2 & XYXYX \end{bmatrix}^T.$$

The proposed construction leads to the tracial Gram matrix $G(2)$ for $S_{10,4}(X,Y)$ of the form

$$G(2) := 5 \begin{bmatrix} 2 & 2 & 2 & 1 & 1 & 1 \\ 2 & 2 & 2 & 1 & 1 & 1 \\ 2 & 2 & 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 2 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

One easily checks that $G(2)$ is positive semidefinite. Calculating the decomposition $G(2) = \sum_i \vec{g_i}\vec{g_i}^T$ we obtain as in Proposition 3.8 that

$$S_{10,4}(X,Y) \overset{\text{cyc}}{\sim} \frac{5}{2}(g_1{}^* g_1 + g_2{}^* g_2 + 2 g_3{}^* g_3)$$

with

$$\begin{aligned} g_1 &= 2YX^2YX + 2YXYX^2 + 2Y^2X^3 + YX^3Y + XY^2X^2 + XYXYX, \\ g_2 &= YX^3Y - XY^2X^2 - XYXYX \text{ and} \\ g_3 &= XY^2X^2. \end{aligned}$$

(b) $m = 14$: The construction described above leads for $S_{14,4}(X,Y)$ to the positive semidefinite tracial Gram matrix

$$G(3) = 7\begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 2 & 2 & 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 2 & 2 & 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 2 & 2 & 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \end{bmatrix}$$

according to the vector $\mathbf{v}$ containing the words YX^4YX, YX^3YX^2, YX^2YX^3, $YXYX^4$, Y^2X^5, YX^5Y, XYX^2YX^2, $XYXYX^3$, XY^2X^4, XYX^3YX, $X^2Y^2X^3$ and X^2YXYX^2 in the same order. Since $G(3)$ is of rank 4 the BMV polynomial $S_{14,4}(X,Y)$ is cyclically equivalent to a sum of four hermitian squares. We will return to the decomposition of $G(3)$ later on.

We proceed to show that the constructed symmetric matrix $G(r)$ is a positive semidefinite tracial Gram matrix of $S_{4r+2,4}(X,Y)$ for any $r \in \mathbb{N}$. To compare the coefficients of $S_{4r+2,4}(X^2,Y^2)$ with the coefficients of $\mathbf{v}^*G(r)\mathbf{v}$ we set

$$f(r) := \mathbf{v}^*G(r)\mathbf{v}.$$

In the first step, we will show that the coefficients in $f(r)$ of cyclically equivalent words sum up to the corresponding order (comparable to the case of even m in the proof of Theorem 3.15). Since this property also holds true for $S_{4r+2}(X,Y)$ and the sum of coefficients are the same in both polynomials, the matrix $G(r)$ is a tracial Gram matrix of $S_{4r+2,4}(X,Y)$.

3.27 Lemma. *For any $r \in \mathbb{N}$ the sum of coefficients in $f(r) = \mathbf{v}^*G(r)\mathbf{v}$ of cyclically equivalent words is bounded by the order of the corresponding words.*

Proof. Let $r \in \mathbb{N}$ be fixed. We set $f := f(r)$ and $G := G(r)$. First we calculate all combinations of exponents for two words w and w' appearing in f to be cyclically equivalent. Let these two words be $w = X^{k_1'}YX^{\ell_1}YX^{k_1}X^{k_2}YX^{\ell_2}YX^{k_2'}$ and $w' = X^{t_1'}YX^{s_1}YX^{t_1}X^{t_2}YX^{s_2}YX^{t_2'}$ for some $k_i, k_i', \ell_i, t_i, t_i', s_i \in \mathbb{N}$. Without loss of generality we assume that $t_1 \le k_1$. Let $G_w \ne 0$ be the coefficient of w and $G_{w'} \ne 0$ be the coefficient of w'. If $w \overset{\text{cyc}}{\sim} w'$, then one of the following cases holds, cf. Remark 3.20:

(a) $\ell_1 = s_1, k_1 + k_2 = t_1 + t_2, \ell_2 = s_2, k_1' + k_2' = t_1' + t_2'$

(b) $\ell_1 = t_1 + t_2, k_1 + k_2 = s_1, \ell_2 = t_1' + t_2', k_1' + k_2' = s_2$

(c) $\ell_1 = s_2, k_1 + k_2 = t_1' + t_2', \ell_2 = s_1, k_1' + k_2' = t_1 + t_2$

(d) $\ell_1 = t_1' + t_2', k_1 + k_2 = s_2, \ell_2 = t_1 + t_2, k_1' + k_2' = s_1$

For these cases we have to calculate all possible combinations separately. By symmetry we can omit case (d) since this case is equal to case (b) after exchanging the indices. Further case (c) has only the trivial solution $k_i = t_i, k_i' = t_i'$ for $i = 1,2$. Thus we are left with the first two cases.

In Case (a) using $\ell_1 + k_1 + k_1' = s_1 + t_1 + t_1'$ and $s_1 = \ell_1$ leads to $k_1 + k_1' = t_1 + t_1'$. Similarly we get $k_2 + k_2' = t_2 + t_2'$ which implies $k_1 = t_1 + t_1' - k_1' = t_1 + t_2 - k_2$ and

$k_1' = t_1 + t_1' - k_1 = t_1' + t_2' - k_2'$. Thus $t_1' - k_1' = t_2 - k_2$ and $t_1 - k_1 = t_2' - k_2' = k_1' - t_1' = k_2 - t_2$. If $t_1 = k_1$ then all exponents are equal and we can omit this trivial solution. If $t_1 = k_1 - i$ for some $i \in \{1, \dots, k_1\}$ we derive a contradiction if $k_1 = k_2$ or $k_2 = k_1 + 1$. For instance, if $k_1 = k_2$ we obtain $k_1 - i + t_2 = t_1 + t_2 = k_1 + k_2 = 2k_1$ which implies $t_2 = k_1 + i = t_1 + 2i$ and hence $t_2 - t_1 \geq 2$ which implies $G_{w'} = 0$. The other case works the same way. Thus let $k_2 = k_1 - 1$ which implies $t_2 = k_1 + i - 1$. Since $t_2 - t_1 = 2i - 1 \leq 2$ we conclude that $i = 1$. Thus $k_2 = k_1 - 1, k_2 < k_2'$ and $t_1 = k_1 - 1, t_2 = t_1 + 1, t_1' = k_1' + 1, t_2' = k_2' - 1$.

Case (b) is more complicated than the other ones but the calculations are simple. From $t_1 + t_2 + k_1 + k_1' = \ell_1 + k_1 + k_1' = s_1 + t_1 + t_1' = k_1' + k_2' + t_1 + t_1'$ and $\ell_2 + k_2 + k_2' = s_2 + t_2 + t_2'$ we obtain $k_2' + t_1' = t_2 + k_1$ and $k_2 + t_2' = t_1 + k_1'$ which implies $t_2 - t_1 \geq k_2 - k_1$.

- If $t_1 = t_2 + 1$, we obtain $k_1 - k_2 \geq 1$. Since $|k_1 - k_2| \leq 1$ we get equality and thus $k_1 - 1 = k_2 = k_2'$ which implies the contradiction $G_w = 0$.
- If $t_1 = t_2$ then $k_1 \geq k_2$. For $k_1 = k_2$ we obtain $k_2' + t_1' = t_1 + k_1$ which implies with $t_1 \leq t_1'$ and $k_1 = k_2 \leq k_2'$ that $t_2' - t_2 = k_1' - k_1, k_1 = k_2 = k_2'$ and $t_1 = t_1' = t_2$.

  For $k_1 = k_2 + 1$, we get $t_1 \leq t_1' \leq t_1 + 1$. If $t_1' = t_1 + 1$ we have $k_1 = k_2' + 1 = k_2 + 1$ which implies the contradiction $G_w = 0$ by construction. Thus we have $t_1 = t_1' = t_2$ and $k_1 = k_2 + 1 = k_2'$.
- Let $t_1 = t_2 - 1$. If $k_1 = k_2 - 1$ we obtain $k_2' + t_1' = k_2 + t_1$ and therefore $t_1 = t_1' = t_2 - 1$ which contradicts $G_{w'} \neq 0$. For $k_1 = k_2$ we obtain $k_2 + 1 \geq k_2' \geq k_2$ which leads to $k_2 = k_2'$ since $k_2 + 1 = k_2'$ gives $G_{w'} = 0$ as above.

  For $k_1 = k_2 + 1$, we get $t_2' - t_2 = k_1' - k_1$. Further, from $k_2' + t_1' = t_2 + k_1 = k_2 + 1 + t_2 = k_2 + t_1 + 1$ we obtain $t_1' - t_1 = k_2 - k_2' + 1 \leq 1$. Taking only possibilities into account for which G_w and $G_{w'}$ are not 0 we derive $t_1 = t_1' - 2 = t_2 - 1$ and $k_1 - 1 = k_2 = k_2' - 1$.

Hence one of the following cases holds, which satisfies additionally $t_2' - t_1 = k_1' - k_2$:

(1) $k_1 = k_2 = k_2', t_1 = t_1' = t_2$

(2) $k_1 - 1 = k_2 = k_2' - 1, t_1 = t_1' = t_2$

(3) $k_1 = k_2 = k_2', t_1 + 1 = t_1' = t_2$

(4) $k_1 - 1 = k_2 = k_2' - 1, t_1 = t_1' - 1 = t_2 - 1$

Finally, we have to take into account that we assumed $t_1 \leq k_1$, and thus we get the following cases:

(I) $k_2 = k_1 - 1, k_2 < k_2', t_1 = k_1 - 1, t_2 = t_1 + 1, t_1' = k_1' + 1, t_2' = k_2' - 1$

(II) $k_2 = k_1 - 1, k_2 < k_2', t_1 = k_1 + 1, t_2 = t_1 + 1, t_1' = k_1' - 1, t_2' = k_2' + 1$

(III) $k_1 = k_2 = k_2', t_1 = t_1' = t_2$ and $t_1 < k_1, t_2' < k_1'$

(IV) $k_1 = k_2 = k_2', t_1 = t_1' = t_2$ and $t_1 > k_1, t_2' > k_1'$

(V) $k_1 - 1 = k_2 = k_2' - 1, t_1 = t_1' = t_2$ and $t_1 < k_1, t_2' < k_1' + 1$

(VI) $k_1 - 1 = k_2 = k_2' - 1, t_1 = t_1' = t_2$ and $t_1 > k_1, t_2' > k_1' + 1$

(VII) $k_1 = k_2 = k_2', t_1 + 1 = t_1' = t_2$ and $t_1 < k_1, t_2' < k_1'$

(VIII) $k_1 = k_2 = k_2', t_1 + 1 = t_1' = t_2$ and $t_1 > k_1, t_2' > k_1'$

(IX) $k_1 - 1 = k_2 = k_2' - 1, t_1 = t_1' - 1 = t_2 - 1$ and $t_1 < k_1, t_2' < k_1' + 1$

(X) $k_1 - 1 = k_2 = k_2' - 1, t_1 = t_1' - 1 = t_2 - 1$ and $t_1 > k_1, t_2' > k_1' + 1$.

By easy observation on the exponents one concludes that for each word only one case (I)–(X) is possible, hence at most two words in f can be pairwise cyclically equivalent. All words satisfying one of the cases (I)–(X) have order m and by construction the coefficient $\frac{m}{2}$. Thus the coefficients add up to m, the order of the corresponding words. In the other cases, where there is no second word cyclically equivalent to w, one gets by construction of $G(r)$ that the corresponding coefficient does not exceed the order of the word. ∎

To conclude the first part of the proof we calculate the sum of all coefficients in f, which should be equal to $\binom{4r+2}{4}$, the sum of coefficients in $S_{4r+2,4}(X,Y)$. If this is the case then $G(r)$ is tracial Gram matrix of $S_{4r+2,4}(X,Y)$ for all $r \in \mathbb{N}$.

3.28 Lemma. *The sum of coefficients in $f(r)$ is $\binom{4r+2}{4}$.*

Proof. The sum of all entries in $G(r)$ is equal to the sum of coefficients of $f(r)$. The former sum is easily calculated. We have in $G(r)$ for each $k = 0, \dots, r-1$ one square block of size $(2r-2k-1)$ with entry $4r+2$, two $(2r-2k-1) \times (2r-2k-1)$ blocks with entry $2r+1$ and one single entry $2r+1$ (corresponding to $k_i = k_j = k_i' = k_j'$). Thus the sum of coefficients of $f(r)$ is given by

$$\begin{aligned}
2(4r+2)\sum_{k=0}^{r-1}(2r-2k-1)^2 + (2r+1)\sum_{k=0}^{r-1} 1 &= (4r+2)\Big(2\sum_{k=1,\text{odd}}^{2r-1} k^2 + \frac{r}{2}\Big) \\
= \frac{4r+2}{24}\Big(4(2r-1)r(2r+1)+3r\Big) &= \frac{1}{24}(4r+2)(4r+1)4r(4r-1) \\
= \binom{4r+2}{4},
\end{aligned}$$

where we used that $\sum_{k=1,odd}^{n} k^2 = \frac{1}{6}n(n+1)(n+2)$. ∎

Hence we have shown that $G(r)$ is in fact a tracial Gram matrix for $S_{4r+2,4}(X,Y)$. To finish the proof we calculate the Cholesky decomposition of $G(r)$, which implies that $G(r)$ is positive semidefinite.

3.29 Lemma. *The tracial Gram matrix $G(r)$ is decomposable as*

$$G(r) = \frac{2r+1}{2} L(r)^T L(r)$$

for some real matrix $L(r)$, i.e. $f(r) = \mathbf{v}^T G(r)\mathbf{v}$ is a sum of hermitian squares.

Proof. We fix the order of the words in the index set of $G(r)$ as in the examples. That is, $\mathbf{v}$ contains first all words of V_0 starting with Y, where the symmetric word $YX^{2r-1}Y$ is the last one among all these words. Since Y appears only twice in each word, there is exactly one symmetric word $X^kYX^\ell YX^k$ in V_0 for each $k = 0, \dots, r-1$. We take all words in V_0 starting with X, again with the symmetric word $XYX^{2r-3}YX$ at the end, and so on until $k = r-1$. Since $G(r)$ is invariant under exchanging two words v_1 and v_1 in $\mathbf{v}$ which are not symmetric and which both start with the same power of X, the order of the non-symmetric words for a fixed power does not matter. Any order of these words gives rise to the same tracial Gram matrix $G(r)$.

Then by construction the tracial Gram matrix $G(r) \in \mathcal{S}\mathbb{R}^{t\times t}$ with $t := r(r+1)$ consists, beside other entries, of r blocks with entry $4r+2$. Each block corresponds to a different power k in the words $X^kYX^\ell YX^{k'}$ labelling $G(r)$ and has size $2r-2k-1$. Now as we have fixed an order, we can label $L(r)$ by indices $i,j \in \mathbb{N}$. We define the entries $L_{i,j}$ of $L(r)$ in the decomposition $G(r) = \frac{2r+1}{2}L(r)^T L(r)$ depending on the exponent $k = 0, \dots, r-1$. Since $L(r)$ is an upper triangular matrix it suffices to define $L_{i,j}$ for $i \geq j$.

For $k = 0$ we set

- $L_{1,j} = 2$ for $j = 1, \dots, 2r-1$,
- $L_{1,j} = 1$ for $j = 2r, \dots, 2r, \dots, 4r-2$,
- $L_{2r,2r} = 1$, $L_{2r,j} = -1$ if $j = 2r+1, \dots, 4r-2$.

The other rows in $L(r)$ can be described simultaneously. They only vary in the size which depends on k. For fixed $k \in \{1, \dots, r-1\}$ the index i, corresponding to words in the index set starting with X^k, lies between

$$s_k := 1 + \sum_{\ell=0}^{k-2} (2r - 2\ell)$$

and $s_{k+1} - 1$. In particular, we have to consider $2r - 2k$ indices. For $i = s_k$ we set

- $L_{s_k,j} = \sqrt{2}$ if $j = s_k, \dots, s_k + 2r - 2k - 2$,
- $L_{s_k,j} = \sqrt{2}$ if $j = s_{k+1}, \dots, s_{k+2} - 1$ and $k+1, k+2 \leq r$.

All other entries in $L(r)$ are 0. We claim that this construction leads to a Cholesky decomposition of $G(r)$. Before we prove our claim, we illustrate this by revisiting Example 3.26.

3.30 Example.

(a) $m = 10$: The tracial Gram matrix $G(2)$ in Example 3.26(a) can be written as $G(2) = \frac{5}{2} L(2)^T L(2)$ with

$$L(2) = \begin{bmatrix} 2 & 2 & 2 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

This is exactly the matrix which we obtain by the proposed construction. In fact, from this decomposition we obtain as in Example 3.26 (a) that

$$S_{10,4}(X,Y) \overset{\text{cyc}}{\sim} \frac{5}{2}({g_1}^* g_1 + {g_2}^* g_2 + 2{g_3}^* g_3)$$

with

$$\begin{aligned} g_1 &= 2YXYX + 2YXYX^2 + 2Y^2X^3 + YX^3Y + XY^2X^2 + XYXYX, \\ g_2 &= YX^3Y - XY^2X^2 - XYXYX \text{ and} \\ g_3 &= XY^2X^2. \end{aligned}$$

(b) $m = 14$: The proposed construction of $L(r)$ leads for $r = 3$ to the matrix

$$L(3) = \begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & \sqrt{2} & 0 & \sqrt{2} & \sqrt{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

One easily calculates that $G(3) = \frac{7}{2}L(3)^T L(3)$ holds, where $G(3)$ is the tracial Gram matrix from Example 3.26(b). This matrix decomposition leads to a representation with four hermitian squares, namely

$$S_{14,4}(X,Y) \overset{\text{cyc}}{\sim} \frac{7}{2}({f_1}^* f_1 + {f_2}^* f_2 + 2{f_3}^* f_3 + 2{f_4}^* f_4)$$

with

$$\begin{aligned} f_1 &= 2YX^4YX + 2YX^3YX^2 + 2YX^2YX^3 + 2YXYX^4 + 2Y^2X^5 \\ &\qquad\qquad + YX^5Y + XYX^4YX + XYXYX^3 + XY^2X^4 + XYX^3YX \\ f_2 &= YX^5Y - XYX^2YX^2 - XYXYX^3 - XY^2X^4 - XYX^3YX \\ f_3 &= YX^5Y + XYX^2YX^2 + XYXYX^3 + X^2Y^2X^3 + X^2YXYX^2 \\ f_4 &= X^2Y^2X^3. \end{aligned}$$

The general case will be proved by induction over r. The induction basis $r = 1$ is easy. We have $S_{6,4}(X,Y) \overset{\text{cyc}}{\sim} \mathbf{v}^T G(1)\mathbf{v}$ with $\mathbf{v} = \begin{bmatrix} Y^2X & YXY \end{bmatrix}^T$ and

$$G(1) = \frac{3}{2}\begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}.$$

Obviously $G(1) = \frac{3}{2}L(1)^T L(1)$ where by construction

$$L(1) = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}.$$

For the induction step assume that the induction hypothesis holds for some $r \in \mathbb{N}$. Let $G'(r) = \frac{2}{2r+1}G(r)$ and $G'(r+1) = \frac{2}{2r+3}G(r+1)$. Then the matrix $G'(r+1)$ is of the same structure as $G'(r)$ with the same entries. It only contains one additional block of size $J_0 = 2r$. Alternatively, we can consider $G'(r+1)$ as an extension of $G'(r)$ where we extended each block (corresponding to some k) from size J_k to size J_{k+1}. By induction hypothesis we have $G'(r) = L(r)^T L(r)$. Hence if we extend the blocks in $L(r)$ (corresponding to some k) in the same way we derive the desired decomposition $G'(r+1) = L(r+1)^T L(r+1)$. ∎

Combining the above statements we have shown that there is a positive semidefinite tracial Gram matrix of $S_{4r+2,4}(X,Y)$. Thus Theorem 3.17 is proved.

4 The tracial moment problem

To show that all matrix-positive polynomials are sums of hermitian squares McCullough solved a non-commutative moment problem [McC], see also [Hel]. The theory of positive polynomials in commuting variables is also intimately connected with its corresponding moment problem by Haviland's theorem [Hav], see e.g. [Mar]. Since Schmüdgen's celebrated solution of the moment problem on compact basic closed semi-algebraic sets [Sch], the moment problem has played a prominent role in Real Algebra, exploiting this duality between positive polynomials and the moment problem, see e.g. [KM, PS, Put, PV]. For more information on the classical moment problem in several variables we refer the reader to Akhiezer [Akh] for the analytic theory, see also [KN, ST], and to the survey of Laurent [Lau1] and references therein for a more algebraic approach.

In this chapter we present the tracial analog of the classical moment problem. This moment problem was first defined by Klep and the author in [BK1] using finitely atomic measures. In the introduction we define the tracial moment problem, including tracial (moment) sequences, tracial Hankel matrices and their truncated analogs. We present some basic properties of the tracial Hankel matrix in relation to its associated linear form and show that the kernel of a tracial Hankel matrix has some real-radical-like properties. We finish Section 4.1 by giving some necessary conditions on a tracial sequence to have a representing measure. Section 4.2 deals with the full tracial moment problem, where we prove that a tracial sequence with positive semidefinite tracial Hankel matrix of finite rank has a tracial moment representation, i.e. the tracial moment problem for this tracial sequence is solvable. Finally, we give several tracial analogs of results concerning the classical truncated moment problem in Section 4.3, e.g. theorems of Stochel [Sto], of Bayer and Teichmann [BT], and of Curto and Fialkow [CF1, CF2]. Several results of this chapter have been published in collaboration with Klep and others and can also be found in [BK1, BCKP].

4.1 Introduction

The classical moment problem deals with the question which linear functionals L on $\mathbb{R}[\underline{x}]$ are integration with respect to some positive measure μ. By Haviland's theorem [Hav] this holds true if and only if L is positive on all polynomials that are positive on $\mathbb{R}^n$. A linear functional L on $\mathbb{R}[\underline{x}]$ can be represented by its moments $y_\alpha := L(\underline{x}^\alpha)$, where $\underline{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ for the multi-index $\alpha \in \mathbb{N}_0^n$. Hence each linear functional L has an associated sequence (y_α) and the moment problem can be reformulated as the question: For which sequences (y_α) does there exist a positive measure μ such that $y_\alpha = \int \underline{x}^\alpha \, d\mu(x)$ for all $\alpha \in \mathbb{N}_0^n$?

In this section we present the tracial analog of this question, namely the tracial moment problem. Furthermore, we define tracial Hankel matrices, which are intimately connected with the tracial moment problem, and present some basic properties of tracial Hankel matrices and of tracial moment sequences.

4.1.1 Tracial moment sequences

We start by defining tracial sequences, which represent tracial linear functionals from $\mathbb{R}\langle \underline{X} \rangle$ to $\mathbb{R}$, and their truncated analog. Further we explain what we understand as tracial moment problem.

4.1 Definition. A sequence of real numbers (y_w) indexed by words $w \in \langle \underline{X} \rangle$ satisfying

$$y_w = y_u \text{ whenever } w \overset{\text{cyc}}{\sim} u, \tag{4.1}$$

$$y_w = y_{w^*} \text{ for all } w, \tag{4.2}$$

and $y_1 = 1$, is called a (normalized) *tracial sequence*.

If we take as index set only words $w \in \langle \underline{X} \rangle$ with $\deg w \leq k$, we write $y = (y_w)_{\leq k}$ for the finite sequence and call y a *truncated tracial sequence* of degree k if y satisfies properties (4.1) and (4.2) for all $w \in \langle \underline{X} \rangle$ with $\deg w \leq k$.

Each such tracial sequence represents a tracial linear functional L_y called the *(tracial) Riesz functional*, which can be defined as

$$\begin{aligned} L_y : \mathbb{R}\langle \underline{X} \rangle &\to \mathbb{R} \\ p = \sum_w p_w w &\mapsto \sum_w p_w y_w. \end{aligned}$$

Since y is a tracial sequence L_y satisfies $L_y(pq - qp) = 0$, hence it is a tracial linear functional. The same holds true for truncated tracial sequences $(y_w)_{\leq k}$ representing a tracial linear functional $L_y : \mathbb{R}\langle \underline{X} \rangle_k \to \mathbb{R}$.

4.2 Example.

(a) Let $\underline{A} = (A_1, \ldots, A_n) \in (\mathcal{S}\mathbb{R}^{s\times s})^n$ be an n-tuple of symmetric matrices. Then the sequence $y^{\underline{A}}$ given by

$$y_w^{\underline{A}} := \operatorname{Tr}(w(\underline{A}))$$

is a tracial sequence since the traces of cyclically equivalent words coincide. Taking in the definition of $(y_w)_w$ only words $w \in \mathbb{R}\langle \underline{X} \rangle_k$ leads to a truncated tracial sequence of degree k and a tracial Riesz functional L_y defined on $\mathbb{R}\langle \underline{X} \rangle_k$.

(b) Any convex combination y of tracial sequences $y^{\underline{A}^{(i)}}$, i.e. y defined by

$$y_w := \sum_{i=1}^{N} \lambda_i \operatorname{Tr}(w(\underline{A}^{(i)}))$$

for some $\lambda_i \in \mathbb{R}_{>0}$ with $\sum_i^N \lambda_i = 1$ and $\underline{A}^{(i)} = (A_1^{(i)}, \ldots, A_n^{(i)}) \in (\mathcal{S}\mathbb{R}^{s\times s})^n$, is again a tracial sequence.

(c) Given $s \in \mathbb{N}$ and a probability measure μ on $(\mathcal{S}\mathbb{R}^{s\times s})^n$, then the sequence y which is given by

$$y_w := \int \operatorname{Tr}(w(\underline{A}))\, d\mu(\underline{A})$$

for all $w \in \langle \underline{X} \rangle$ is a tracial sequence.

The tracial moment problem deals with the question which tracial linear functionals L_y on $\mathbb{R}\langle \underline{X} \rangle$ can be expressed as in Example 4.2(c). To be more specific, we distinguish three kinds of tracial moment problems – the (full) tracial moment problem, the truncated tracial tracial moment problem and the tracial K-moment problem – which are all connected.

4.3 Definition.

1. The *tracial moment problem* consists of the characterization of tracial sequences $y = (y_w)$ for which there is some $s \in \mathbb{N}$ and a probability measure μ on $(\mathcal{S}\mathbb{R}^{s\times s})^n$, such that for all $w \in \langle \underline{X} \rangle$,
$$y_w = \int \operatorname{Tr}(w(\underline{A}))\, d\mu(\underline{A}). \tag{4.3}$$
We then say that y has a *tracial moment representation* and call it a *tracial moment sequence*. The measure μ is then a *representing measure* for y.
2. The *truncated tracial moment problem* consists of the characterization of truncated tracial sequences which have a tracial moment representation, i.e. when does there exist a representation of the values y_w as in (4.3) for all $w \in \langle \underline{X} \rangle$ with $\deg w \leq k$? If this is the case, then the sequence $(y_w)_{\leq k}$ is called a *truncated tracial moment sequence*.
3. If we restrict the support of the measure μ to a measurable closed set $K \subseteq \mathcal{S}^n$, we derive the *tracial K-moment problem*: For which sequences $y = (y_w)$ does there exist an $s \in \mathbb{N}$ and a probability measure μ supported in $K \cap (\mathcal{S}\mathbb{R}^{s\times s})^n$ such that for all $w \in \langle \underline{X} \rangle$ a representation (4.3) holds? Then we call y a *tracial K-moment sequence* or a tracial moment sequence with representing measure supported in K. The measure μ is also called a *K-representing measure* for y. Since we assume that $y_1 = 1$ the set K cannot be empty.

We defined the tracial moment problem in a seemingly more general way using integrals over measures on $(\mathcal{S}\mathbb{R}^{s\times s})$ as opposed to finitely atomic measures on $(\mathcal{S}\mathbb{R}^{s\times s})$ as is originally done in [BK1]. In the truncated case both definitions turn out to be equivalent, which will be shown in Theorem 4.23.

4.4 Remark.

1. The tracial moment problem is a natural extension of the classical moment problem in the following sense. Let y be a tracial moment sequence with a probability measure μ on real 1×1 matrices, and let α_w be the multi-index of the commutative collapse of w. Then we have for all $w \in \langle \underline{X} \rangle$
$$y_w = \int \operatorname{Tr}(w(\underline{a}))\, d\mu(\underline{a}) = \int w(\underline{a})\, d\mu(\underline{a}) = \int \underline{x}^{\alpha_w}\, d\mu(\underline{x}).$$
2. Definition 4.3 can be transferred to the complex case, where one considers tracial linear functionals $L_y : \mathbb{C}\langle \underline{X} \rangle \to \mathbb{C}$. The tracial sequences are then sequences of complex numbers satisfying (4.1) and $y_{w^*} = \overline{y_w}$ for all $w \in \langle \underline{X} \rangle$. Here $\overline{y_w}$ denotes the conjugate transpose of y_w. The complex tracial moment problem is then given by the question for which sequences does there exist a representation (4.3) with tuples $\underline{A} \in (\mathbb{C}^{s\times s})^n$ of hermitian matrices.
3. We assume in Definition 4.3 that there is a fixed matrix size s for the symmetric matrices in $\underline{A}$ in a representation (4.3). One can generalize this to the case that there is an upper bound $s \in \mathbb{N}$ for the matrix size of the tuples $\underline{A}$. In this case one has to define the evaluation of the empty word 1 in a different way. Assume that there is a tracial sequence y given by
$$y_w = \sum_{i=1}^{N} \lambda_i \operatorname{Tr}(w(\underline{A}^{(i)}))$$
with $\underline{A}^{(i)} \in (\mathcal{S}\mathbb{R}^{s_i \times s_i})^n$ for some $s_i \in \mathbb{N}$ with $s_i \leq s$ for all $i = 1, \ldots, N$. One can embed the matrices $A_j^{(i)} \in \mathcal{S}\mathbb{R}^{s_i \times s_i}$ into the vector space
$$\mathcal{S}_{s_i,s} := \{B \in \mathcal{S}\mathbb{R}^{s\times s} \mid B = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \text{ for some } A \in \mathcal{S}\mathbb{R}^{s_i \times s_i}\} \subseteq \mathcal{S}\mathbb{R}^{s\times s}$$

with identity $\mathbf{1}_{s_i,s} := \begin{bmatrix} \mathbf{1}_{s_i} & 0 \\ 0 & 0 \end{bmatrix}$. To obtain a representation

$$y_w = \sum_{i=1}^{N} \mu_i \operatorname{Tr}(w(\underline{B}^{(i)}))$$

for y with $\underline{B}^{(i)} \in (\mathcal{S}\mathbb{R}^{s\times s})^n$ we replace the tuples $\underline{A}^{(i)}$ by their images $\underline{B}^{(i)}$ in $\mathcal{S}^n_{s_i,s}$ and the weight λ_i by $\mu_i := \frac{s}{s_i}\lambda_i$. Further, we consider that the trace Tr acts as trace on $\mathcal{S}^n_{s_i,s}$ if we plug in the tuples $\underline{B}^{(i)}$ that we derived from the $\underline{A}^{(i)}$. Hence

$$\mu_i \operatorname{Tr}(w(\underline{B}^{(i)})) = \frac{1}{s_i}\lambda_i \operatorname{tr}(w(\underline{B}^{(i)})) = \frac{1}{s_i}\lambda_i \operatorname{tr}(w(\underline{A}^{(i)})) = \lambda_i \operatorname{Tr}(w(\underline{A}^{(i)}))$$

for all $w \neq 1$ and $\mu_i \operatorname{Tr}(1(\underline{B}^{(i)})) = \frac{s}{s_i}\lambda_i \operatorname{Tr}(\mathbf{1}_{s_i,s}) = \lambda_i = \lambda_i \operatorname{Tr}(1(\underline{A}^{(i)}))$.

This construction works similarly for a representation (4.3). For simplicity in notation we consider in this work that the matrices of all tuples $\underline{A}$ in a representation (4.3) are of the same size.

4.5 Remark. A natural extension of the tracial moment problem with respect to matrices would be to consider moments obtained by traces in finite *von Neumann algebras* as done for example by Hadwin [Had]. He proposed the moment problem for linear functionals on free powers of the C^*-algebra $C[-1,1]$. However, our primary motivation are trace-positive polynomials defined via traces of matrices (of finite size). Understanding these is one of the approaches to the BMV Conjecture 2.1 and Connes' embedding Conjecture 2.4. As we will show in Chapter 5, the notion dual to that of trace-positive polynomials is the tracial moment problem as defined above. Hence the characterization of trace-positive polynomials might be intimately connected with the tracial moment problem, in analogy to the commutative case. Therefore we do not generalize the tracial moment problem to the possible extension to moments obtained by traces in finite von Neumann algebras.

4.1.2 Tracial Hankel matrices and bilinear forms

Given a tracial sequence y and its Riesz functional L_y, the induced bilinear form on $\mathbb{R}\langle\underline{X}\rangle \times \mathbb{R}\langle\underline{X}\rangle$ is given by

$$(f,g) \mapsto L_y(f^*g).$$

It is represented by the tracial Hankel matrix of y with respect to a basis given by words $w \in \langle\underline{X}\rangle$.

4.6 Definition. The (infinite) *tracial Hankel matrix* $M(y)$ of a tracial sequence $y = (y_w)$ is defined by

$$M(y)_{u,v} := L(u^*v) = y_{u^*v}.$$

The tracial Hankel matrix of *order* k is the (finite) tracial Hankel matrix $M_k(y)$ that is defined similar as $M(y)$ but indexed by words u, v with $\deg u, \deg v \leq k$.

According to a truncated tracial sequence $(y_w)_{\leq 2k}$ we only have tracial Hankel matrices of order ℓ with $\ell \leq k$. These are defined as for infinite tracial sequences. In fact, if we truncate an infinite tracial sequence y to $(y_w)_{\leq 2k}$ we have $M_\ell(y) = M_\ell((y_w)_{\leq 2k})$ for all $\ell \leq k$.

Further, for a given symmetric polynomial $g = \sum_w g_w w \in \mathcal{S}\mathbb{R}\langle\underline{X}\rangle$ and a tracial sequence $y = (y_w)$ we define the *tracial localizing matrix* $M[gy]$ by

$$M[gy]_{u,v} := L_y(u^*gv) = \sum_w g_w y_{u^*wv}.$$

Since by definition $y_w = y_{w^*}$ for all $w \in \langle \underline{X} \rangle$ the tracial Hankel matrix is symmetric. The same holds true for the tracial localizing matrix since g is symmetric. Note that the sequence gy which one could try to define by $(gy)_w = \sum_v g_v y_{vw}$ in analogy to the commutative case is in general not a tracial sequence and thus would not lead to the tracial Hankel matrix $M(gy)$ as defined in Definition 4.6.

Recall that $\vec{p}$ is the vector of coefficients of p labelled by $w \in \langle \underline{X} \rangle$.

4.7 Example. We consider the truncated tracial sequence $y = (y_1, y_\mathsf{X}, y_{\mathsf{X}^2}, y_{\mathsf{X}^3}, y_{\mathsf{X}^4})$ which is given by $y := (1, 2, 4, 0, 3)$. Then the tracial Hankel matrix $M_2(y)$ of order 2 is with respect to the basis $b = (1, X, X^2)$ of the form

$$M_2(y) = \begin{bmatrix} y_1 & y_\mathsf{X} & y_{\mathsf{X}^2} \\ y_\mathsf{X} & y_{\mathsf{X}^2} & y_{\mathsf{X}^3} \\ y_{\mathsf{X}^2} & y_{\mathsf{X}^3} & y_{\mathsf{X}^4} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 0 \\ 4 & 0 & 3 \end{bmatrix}.$$

The bilinear form on $\mathbb{R}\langle \underline{X} \rangle_2 \times \mathbb{R}\langle \underline{X} \rangle_2$ given by $(f, g) \mapsto L_y(f^*g)$ is represented by $M_2(y)$ with respect to b. Let $f = f_1 + f_\mathsf{X} X + f_{\mathsf{X}^2} X^2$ and $g = g_1 + g_\mathsf{X} X + g_{\mathsf{X}^2} X^2$ in $\mathbb{R}\langle X \rangle_2$ be given. Then one easily checks that

$$\vec{f}^T M_2(y) \vec{g} = \begin{bmatrix} f_1 & f_\mathsf{X} & f_{\mathsf{X}^2} \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 0 \\ 4 & 0 & 3 \end{bmatrix} \begin{bmatrix} g_1 \\ g_\mathsf{X} \\ g_{\mathsf{X}^2} \end{bmatrix}$$

is equal to

$$L_y(f^*g) = f_1 g_1 + 2(f_1 g_\mathsf{X} + f_\mathsf{X} g_1) + 4(f_1 g_{\mathsf{X}^2} + f_\mathsf{X} g_\mathsf{X} + f_{\mathsf{X}^2} g_1) + 3 f_{\mathsf{X}^2} g_{\mathsf{X}^2}.$$

For tracial sequences y the relation between the bilinear form $(f, g) \mapsto L_y(f^*g)$ and $M(y)$ can be extended. The Riesz functional L_y itself can then be expressed by $M(y)$ in the way given by the next lemma, see also [BK1, Lemma 3.6].

4.8 Lemma. *Let y be a tracial sequence with associated Riesz functional L_y and tracial Hankel matrix $M = M(y)$. Then for all $p, q, g \in \mathbb{R}\langle \underline{X} \rangle$ the following holds:*

(1) $L_y(p^*q) = \vec{p}^T M \vec{q} = \vec{1}^T M \overrightarrow{p^*q}$ *which implies* $L_y(p) = \vec{1}^T M \vec{p}$.
In particular, $\vec{1}^T M \vec{p} = \vec{1}^T M \vec{q}$ *if* $p \overset{\text{cyc}}{\sim} q$.

(2) $L_y(p^*gq) = \vec{p}^T M(y) \overrightarrow{gq} = \overrightarrow{g^*p}^T M(y) \vec{q} = \vec{p}^T M[gy] \vec{q}$.

This is in perfect analogy to the commutative case. In fact, if one considers a tracial sequence $y = (y_w)$ which satisfies $y_u = y_v$ if $\check{u} = \check{v}$, then one obtains for all $p, q, g \in \mathbb{R}[\underline{x}]$ naturally the corresponding statement in the commutative context [Lau1, Lemma 4.1].

Proof. Let $p, q \in \mathbb{R}\langle \underline{X} \rangle$. Since $\vec{p}, \vec{q}$ have only finitely many entries $\neq 0$, the following calculations are well defined. In fact, one could also consider the finite tracial Hankel matrix $M_k(y)$ instead of $M(y)$ for $k := \max\{\deg p, \deg q\}$. Statement (1) follow by direct calculation. Indeed,

$$L_y(p^*q) = \sum_{u,v} p_u q_v y_{u^*v} = \vec{p}^T M(y) \vec{q}.$$

For statement (2) we use that $p^*(gq) = (g^*p)^*q$ which implies by (1) that

$$L_y(p^*gq) = \vec{p}^T M(y) \overrightarrow{gq} = \overrightarrow{g^*p}^T M(y) \vec{q} = \sum_{u,v,w} p_u g_v q_w y_{u^*vw}.$$

Further, we have

$$\vec{p}^T M[gy]\vec{q} = \sum_{u,w} p_u q_w (\sum_v g_v y_{u^*vw}) = \sum_{u,v,w} p_u g_v q_w y_{u^*vw}$$

which implies (2). ■

The tracial Hankel matrix $M = M(y)$ induces the linear map

$$\begin{aligned} \varphi_M : \mathbb{R}\langle \underline{X} \rangle &\to \operatorname{ran} M \\ p &\mapsto M\vec{p}. \end{aligned}$$

For the tracial Hankel matrices $M_k(y)$ of degree k we have a similar map φ_{M_k} which satisfies $\varphi_{M_k} = \varphi_M|_{\mathbb{R}\langle \underline{X} \rangle_k}$ if y is a full tracial sequence (and hence $M(y)$ is well defined).

Let $M = M(y)$ be fixed. The kernel of φ_M is then given by

$$I_M := \{p \in \mathbb{R}\langle \underline{X} \rangle \mid M\vec{p} = 0\} \tag{4.4}$$

and is isomorphic to the kernel of M. Hence the results on I_M can also be formulated for $\ker M$. The kernel I_M has some nice properties if M is positive semidefinite. In fact, it will turn out that in this case I_M is a two-sided ideal, see also [BK1, Prop. 3.7].

4.9 Proposition. *Let M be a positive semidefinite tracial Hankel matrix. Then*

$$I_M = \{p \in \mathbb{R}\langle \underline{X} \rangle \mid \langle M\vec{p}, \vec{p}\rangle = 0\}. \tag{4.5}$$

Further, I_M is a two-sided ideal of $\mathbb{R}\langle \underline{X} \rangle$ invariant under the involution $*$.

Proof. Let $J := \{p \in \mathbb{R}\langle \underline{X} \rangle \mid \langle M\vec{p}, \vec{p}\rangle = 0\}$. The implication $I \subseteq J$ is obvious. Let $p \in J$ be given and $k = \deg p$. Since M and thus its truncated tracial Hankel matrices M_k are positive semidefinite for each $k \in \mathbb{N}$, the square root $\sqrt{M_k}$ of M_k exists. Then

$$0 = \langle M_k\vec{p}, \vec{p}\rangle = \langle \sqrt{M_k}\vec{p}, \sqrt{M_k}\vec{p}\rangle$$

implies $\sqrt{M_k}\vec{p} = 0$. This leads to $M_k\vec{p} = M\vec{p} = 0$, thus $p \in I_M$.

To prove that I_M is a two-sided ideal, it suffices to show that I_M is a right-ideal which is closed under $*$. To do this, consider the bilinear map

$$\langle p, q\rangle_M := \langle M\vec{p}, \vec{q}\rangle$$

on $\mathbb{R}\langle \underline{X} \rangle$, which is a semi-scalar product. By Lemma 4.8, we get that

$$\langle pq, pq\rangle_M = ((pq)^*pq)(y) = (qq^*p^*p)(y) = \langle pqq^*, p\rangle_M.$$

Then by the Cauchy-Schwarz inequality it follows that for $p \in I_M$, we have

$$0 \leq \langle pq, pq\rangle_M^2 = \langle pqq^*, p\rangle_M^2 \leq \langle pqq^*, pqq^*\rangle_M \langle p, p\rangle_M = 0.$$

Hence $pq \in I_M$, i.e. I_M is a right-ideal. Since $p^*p \overset{\text{cyc}}{\sim} pp^*$, Lemma 4.8 implies

$$\langle M\vec{p}, \vec{p}\rangle = \langle p, p\rangle_M = (p^*p)(y) = (pp^*)(y) = \langle p^*, p^*\rangle_M = \langle M\vec{p}^*, \vec{p}^*\rangle.$$

Thus if $p \in I_M$ then also $p^* \in I_M$. ■

In the commutative context, the kernel of M is a real radical ideal if M is positive semidefinite as observed by Scheiderer, see [Lau2, p. 2974]. The next proposition, which can also be found in [BK1, Prop. 3.8], gives a description of I_M, or equivalently the kernel of M, in the non-commutative setting and could be helpful in defining a non-commutative real radical ideal.

4.10 Proposition. *Let M be a positive semidefinite tracial Hankel matrix. For the ideal I_M in* (4.4) *we have*

$$I_M = \{f \in \mathbb{R}\langle \underline{X}\rangle \mid (f^*f)^k \in I_M \text{ for some } k \in \mathbb{N}\}.$$

Further,

$$I_M = \{f \in \mathbb{R}\langle \underline{X}\rangle \mid (f^*f)^{2k} + \sum g_i^* g_i \in I_M \text{ for some } k \in \mathbb{N}, g_i \in \mathbb{R}\langle \underline{X}\rangle\}.$$

Proof. If $f \in I_M$ then also $f^*f \in I_M$ since I_M is an ideal. If $f^*f \in I_M$ we have $M\overrightarrow{f^*f} = 0$ which implies $f \in I_M$ as by Lemma 4.8 we get

$$0 = \vec{1}^T M\overrightarrow{f^*f} = \vec{f}^T M\vec{f} = \langle M\vec{f}, \vec{f}\rangle.$$

If $(f^*f)^k \in I_M$ then also $(f^*f)^{k+1} \in I_M$. Thus without loss of generality let k be even. From $(f^*f)^k \in I_M$ we obtain

$$0 = \vec{1}^T M\overrightarrow{(f^*f)^k} = \overrightarrow{(f^*f)^{k/2}}^T M\overrightarrow{(f^*f)^{k/2}},$$

implying $(f^*f)^{k/2} \in I_M$. This leads to $f \in I_M$ by induction.

To show the second statement let $(f^*f)^{2k} + \sum_i g_i^* g_i \in I_M$. This implies

$$0 = \vec{1}^T M\overrightarrow{(f^*f)^{2k} + \sum g_i^* g_i} = \vec{1}^T M\overrightarrow{(f^*f)^{2k}} + \sum_i \vec{1}^T M\overrightarrow{g_i^* g_i}$$

and Lemma 4.8 leads to

$$\overrightarrow{(f^*f)^k}^T M\overrightarrow{(f^*f)^k} + \sum_i \vec{g_i}^T M\vec{g_i} = 0.$$

Since $M(y) \succeq 0$ we have $\overrightarrow{(f^*f)^k}^T M\overrightarrow{(f^*f)^k} \geq 0$ and $\vec{g_i}^T M\vec{g_i} \geq 0$. Thus each summand has to be zero, in particular

$$\overrightarrow{(f^*f)^k}^T M\overrightarrow{(f^*f)^k} = 0$$

which implies $f \in I_M$ as in the first statement. ∎

4.1.3 Necessary conditions for tracial moment sequences

If a tracial sequence y has a representation (4.3) it satisfies some additional properties like positive semidefiniteness of the tracial Hankel matrix $M(y)$. These necessary conditions on a (truncated) tracial sequence y to be a tracial moment sequence are the topic of this section. The specific case of (commutative) moment sequences is indirectly included in the following. In fact, one derives from the following statements naturally several well-known necessary conditions for moment sequences.

The *zero-set* of a given set $P \subseteq \mathbb{R}\langle \underline{X}\rangle$ of polynomials is defined as

$$V(P) = \{\underline{A} \in \mathcal{S}^n \mid p(\underline{A}) = 0 \text{ for all } p \in P\}.$$

We want to bound the matrix size of the zeros of $p \in P$, therefore we set

$$V_s(P) = V(P) \cap (\mathcal{S}\mathbb{R}^{s\times s})^n.$$

4.11 Proposition. *Let $y = (y_w)_{\le 2k}$ be a truncated tracial moment sequence with representing measure μ on $(\mathcal{S}\mathbb{R}^{s\times s})^n$ for some $s \in \mathbb{N}$. Then the following statements hold.*

(i) $M_k(y) \succeq 0$,

(ii) $\operatorname{supp}\mu \subseteq V_s(I_{M_k(y)})$ *and*

(iii) $\operatorname{rank} M_k(y) \le |\operatorname{supp}\mu| s^2$.

The equivalent necessary conditions for commutative moment sequences are included in Proposition 4.11 since a moment sequence has a representing measure on $\mathbb{R}^n = (\mathbb{R}^{1\times 1})^n$. In fact, for $s = 1$ we derive [Lau1, Lemma 4.2.(i)].

Proof. Let $p \in \mathbb{R}\langle \underline{X}\rangle_k$. Then

$$\vec{p}^T M_k(y)\vec{p} = \sum_{u,v} p_u p_v y_{u^*v} = \int \operatorname{Tr}(p^* p(\underline{A}))\, d\mu(\underline{A}) \ge 0$$

since hermitian squares are trace-positive. Thus $M_k(y) \succeq 0$.

For statement (ii) let $q \in I_{M_k(y)} \setminus \{0\}$, hence $M_k(y)\vec{q} = 0$. Let $C_q := (\mathcal{S}\mathbb{R}^{s\times s})^n \setminus V_s(q)$. Since $q \ne 0$ we have $C_q \ne \varnothing$. We will show that $\mu(C_q) = 0$. For this, let $\|\_\|$ denote the operator norm and let

$$U_\ell := \{\underline{A} \in (\mathcal{S}\mathbb{R}^{s\times s})^n \mid \|q^* q(\underline{A})\| \ge \frac{1}{\ell}\}$$

for $\ell \in \mathbb{N}$. Then $C_q = \bigcup_\ell U_\ell$. Therefore, since $q \in I_{M_k(y)}$ implies $\vec{q}^T M_k(y)\vec{q} = 0$ by Proposition 4.9, and since $\operatorname{tr}(A) \ge \|A\|$ for any positive semidefinite matrix A, we get

$$\begin{aligned} 0 = \vec{q}^T M_k(y)\vec{q} &= \int \operatorname{Tr}(q^* q)\, d\mu \\ &= \int_{C_q} \operatorname{Tr}(q^* q)\, d\mu \ge \int_{U_\ell} \operatorname{Tr}(q^* q)\, d\mu \ge \frac{1}{s\ell}\,\mu(U_\ell). \end{aligned}$$

This implies $\mu(U_\ell) = 0$ for all $\ell \ge 1$ and hence $\mu(C_q) = 0$. Thus $\operatorname{supp}\mu \subseteq V_s(q)$ for all $q \in I_{M_k(y)}$, which shows $\operatorname{supp}\mu \subseteq V_s(I_{M_k(y)})$.

Statement (iii) is clear if μ has infinite support. Therefore we assume y has a finite representation

$$y_w = \sum_{i=1}^{|\operatorname{supp}\mu|} \lambda_i \operatorname{Tr}(w(\underline{A}^{(i)}))$$

for some $\lambda_i \in \mathbb{R}_{\ge 0}$ with $\sum_i \lambda_i = 1$ and $\underline{A}^{(i)} \in (\mathcal{S}\mathbb{R}^{s\times s})^n$. For the statement it suffices to consider the truncated tracial moment sequence $y^{(i)}$ given by $y^{(i)}_w = \operatorname{Tr}(w(\underline{A}^{(i)}))$ for $w \in \langle \underline{X}\rangle$ with $\deg w \le 2k$ and to show that

$$\operatorname{rank} M_k(y^{(i)}) \le s^2.$$

Then the convex combination $M_k(y) = \sum_i \lambda_i M_k(y^{(i)})$ has rank at most $|\operatorname{supp}\mu| s^2$. The truncated tracial Hankel matrix $M := M_k(y^{(i)})$ induces a linear map

$$\Phi : \mathbb{R}\langle \underline{X}\rangle_k \to \mathbb{R}\langle \underline{X}\rangle_k^*, \quad p \mapsto \Big(q \mapsto \operatorname{Tr}\big((q^* p)(\underline{A}^{(i)})\big)\Big),$$

where $\mathbb{R}\langle \underline{X}\rangle_k^*$ is the algebraic dual space of $\mathbb{R}\langle \underline{X}\rangle_k$. This implies

$$\operatorname{rank} M = \dim(\operatorname{ran}\Phi) = \dim(\mathbb{R}\langle \underline{X}\rangle_k / \ker\Phi).$$

The kernel of the evaluation map $\varepsilon_{\underline{A}^{(i)}} : \mathbb{R}\langle \underline{X}\rangle_k \to \mathbb{R}^{s\times s}$, $p \mapsto p(\underline{A}^{(i)})$ is a subset of $\ker\Phi$. In particular,

$$\dim(\mathbb{R}\langle \underline{X}\rangle_k / \ker\Phi) \le \dim(\mathbb{R}\langle \underline{X}\rangle_k / \ker\varepsilon_{\underline{A}^{(i)}}) = \dim(\operatorname{ran}\varepsilon_{\underline{A}^{(i)}}) \le s^2. \qquad \blacksquare$$

For a tracial moment sequence y, Proposition 4.11 holds true for all $k \in \mathbb{N}$. We can derive the following proposition with the same line of reasoning as in Proposition 4.11. For $s = 1$ we obtain the corresponding statements for the commutative context [Lau1, Lemma 4.2.(iii)].

4.12 Proposition. *Let y be a tracial moment sequence with representing measure μ on $(\mathcal{S}\mathbb{R}^{s\times s})^n$. Then $M(y) \succeq 0$ and $\operatorname{supp}\mu \subseteq V_s(I_{M(y)})$. Furthermore, $\operatorname{rank} M(y) \leq |\operatorname{supp}\mu| s^2$. In particular, if μ is a representing measure with finite support, the tracial Hankel matrix $M(y)$ is of finite rank.*

If we consider the tracial K-moment problem and restrict the support of the representing measure μ to be contained in a set K, we derive an additional condition on y concerning localizing matrices, which provides for $s = 1$ the well-known necessary conditions for K-moment sequences with $K \subseteq \mathbb{R}^n$, see [Lau1, Chapter 4].

4.13 Proposition. *Let y be a tracial moment sequence with representing measure μ supported in K, where $K = \{\underline{A} \in (\mathcal{S}\mathbb{R}^{s\times s})^n \mid g(\underline{A}) \succeq 0\}$ for some $g \in \mathbb{R}\langle \underline{X}\rangle$ of degree d_g and some $s \in \mathbb{N}$. Then $M[gy] \succeq 0$. In particular, $M_{k-d_g}[gy] \succeq 0$ for any $k \geq d_g$ considering y as a truncated tracial moment sequence of degree $2k$.*

Proof. Let y be a tracial moment sequence with representing measure μ as in the assumption and let $p \in \mathbb{R}\langle \underline{X}\rangle$. Then, since $pp^*(\underline{A})$ and $g(\underline{A})$ are positive semidefinite for all $\underline{A} \in K$, we obtain

$$\vec{p}^{\,T} M[gy]\vec{p} = \sum_{u,w} p_u p_w \Big(\sum_v g_v y_{u^*vw}\Big) = \int \operatorname{Tr}(p^* g p(\underline{A}))\, d\mu(\underline{A}) \geq 0$$

Hence $M[gy] \succeq 0$. If we have a truncated tracial sequence $(y_w)_{\leq 2k}$ we can define the tracial localizing matrix $M_\ell[gy]$ only for $\ell \leq k - d_g$. The proof of positive-semidefiniteness then works the same way. ∎

Here are two easy examples showing that the necessary condition of positive semidefiniteness of the tracial Hankel matrix $M_k(y)$ is not sufficient in general for a truncated tracial sequence $y = (y_w)_{\leq 2k}$ to have a representation (4.3). The first example is a classical example for sequences in the commutative context, see for instance [CF3, Example 2.1]. The second example is an extended version of [BK1, Example 3.5].

4.14 Example.

(a) We consider the truncated (tracial) sequence $y = (y_1, y_{\mathtt{X}}, y_{\mathtt{X}^2}, y_{\mathtt{X}^3}, y_{\mathtt{X}^4})$ which is given by $y := (1, 1, 1, 1, 2)$. Then the tracial Hankel matrix

$$M_2(y) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

with respect to the basis $(1, X, X^2)$ is positive semidefinite, but y does not have a tracial representation. This follows from Proposition 4.11(ii). One easily sees $I_{M_2} = \operatorname{span}\{1 - X\}$. Hence for any $s \in \mathbb{N}$ we have $\operatorname{supp}\mu \subseteq \{\mathbf{1}_s\}$. However the tracial sequence $y^{\mathbf{1}_s}$, which is given by $y_w = \operatorname{Tr}(w(\mathbf{1}_s))$, does not satisfy $y_{\mathtt{X}^4} = 2$. Hence for any matrix size $s \in \mathbb{N}$ we cannot find a tracial representation for y.

(b) To obtain a compact description of the truncated tracial sequence y we fix the order of words of degree 4 in two variables in the index set as follows:

$$1, X, Y, X^2, XY, Y^2, X^3, X^2Y, XY^2, Y^3, X^4, X^3Y, X^2Y^2, XYXY, XY^3, Y^4.$$

The sequence y will be expressed as $y = (y_{w_1}, y_{w_2}, \dots, y_{w_{16}})$ where w_i is the i-th word in our list of words in the fixed order. Using this description the truncated tracial sequence

$$y = (1, 0, 0, 1, 1, 1, 0, 0, 0, 0, 4, 0, 2, 1, 0, 4)$$

yields the positive semidefinite tracial Hankel matrix

$$M := M_2(y) = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 4 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 2 & 0 \\ 1 & 0 & 0 & 2 & 0 & 0 & 4 \end{bmatrix}$$

with respect to the basis $(1, X, Y, X^2, XY, YX, Y^2)$. However y does not have a tracial representation.

One easily obtains, that there cannot be a (tracial) representation of y on 1×1 matrices. In fact $V_1(I_M) = \{(-1, -1), (1, 1)\}$. Hence by Proposition 4.11(iii), if y is a moment sequence then $\operatorname{rank} M = 5$ should be bounded by $|V_1(I_M)| = 2$, which is not possible. On the other hand, $|V_2(I_M)| = 4$, thus $\operatorname{rank} M = 5 \leq 4\,|V_2(I_M)| = 16$ is satisfied. Therefore we need another argument to show that there is not a tracial representation for y.

We will see later (Theorem 4.23) that each truncated tracial sequence which has a representation (4.3) has also a representation with finite support. Therefore we assume that y has a representation

$$y_w = \sum_{i=1}^{N} \lambda_i \operatorname{Tr}(w(A^{(i)}, B^{(i)}))$$

for some symmetric matrices $A^{(i)}, B^{(i)} \in \mathcal{S}\mathbb{R}^{s\times s}$ and $\lambda_i \in \mathbb{R}_{>0}$ with $\sum_i \lambda_i = 1$. Setting

$$T^{(i)} := \left[\operatorname{Tr}(u^* v(A^{(i)}, B^{(i)}))\right]_{u,v}$$

we have $M_2(y) = \sum_{i=1}^{N} \lambda_i T^{(i)}$. Each $T^{(i)}$ is positive semidefinite, thus in particular for all $i = 1, \dots, N$ we have $T^{(i)}_{22} = T^{(i)}_{33} = T^{(i)}_{23} =: t_i$. From

$$\begin{aligned} \frac{1}{s^2}\langle A^{(i)}, A^{(i)}\rangle\langle B^{(i)}, B^{(i)}\rangle &= \operatorname{Tr}({A^{(i)}}^2)\operatorname{Tr}({B^{(i)}}^2) \\ &= t_i^2 \\ &= (\operatorname{Tr}(A^{(i)}B^{(i)}))^2 = \frac{1}{s^2}\langle A^{(i)}, B^{(i)}\rangle^2 \end{aligned}$$

we obtain by the Cauchy-Schwarz inequality that $A^{(i)} = \alpha_i B^{(i)}$ for some $\alpha_i \in \mathbb{R}$ for all $i = 1, \dots, N$. This leads to the contradiction

$$\begin{aligned} 2 = M_2(y)_{55} &= \sum_i \lambda_i \operatorname{Tr}({A^{(i)}}^2 {B^{(i)}}^2) \\ &= \sum_i \lambda_i \alpha_i^2 \operatorname{Tr}({B^{(i)}}^4) \\ &= \sum_i \lambda_i \operatorname{Tr}(A^{(i)}B^{(i)}A^{(i)}B^{(i)}) = M_2(y)_{45} = 1. \end{aligned}$$

4.2 The full tracial moment problem

This section deals with the full moment problem, i.e. we consider infinite tracial sequences y. First, we will show an auxiliary proposition which characterizes tracial states on matrix $*$-algebras. This will enable us to prove the existence of a tracial moment representation for tracial sequences with a positive semidefinite tracial Hankel matrix of finite rank, which is the tracial analog of the result of Curto and Fialkow [CF1, Theorem 4.7] on sequences with positive semidefinite Hankel matrices of finite rank.

4.2.1 A tracial representation theorem

In this section we shall characterize tracial states on matrix $*$-algebras with the aid of the Artin-Wedderburn theorem and the Riesz representation theorem for positive linear functionals on a finite-dimensional Hilbert space H, stating that H is isometrically isomorphic to its algebraic dual space H^* via $x \mapsto (y \mapsto \langle y, x\rangle)$.

4.15 Remark. The only central simple algebras over $\mathbb{R}$ are the full matrix algebras over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$. This follows by combining the Frobenius theorem [Lam, (13.12)] with the Artin-Wedderburn theorem [Lam, (3.5)]. In order to understand ($\mathbb{R}$-linear) tracial states, see Definition 1.21, on these, we recall some basic Galois theory. Let

$$\mathrm{Trd}_{\mathbb{C}/\mathbb{R}} : \mathbb{C} \to \mathbb{R}, \quad z \mapsto \frac{1}{2}(z + \bar{z})$$

denote the *field trace* and

$$\mathrm{Trd}_{\mathbb{H}/\mathbb{R}} : \mathbb{H} \to \mathbb{R}, \quad z \mapsto \frac{1}{2}(z + \bar{z})$$

the *reduced trace* [KMRT, p. 5]. Here the Hamilton quaternions $\mathbb{H}$ are endowed with the *standard involution*

$$z = a + \mathbb{i}b + \mathbb{j}c + \mathbb{k}d \mapsto a - \mathbb{i}b - \mathbb{j}k - \mathbb{k}d = \bar{z}$$

for $a, b, c, d \in \mathbb{R}$. We extend the canonical involution on $\mathbb{C}$ and $\mathbb{H}$ to the conjugate transpose involution * on matrices over $\mathbb{C}$ and $\mathbb{H}$, respectively.

Composing the field trace and reduced trace, respectively, with the normalized trace, yields an $\mathbb{R}$-linear map from $\mathbb{C}^{s\times s}$ and $\mathbb{H}^{s\times s}$, respectively, to $\mathbb{R}$. We will denote it simply by Tr. A word of caution: In this context $\mathrm{Tr}(A)$ does not denote the normalized matricial trace over $\mathbb{K}$ if $A \in \mathbb{K}^{s\times s}$ and $\mathbb{K} \in \{\mathbb{C}, \mathbb{H}\}$.

An alternative description is given by the following lemma [BK1, Lemma 3.11].

4.16 Lemma. *Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$. Then the only ($\mathbb{R}$-linear) tracial state on $\mathbb{K}^{s\times s}$ is* Tr.

Proof. An easy calculation shows that Tr is indeed a tracial state.

Let L be a tracial state on $\mathbb{R}^{s\times s}$. By the Riesz representation theorem there exists a positive semidefinite matrix $B \in \mathbb{R}^{s\times s}$ with $\mathrm{Tr}(B) = 1$ such that

$$L(A) = \mathrm{Tr}(BA)$$

for all $A \in \mathbb{R}^{s\times s}$. Write $B = \begin{bmatrix} b_{ij} \end{bmatrix}$. Let $i \neq j$. The matrix $A = \lambda E_{ij}$ has zero trace for every $\lambda \in \mathbb{R}$ and is thus a sum of commutators. (Here E_{ij} denotes the $s \times s$ matrix unit with a one in the (i, j)-position and zeros elsewhere.) Hence

$$\lambda b_{ij} = L(A) = 0.$$

Since $\lambda \in \mathbb{R}$ was arbitrary, $b_{ij} = 0$. Now let $A = \lambda(E_{ii} - E_{jj})$. Clearly, $\operatorname{Tr}(A) = 0$ and hence

$$\lambda(b_{ii} - b_{jj}) = L(A) = 0.$$

As before, this gives $b_{ii} = b_{jj}$. So B is scalar, and $\operatorname{Tr}(B) = 1$. Hence it is the identity matrix. In particular, $L = \operatorname{Tr}$.

If L is a tracial state on $\mathbb{C}^{s\times s}$, then L induces a tracial state on $\mathbb{R}^{s\times s}$, so $L_0 := L|_{\mathbb{R}^{s\times s}} = \operatorname{Tr}$ by the above. Extend L_0 to

$$L_1 : \mathbb{C}^{s\times s} \to \mathbb{R}, \quad A + \mathbb{i}B \mapsto L_0(A) = \operatorname{Tr}(A) \quad \text{for } A, B \in \mathbb{R}^{s\times s}.$$

L_1 is a tracial state on $\mathbb{C}^{s\times s}$ as a straightforward computation shows. As $\operatorname{Tr}(A) = \operatorname{Tr}(A + \mathbb{i}B)$, all we need to show is that $L_1 = L$.

Clearly, L_1 and L agree on the vector space spanned by all commutators in $\mathbb{C}^{s\times s}$. This space is (over $\mathbb{R}$) of co-dimension 2. By construction, $L_1(1) = L(1) = 1$ and $L_1(\mathbb{i}) = 0$. On the other hand,

$$L(\mathbb{i}) = L(\mathbb{i}^*) = -L(\mathbb{i})$$

implying $L(\mathbb{i}) = 0$. This shows $L = L_1 = \operatorname{Tr}$.

The remaining case of tracial states over $\mathbb{H}$ is dealt similarly. ∎

4.17 Remark. Every complex number $z = a + \mathbb{i}b$ can be represented as a 2×2 real matrix

$$z' = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

This gives rise to an $\mathbb{R}$-linear $*$-map $\mathbb{C}^{s\times s} \to \mathbb{R}^{(2s)\times(2s)}$ that commutes with Tr. A similar property holds if quaternions $a + \mathbb{i}b + \mathbb{j}c + \mathbb{k}d$ are represented by the 4×4 real matrix

$$\begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{bmatrix}.$$

Now we are ready to prove the proposed characterization for tracial states on matrix-$*$-algebras. The following proposition can also be found in [BK1, Prop. 3.13].

4.18 Proposition. *Let $\mathcal{A}$ be a $*$-subalgebra of $\mathbb{R}^{s\times s}$ for some $s \in \mathbb{N}$ and $L : \mathcal{A} \to \mathbb{R}$ a tracial state. Then there exist full matrix algebras $\mathcal{A}^{(i)}$ over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$, a $*$-isomorphism*

$$\mathcal{A} \to \bigoplus_{i=1}^{N} \mathcal{A}^{(i)}, \tag{4.6}$$

and $\lambda_1, \ldots, \lambda_N \in \mathbb{R}_{\geq 0}$ with $\sum_i \lambda_i = 1$, such that for all $A \in \mathcal{A}$,

$$L(A) = \sum_{i=1}^{N} \lambda_i \operatorname{Tr}(A^{(i)}), \tag{4.7}$$

where the $A^{(i)}$ come from $\bigoplus_i A^{(i)}$, the image of A under the isomorphism (4.6). The size of (the real representation of) $\bigoplus_i A^{(i)}$ is at most s.

Proof. Since L is tracial, $L(U^*AU) = L(A)$ for all orthogonal $U \in \mathbb{R}^{s\times s}$. Hence we can apply orthogonal transformations to $\mathcal{A}$ without changing the values of L. So $\mathcal{A}$ can be transformed into block diagonal form as in (4.6) according to its invariant subspaces. That is, each of the blocks $\mathcal{A}^{(i)}$ acts irreducibly on a subspace of $\mathbb{R}^s$ and is thus a central simple algebra (with involution) over $\mathbb{R}$. The involution on $\mathcal{A}^{(i)}$ is induced by the conjugate transpose involution. (Equivalently, by the transpose on the real matrix representation in the complex or quaternion case.)

Now L induces (after a possible normalization) a tracial state on the block $\mathcal{A}^{(i)}$ and hence by Lemma 4.16, we have $L_i := L|_{\mathcal{A}^{(i)}} = \lambda_i \operatorname{Tr}$ for some $\lambda_i \in \mathbb{R}_{\geq 0}$. Then

$$L(A) = L\big(\bigoplus_i A^{(i)}\big) = \sum_i L_i\big(A^{(i)}\big) = \sum_i \lambda_i \operatorname{Tr}\big(A^{(i)}\big)$$

and $1 = L(1) = \sum_i \lambda_i$. ∎

4.2.2 Tracial Hankel matrices of finite rank

We will prove the tracial version of the theorem of Curto and Fialkow for positive semidefinite Hankel matrices of finite rank [CF1, Theorem 4.7], see also [Lau1, Theorem 5.1]. The following theorem, in combination with Proposition 4.12, characterizes the infinite tracial sequences having a representing measure with finite support. It can also be found in [BK1, Theorem 3.14].

4.19 Theorem. *Let y be a tracial sequence with positive semidefinite tracial Hankel matrix $M(y)$ of finite rank s. Then y is a tracial moment sequence with finite support, i.e. there exist an $N \in \mathbb{N}_0$ and tuples $\underline{A}^{(i)} = (A_1^{(i)}, \dots, A_n^{(i)})$ of symmetric matrices $A_j^{(i)}$ of size at most s and $\lambda_i \in \mathbb{R}_{\geq 0}$ with $\sum_i \lambda_i = 1$ such that for all $w \in \langle \underline{X} \rangle$,*

$$y_w = \sum_{i=1}^{N} \lambda_i \operatorname{Tr}(w(\underline{A}^{(i)})).$$

The proof works with a similar line of reasoning as in [Lau1, Theorem 5.1].

Proof. Let $M := M(y)$. We equip $\mathbb{R}\langle \underline{X} \rangle$ with the bilinear form given by

$$\langle p, q\rangle_M := \langle M\vec{p}, \vec{q}\rangle = \vec{q}^{\,T} M \vec{p}.$$

Let $I := I_M = \{p \in \mathbb{R}\langle \underline{X} \rangle \mid \langle p, p\rangle_M = 0\}$. Then by Proposition 4.9, I is an ideal of $\mathbb{R}\langle \underline{X} \rangle$, in fact, $I = \ker \varphi_M$ for

$$\varphi_M : \mathbb{R}\langle \underline{X} \rangle \to \operatorname{ran} M, \quad p \mapsto M\vec{p}.$$

Thus if we define $E := \mathbb{R}\langle \underline{X} \rangle / I$, the induced linear map

$$\overline{\varphi}_M : E \to \operatorname{ran} M, \quad \overline{p} \mapsto M\vec{p}$$

is an isomorphism and

$$\dim E = \dim(\operatorname{ran} M) = \operatorname{rank} M = s < \infty.$$

Hence $(E, \langle \_, \_ \rangle_E)$ is a finite-dimensional Hilbert space for $\langle \overline{p}, \overline{q}\rangle_E = \vec{q}^{\,T} M \vec{p}$.

Let $\hat{X}_i$ be the right multiplication with X_i on E, i.e. $\hat{X}_i \overline{p} := \overline{pX_i}$. Since I is a right ideal of $\mathbb{R}\langle \underline{X} \rangle$, the operator $\hat{X}_i$ is well defined. Further, $\hat{X}_i$ is symmetric since

$$\begin{aligned}\langle \hat{X}_i \overline{p}, \overline{q}\rangle_E &= \langle M\overrightarrow{pX_i}, \vec{q}\rangle = (X_i p^* q)(y) \\ &= (p^* q X_i)(y) = \langle M\vec{p}, \overrightarrow{qX_i}\rangle = \langle \overline{p}, \hat{X}_i \overline{q}\rangle_E.\end{aligned}$$

Thus each $\hat{X}_i$, acting on an s-dimensional vector space, has a representation matrix $A_i \in \mathcal{S}\mathbb{R}^{s\times s}$. Let $\mathcal{B} = B(\hat{X}_1,\ldots,\hat{X}_n) = B(A_1,\ldots,A_n)$ be the operator algebra generated by $\hat{X}_1,\ldots,\hat{X}_n$. These operators can be written as

$$\hat{p} = \sum_{w\in\langle\underline{X}\rangle} p_w \hat{w}$$

for some $p_w \in \mathbb{R}$, where $\hat{w} = \hat{X}_{w_1}\cdots\hat{X}_{w_s}$ for $w = X_{w_1}\cdots X_{w_s}$. Observe that $\hat{w}$ is similar to $w(A_1,\ldots,A_n)$. We define the linear functional

$$\begin{aligned} L : \mathcal{B} &\to \mathbb{R} \\ \hat{p} &\mapsto \vec{1}^T M \vec{p} = L_y(p), \end{aligned}$$

which is a state on $\mathcal{B}$. Since $y_w = y_u$ for $w \overset{\text{cyc}}{\sim} u$, it follows that L is tracial. Thus by Proposition 4.18 (and Remark 4.17), there exist $\lambda_1,\ldots,\lambda_N \in \mathbb{R}_{\geq 0}$ with $\sum_i \lambda_i = 1$ and real symmetric matrices $A_j^{(i)}$ $(i = 1,\ldots,N)$ for each $A_j \in \mathcal{S}\mathbb{R}^{s\times s}$, such that for all $w \in \langle\underline{X}\rangle$,

$$y_w = L(\hat{w}) = \sum_i \lambda_i \operatorname{Tr}(w(\underline{A}^{(i)})),$$

as desired. ∎

By Proposition 4.12, we have that $\operatorname{supp}\mu \subseteq V_s(I_{M(y)})$, hence the tuples $\underline{A}^{(i)}$ satisfy $p(\underline{A}) = 0$ for all $p \in I_{M(y)}$. Further, the converse of Theorem 4.19 also holds true by Proposition 4.12. If y has a tracial moment representation with finite support, i.e. $y_w = \sum_{i=1}^N \lambda_i \operatorname{Tr}(w(\underline{A}^{(i)}))$, then $M(y)$ has finite rank, see also [BK1, Corollary 3.15].

4.3 The truncated tracial moment problem

In this section we investigate the truncated tracial moment problem. This is motivated by the fact that the truncated tracial moment problem is more general than the full tracial moment problem as explained in Section 4.3.1. In fact, we present the tracial analog of the result of Stochel [Sto, Theorem 4] for sequences in the context of commuting variables. Theorem 4.19 characterizes infinite tracial sequences which have a representing measure with finite support. For a truncated tracial sequence the existence of a representing measure implies the existence of another representing measure with finite support which is the tracial analogy of a result due to Bayer and Teichmann [BT, Theorem 2]. In other words, for truncated tracial sequences we can replace the representing measure μ by a cubature formula. This will be shown in Section 4.3.2. Then we present a tracial analog of the Riesz-Haviland theorem for the truncated moment problem as proposed by Curto and Fialkow [CF3] involving results on positive extensions of tracial Riesz functionals. Moreover, we show that truncated tracial sequences with strictly positive Riesz functionals are tracial moment sequences which is in analogy to results of Fialkow and Nie [FN]. We close this section by proving the existence of representing measures for truncated tracial sequences whose positive semidefinite tracial Hankel matrix admits a flat extension.

4.3.1 A variant of Stochel's theorem

The truncated tracial moment problem is more general than the full tracial moment problem in the sense explained below. This result, which can also be found in [BCKP, Theorem 3.6], is in analogy to a result of Stochel [Sto, Theorem 4] for commutative sequences. This classical result follows directly if we set $s = 1$.

4.20 Theorem. *Suppose y is a tracial sequence and K a closed subset of $(\mathcal{S}\mathbb{R}^{s\times s})^n$ for some $s \in \mathbb{N}$. If there is for all $k \in \mathbb{N}$ a measure μ_k supported in K such that*

$$y_w = \int \operatorname{Tr}(w(\underline{A}))\, d\mu_k(\underline{A})$$

for all $w \in \langle \underline{X} \rangle_k$, then y is a tracial moment sequence. Furthermore, there exists a representing measure supported in K.

For the proof we refer to Section 1.6 for notation concerning measures and the spaces of continuous functions vanishing at infinity. We start by a preliminary lemma.

4.21 Lemma. *For $u \in \langle \underline{X} \rangle$ the map $\varphi_u : K \to \mathbb{R}$ defined by*

$$\varphi_u(\underline{A}) := \frac{\operatorname{Tr}\big(u(\underline{A})\big)}{1 + \sum_{i=1}^n \operatorname{Tr}(A_i^{2\deg(u)+2})}$$

lies in $\mathcal{C}_0(K)$.

Proof. Let $u \in \mathbb{R}\langle \underline{X} \rangle$ be fixed with $\deg(u) := d$. If K is compact, we are done since φ_u is continuous. If K is not compact, let $\underline{A} \in K$ be such that $\sum_{i=1}^n \operatorname{Tr}(A_i^2) > \ell^2$ for some $\ell \in \mathbb{N}$. Choose the index i_A such that $\operatorname{Tr}(A_{i_A}^2) \geq \operatorname{Tr}(A_i^2)$ for all $i = 1, \ldots, n$. Then

$$\operatorname{Tr}(A_{i_A}^2) \geq \frac{\sum_i \operatorname{Tr}(A_i^2)}{n} > \frac{\ell^2}{n}.$$

Since the matrices A_i^2 are positive semidefinite we have $\operatorname{Tr}(A_i^{2d+2}) = \|A_i^2\|_{d+1}^{d+1}$, where $\|\llcorner\!\lrcorner\|_p$ denotes the normalized p-Schatten norm on $\mathcal{S}\mathbb{R}^{s\times s}$, which generalizes the Hilbert-Schmidt norm ($p = 2$) and is given by

$$\|T\|_p^p = \operatorname{Tr}(|T|^p) \text{ with } |T| = \sqrt{T^2} \text{ for all } T \in \mathcal{S}\mathbb{R}^{s\times s}.$$

Since $\mathcal{S}\mathbb{R}^{s\times s}$ is finite dimensional, the $(d+1)$-Schatten norm is equivalent to the 1-Schatten norm, also known as trace-norm, on $\mathcal{S}\mathbb{R}^{s\times s}$. Hence there is a $c \in \mathbb{R}_{>0}$ such that

$$c \operatorname{Tr}(A_i^2)^{d+1} = c\|A_i^2\|_1^{d+1} \leq \|A_i^2\|_{d+1}^{d+1} = \operatorname{Tr}(A_i^{2d+2})$$

for all $A_i \in \mathcal{S}\mathbb{R}^{s\times s}$. Further, for the numerator of φ_u we have

$$(\operatorname{Tr}(u(\underline{A})))^2 \leq s^{d-2} u(\operatorname{Tr}(A_1^2), \ldots, \operatorname{Tr}(A_n^2)) \leq s^{d-2} (\operatorname{Tr}(A_{i_A}^2))^d$$

by the Cauchy-Schwarz inequality. All together this implies

$$\begin{aligned}
\varphi_u(\underline{A})^2 = \frac{\big(\operatorname{Tr}(u(\underline{A}))\big)^2}{\big(1 + \sum_{i=1}^n \operatorname{Tr}(A_i^{2d+2})\big)^2} &\leq \frac{s^{d-2}(\operatorname{Tr}(A_{i_A}^2))^d}{\big(1 + \sum_{i=1}^n \operatorname{Tr}(A_i^{2d+2})\big)^2} \\
&\leq \frac{s^{d-2}(\operatorname{Tr}(A_{i_A}^2))^d}{\big(1 + c\sum_{i=1}^n (\operatorname{Tr}(A_i^2))^{d+1}\big)^2} < \frac{s^{d-2}(\operatorname{Tr}(A_{i_A}^2))^d}{c^2(\operatorname{Tr}(A_{i_A}^2))^{2d+2}} \\
&\leq \frac{s^{d-2}}{c^2 \operatorname{Tr}(A_{i_A}^2)^{d+2}} < \frac{s^{d-2} n^{d+2}}{c^2 \ell^{2d+4}}
\end{aligned}$$

which goes to zero for large ℓ. Hence $\varphi_u \in \mathcal{C}_0(K)$. ∎

With Lemma 4.21 we are able to show Theorem 4.20. The proof works with a similar line of reasoning as the classical statement [Sto, Theorem 4] but uses different functions.

Proof of Theorem 4.20. Let $\mathcal{C}_0(K)$ be denoted with the supremum norm $\|\llcorner\!\lrcorner\|_\infty$. For each $k \in \mathbb{N}$ let $\widehat{\mu}_k$ be the linear functional associated to μ_k, cf. Section 1.6. Due to our normalization of tracial sequences, for all $k \in \mathbb{N}$ we have

$$|\widehat{\mu}_k(f)| \leq \int \|f\|_\infty \, d\mu_k = \|f\|_\infty \text{ for all } f \in \mathcal{C}_0(K)$$

so all the $\widehat{\mu}_k$ belong to $\mathbb{B}$, the closed unit ball in the dual space $\mathcal{C}_0(K)'$. By the Banach-Alaoglu theorem [Ber, (44.12)], there is a subsequence $(\widehat{\mu}_{k_\ell})_\ell$ of $(\widehat{\mu}_k)_k$ converging in the weak-$*$-topology to some $\psi \in \mathbb{B}$. For simplicity of notation, we omit the subindex ℓ in the sequel and assume that $(\widehat{\mu}_k)_k$ converges to ψ. If $f \in \mathcal{C}_0(K)$ and $f \geq 0$, then

$$\psi(f) = \lim_{k\to\infty} \widehat{\mu}_k(f) \geq 0.$$

Hence by the Riesz representation theorem, there is a finite measure μ on K with $\widehat{\mu} = \psi$, that is,

$$\lim_{k\to\infty} \int f \, d\mu_k = \int f \, d\mu \text{ for all } f \in \mathcal{C}_0(K).$$

Since $\widehat{\mu}(1) = 1$, μ is a probability measure.

Let $u \in \langle \underline{X} \rangle$ be fixed and

$$\varrho_u(\underline{A}) := 1 + \sum_{i=1}^n \operatorname{Tr}\left(A_i^{2\deg(u)+2}\right).$$

The assumption that $(y_w)_{\leq k}$ is a truncated tracial moment sequence with corresponding measure μ_k, implies that for all $k \geq 2\deg(u) + 2$

$$\begin{aligned}
\int \varrho_u \, d\mu_k &= \int \Big(1 + \operatorname{Tr}\Big(\sum_{i=1}^n A_i^{2\deg(u)+2}\Big)\Big) \, d\mu_k(\underline{A}) \\
&= \int \Big(1 + \sum_{i=1}^n \operatorname{Tr}\left(X_i^{2\deg(u)+2}(\underline{A})\right)\Big) \, d\mu_k(\underline{A}) \\
&= 1 + \sum_{i=1}^n y_{\mathrm{X}_i^{2\deg(u)+2}}.
\end{aligned}$$

Thus $\sup_{k\geq 2\deg(u)+2} \int \varrho_u \, d\mu_k < \infty$.

Hence we can apply Proposition 1.24 and obtain for $\varphi_u \in C_0(K)$ from Lemma 4.21 that

$$\begin{aligned}
y_u &= \lim_{k\to\infty} \int \operatorname{Tr}(u(\underline{A})) \, d\mu_k(\underline{A}) \\
&= \lim_{k\to\infty} \int \varphi_u \varrho_u \, d\mu_k = \int \varphi_u \varrho_u \, d\mu = \int \operatorname{Tr}\left(u(\underline{A})\right) d\mu(\underline{A}).
\end{aligned}$$

■

4.3.2 Cubature formulas

In this section we show that every truncated tracial moment sequence y that admits a representing measure, also admits a representing measure with finite support. That is, the corresponding tracial Riesz functional L_y can be expressed by a cubature formula. This result is the tracial version of the result of Bayer and Teichmann [BT, Theorem 2] for truncated moment sequences.

4.22 Definition. Let μ be a measure on $(\mathcal{S}\mathbb{R}^{s\times s})^n$ for some $s \in \mathbb{N}$ and let $m \in \mathbb{N}$ be such that all tracial moments $\int \operatorname{Tr}(w) d\mu$ exist for all $w \in \langle \underline{X} \rangle_m$. A *(tracial) cubature formula* of degree m for μ is given by an integer $N \in \mathbb{N}$, points $\underline{A}^{(1)}, \ldots, \underline{A}^{(N)} \in \operatorname{supp} \mu$ and weights $\lambda_1, \ldots, \lambda_N \in \mathbb{R}_{\geq 0}$ such that for all polynomials $p \in \mathbb{R}\langle \underline{X} \rangle_m$,

$$\int \operatorname{Tr}(p(\underline{A}))\, d\mu(\underline{A}) = \sum_{i=1}^{N} \lambda_i \operatorname{Tr}(p(\underline{A}^{(i)})).$$

The following theorem can also be found [BCKP, Theorem 3.7]. The proof is an adaptation of the proof of the theorem of Bayer and Teichmann [BT, Theorem 2] presented by Laurent [Lau1, Theorem 5.9].

4.23 Theorem. *If $y = (y_w)_{\leq k}$ is a truncated tracial moment sequence with measure μ on $(\mathcal{S}\mathbb{R}^{s\times s})^n$ for some $s \in \mathbb{N}$, then μ has a cubature formula of degree k. In fact, there exist an integer $N \in \mathbb{N}$, weights $\lambda_i \in \mathbb{R}_{\geq 0}$ with $\sum_i^N \lambda_i = 1$ and n-tuples $\underline{A}^{(i)} = (A_1^{(i)}, \ldots, A_n^{(i)}) \in \operatorname{supp} \mu$, such that for all $w \in \langle \underline{X} \rangle_k$:*

$$y_w = \sum_{i=1}^{N} \lambda_i \operatorname{Tr}(w(\underline{A}^{(i)})). \tag{4.8}$$

In other words, Theorem 4.23 states that a truncated tracial moment sequence is a convex combination of truncated tracial moment sequences $y^{\underline{A}}$ as in Example 4.2(b). For $s = 1$ we obtain the original statement [BT, Theorem 2] but without an explicit upper bound on the numbers of points needed in the cubature formula.

Proof. Let R be the finite dimensional $\mathbb{R}$-vector space of all truncated tracial sequences in n variables of order k. Let C be the convex cone of truncated tracial moment sequences $y^{\underline{A}} \in R$, given by $y_w^{\underline{A}} = \operatorname{Tr}(w(\underline{A}))$ for $w \in \langle \underline{X} \rangle_k$, with $\underline{A} \in \operatorname{supp} \mu$, see also Example 4.2(a). That is,

$$C = \operatorname{cone}\{y^{\underline{A}} = (y_w^{\underline{A}})_{\leq k} \mid \underline{A} \in \operatorname{supp} \mu\} \subseteq R,$$

where $\operatorname{cone}(D)$ denote the convex cone which is generated by all elements $d \in D$ for some set D. If $y \in C$, then y has representing measure with finite support in $\operatorname{supp} \mu$, hence a cubature formula.

The closure $\overline{C}$ can be written as the intersection of supporting hyperplanes, i.e.

$$\overline{C} = \{z = (z_w)_{\leq k} \mid \forall h \in H : L_h(z) \geq 0\}$$

for some set $H \subseteq R$ where the linear map $L_h : R \to \mathbb{R}$ is given by $L_h(z) = \sum_{w \in \langle \underline{X} \rangle_k} h_w z_w$. Thus $y \in \overline{C}$ since for any $h \in H$ we have

$$L_h(y) = \sum_w h_w y_w = \int \operatorname{Tr}(h(\underline{A}))\, d\mu(\underline{A}) \geq 0.$$

We now proceed to show that $y \in \operatorname{rel\,int} \overline{C}$. For this, suppose $y \in \overline{C} \setminus \operatorname{rel\,int} C$. Then there is a supporting hyperplane $H_h := \{z = (z_w)_{\leq k} \mid L_h(z) = 0\}$ that contains y but does not contain $\overline{C}$. Let $U = \{\underline{A} \in \operatorname{supp} \mu \mid L_h(y^{\underline{A}}) > 0\}$ and for $\ell \geq 1$, $U_\ell = \{\underline{A} \in U \mid L_h(y^{\underline{A}}) \geq \frac{1}{\ell}\}$. Then $U = \bigcup_\ell U_\ell$ and $U \neq \varnothing$ since $\overline{C} \not\subseteq H_h$. Hence there is some ℓ with $\mu(U_\ell) > 0$. Thus

$$0 = L_h(y) = \int_U \operatorname{Tr}(h)\, d\mu \geq \int_{U_\ell} \operatorname{Tr}(h)\, d\mu \geq \frac{1}{\ell} \mu(U_\ell) > 0,$$

a contradiction. This shows $L_h(y) > 0$, hence $y \in \operatorname{rel\,int} \overline{C} \subseteq C$ as desired. ∎

4.24 Remark. Using Carathéodory's theorem [Bar, p. 10] , we deduce that y from Theorem 4.23 can be written as a conic combination of at most $N \leq 1 + \sum_{\ell=1}^{k} B_n(\ell)$ tracial sequences $y^{\underline{A}}$, where

$$B_n(\ell) = \begin{cases} \frac{1}{2}N_n(\ell) + \frac{1}{4}(n+1)n^{\ell/2}; & \text{if } \ell \text{ even} \\ \frac{1}{2}N_n(\ell) + \frac{1}{2}n^{(\ell+1)/2}; & \text{if } \ell \text{ odd} \end{cases}$$

is the bracelet number,

$$N_n(\ell) = \frac{1}{\ell}\sum_{d|\ell} \phi\left(\frac{\ell}{d}\right) n^d$$

is the necklace number, and ϕ is the Euler function, cf. [Pól].

The next example [BK1, Example 3.4] shows that finite convex combinations of (truncated) tracial moment sequences $y^{\underline{A}^{(i)}}$ in general can not be written as $y^{\underline{A}}$ for an $\underline{A}$, i.e. there is in general no tuple $\underline{A}$ of matrices such that $y_w = \operatorname{Tr}(w(\underline{A}))$ for all $w \in \langle \underline{X} \rangle$.

4.25 Example. Let X be a single variable. We take the index set $1, X, X^2, X^3, X^4, \ldots$ and the (tracial) sequence $y = (1, 1-\sqrt{2}, 1, 1-\sqrt{2}, 1, \ldots)$. Then

$$y_w = \frac{\sqrt{2}}{2}w(-1) + (1 - \frac{\sqrt{2}}{2})w(1),$$

i.e. y has a representation (4.8) where $\lambda_1 = \frac{\sqrt{2}}{2}$, $\lambda_2 = 1 - \lambda_1$ and $A^{(1)} = -1$, $A^{(2)} = 1$. However there does not exist a symmetric matrix $A \in \mathcal{S}\mathbb{R}^{s \times s}$ for any $s \in \mathbb{N}$ such that $y_w = \operatorname{Tr}(w(A))$ for all $w \in \langle \underline{X} \rangle$. To show this, assume that one can find such an A. Without loss of generality we can choose A to be diagonal with diagonal elements $a_1, \ldots, a_s$. Then $y_w = \operatorname{Tr}(w(A))$ for all $w \in \langle \underline{X} \rangle$ only if the following equations hold:

$$\sum_{i=1}^{s} a_i = \sum_{i=1}^{s} a_i^3 = (1-\sqrt{2})s, \tag{4.9}$$

$$\sum_{i=1}^{s} a_i^2 = \sum_{i=1}^{s} a_i^4 = s. \tag{4.10}$$

In the general mean inequality, which follows from the Hölder inequality,

$$\frac{\sum_{i=1}^{s} x_i}{s} \leq \sqrt{\frac{\sum_{i=1}^{s} x_i^2}{s}}$$

for the arithmetic and the quadratic mean of $x = (x_1, \ldots, x_s) \in \mathbb{R}^s_{\geq 0}$, equality holds if and only if all the x_i are the same. Hence (4.10) rewritten as

$$\frac{\sum a_i^2}{s} = 1 = \sqrt{\frac{\sum a_i^4}{s}},$$

gives $a_1^2 = \cdots = a_s^2 = 1$. Therefore,

$$\sum_{i=1}^{s} a_i = \sum_{i=1}^{s} a_i^3 \in \mathbb{Z}.$$

Since $(1-\sqrt{2})s \notin \mathbb{Z}$, this contradicts (4.10), thus there is no representation involving only one matrix A.

4.3.3 Extensions of Riesz functionals

The aim of this section is to present a sufficient condition, dealing with positive extensions of Riesz functionals, for a truncated tracial sequence y to be a truncated tracial moment sequence. As consequence of this characterization and Theorem 4.20 we obtain further a tracial analog of classical theorems of Riesz and Haviland [Rie, Hav].

Let y be a truncated tracial sequence of degree k and let K be a non-empty closed subset of $(\mathcal{S}\mathbb{R}^{s\times s})^n$ for some $s \in \mathbb{N}$.

4.26 Definition. We say that the tracial Riesz functional L_y is *positive* (denoted by $L_y \geq 0$) if

$$L_y(p) \geq 0 \text{ for all trace-positive } p \in \mathbb{R}\langle \underline{X} \rangle_k.$$

The functional L_y is called *strictly positive* (denoted by $L_y > 0$), if

$$L_y(p) > 0 \text{ for all } p \in \mathbb{R}\langle \underline{X} \rangle_k, p \overset{\text{cyc}}{\not\sim} 0.$$

Further we call L_y *K-positive* (denoted by $L_y|_K \geq 0$) if

$$L_y(p) \geq 0 \text{ for all } p \in \mathbb{R}\langle \underline{X} \rangle_k \text{ trace-positive on } K,$$

and call it *strictly K-positive* (denoted by $L_y|_K > 0$), if

$$L_y(p) > 0 \text{ for all } p \in \mathbb{R}\langle \underline{X} \rangle_k \text{ trace-positive on } K, \operatorname{Tr} \circ p|_K \neq 0.$$

For a tracial sequence y we call L_y (strictly) positive if $L_y|_{\mathbb{R}\langle \underline{X} \rangle_k}$ is (strictly) positive for all $k \in \mathbb{N}$.

Equivalently, a tracial Riesz functional L_y is positive (respectively, strictly positive) if and only if the map $\bar{L}_y$ it induces on $\mathbb{R}\langle \underline{X} \rangle_k/_{\overset{\text{cyc}}{\sim}}$ is positive (respectively, strictly positive) on the non-zero images of trace-positive polynomials in $\mathbb{R}\langle \underline{X} \rangle_k/_{\overset{\text{cyc}}{\sim}}$. We mention that positive Riesz functionals are states as defined in Definition 1.21.

4.27 Remark. If y is a tracial moment sequence with representing measure μ supported in K, then L_y is K-positive. In fact, for all $p \in \mathbb{R}\langle \underline{X} \rangle$ that are trace-positive on K we have

$$L_y(p) = \int \operatorname{Tr}(p(\underline{A}))\, d\mu(\underline{A}) \geq 0,$$

since $\operatorname{Tr}(p(\underline{A})) \geq 0$ on K implies $\operatorname{Tr}(p(\underline{A})) \geq 0$ on $\operatorname{supp} \mu$. The same holds true analogously for truncated tracial moment sequences.

Hence K-positivity of the tracial Riesz functional L_y is a necessary condition for y to be a tracial K-moment sequence. In Theorem 4.30 we will see, that this condition is also sufficient for infinite tracial sequences. Namely, a tracial sequence y is a tracial K-moment sequence if and only if there is some $s \in \mathbb{N}$ such that L_y is K-positive for some closed set $K \subseteq (\mathcal{S}\mathbb{R}^{s\times s})^n$. The same statement is false in general for truncated tracial sequences as in the commutative case [CF3, Example 2.1]. Therefore we present the tracial version of the appropriate analog of the Riesz-Haviland theorem as proposed by Curto and Fialkow [CF3, Theorem 1.2], which deals with positive extensions of Riesz functionals.

The next result is the main tool to prove the analog of the Riesz-Haviland theorem. If a truncated tracial sequence of degree $2k$ has a K-positive tracial Riesz functional, then we can find a measure μ such that representation (4.3) holds for all $w \in \langle \underline{X} \rangle_{2k-1}$. This theorem resembles [CF3, Theorem 2.4], and in fact, for $s = 1$ we derive exactly the corresponding classical result.

4.28 Theorem. *Let K be a non-empty closed subset of $(\mathcal{S}\mathbb{R}^{s\times s})^n$ for some $s \in \mathbb{N}$, further let $y = (y_w)_{\leq 2k}$ be a truncated tracial sequence. If L_y is K-positive then $(y_w)_{\leq 2k-1}$ is a tracial K-moment sequence.*

The proof follows the proof of [CF3, Theorem 2.3] and works as follows. First, by the Hahn-Banach separation theorem it follows from the K-positivity of L_y, that y lies in the closure $\overline{C}$ of all truncated tracial moment sequences of order $2k$ with representing measure μ supported in K. With a similar argumentation as in the proof of Theorem 4.20 we then deduce that $(y_w)_{\leq 2k-1}$ lies in the interior of $\overline{C}$, hence is a truncated tracial K-moment sequence.

Proof. Let R be the finite dimensional $\mathbb{R}$-vector space of all truncated tracial sequences in n variables of order $2k$. Let C be the convex cone of truncated tracial moment sequences $y^{\underline{A}} \in R$ with $\underline{A} \in K$, given by $y_w^{\underline{A}} = \operatorname{Tr}(w(\underline{A}))$ for $w \in \langle \underline{X} \rangle_{2k}$. That is,

$$C = \operatorname{cone}\{(y_w^{\underline{A}})_{\leq 2k} \mid \underline{A} \in K\} \subseteq R.$$

First, we will show that $y \in \overline{C}$, the closure of C. This will then imply that $(y_w)_{\leq 2k-1}$ admits a K-representing measure.

Assume that $y \notin \overline{C}$. Then by the Hahn-Banach separation theorem there exists a polynomial $p = \sum_w p_w w \in \mathbb{R}\langle \underline{X} \rangle_{2k} = \mathbb{R}\langle \underline{X} \rangle_{2k}^{**}$ such that $L_p(z) := \sum_w p_w z_w \geq 0$ for all $z \in \overline{C}$ and $L_p(y) = \sum_w p_w y_w < 0$. In particular, $\operatorname{Tr}(p(\underline{A})) \geq 0$ for all $\underline{A} \in K$, hence p is trace-positive on K. Since L_y is K-positive, we derive the contradiction

$$0 \leq L_y(p) = \sum_w p_w y_w = L_p(y) < 0.$$

Therefore $y \in \overline{C}$, and we can write y as limit of elements in C. In fact, we find $\lambda_{i_\ell} \in \mathbb{R}_{>0}$ and $\underline{A}^{(i_\ell)} \in K$ such that $y = \lim_{\ell\to\infty} \sum_{i_\ell=1}^{N_\ell} \lambda_{i_\ell} y^{\underline{A}^{(i_\ell)}}$. This can be written as

$$y_w = \lim_{\ell\to\infty} \int \operatorname{Tr}(w)\, d\mu_\ell \tag{4.11}$$

with measures $\mu_\ell = \sum_{i_\ell=1}^{N_\ell} \lambda_{i_\ell} \delta_{\underline{A}^{(i_\ell)}}$. Let $\widehat{\mu}_\ell$ be the corresponding linear functional associated to μ_ℓ, see Section 1.6. Then we have

$$|\widehat{\mu}_\ell(f)| \leq \int \|f\|_\infty \, d\mu_\ell \leq \|f\|_\infty \text{ for all } f \in \mathcal{C}_0(K),$$

so all $\widehat{\mu}_\ell \in \mathcal{C}_0(K)'$ belong to the closed unit ball $\mathbb{B}$ in the dual space $\mathcal{C}_0(K)'$. Since $\mathbb{B}$ is weak-$*$-closed by the Banach-Alaoglu theorem [Ber, (44.12)], there is a subsequence $(\widehat{\mu}_{\ell_j})_j$ of $(\widehat{\mu}_\ell)$ converging in the weak-$*$-topology to some $\psi \in \mathbb{B}$. For simplicity of notation, we omit the subindex j in the sequel and assume that $(\widehat{\mu}_\ell)$ converges to ψ. If $f \in \mathcal{C}_0(K)$ and $f \geq 0$ on K, then $\psi(f) = \lim_{\ell\to\infty} \widehat{\mu}_\ell(f) \geq 0$. Hence by the Riesz representation theorem, there is a finite measure μ supported in K with $\widehat{\mu} = \psi$, that is,

$$\lim_{\ell\to\infty} \int f\, d\mu_\ell = \int f\, d\mu \text{ for all } f \in \mathcal{C}_0(K). \tag{4.12}$$

Let μ be the associated measure to $\widehat{\mu}$. We proceed to show that μ is a K-representing measure for $(y_w)_{\leq 2k-1}$. Let

$$\varrho(\underline{A}) := 1 + \sum_{i=1}^{n} \operatorname{Tr}(A_i^{2k}).$$

By (4.11) we have that $\int \varrho \, d\mu_\ell$ converges to $y_\varrho := y_1 + \sum_{i=1}^n y_{X_i^{2k}}$ if ℓ tends to infinity, hence $\sup_\ell \int \varrho \, d\mu_\ell < \infty$. With the same line of reasoning as in Lemma 4.21, we obtain that for all $u \in \langle \underline{X} \rangle_{2k-1}$ the map

$$\varphi_u(\underline{A}) := \frac{\operatorname{Tr}\left(u(\underline{A})\right)}{1 + \sum_{i=1}^n \operatorname{Tr}(A_i^{2k})}$$

lies in $\mathcal{C}_0(K)$. Hence we can apply Proposition 1.24, which implies

$$\begin{aligned} \int \operatorname{Tr}(u) \, d\mu &= \int \varphi_u \varrho \, d\mu \\ &= \lim_{\ell \to \infty} \int \varphi_u \varrho \, d\mu_\ell = \lim_{\ell \to \infty} \int \operatorname{Tr}(u) \, d\mu_\ell = y_u. \end{aligned}$$

■

If we combine Theorem 4.28 with Theorem 4.23, we derive the following corollary, which is the proposed tracial analog of [CF3, Theorem 1.2]. Again, the classical result is obtained by setting $s = 1$.

4.29 Corollary. *Let K be a non-empty closed subset of $(\mathcal{S}\mathbb{R}^{s \times s})^n$ for some $s \in \mathbb{N}$. Further let $y = (y_w)_{\leq 2k}$ be a truncated tracial sequence. Then y is a truncated tracial K-moment sequence if and only if L_y admits a K-positive tracial linear extension $L_{\tilde{y}} : \mathbb{R}\langle \underline{X} \rangle_{2k+2} \to \mathbb{R}$.*

Proof. Let $y = (y_w)_{\leq 2k}$ be a truncated tracial moment sequence with representing measure supported in K. By Theorem 4.23, y admits a cubature formula hence a representing measure with finite support in K, i.e. there exist $\lambda_i \in \mathbb{R}_{\geq 0}$ and $\underline{A}^{(i)} \in K$ such that $y_w = \sum_i \lambda_i \operatorname{Tr}(w(\underline{A}^{(i)}))$ for all $w \in \langle \underline{X} \rangle_{2k}$. In particular, all moments of all orders are finite. If we extend y canonically by

$$\tilde{y}_w := \sum_i \lambda_i \operatorname{Tr}(w(\underline{A}^{(i)}))$$

for $w \in \langle \underline{X} \rangle_{2k+2}$, then $L_{\tilde{y}}$ is a K-positive tracial linear extension of L_y.

Conversely, let $y = (y_w)_{\leq 2k}$ be a tracial sequence which admits a K-positive tracial linear extension $L_{\tilde{y}} : \mathbb{R}\langle \underline{X} \rangle_{2k+2} \to \mathbb{R}$. Then by Theorem 4.28, there is a positive measure μ supported in K such that $y_w = \int \operatorname{Tr}(w) \, d\mu$ for all $w \in \langle \underline{X} \rangle$ of degree at most $2k+1$, hence μ is a K-representing measure for y. ■

We note that Corollary 4.29 remains true if we replace $(y_w)_{\leq 2k}$ by $(y_w)_{\leq 2k+1}$ as is clear from the proof.

The following theorem is a consequence of Corollary 4.29 and Theorem 4.20. It is the tracial analog of the results of Riesz and Haviland [Rie, Hav] stating that a sequence is a K-moment sequence if and only if its Riesz functional is K-positive. We obtain this classical result from Theorem 4.30 in the case $s = 1$.

4.30 Theorem. *Let K be a non-empty closed subset of $(\mathcal{S}\mathbb{R}^{s \times s})^n$ for some $s \in \mathbb{N}$ and let y be a tracial sequence. Then y is a tracial K-moment sequence if and only if L_y is K-positive.*

Proof. If y is a tracial K-moment sequence then L_y is K-positive by Remark 4.27. For the other implication it suffices to show by Theorem 4.20 that for all $k \in \mathbb{N}$ the truncated tracial sequence $(y_w)_{\leq 2k}$ has a representing measure supported in K. Let $k \in \mathbb{N}$ be fixed. Then the tracial Riesz functional $L := L_y|_{\mathbb{R}\langle \underline{X} \rangle_{2k+2}}$ obtained from L_y by restriction is a K-positive extension of $L|_{\mathbb{R}\langle \underline{X} \rangle_{2k}}$. Hence by Corollary 4.29, the tracial sequence $(y_w)_{\leq 2k}$ is a truncated tracial moment sequence with representing measure supported in K. ■

4.31 Remark. Let $K_s := (\mathcal{S}\mathbb{R}^{s\times s})^n$ for $s \in \mathbb{N}$. Obviously, a tracial Riesz functional L_y which is K_s-positive is also K_r-positive but in general not K_t-positive for any $r, s, t \in \mathbb{N}$ with $t \leq s \leq r$. Hence, if y has a tracial representation with a representing measure μ on $(\mathcal{S}\mathbb{R}^{s\times s})^n$ then it has also a representing measure ν on $(\mathcal{S}\mathbb{R}^{r\times r})^n$ for any $r \geq s$, see also Remark 4.4 3. On the other hand, if L_y is a positive tracial Riesz functional, meaning that it take positive values on all trace-positive polynomials, then it might not have a tracial representation.

4.3.4 Strictly positive Riesz functionals

Theorem 4.30 does not hold true in general if we replace the tracial sequence by a truncated tracial sequence. That is, positivity of L_y would not suffice for the existence of a *truncated* tracial moment representation (4.3) for y. This is, for instance, a consequence of Example 4.14.

4.32 Example (Example 4.14 revisited)**.**

(a) We considered the truncated tracial sequence $y = (y_1, y_{\mathtt{X}}, y_{\mathtt{X}^2}, y_{\mathtt{X}^3}, y_{\mathtt{X}^4})$ which is given by $y = (1,1,1,1,2)$. Since univariate positive polynomials are sums of (two) squares [Mar, Prop. 1.2.1], the positive semidefiniteness of the tracial Hankel matrix $M_2(y)$ implies the K-positivity of L_y for $K = \mathbb{R}$. In fact, let $p \in \mathbb{R}\langle X\rangle_4$ be positive, i.e. p is trace-positive on 1×1-matrices. Then $p = f^2 + g^2 = f^*f + g^*g$ for some $f, g \in \mathbb{R}\langle X\rangle_2$. Hence

$$L_y(p) = L_y(f^2+g^2) = \vec{f}^T M_2(y)\vec{f} + \vec{g}^T M_2(y)\vec{g} \geq 0.$$

However we have seen in Example 4.14(a) that y does not have a representing measure supported on 1×1-matrices.

(b) In Example 4.14(b) we have shown that the truncated tracial moment sequence

$$y = (1,0,0,1,1,1,0,0,0,0,4,0,2,1,0,4),$$

written as $y = (y_{w_1},\dots,y_{w_{16}})$, where we fixed the order of the words w_i as $1, X, Y, X^2, XY, Y^2, X^3, X^2Y, XY^2, Y^3, X^4, X^3Y, X^2Y^2, XYXY, XY^3, Y^4$, has a positive semidefinite tracial Hankel matrix but y is not a tracial moment sequence.

By Theorem 3.10, bivariate quartic polynomials that are trace-positive on symmetric 2×2-matrices are sums of hermitian squares and commutators. Hence as in Example 4.32(a) the positive semidefiniteness of $M_2(y)$ implies K-positivity of L_y for $K = (\mathcal{S}\mathbb{R}^{2\times 2})^2$. Therefore y is a truncated tracial sequence with K-positive tracial Riesz functional L_y, but without a representing measure.

Although Theorem 4.30 is false in general if we consider a truncated tracial sequences, it becomes true if we assume strict K-positivity of the Riesz functional L_y for some appropriate set K. These results are motivated by and resemble the results of Fialkow and Nie [FN, Section 2] in the commutative context.

We will see that strictly positive Riesz functionals lie in the interior of the cone of positive Riesz functionals. This follows from the fact, that this holds true for strictly K-positive Riesz functionals if K has non-empty interior or if K is a so-called *determining set*, meaning that $(\operatorname{Tr}\circ p)|_K = 0$ implies $p \overset{\text{cyc}}{\sim} 0$.

Before we state the proposed theorem, we first prove an auxiliary lemma, which is an extension of [FN, Lemma 2.3] from $K \subseteq \mathbb{R}^n$ to $K \subseteq \mathcal{S}^n$. The case $K = \mathcal{S}^n$ can also be found in [BK1, Lemma 4.7].

4.33 Lemma. *Let $K \subseteq \mathcal{S}^n$ be a closed determining set and let y be a truncated tracial sequence of degree k. If $L_y|_K > 0$ then there exists an $\varepsilon > 0$ such that $L_{\tilde{y}}|_K > 0$ for all $\tilde{y}$ with $\|y - \tilde{y}\|_1 < \varepsilon$.*

Proof. Let $\pi : \mathbb{R}\langle \underline{X}\rangle_k \to \mathbb{R}\langle \underline{X}\rangle_k/_{\overset{\text{cyc}}{\sim}}$ denote the quotient map and let $\mathbb{R}\langle \underline{X}\rangle_k/_{\overset{\text{cyc}}{\sim}}$ be equipped with the quotient norm as in (1.3). Then

$$S := \{\pi(p) \in \mathbb{R}\langle \underline{X}\rangle_k/_{\overset{\text{cyc}}{\sim}} \mid p \text{ is trace-positive on } K,\ \|\pi(p)\| = 1\}$$

is compact. For any $\pi(p) \in S$ we have $p \overset{\text{cyc}}{\nsim} 0$. Since K is a determining set we obtain $\operatorname{Tr}(\pi(p))|_K \neq 0$. Hence by a scaling argument, it suffices to show that $\bar{L}_{\tilde{y}}|_K > 0$ on S for $\tilde{y}$ close to y. The map $y \mapsto \bar{L}_y$ is a linear map between finite-dimensional vector spaces. Thus

$$|\bar{L}_{y'}(\pi(p)) - \bar{L}_{y''}(\pi(p))| \leq C\|y' - y''\|_1$$

for all $\pi(p) \in S$, truncated tracial moment sequences y', y'', and some $C \in \mathbb{R}_{>0}$.

Since $\bar{L}_y$ is continuous and strictly positive on S, there exists an $\varepsilon > 0$ such that $\bar{L}_y(\pi(p)) \geq 2\varepsilon$ for all $\pi(p) \in S$. Let $\tilde{y}$ satisfy $\|y - \tilde{y}\|_1 < \varepsilon/C$. Then

$$\bar{L}_{\tilde{y}}(\pi(p)) \geq \bar{L}_y(\pi(p)) - C\|y - \tilde{y}\|_1 \geq \varepsilon > 0. \qquad \blacksquare$$

Now we are ready to prove the main theorem of this section, which is the tracial analog of [FN, Theorem 2.4] in the commutative context. The proof works the same way as for the classical result. However, in this case the analogue statement [FN, Theorem 2.4] from the commutative context cannot be obtained directly from Theorem 4.34 since $\mathbb{R}^n$ is not a determining set in our context.

4.34 Theorem. *Let $K \subseteq \mathcal{S}^n$ be a closed determining set and $y = (y_w)_{\leq k}$ be a truncated tracial sequence of degree k. If $L_y|_K > 0$, then y is a truncated tracial K-moment sequence.*

Proof. Let

$$C := \{(y_w)_k \mid \exists s \in \mathbb{N}, \underline{A}^{(i)} \in K \cap (\mathcal{S}\mathbb{R}^{s\times s})^n, \lambda_i \in \mathbb{R}_{\geq 0} \,:\, y_w = \sum_i \lambda_i \operatorname{Tr}(w(\underline{A}^{(i)}))\}$$

denote the convex cone of truncated tracial sequences y of degree k which are supported in K. We show first that $y \in \overline{C}$, the closure of C.

Assume that $L_y|_K > 0$ but $y \notin \overline{C}$. Since $\overline{C}$ is a closed convex cone in a finite dimensional vector space, by the Hahn-Banach separation theorem there exists a polynomial $p \in \mathbb{R}\langle \underline{X}\rangle_k$ such that $L_p(y) = \sum_w p_w y_w < 0$ and $L_p(z) \geq 0$ for all $z \in \overline{C}$. In particular, p is trace-positive on K, implying the contradiction

$$0 \leq L_y(p) = L_p(y) < 0.$$

By Lemma 4.33, $y \in \operatorname{int}(\overline{C})$. Thus $y \in \operatorname{int}(\overline{C}) \subseteq C$ [Ber, Theorem 25.20]. $\blacksquare$

From the proof of Theorem 4.34 we obtain that a truncated tracial sequence with K-positive Riesz functional lies in the closure of C, the cone of all tracial moment sequences with representing measure with finite support in K.

Since $K = \mathcal{S}^n$ is a closed determining set, we derive immediately the following Corollary, see also [BK1, Theorem 4.8].

4.35 Corollary. *Let $y = (y_w)_{\leq k}$ be a truncated tracial sequence of degree k. If L_y is strictly positive, then y is a truncated tracial moment sequence.*

4.3.5 Flat tracial Hankel matrices

In this last section we continue investigating extensions of truncated tracial sequences. We present the tracial analog of a theorem of Curto and Fialkow [CF1, Theorem 5.13] stating that a truncated sequence with positive semidefinite Hankel matrix which has a flat extension is a truncated moment sequence.

4.36 Definition. Let $A \in \mathcal{S}\mathbb{R}^{s\times s}$. A (symmetric) extension of A is a matrix $\tilde{A} \in \mathcal{S}\mathbb{R}^{(t+s)\times(t+s)}$ of the form

$$\tilde{A} = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \tag{4.13}$$

for some $B \in \mathbb{R}^{t\times s}$ and $C \in \mathbb{R}^{s\times s}$. Such an extension is *flat* if $\operatorname{rank} A = \operatorname{rank} \tilde{A}$, or, equivalently, if $B = AW$ and $C = W^T AW$ for some matrix W.

We need some basic properties for flat matrix extensions, see for instance [CF1, Lemma 5.2] or [CF2, Prop. 2.1]..

4.37 Lemma. *Let $\tilde{A}$ as in (4.13) be a flat extension of A. Then the following statements hold.*

(i) $\ker \tilde{A} = \ker[A \quad B]$;

(ii) $x \in \ker A \implies [x \quad 0]^T \in \ker \tilde{A}$;

(iii) $A \succeq 0$ *if and only if* $\tilde{A} \succeq 0$.

Proof. We have $\operatorname{rank} \tilde{A} \geq \operatorname{rank}[A \quad B] \geq \operatorname{rank} A$. Since $\operatorname{rank} A = \operatorname{rank} \tilde{A}$, equality holds, which implies (i). To show (ii) let $x \in \ker A$. Since $\tilde{A}$ is a flat extension of A there is a matrix W such that $B = AW$. Hence we have $B^T x = 0$, which implies $[x \quad 0]^T \in \ker \tilde{A}$. For the last statement let $W \in \mathbb{R}^{s\times t}$ be given such that $B = AW$ and $C = W^T AW$. Let $v = \begin{bmatrix} a & b \end{bmatrix}^T \in \mathbb{R}^{s+t}$ with $a \in \mathbb{R}^s$ and $b \in \mathbb{R}^t$ be given. Then one easily verifies that

$$v^T \tilde{A} v = (a + Wb)^T A(a + Wb),$$

which implies (iii). ∎

Recall that $I_{M_k} = \{p \in \mathbb{R}\langle \underline{X}\rangle_k \mid M_k \vec{p} = 0\}$ for a tracial Hankel matrix M_k of order k. This set I_{M_k} is in general not an ideal. However if the tracial Hankel matrix M_k is a flat extension of M_{k-1} then I_{M_k} contains $I_{M_{k-1}}$ and has also some truncated *ideal-like properties* in the sense of the following lemma. A variant of Lemma 4.38, which resembles [Lau1, Lemma 5.7] in the commutative case, can also be found in [BK1, Lemma 3.17].

4.38 Lemma. *Let $f \in \mathbb{R}\langle \underline{X}\rangle_{k-1}$ be an element of $I_{M_{k-1}}$ and let the tracial Hankel matrix M_k be a flat extension of M_{k-1}. Then $f \in I_{M_k}$. Furthermore we have $fX_i, X_i f \in I_{M_k}$.*

Proof. The first statement is clear by Lemma 4.37(ii). For the second statement let $f = \sum_w f_w w$. Then for $v \in \langle \underline{X}\rangle_{k-1}$, we have

$$[M_k \overrightarrow{fX_i}]_v = \sum_w f_w y_{v^* w X_i} = \sum_w f_w y_{(vX_i)^* w} = [M_k \vec{f}\,]_{vX_i} = 0. \tag{4.14}$$

The matrix M_k is of the form

$$M_k = \begin{bmatrix} M_{k-1} & B \\ B^T & C \end{bmatrix}.$$

Since M_k is a flat extension of M_{k-1}, Lemma 4.37(i) implies $I_{M_k} = I_{M'}$ for $M' := [M_{k-1} \quad B]$. Thus from (4.14) it follows, $fX_i \in I_{M'} = I_{M_k}$. For $X_i f$ we obtain in the same way

$$[M_k \overrightarrow{X_i f}]_v = \sum_w f_w y_{v^* X_i w} = \sum_w f_w y_{(X_i v)^* w} = [M_k \vec{f}]_{X_i v} = 0$$

for $v \in \langle \underline{X} \rangle_{k-1}$, which implies $X_i f \in I_{M_k}$. ∎

We are now ready to prove the tracial version of the flat extension theorem of Curto and Fialkow [CF1, Theorem 5.4]. This result can also be found in [BK1, Theorem 3.18].

4.39 Theorem. *Let y be a truncated tracial sequence of degree $2k$. If* $\operatorname{rank} M_k(y) = \operatorname{rank} M_{k-1}(y)$ *then there exists a unique tracial extension $\tilde{y} = (\tilde{y}_w)_{\leq 2k+2}$ of y such that $M_{k+1}(\tilde{y})$ is a flat extension of $M_k(y)$.*

The proof is an adaptation of the elementary proof of [CF1, Theorem 5.4]. We will define a flat extension M_{k+1} of $M_k(y)$, which will turn out to be a tracial Hankel matrix. The uniqueness of M_{k+1} follows by construction.

Proof. Let $M_k := M_k(y)$. We will construct a flat extension

$$M_{k+1} = \begin{bmatrix} M_k & B \\ B^T & C \end{bmatrix}$$

such that M_{k+1} is a tracial Hankel matrix. Since M_k is a flat extension of $M_{k-1}(y)$ we can find a subset $V \subseteq \langle \underline{X} \rangle_{k-1}$ of words labelling a maximum set of linearly independent columns of M_k. Then any column of M_k can be expressed (in a unique way) as a linear combination of columns labelled by V. That is, for each $p \in \mathbb{R}\langle \underline{X} \rangle$ with $\deg p \leq k$ there exists a *unique* $r \in \operatorname{span}(V) \subseteq \mathbb{R}\langle \underline{X} \rangle_{k-1}$ such that $M_k \vec{p} = M_k \vec{r}$, or equivalently, $p - r \in I_{M_k}$.

Let $v \in \langle \underline{X} \rangle$ of degee $k+1$ be given. We write $v = v' X_i$ for some $i \in \{1, \ldots, n\}$ and $v' \in \langle \underline{X} \rangle_k$. For v' there exists an $r \in \operatorname{span}(V)$ such that $v' - r \in I_{M_k}$. According to Lemma 4.38, a flat extension of M_k should satisfy $M_{k+1} \vec{v} = M_{k+1} \overrightarrow{r X_i} = M_k \overrightarrow{r X_i}$, such that $v' - r \in I_{M_k}$ implies $(v' - r) X_i \in I_{M_{k+1}}$. Therefore we define B such that for all $v \in \langle \underline{X} \rangle_k$:

$$[M_k \quad B] \overrightarrow{v X_i} = [M_k \quad B] \overrightarrow{r_v X_i}, \tag{4.15}$$

where r_v denotes be the unique element in $\operatorname{span}(V)$ with $v - r_v \in I_{M_k}$. More precisely, let $(w_1, \ldots, w_\ell)$ be the elements of the basis of M_k of degree exactly k. We define $B := M_k W$ with $W = (\overrightarrow{r_{w_1} X_1}, \overrightarrow{r_{w_1} X_2} \ldots, \overrightarrow{r_{w_\ell} X_n})$. Then B satisfies (4.15). Further we set $C := W^T M_k W$. Since the r_{w_i} are uniquely determined, the matrix

$$M_{k+1} = \begin{bmatrix} M_k & B \\ B^T & C \end{bmatrix} \tag{4.16}$$

is well-defined. The constructed M_{k+1} is a flat extension of M_k, and $M_{k+1} \succeq 0$ if and only if $M_k \succeq 0$ by Lemma 4.37(iii). Moreover, once B is chosen, there is only one C making M_{k+1} as in (4.16) a flat extension of M_k. This follows from general linear algebra, see e.g. [CF2, p. 11]. Hence M_{k+1} is the *only* candidate for a flat extension.

Therefore we are done if M_{k+1} is a tracial Hankel matrix, i.e. we have to show that

$$[M_{k+1}]_{v,w} = [M_{k+1}]_{v_1,w_1} \quad \text{whenever } v^* w \overset{\text{cyc}}{\sim} v_1^* w_1. \tag{4.17}$$

First we will prove that $[M_{k+1}]_{u,vX_i} = [M_{k+1}]_{uX_i,v}$ for $u, v \in \langle \underline{X} \rangle_k$. This implies recursively the tracial Hankel property of $[M_k \quad B]$. If $\deg vX_i \leq k$ there is nothing to show since M_k is a tracial Hankel matrix. If $\deg u \leq k$ and $\deg vX_i = k+1$ there exists an $r \in \operatorname{span}(V)$ such that $r - v \in I_{M_{k-1}}$, and by construction also $vX_i - rX_i \in I_{M_k}$. Then we get

$$\begin{aligned}[M_{k+1}]_{u,vX_i} &= \vec{u}^T M_{k+1}\overrightarrow{vX_i} = \vec{u}^T M_{k+1}\overrightarrow{rX_i} = \vec{u}^T M_k \overrightarrow{rX_i} \\ &= [M_k]_{u^* rX_i} = [M_k]_{X_i u^* r} = [M_k]_{(uX_i)^* r} \\ &\overset{(*)}{=} \overrightarrow{uX_i}^T M_{k+1}\vec{v} = [M_{k+1}]_{uX_i,v},\end{aligned}$$

where equality $(*)$ holds by (4.15). By symmetry we get also the tracial Hankel property of $\begin{bmatrix} M_k \\ B^T \end{bmatrix}$.

In the second step we prove that $[M_{k+1}]_{X_j u, vX_i} = [M_{k+1}]_{uX_i, X_j v}$ for $u, v \in \langle \underline{X} \rangle$ of degree k, implying recursively the tracial Hankel property for block C. Let $u, v \in \langle \underline{X} \rangle$ of degree k be given. There exist $s, r \in \operatorname{span}(V)$ with $u - s \in I_{M_{k-1}}$ and $r - v \in I_{M_{k-1}}$. Then

$$\begin{aligned}[M_{k+1}]_{X_j u, vX_i} &= \overrightarrow{X_j u}^T M_{k+1}\overrightarrow{vX_i} = \overrightarrow{X_j s}^T M_k \overrightarrow{rX_i} \\ &= [M_k]_{s^* X_j r X_i} = [M_k]_{(sX_i)^*(X_j r)} \\ &\overset{(*)}{=} \overrightarrow{uX_i}^T M_{k+1}\overrightarrow{X_j v} = [M_{k+1}]_{uX_i, X_j v}.\end{aligned}$$

Combining this with the first step we obtain Property (4.17). Hence the flat extension M_{k+1} of M_k is a tracial Hankel matrix. The construction of $\tilde{y}$ from M_{k+1} is clear. ∎

From Theorem 4.39 on flat extensions of truncated tracial sequences in combination with Theorem 4.19 on tracial sequences with tracial Hankel matrices of finite rank we derive the proposed theorem on truncated tracial sequences which admit a flat extension, see also [BK1, Corollary 3.19], and [CF1, Theorem 5.13] for the appropriate result in the commutative case.

4.40 Theorem. *Let $y = (y_w)_{\leq 2k}$ be a truncated tracial sequence. If $M_k(y)$ is positive semidefinite and $M_k(y)$ is a flat extension of $M_{k-1}(y)$, then y is a truncated tracial moment sequence.*

Proof. By Theorem 4.39 we can extend $M_k(y)$ inductively to an infinite positive semidefinite tracial Hankel matrix $M(\tilde{y})$ with $\operatorname{rank} M(\tilde{y}) = \operatorname{rank} M_k(y) < \infty$. Thus $M(\tilde{y})$ has finite rank and by Theorem 4.19, there exists a tracial moment representation of $\tilde{y}$. Therefore y is a truncated tracial moment sequence. ∎

From Theorem 4.40 one derives naturally the following corollary.

4.41 Corollary. *Let $y = (y_w)_{\leq 2k}$ be a truncated tracial sequence. Then y is a truncated tracial moment sequence if there is an $\ell \geq k$ and an extension $\tilde{y} = (\tilde{y})_{\leq 2\ell}$ of y such that $M_\ell(\tilde{y})$ is positive semidefinite and* $\operatorname{rank} M_\ell(\tilde{y}) = \operatorname{rank} M_{\ell-1}(\tilde{y})$.

In Chapter 6 we present an application of Theorem 4.40, where we will use the methods of this section together with the proof of Theorem 4.19 to construct numerically a tracial moment representation for tracial sequences whose tracial Hankel matrix admits a flat extension.

5 Duality

It is well known that the dual cones of the cone of positive polynomials and the cone of sums of squares are described by moment sequences, and by infinite sequences with a positive semidefinite Hankel matrix, respectively. Hence the theory of positive polynomials in commuting variables is intimately connected with the moment problem, e.g. in Haviland's theorem [Hav] or Schmüdgen's solution to the moment problem on compact basic closed semialgebraic sets which implies Schmüdgen's Positivstellensatz [Sch], see also [KM, Mar, PS, Put, PV] for further applications.

This chapter recalls some results from the previous Chapters 3 and 4 in terms of convex cones. Further it presents the duality properties of these cones. Namely, the dual cones of the cone of polynomials that are trace-positive on $K \subseteq (\mathcal{S}\mathbb{R}^{s\times s})^n$ and of the cone Θ^2 can be described by the cone of tracial moment sequences and the cone of tracial sequences with a positive semidefinite tracial Hankel matrix, respectively. We also describe the connection of the corresponding cones in the truncated case. Finally, we present some consequences of these dualities concerning trace-positive polynomials and sums of hermitian squares and commutators.

5.1 Positivity cones

In Chapter 3 we dealt with the question of which trace-positive polynomials can be written as a sum of hermitian squares and commutators, i.e. which trace-positive polynomials lie in the convex cone Θ^2. Since the set of trace-positive polynomials is also a convex cone, we can reformulate the obtained results in terms of convex cones.

5.1 Definition. 1. Let Θ^2_n denote the convex cone of sums of hermitian squares and commutators in $\mathbb{R}\langle\underline{X}\rangle$, where n denotes the number of variables in $\mathbb{R}\langle\underline{X}\rangle$. If n is arbitrary, we simply write Θ^2 as in the previous chapters. Further, let

$$\Theta^2_{n,k} = \{f \in \mathbb{R}\langle\underline{X}\rangle_{2k} \mid f \overset{\mathrm{cyc}}{\sim} \sum_{i=1}^{r} {g_i}^* g_i \text{ for some } g_i \in \mathbb{R}\langle\underline{X}\rangle_k, r \in \mathbb{N}_0\}$$

denote the convex cone of sums of hermitian squares and commutators of degree at most $2k$. These convex cones have already been introduced in Section 1.4.1.

2. Let

$$\mathcal{P}_n := \{p \in \mathbb{R}\langle\underline{X}\rangle \mid p \text{ trace-positive}\}$$

denote the convex cone of trace-positive polynomials in $\mathbb{R}\langle\underline{X}\rangle$. Again, n denotes the number variables in $\mathbb{R}\langle\underline{X}\rangle$. The cone of trace-positive polynomials of degree at most $2k$ will be

$$\mathcal{P}_{n,2k} := \mathcal{P}_n \cap \mathbb{R}\langle\underline{X}\rangle_{2k}.$$

3. Let K be a non-empty closed subset of $\mathcal{S}^n$. The convex cone of polynomials in $\mathbb{R}\langle\underline{X}\rangle$ being trace-positive on K is denoted by $\mathcal{P}_n(K)$ and by $\mathcal{P}_{n,2k}(K)$ if we only consider polynomials in n variables of degree at most $2k$. For $K = (\mathcal{S}\mathbb{R}^{s\times s})^n$ we simply write $\mathcal{P}_n(s)$ or $\mathcal{P}_{n,k}(s)$, respectively. Hence

$$\mathcal{P}_n(s) = \{p \in \mathbb{R}\langle\underline{X}\rangle \mid p \text{ trace-positive on } (\mathcal{S}\mathbb{R}^{s\times s})^n\}.$$

and

$$\mathcal{P}_{n,2k}(s) = \{p \in \mathbb{R}\langle\underline{X}\rangle_{2k} \mid p \text{ trace-positive on } (\mathcal{S}\mathbb{R}^{s\times s})^n\}.$$

Obviously, we have $\Theta^2_n \subseteq \mathcal{P}_n$ and $\Theta^2_{n,k} \subseteq \mathcal{P}_{n,2k}$ for all $n,k \in \mathbb{N}_0$. A natural question is: When does equality hold and when are the inclusions strict? The remaining part of this section gives an overview of the results of Chapter 3 dealing with this question.

By Section 3.1.1, any positive univariate polynomial is a sum of hermitian squares and commutators, hence also trace-positive. In other words

$$\mathcal{P}_1 = \mathcal{P}_1(1) = \Theta^2_1. \tag{5.1}$$

The same statement holds for quadratic polynomials, i.e. for all $n \in \mathbb{N}$,

$$\mathcal{P}_{n,2} = \mathcal{P}_{n,2}(1) = \Theta^2_{n,1}. \tag{5.2}$$

Further, Theorem 3.10 in Section 3.3 states that any polynomial $f \in \mathbb{R}\langle X,Y\rangle_4$ which is trace-positive on $(\mathcal{S}\mathbb{R}^{2\times 2})^2$ lies in Θ^2_2. Since $\Theta^2_{n,k} = \Theta^2_n \cap \mathbb{R}\langle \underline{X}\rangle_{2k}$, as well as directly from the proof of Theorem 3.10, we obtain that $f \in \Theta^2_{2,2}$ holds, hence

$$\mathcal{P}_{2,4} = \mathcal{P}_{2,4}(2) = \Theta^2_{2,2}. \tag{5.3}$$

On the other hand, the Motzkin polynomials from Example 3.5 show that $\mathcal{P}_{2,6} \neq \Theta^2_{2,3}$, which implies that

$$\mathcal{P}_{n,2k} \neq \Theta^2_{n,k} \text{ for all } n \geq 2, k \geq 3. \tag{5.4}$$

It is not clear if in the remaining cases $n \geq 3, k = 2$ equality holds. This seems to be unlikely since it is false in the commutative context, but we have no proof.

So far, all results agree with similar statements for positive polynomials in commuting variables, see Remark 1.12. By homogenization one obtains in the commutative context also statements for forms, i.e. homogeneous polynomials. However these statements do not hold in general in their tracial analog. Trivially, if we consider univariate forms or bivariate forms of degree ≤ 4, then any such form which is trace-positive is also cyclically equivalent to a sum of hermitian squares of appropriate forms by the above results. If we consider instead trace-positive bivariate forms of any even degree, then it becomes false. In fact, the bivariate form $S_{6,3}(X^2,Y^2)$ of degree 12 is trace-positive but cannot be written as sum of hermitian squares and commutators, see Section 2.1.2. The case of homogeneous polynomials of degree four in three variables remains open.

All cones defined in Definition 5.1 can also be considered as cones in $\mathbb{R}\langle \underline{X}\rangle/_{\underset{\sim}{\text{cyc}}}$. To describe their dual cones, considered as convex cones in the algebraic dual space $(\mathbb{R}\langle \underline{X}\rangle/_{\underset{\sim}{\text{cyc}}})^*$, we introduce moment cones, which are intimately connected to the tracial moment problem.

5.2 Moment cones

Chapter 4 presented results concerning the tracial moment problem. Since the set of tracial sequences as well as the set of tracial moment sequences are convex cones, we can reformulate the tracial moment problem and several results on it in terms of convex cones.

5.2 Definition. Let

$$\mathcal{T}_n := \{y \mid y \text{ tracial sequence}\}$$

denote the convex cone of tracial sequences $y = (y_w)$ with index set $\langle \underline{X}\rangle$, where n denotes the number of variables in $\langle \underline{X}\rangle$. Further, let $\mathcal{T}_{n,k}$ denote the convex cone of truncated tracial sequences y of degree k.

As we are interested in the dual cones of the positivity cones from Definition 5.1 we consider convex cones in $(\mathbb{R}\langle \underline{X}\rangle/_{\underset{\sim}{\text{cyc}}})^*$, the algebraic dual space of $\mathbb{R}\langle \underline{X}\rangle/_{\underset{\sim}{\text{cyc}}}$. Since

$$\begin{aligned}(\mathbb{R}\langle \underline{X}\rangle/_{\underset{\sim}{\text{cyc}}})^* &= \{L : (\mathbb{R}\langle \underline{X}\rangle/_{\underset{\sim}{\text{cyc}}}) \to \mathbb{R} \mid L \text{ linear}\}\\ &\cong \{L : \mathbb{R}\langle \underline{X}\rangle \to \mathbb{R} \mid L \text{ tracial }\} = \{L_y \mid y \in \mathcal{T}_n\}\end{aligned}$$

we consider convex cones consisting of tracial Riesz functionals instead of the tracial sequences themselves.

5.3 Definition. 1. Let

$$\mathcal{H}_n := \{L_y : \mathbb{R}\langle \underline{X}\rangle \to \mathbb{R} \mid y \in \mathcal{T}_n,\ M(y) \succeq 0\}$$

be the convex cone of tracial Riesz functionals with positive semidefinite tracial Hankel matrix. Further, let $\mathcal{H}_{n,k}$ denote the convex cone of truncated tracial sequences of degree $2k$ with positive semidefinite tracial Hankel matrix of order k. That is,

$$\mathcal{H}_{n,k} = \{L_y : \mathbb{R}\langle \underline{X}\rangle_{2k} \to \mathbb{R} \mid y \in \mathcal{T}_{n,2k},\ M_k(y) \succeq 0\}.$$

2. Let

$$\mathcal{M}_n := \{L_y : \mathbb{R}\langle \underline{X}\rangle \to \mathbb{R} \mid y \in \mathcal{T}_n \text{ tracial moment sequence}\}$$

denote the convex cone of tracial Riesz functionals corresponding to tracial moment sequences, i.e. for each tracial Riesz functional L_y exists an $s \in \mathbb{N}$ and a measure μ on $(\mathcal{S}\mathbb{R}^{s\times s})^n$ such that $L_y(f) = \int \operatorname{Tr}(f(\underline{A}))\, d\mu(\underline{A})$ for all $f \in \mathbb{R}\langle \underline{X}\rangle$. The truncated equivalent of $\mathcal{M}_n$ is for any $k \in \mathbb{N}_0$ given by

$$\mathcal{M}_{n,k} := \{L_y : \mathbb{R}\langle \underline{X}\rangle_k \to \mathbb{R} \mid y \in \mathcal{T}_{n,k} \text{ truncated tracial moment sequence}\}.$$

3. Let K be a non-empty closed subset of $\mathcal{S}^n$. The convex cone of tracial Riesz functionals L_y of tracial moment sequences $y \in \mathcal{T}_n$ for which there exists an $s \in \mathbb{N}$ and a representing measure supported in $K \cap (\mathcal{S}\mathbb{R}^{s\times s})^n$ is denoted by $\mathcal{M}_n(K)$, and by $\mathcal{M}_{n,k}(K)$ if we only consider truncated tracial moment sequence of degree k. For $K = (\mathcal{S}\mathbb{R}^{s\times s})^n$, we simply write $\mathcal{M}_n(s)$ and $\mathcal{M}_{n,k}(s)$. Hence

$$\mathcal{M}_n(s) = \{L_y : \mathbb{R}\langle \underline{X}\rangle \to \mathbb{R} \mid y \in \mathcal{M}_n, \operatorname{supp}\mu \subseteq (\mathcal{S}\mathbb{R}^{s\times s})^n\}.$$

By Theorem 4.23, the cone $\mathcal{M}_{n,k}$ can also be written as

$$\mathcal{M}_{n,k} = \{L_y \mid y \in \mathcal{T}_{n,k} \text{ truncated tracial moment sequence with finite support}\}.$$

These convex cones are called moment cones since they are intimately connected to the tracial moment problem. By Corollary 4.12, we have $\mathcal{M}_n \subseteq \mathcal{H}_n$ but we do not have equality in general. The same holds true for the truncated case, i.e. $\mathcal{M}_{n,2k} \subseteq \mathcal{H}_{n,k}$. The (truncated) tracial moment problem deals with the question when equality holds, or which subsets of $\mathcal{H}_n$, respectively $\mathcal{H}_{n,k}$, are contained in $\mathcal{M}_n$, respectively $\mathcal{M}_{n,2k}$. The results of Chapter 4 give some partial answers.

Let

$$\mathcal{H}_{\text{finite}} := \{L_y \in \mathcal{H}_n \mid \operatorname{rank} M(y) < \infty\}.$$

For infinite tracial sequences we know by Theorem 4.19 that $L_y \in \mathcal{M}_n$ holds if the tracial Hankel matrix $M(y)$ is positive semidefinite and of finite rank. Hence

$$\mathcal{H}_{\text{finite}} \subseteq \mathcal{M}_n. \tag{5.5}$$

By Theorem 4.20 we have that y is a tracial moment sequence with representing measure supported in $(\mathcal{S}\mathbb{R}^{s\times s})^n$ if $(y_w)_{\le k}$ is for all $k \in \mathbb{N}$ a truncated tracial moment sequence with representing measure supported in $(\mathcal{S}\mathbb{R}^{s\times s})^n$. That is,

$$L_y|_{\mathbb{R}\langle \underline{X}\rangle_k} \in \mathcal{M}_{n,k}(s) \text{ for all } k \in \mathbb{N} \implies y \in \mathcal{M}_n(s). \tag{5.6}$$

Further, if $L_y \in \mathcal{H}_{n,k}$ admits an extension $L_{\tilde{y}} \in \mathcal{H}_{n,k+i}$ for some $i \in \mathbb{N}$ such that $\operatorname{rank} M_{k+i}(\tilde{y}) = \operatorname{rank} M_{k+i-1}(\tilde{y})$, then $y \in \mathcal{M}_{n,2k}$ by Corollary 4.41.

The moment cones and the positivity cones are related to each other by conic duality. This is the topic of the next section.

5.3 Duality of positivity cones and moment cones

In this section we show that the well-known duality between positive polynomials and the moment problem can be extended to the tracial case. We will prove that the cone Θ_n^2 of sums of hermitian squares and commutators is dual to the cone $\mathcal{H}_n$ of tracial sequences with positive semidefinite tracial Hankel matrix. Further, the cones $\mathcal{P}_n(K)$ and $\mathcal{M}_n(K)$ are dual to each other for certain interesting sets K. Finally, we show which duality properties hold for the corresponding cones in $\mathbb{R}\langle \underline{X}\rangle_{2k}/\underset{\sim}{\text{cyc}}$.

First, we recall the notion of dual cones needed in this section. The (algebraic) dual space $\mathcal{A}^*$ of an $\mathbb{R}$-vector space $\mathcal{A}$ consists of all linear maps $L : \mathcal{A} \to \mathbb{R}$. For any $a \in \mathcal{A}$ the map

$$\Lambda_a : \mathcal{A}^* \to \mathbb{R},\ \Lambda_a(L) = L(a)$$

is linear. Hence $\Lambda_a \in \mathcal{A}^{**} := (\mathcal{A}^*)^*$ for any $a \in \mathcal{A}$. That is, there is a natural embedding of $\mathcal{A}$ into $\mathcal{A}^{**}$. We consider conic duality for convex cones regarded as cones in $\mathcal{A}$ and $\mathcal{A}^*$. That is, we intersect cones in the double dual space $\mathcal{A}^{**}$ with the subspace $\mathcal{A}$. More specific, for a given convex cone $C \in \mathcal{A}$ its *dual cone* is defined as

$$C^* := \{L \in \mathcal{A}^* \mid L(c) \ge 0 \text{ for all } c \in C\}$$

and the double dual cone C^{**} of C is given by $C^{**} := (C^*)^* \cap \mathcal{A}$. Hence

$$C^{**} = \{f \in \mathcal{A} \mid L(f) \ge 0 \text{ for all } L \in C^*\}.$$

Clearly $C \subseteq C^{**}$. Furthermore, if $\mathcal{A}$ is a countable dimensional vector space then C^{**} is equal to $\overline{C}$, the closure of C with respect to the finest locally convex topology on $\mathcal{A}$, see e.g. [Mar, Cor. 3.6.3]. In particular, for $\mathcal{A} = \mathbb{R}\langle \underline{X}\rangle$ we have $C = C^{**}$ if C is closed.

In the sequel we set $\mathcal{A} = \mathbb{R}\langle \underline{X}\rangle/\underset{\sim}{\text{cyc}}$ or $\mathcal{A} = \mathbb{R}\langle \underline{X}\rangle_{2k}/\underset{\sim}{\text{cyc}}$. The dual cones of the moment cones $\mathcal{H}_n$ or $\mathcal{M}_n(K)$ (respectively $\mathcal{H}_{n,k}$ or $\mathcal{M}_{n,k}(K)$) are considered as cones in $\mathbb{R}\langle \underline{X}\rangle/\underset{\sim}{\text{cyc}}$ (respectively $\mathbb{R}\langle \underline{X}\rangle_{2k}/\underset{\sim}{\text{cyc}}$), e.g.

$$\mathcal{H}_n^* := \{p \in \mathbb{R}\langle \underline{X}\rangle/\underset{\sim}{\text{cyc}} \mid L_y(p) \ge 0 \text{ for all } L_y \in \mathcal{H}_n\}.$$

We start by proving the duality of the cones Θ_n^2 and $\mathcal{H}_n$. This is in perfect analogy to the commutative case, see [Lau1, Prop. 4.5].

5.4 Proposition. *For any $n \in \mathbb{N}_0$, the convex cones $\mathcal{H}_n$ and Θ_n^2 are dual to each other, that is $\mathcal{H}_n^* = \Theta_n^2$ and $(\Theta_n^2)^* = \mathcal{H}_n$.*

This follows essentially from the closedness of Θ_n^2, see Corollary 1.20.

Proof. Let $n \in \mathbb{N}_0$ be fixed. The dual cone of Θ_n^2 is given by

$$(\Theta_n^2)^* = \{L_y : \mathbb{R}\langle \underline{X} \rangle \to \mathbb{R} \mid y \in \mathcal{T}_n, L_y(p^*p) \geq 0 \text{ for all } p \in \mathbb{R}\langle \underline{X} \rangle\}.$$

For any tracial Riesz functional L_y with tracial sequence $y \in \mathcal{T}_n$ we have $M(y) \succeq 0$ if and only if

$$\vec{p}^{\,T} M(y) \vec{p} = L_y(p^*p) \geq 0$$

for all $p \in \mathbb{R}\langle \underline{X} \rangle$. Hence for $L_y \in \mathcal{H}_n$ and $f \in \Theta_n^2$, written as $f = \sum_i {g_i}^* g_i + \sum_j [p_j, q_j]$ for some $g_i, p_j, q_j \in \mathbb{R}\langle \underline{X} \rangle$, we have

$$L_y(f) = \sum_i L_y({g_i}^* g_i) + \sum_j L_y([p_j, q_j]) = \sum_i L_y({g_i}^* g_i) \geq 0,$$

which implies $\mathcal{H}_n = (\Theta_n^2)^*$. Since Θ_n^2 is closed by Corollary 1.20, and $\mathbb{R}\langle \underline{X} \rangle/\underset{\sim}{\text{cyc}}$ has countable dimension, we have $(\Theta_n^2)^{**} = \overline{\Theta_n^2} = \Theta_n^2$, which implies the dual statement $\mathcal{H}_n^* = \Theta_n^2$. ∎

Let K be a non-empty subset of $(\mathcal{S}\mathbb{R}^{s\times s})^n$ for some $s \in \mathbb{N}$. Then we also have duality of the cone $\mathcal{P}_n(K)$ of polynomials $p \in \mathbb{R}\langle \underline{X} \rangle$ which are trace-positive on K and the cone $\mathcal{M}_n(K)$ of tracial moment sequences with representing measure supported in K. This follows essentially from Theorem 4.30, the tracial analog of the theorems of Riesz and of Haviland. This duality is in analogy to the commutative case [Lau1, Section 4.4], which follows from Proposition 5.5 in the case $s = 1$.

5.5 Proposition. *Let K be a non-empty closed set in $(\mathcal{S}\mathbb{R}^{s\times s})^n$ for some $s \in \mathbb{N}$. Then the convex cones $\mathcal{M}_n(K)$ and $\mathcal{P}_n(K)$ are dual to each other. In particular, the convex cones $\mathcal{M}_n(s)$ and $\mathcal{P}_n(s)$ are dual to each other.*

Proof. By definition, we have

$$\mathcal{P}_n(K)^* = \{L_y : \mathbb{R}\langle \underline{X} \rangle \to \mathbb{R} \mid y \in \mathcal{T}_n, L_y(p) \geq 0 \text{ for all } p \in \mathcal{P}_n(K)\},$$

hence $\mathcal{P}_n(K)^*$ consists of all tracial Riesz functionals which are K-positive, see Definition 4.26. Thus the first statement $\mathcal{P}_n(K)^* = \mathcal{M}_n(K)$ is exactly Theorem 4.30.

The dual cone of $\mathcal{M}_n(K)$ is by definition

$$\mathcal{M}_n(K)^* = \{p \in \mathbb{R}\langle \underline{X} \rangle/\underset{\sim}{\text{cyc}} \mid L_y(p) \geq 0 \text{ for all } L_y \in \mathcal{M}_n(K)\}.$$

To show that $\mathcal{P}_n(K) = \mathcal{M}_n(K)^*$, let $p \in \mathcal{P}_n(K)$ and $L_y \in \mathcal{M}_n(K)$ be given. By Remark 4.27 we have $L_y(p) = \sum_w p_w y_w \geq 0$, thus $\mathcal{P}_n(K) \subseteq \mathcal{M}_n(K)^*$. If $p \in \mathcal{M}_n(K)^*$ then we have for all $\underline{A} \in K$ that $L_{y^{\underline{A}}}(p) = \operatorname{Tr}(p(\underline{A})) \geq 0$, which implies $p \in \mathcal{P}_n(K)$. ∎

From Proposition 5.4 we obtain a new condition for f to be a sum of hermitian squares and commutators. Unfortunately, this condition is in general difficult to verify.

5.6 Corollary. *Let $f \in \mathbb{R}\langle \underline{X} \rangle$. Then $f \in \Theta_n^2$ if and only if $L_y(f) \geq 0$ for all $L_y \in \mathcal{H}_n$.*

Propositions 5.4 and 5.5 in combination with (5.1) give rise to the solution to the (tracial) univariate moment problem, also known as Hamburger moment problem, see for instance [Lau1, Theorem 4.6]. Indeed, from $\mathcal{P}_1 = \Theta_1^2$ follows ${\mathcal{P}_1}^* = (\Theta_1^2)^*$, and with Propositions 5.4 and 5.5 we obtain

$$\mathcal{M}_1 = {\mathcal{P}_1}^* = (\Theta_1^2)^* = \mathcal{H}_1.$$

That is, any univariate tracial sequence y with positive semidefinite tracial Hankel matrix has a representing measure.

One may ask if the conic dualities in Proposition 5.4 and Proposition 5.5 also hold for the appropriate cones in $\mathbb{R}\langle \underline{X} \rangle_{2k}/\underset{\sim}{\text{cyc}}$. The answer is given in the following two propositions.

5.7 Proposition. *For any $k \in \mathbb{N}_0$ the cones $\mathcal{H}_{n,k}$ and $\Theta^2_{n,k}$ are dual to each other.*

Proof. Since $L_y(p^*p)$ is positive for all $p \in \mathbb{R}\langle \underline{X} \rangle_k$ if and only if $M_k(y)$ is positive semidefinite, we get $\mathcal{H}_{n,k} = (\Theta^2_{n,k})^*$ as in Proposition 5.4. The dual statement $\mathcal{H}_{n,k}{}^* = \Theta^2_{n,k}$ follows by the closedness of $\Theta^2_{n,k}$, which has been shown in Proposition 1.19. ∎

5.8 Proposition. *Let K be a non-empty closed set in $(\mathcal{S}\mathbb{R}^{s \times s})^n$ for some $s \in \mathbb{N}$. Then for any $k \in \mathbb{N}_0$ we have $\mathcal{M}_{n,k}(K) \subseteq \mathcal{P}_{n,k}(K)^*$ and $\mathcal{M}_{n,k}(K)^* = \mathcal{P}_{n,k}(K)$.*

Proof. The first statement $\mathcal{M}_{n,k}(K) \subseteq \mathcal{P}_{n,k}(K)^*$ is easy. For any $L_y \in \mathcal{M}_{n,k}(K)$ there exists a measure μ supported in K such that for all $p \in \mathbb{R}\langle \underline{X} \rangle_k$,

$$L_y(p) = \int \operatorname{Tr}(p(\underline{A}))\, d\mu(\underline{A}).$$

Hence $L_y(p) \geq 0$ for any $p \in \mathcal{P}_{n,k}(K)$. This also implies $\mathcal{P}_{n,k}(K) \subseteq \mathcal{M}_{n,k}(K)^*$. Since $(y^{\underline{A}})_{\leq k} \in \mathcal{M}_{n,k}(K)$ for any $\underline{A} \in K$, we get as in Proposition 5.5 that $\mathcal{M}_{n,k}(K)^* \subseteq \mathcal{P}_{n,k}(K)$. ∎

Since Theorem 4.30 does not hold in the truncated case, the inclusion $\mathcal{M}_{n,k}(K) \subseteq \mathcal{P}_{n,k}(K)^*$ may be strict, see for instance Example 4.32. However these duality properties are still in analogy to the duality properties in the commutative case. In fact, Example 4.32(a) is a well-known example in the commutative context of a Riesz functional which is positive on $\mathbb{R}$ but which does not correspond to a truncated moment sequence, see [CF3, Example 2.1].

Although in general $\mathcal{M}_{n,k}(K) \subsetneq \mathcal{P}_{n,k}(K)^*$, the results of Section 4.3.3 give rise to interesting subsets of $\mathcal{P}_{n,k}(K)^*$ which lie in $\mathcal{M}_{n,k}(K)$. In fact, Theorem 4.28 and Corollary 4.29 can be reformulated as the following statements on convex cones.

5.9 Theorem. *Let K be a closed subset of $(\mathcal{S}\mathbb{R}^{s \times s})^n$ for some $s \in \mathbb{N}$ and let $k \in \mathbb{N}_0$. Then $L_y \in \mathcal{P}_{n,2k}(K)^*$ implies $L_y|_{\mathbb{R}\langle \underline{X} \rangle_{2k-1}} \in \mathcal{M}_{n,2k-1}(K)$.*

5.10 Corollary. *Let K be a closed subset of $(\mathcal{S}\mathbb{R}^{s \times s})^n$ for some $s \in \mathbb{N}$ and let $k \in \mathbb{N}_0$. Then any functional $L_y \in \mathcal{P}_{n,2k}(K)^*$ which admits an extension $L_{\tilde{y}} \in \mathcal{P}_{n,2k+2}(K)^*$ lies in $\mathcal{M}_{n,2k}(K)$.*

If $\mathcal{M}_{n,2k} = \mathcal{H}_{n,k}$ then by duality it would follow that $\Theta^2_{n,k} = \mathcal{H}_{n,k}{}^* = \mathcal{M}_{n,2k}{}^* = \mathcal{P}_{n,2k}$. Unfortunately, for $n \geq 2$ we have in general $\mathcal{M}_{n,2k} \subsetneq \mathcal{H}_{n,k}$. However, since $\Theta^2_{n,k}$ is closed, we can obtain with Corollary 4.35 the following necessary and sufficient condition for $\mathcal{P}_{n,2k} = \Theta^2_{n,k}$ in terms of moment cones. For this, let

$$\mathcal{H}^+_{n,k} := \{L_y : \mathbb{R}\langle \underline{X} \rangle_{2k} \to \mathbb{R} \mid y \in \mathcal{T}_{n,2k},\, M_k(y) \succ 0\} \subseteq \mathcal{H}_{n,k}.$$

Theorem 5.11 can also be found in [BK1, Theorem 4.4].

5.11 Theorem. *For any $k \in \mathbb{N}_0$, the following statements are equivalent:*

(i) $\mathcal{H}^+_{n,k} \subseteq \mathcal{M}_{n,2k}$;

(ii) $\mathcal{P}_{n,2k} = \Theta^2_{n,k}$.

In other words, if all truncated tracial sequences y of degree $2k$ with *positive definite* tracial Hankel matrix $M_k(y)$ have a tracial moment representation, then all trace-positive polynomials in $\mathbb{R}\langle \underline{X} \rangle_{2k}$ are sums of hermitian squares and commutators, and vice versa.

Proof. To show (i) $\implies$ (ii), assume that $f = \sum_w f_w w \in \mathbb{R}\langle \underline{X} \rangle_{2k}$ is trace-positive but $f \notin \Theta^2_{n,k}$. By Lemma 1.19, $\Theta^2_{n,k}$ is a closed convex cone in $\mathbb{R}\langle \underline{X} \rangle_{2k}$. Hence by the Hahn-Banach separation theorem we find a hyperplane which separates f and $\Theta^2_{n,k}$. In other words, there is a linear form $L : \mathbb{R}\langle \underline{X} \rangle_{2k} \to \mathbb{R}$ such that $L(f) < 0$ and $L(p) \geq 0$ for $p \in \Theta^2_{n,k}$. In particular, $L(f) = 0$ for all $f \overset{\text{cyc}}{\sim} 0$, i.e. without loss of generality, L is tracial. Since there are tracial states strictly positive on $(\Sigma^2 \cap \mathbb{R}\langle \underline{X} \rangle_{2k}) \setminus \{0\}$, we may assume $L(p) > 0$ for all $p \in \Theta^2_{n,k}$, $p \overset{\text{cyc}}{\nsim} 0$. Hence the bilinear form $(p, q) \mapsto L(q^* p)$ can be written as $L(q^* p) = \vec{q}^T M \vec{p}$ for some truncated tracial Hankel matrix $M \succ 0$. By assumption, the corresponding truncated tracial sequence y has a tracial moment representation (4.3). By Theorem 4.23 we can also find a finite representation (4.8), i.e.

$$y_w = \sum \lambda_i \operatorname{Tr}(w(\underline{A}^{(i)}))$$

for some tuples $A^{(i)}$ of symmetric matrices $A^{(i)}_j$ and $\lambda_i \in \mathbb{R}_{\geq 0}$ which implies the contradiction

$$0 > L(f) = \sum_i \lambda_i \operatorname{Tr}(f(\underline{A}^{(i)})) \geq 0.$$

Conversely, if (ii) holds, then $L_y > 0$ on $\mathbb{R}\langle \underline{X} \rangle_{2k}$ if and only if $M_k(y) \succ 0$. Thus a positive definite tracial Hankel matrix $M_k(y)$ defines a strictly positive functional L_y on $\mathbb{R}\langle \underline{X} \rangle_{2k}$ which by Corollary 4.35 has a tracial representation. ∎

Theorem 5.11 and (5.2) imply that any truncated tracial sequence y of degree 2 with positive definite tracial Hankel matrix is a tracial moment sequence. We can also use the other implication to show that $\mathcal{P}_{n,2} = \Theta^2_{n,1}$ holds. Since $\mathcal{H}_{n,1} = \mathcal{M}_{n,2}$, which follows from the spectral theorem as in the commutative case, see e.g. [CF1, Theorem 6.1], Theorem 5.11 implies $\mathcal{P}_{n,2} = \Theta^2_{n,1}$.

Less obvious is the fact that all tracial sequences y of degree 4 in two variables with positive definite tracial Hankel matrix have a representing measure. This follows with Theorem 5.11.

On the other hand, the Motzkin polynomial M_{nc} of Example 3.5 is trace-positive but does not lie in $\Theta^2_{2,3}$. Thus by Theorem 5.11 there exist tracial Riesz functionals with positive definite tracial Hankel matrices of order 3 but without a representing measure for the corresponding truncated tracial sequence.

5.12 Example. To represent a truncated tracial Hankel matrix of order three in two variables we choose the basis of $\mathbb{R}\langle X, Y \rangle_3$ that is given by the words $1, X, Y, X^2, XY, YX, Y^2, X^2Y, XY^2, YX^2, Y^2X, X^3, Y^3, XYX, YXY$ in this order. Then the following matrix

$$M_3 := \begin{bmatrix}
1 & 0 & 0 & \frac{7}{4} & 0 & 0 & \frac{7}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{7}{4} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{19}{16} & 0 & \frac{19}{16} & \frac{21}{4} & 0 & 0 & 0 \\
0 & 0 & \frac{7}{4} & 0 & 0 & 0 & 0 & \frac{19}{16} & 0 & \frac{19}{16} & 0 & 0 & \frac{21}{4} & 0 & 0 \\
\frac{7}{4} & 0 & 0 & \frac{21}{4} & 0 & 0 & \frac{19}{16} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{19}{16} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{19}{16} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{7}{4} & 0 & 0 & \frac{19}{16} & 0 & 0 & \frac{21}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{19}{16} & 0 & 0 & 0 & 0 & \frac{9}{8} & 0 & \frac{5}{6} & 0 & 0 & \frac{9}{8} & 0 & 0 \\
0 & \frac{19}{16} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{9}{8} & 0 & \frac{5}{6} & \frac{9}{8} & 0 & 0 & 0 \\
0 & 0 & \frac{19}{16} & 0 & 0 & 0 & 0 & \frac{5}{6} & 0 & \frac{9}{8} & 0 & 0 & \frac{9}{8} & 0 & 0 \\
0 & \frac{19}{16} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{5}{6} & 0 & \frac{9}{8} & \frac{9}{8} & 0 & 0 & 0 \\
0 & \frac{21}{4} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{9}{8} & 0 & \frac{9}{8} & 51 & 0 & 0 & 0 \\
0 & 0 & \frac{21}{4} & 0 & 0 & 0 & 0 & \frac{9}{8} & 0 & \frac{9}{8} & 0 & 0 & 51 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{5}{6} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{5}{6}
\end{bmatrix}.$$

is a positive definite tracial Hankel matrix of order 3 with corresponding truncated tracial sequence y of degree 6. Since

$$L_y(M_{\mathrm{nc}}) = M_{\mathrm{nc}}(y) = -\frac{5}{16} < 0,$$

the corresponding truncated tracial sequence y is not a truncated tracial moment sequence. Otherwise L_y would be positive for all trace-positive polynomials $p \in \mathbb{R}\langle X, Y\rangle_6$ by Remark 4.27.

We remark that the (free) non-commutative moment problem is always solvable for positive semidefinite Hankel matrices [McC, Theorem 2.1]. In Example 5.12 this means there are symmetric matrices $A, B \in \mathbb{R}^{15\times 15}$ and a vector $v \in \mathbb{R}^{15}$ such that

$$y_w = \langle w(A, B)v, v\rangle$$

for all $w \in \langle X, Y\rangle_6$.

6 A relaxation for numerical trace-optimization

In this chapter we present an application of the duality of trace-positive polynomials and the tracial moment problem. After a short introduction to semidefinite programming, we propose a sum of hermitian squares relaxation for trace-minimization of polynomials, which can be implemented by semidefinite programming. We prove that strong duality holds for this relaxation and its dual. Further, we give a sufficient condition under which the solution of this relaxation is equal to the optimum value. Finally, we show how one can in this case extract a trace-minimizer. This part is influenced by the method of Henrion and Lasserre [HL] for the commutative case, which has been implemented in GloptiPoly [HLL], see also [La2]. For a similar investigation in the free non-commutative setting see [PNA].

This application, which has been implemented in `NCSOStools` [CKP1], a software package for Matlab, can also been found in [BCKP, Section 3.1].

6.1 Semidefinite programming

Semidefinite programming (SDP) is a generalization of linear programming (LP). It is a subfield of convex optimization concerned with the optimization of a linear objective function over the intersection of the cone of positive semidefinite matrices with an affine space. More precisely, given symmetric matrices $C, A_1, \ldots, A_m \in \mathcal{S}\mathbb{R}^{s\times s}$ and a vector $b \in \mathbb{R}^m$, we formulate a *semidefinite program in standard primal form* as follows:

$$\begin{array}{lrcl} \inf & \langle C, G\rangle & & \\ \text{s.t.} & \langle A_i, G\rangle & = & b_i, \quad i = 1, \ldots, m \\ & G & \succeq & 0. \end{array} \tag{PSDP}$$

Here $\langle \_, \_ \rangle$ stands for the standard scalar product of matrices: $\langle A, B\rangle = \operatorname{Tr}(B^T A)$. The dual problem to (PSDP) is the *semidefinite program in the standard dual form*, with $y \in \mathbb{R}^m$:

$$\begin{array}{ll} \sup & \langle b, y\rangle \\ \text{s.t.} & \sum_{i=1}^m y_i A_i \preceq C. \end{array} \tag{DSDP}$$

The relevance of SDPs increased with the ability to solve these problems efficiently in theory and in practice. Given an $\varepsilon > 0$ we can extend most interior point methods for linear programming to polynomial time algorithms giving an ε-optimal solution for SDPs [Ali, NN]. However, the complexity to obtain solutions of an SDP is still an open question in semidefinite optimization. One of the problems is, that an SDP may have a feasible region at a double exponential distance from the origin or it may have rational outputs which require a double exponential bit length. See [Ram] for details. Further, the SDP feasibility problem SDFP, i.e. the decision problem of whether there exists a feasible solution of an SDP is in the Turing machine complexity model neither NP-complete nor co-NP-complete unless NP=co-NP [Ram, Theorem 6].

There exist several open source packages (e.g., SeDuMi [Stu], SDPT3 [TTT], SDPA [YFK]) which in practice often find efficiently ε-optimal solutions. If the problem is of medium size (i.e. $s \leq 1000$ and $m \leq 10.000$), these packages are based on interior point methods, while packages for larger semidefinite programs use some variant of the first order methods. Nevertheless, once $s \geq 3000$ or $m \geq 250000$, the problem must share some special property otherwise state-of-the-art solvers will fail to solve it for complexity reasons.

6.2 Sums of hermitian squares relaxation for global trace-optimization

In this section we propose a sum of hermitian squares and commutators relaxation for minimizing the trace of a given polynomial. We also present its dual problem and prove strong duality. Finally, we give an optimality criterion, under which the relaxed solution is the exact trace-minimum.

Let $f \in \mathcal{S}\mathbb{R}\langle \underline{X} \rangle$ be given. Since $\operatorname{Tr}(f(\underline{A})) = \operatorname{Tr}(f^*(\underline{A}))$ for all tuples $\underline{A} \in \mathcal{S}^n$ there is no harm in assuming f to be symmetric. We are interested in the *trace-infimum* of f, that is,

$$f_{\inf} := \inf\{\operatorname{Tr}\left(f(\underline{A})\right) \mid s \in \mathbb{N},\ \underline{A} \in (\mathcal{S}\mathbb{R}^{s\times s})^n\}.$$

To tackle this problem one replaces the trace-positivity condition by some simpler condition involving sums of hermitian squares and commutators, which can then be handled by semidefinite programming. This idea resembles the approach of SOS relaxation in the commutative case, which goes back to the ideas of Shor [Sho] and Nesterov [Nes] and has been invented by Lasserre [La1] and Parrilo [Par], see also [La2, PaS]. We propose the following relaxation of trace-minimization of polynomials in non-commuting variables:

$$f_{\mathrm{sos}} := \sup\{a \in \mathbb{R} \mid f - a \in \Theta^2\}. \tag{6.1}$$

Relaxation (6.1) gives obviously an lower bound on $f_{\inf}$.

6.1 Lemma. *Let $f \in \mathcal{S}\mathbb{R}\langle \underline{X} \rangle$. Then $f_{\mathrm{sos}} \le f_{\inf}$.*

In general, this relaxation is not exact, that is, we do not have equality in Lemma 6.1. For instance, the non-commutative Motzkin polynomial

$$p := M_{\mathrm{nc}} = XY^4X + YX^4Y - 3XY^2X + 1 \in \mathcal{S}\mathbb{R}\langle X, Y \rangle$$

from Example 3.5 satisfies $p_{\inf} = 0$, as it is trace-positive, and $p_{\mathrm{sos}} = \sup \varnothing := -\infty$, cf. [Lau1, Example 3.7]. Nevertheless, f_{sos} gives a solid approximation of $f_{\inf}$ for most of the examples and is easier to compute. It is obtained by solving the SDP

$$\begin{array}{ll} \sup & a \\ \text{s.t.} & f - a \in \Theta^2. \end{array} \tag{SDP$_{\min}$}$$

To see that (SDP$_{\min}$) is a semidefinite program, we reformulate (SDP$_{\min}$) with the help of Proposition 3.8. In fact, we have $f - a \in \Theta^2$ if and only if there is a positive semidefinite tracial Gram matrix G for $f - a$. Hence the problem (SDP$_{\min}$) can be written as

$$\begin{array}{llll} \sup & a & & \\ \text{s.t.} & f - a & \overset{\mathrm{cyc}}{\sim} & \mathbf{v}^* G \mathbf{v} \\ & G & \succeq & 0. \end{array} \tag{SDP$_{\min'}$}$$

The cyclic equivalence translates into a set of linear constraints, see Remark 1.7, and we obtain an SDP in primal standard form.

By Proposition 5.7, the dual cone $(\Theta^2_{n,k})^*$ of $\Theta^2_{n,k}$ is equal to the cone $\mathcal{H}_{n,k}$ of tracial Riesz functionals corresponding to truncated tracial sequences y of degree $2k$ with positive semidefinite tracial Hankel matrix of order k. Hence the SDP (SDP$_{\min}$) is equivalent to the following SDP.

$$\begin{array}{ll} \inf & L_y(f) \\ \text{s.t.} & y \in \mathcal{T}_{n,2k} \\ & M_k(y) \succeq 0. \end{array} \tag{DSDP$_{\min}$}$$

This can also be written as

$$\begin{array}{ll} \inf & L(f) \\ \text{s.t.} & L : \mathbb{R}\langle \underline{X} \rangle_{2k} \to \mathbb{R} \text{ linear } *\text{-map} \\ & L(1) = 1 \\ & L(p) \geq 0 \text{ for all } p \in \Theta^2_{n,k}. \end{array} \tag{DSDP$_{\min'}$}$$

The constraints enforce that L is a tracial state on $\mathbb{R}\langle \underline{X} \rangle_{2k}$, see Definition 1.21. We continue with the duality properties of the SDPs (SDP$_{\min}$) and (DSDP$_{\min'}$).

6.2.1 Duality

Let f^{sos} denote the optimal value of (DSDP$_{\min'}$). As for every SDP we have *weak duality*:

6.2 Lemma. $f_{\text{sos}} \leq f^{\text{sos}}$.

Proof. Let a be a feasible solution for (SDP$_{\min}$). Then any L of (DSDP$_{\min'}$) satisfies $L(f-a) \geq 0$, hence $L(f) \geq a$. ∎

In general (SDP$_{\min}$) does not satisfy the Slater condition, i.e. it does not admit a strictly feasible solution with $G \succ 0$. Nevertheless we have *strong duality*, see also [BCKP, Theorem 3.3]. The statement follows essentially from the closedness of $\Theta^2_{n,k}$ and works with the same line of reasoning as [Lau1, Theorem 6.1] from the commutative case, see also [Sw].

6.3 Theorem. $f_{\text{sos}} = f^{\text{sos}}$.

Proof. We first mention, that (DSDP$_{\min'}$) is always feasible, since the tracial linear map

$$L_1 : \mathbb{R}\langle \underline{X} \rangle_{2k} \to \mathbb{R}, f \mapsto f_1$$

satisfies the constraints in (SDP$_{\min}$) for any d. Hence $f^{\text{sos}} < \infty$. By Lemma 6.2 we always have $f_{\text{sos}} \leq f^{\text{sos}}$. To show $f_{\text{sos}} \geq f^{\text{sos}}$, suppose that (SDP$_{\min}$) is feasible, i.e. $f_{\text{sos}} > -\infty$. Since $L_y(f - f^{\text{sos}}) \geq 0$ for all truncated tracial sequences $y \in \mathcal{H}_{n,k}$, we have

$$f - f^{\text{sos}} \in \mathcal{H}_{n,k}{}^* = (\Theta^2_{n,k})^{**} = \Theta^2_{n,k}$$

by Proposition 5.7. Hence $f_{\text{sos}} \geq f^{\text{sos}}$.

If (SDP$_{\min}$) is not feasible, then (DSDP$_{\min'}$) is unbounded from below and strong duality holds as well. In fact, if (SDP$_{\min}$) is not feasible we can for any $a \in \mathbb{R}$ strictly separate $f - a$ from the closed convex cone $\Theta^2_{n,k}$ in $\mathbb{R}\langle \underline{X} \rangle_{2k}/\underset{\sim}{\text{cyc}}$.

In fact, for any $a \in \mathbb{R}$ we find a tracial linear map $L' : \mathbb{R}\langle \underline{X} \rangle_{2k}/\underset{\sim}{\text{cyc}} \to \mathbb{R}$ with $L'(p^*p) \geq 0$ for all $p \in \mathbb{R}\langle \underline{X} \rangle_k$ and $L'(f-a) < 0$. If $L'(1) > 0$ we can normalize L' to obtain a linear functional $L \in \mathcal{H}_{n,k}$ with $L(f) < a$. If $L'(1) = 0$ we replace L' by $L'' = L' + \varepsilon L_1$ with $\varepsilon := \frac{|L(f-a)|}{2}$. Then $L''(1) = \varepsilon > 0$, $L''(p^*p) \geq 0$ for all $p \in \mathbb{R}\langle \underline{X} \rangle_k$ and $L''(f-a) < 0$. Hence we can normalize L'' and obtain also a linear functional $L \in \mathcal{H}_{n,k}$ with $L(f) < a$.

This implies $f^{\text{sos}} \leq a$. Since a was arbitrary, we get that (DSDP$_{\min'}$) is unbounded and strong duality holds as well. ∎

6.2.2 Optimality

By Theorem 6.3, we have $f_{\text{sos}} = f^{\text{sos}}$. The question is, does $f_{\text{sos}} = f^{\text{sos}} = f_{\text{inf}}$ hold? If so, can we detect this using the above SDPs? In the sequel we explain how the results of Section 4.3.5 on the truncated tracial moment problem can be used to answer this question.

In fact, let f^{sos} be attained and let L^{sos} be the optimizing tracial state satisfying $L^{\text{sos}}(f) = f^{\text{sos}}$. If the truncated tracial sequence y corresponding to L^{sos} can be written as $y^{\underline{A}}$ for some n-tuple $\underline{A} \in (\mathcal{S}\mathbb{R}^{s\times s})^n$, confer Example 4.2(a), then

$$f^{\text{sos}} = L_y(f) = \operatorname{Tr}(f(\underline{A})) \geq f_{\text{inf}}.$$

Thus by Theorem 6.3 and $f_{\text{sos}} \leq f_{\text{inf}}$ follows $f^{\text{sos}} = f_{\text{sos}} = f_{\text{inf}}$ and $\underline{A}$ is a trace-minimizer of f. In particular, the trace-infimum of f is attained. This fact holds true if L^{sos} can be expressed by a cubature formula, see Definition 4.22.

6.4 Proposition. *Let f^{sos} be attained and L^{sos} be a tracial state satisfying $L^{\text{sos}}(f) = f^{\text{sos}}$. If L^{sos} of (DSDP$_{\text{min}'}$) can be expressed by a cubature formula, then the relaxation f_{sos} is exact and its points $\underline{A}^{(i)} \in (\mathcal{S}\mathbb{R}^{s\times s})^n$ are trace-minimizer.*

Proof. By assumption there exist an integer $N \in \mathbb{N}$, positive weights $\lambda_i \in \mathbb{R}_{\geq 0}$ with $\sum_i \lambda_i = 1$ and tuples $\underline{A}^{(i)} \in (\mathcal{S}\mathbb{R}^{s\times s})^n$, such that

$$L^{\text{sos}}(f) = \sum_{i=1}^{N} \lambda_i \operatorname{Tr}(f(\underline{A}^{(i)})).$$

Since $\operatorname{Tr}(f(\underline{A}^{(i)})) \geq L^{\text{sos}}(f) = f^{\text{sos}} = f_{\text{sos}}$ for each $i = 1, \ldots, N$, we get equality and hence

$$f_{\text{inf}} \leq \operatorname{Tr}(f(\underline{A}^{(i)})) = f_{\text{sos}} \leq f_{\text{inf}}.$$

Thus the trace-minimum $f_{\text{inf}} = f_{\text{sos}}$ is attained at each of the $\underline{A}^{(i)}$. ■

In general, it is difficult to decide whether L^{sos} can be expressed by a cubature formula, and to find an explicit trace-optimizer $\underline{A}$. However if the corresponding truncated tracial sequence y admits a flat extension, the proposed Θ^2-relaxation (6.1) is exact and we can even extract numerically global trace-minimizers of f. This is based on Corollary 4.41, uses the Gelfand-Naimark-Segal construction and the Artin-Wedderburn theorem. To verify numerically whether there is a flat extension one needs to implement a method to find flat extensions of matrices which also will be tracial Hankel matrices. This implementation seems to be hard. Therefore we only prove exactness of the Θ^2-relaxation for the following weaker flatness condition which is based on Theorem 4.40 and can easily be checked numerically.

6.5 Definition. Let f^{sos} be attained, L^{sos} be the optimizer of (DSDP$_{\text{min}'}$) and y its corresponding truncated tracial sequence of degree $2k$. We say that L^{sos} satisfies the *flatness condition* if the tracial Hankel matrix $M_{\text{sos}} := M_k(y)$ is flat over $M_{k-1}(y)$.

The following theorem can also be found in [BCKP, Theorem 3.11]. It resembles the appropriate optimality result in the commutative context [La2, Theorem 5.5], but without an explicit lower bound on the number of minimizers.

6.6 Theorem. *If f^{sos} is attained and the optimizer L^{sos} of (DSDP$_{\text{min}'}$) satisfies the flatness condition, then the Θ^2-relaxation (6.1) is exact: $f_{\text{sos}} = f^{\text{sos}} = f_{\text{inf}}$.*

Proof. By assumption the tracial Hankel matrix $M_{\text{sos}} = M_k(y)$ is a flat extension of $M_{k-1}(y)$, where y is the corresponding tracial sequence of L^{sos} of degree $2k$. Since $L^{\text{sos}} \in \mathcal{H}_{n,k}$ we have that M_{sos} is positive semidefinite. Thus y is a truncated tracial moment sequence by Theorem 4.40. Hence by Theorem 4.23, L^{sos} can be expressed by a cubature formula. By Proposition 6.4, it then follows that the Θ^2-relaxation is exact. ∎

Extracting trace-optimizers

For the rest of this section we assume that $f \in \mathcal{S}\mathbb{R}\langle \underline{X}\rangle_{2k}$ is such that the optimizer $L := L^{\text{sos}}$ of ($\text{DSDP}_{\min'}$) exists and satisfies the flatness condition. Let y be the corresponding truncated tracial sequence. We will now explain how to construct under this condition concrete trace-minimizing tuples $\underline{A}^{(i)}$ for f. This construction uses the same methods as needed for the proof of Theorem 4.40. This procedure has also been published in [BCKP, Section 3.3].

First, we use the Gelfand-Naimark-Segal (GNS) construction to associate a matrix $*$-algebra $\mathcal{A}$ to L. Since $M := M_k(y)$ is flat over $M_{k-1} := M_{k-1}(y)$, there exist $s = \operatorname{rank} M_{k-1}$ linear independent columns of M labelled by words $w \in \langle \underline{X}\rangle$ with $\deg w \leq k-1$ which form a basis b of $E = \operatorname{ran} M$. Now L (or M) induces a positive definite bilinear form (i.e. a scalar product) $\langle \_, \_ \rangle_E$ on E.

Let $\hat{X}_i$ be the right multiplication with X_i on E, i.e. if $\overline{w}$ denotes the column of M labelled by $w \in \langle \underline{X}\rangle_k$, then $\hat{X}_i \overline{u} := \overline{uX_i}$ for $u \in \langle \underline{X}\rangle_{k-1}$. The operator $\hat{X}_i$ is well defined and symmetric by the tracial property of L:

$$\langle \hat{X}_i \overline{p}, \overline{q}\rangle_E = L(X_i p^* q) = L(p^* q X_i) = \langle \overline{p}, \hat{X}_i \overline{q}\rangle_E.$$

Therefore we can construct matrix representations $A_i \in \mathcal{S}\mathbb{R}^{s\times s}$ of these multiplication operators $\hat{X}_i$ by calculating their image according to our chosen basis b. To be more specific, $\hat{X}_i \overline{u}_1$ for $u_1 \in \langle \underline{X}\rangle_{k-1}$ being the first label in b, can be written as a unique linear combination $\sum_{j=1}^s \lambda_j \overline{u}_j$ with words u_j labelling b such that

$$L\big((u_1 X_i - \sum \lambda_j u_j)^*(u_1 X_i - \sum \lambda_j u_j)\big) = 0.$$

Then

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_s \end{bmatrix}$$

will be the first column of A_i.

Let $\mathcal{A}$ denote the unital ($*$-)subalgebra of $\mathbb{R}^{s\times s}$ which is generated by $A_1, \ldots, A_n$.

6.7 Remark. We note that for the general case, where the optimizer L admits a flat extension, one can use, in theory, the following more abstract approach to the construction of the $\hat{X}_i$ based upon properties of flat Hankel matrices. Let $\tilde{L} : \mathbb{R}\langle \underline{X}\rangle \to \mathbb{R}$ be the linear functional corresponding to the unique flat extension $\tilde{y}$ of y, see Theorem 4.39. Since $\tilde{L}|_{\mathbb{R}\langle \underline{X}\rangle_{2k}} = L$ we write L instead of $\tilde{L}$. Equip $\mathbb{R}\langle \underline{X}\rangle$ with the bilinear form given by

$$\langle p, q\rangle := L(p^* q).$$

Let $I = \{p \in \mathbb{R}\langle \underline{X}\rangle \mid L(p^* p) = 0\}$. By Proposition 4.9, I is an ideal of $\mathbb{R}\langle \underline{X}\rangle$. Thus the vector space $E := \mathbb{R}\langle \underline{X}\rangle / I$ with the induced scalar product is a Hilbert space of dimension $\operatorname{rank} M_k(y) < \infty$. Let $\hat{X}_i$ be the right regular representation of X_i on E, i.e. $\hat{X}_i \overline{p} := \overline{pX_i}$ for $\overline{p} = p + I \in E$. The operator $\hat{X}_i$ is well defined and symmetric with respect to the scalar product induced by L. The construction of the matrices A_i and the $*$-subalgebra $\mathcal{A}$ is now similar as above.

Further, one uses the Artin-Wedderburn block decomposition of the semisimple matrix $*$-algebra $\mathcal{A}$ as in Proposition 4.18; each of the blocks obtained will yield a trace-minimizer of f.

Elements of $\mathcal{A}$ can be presented as $\hat{p} := p(A_1, \dots, A_n)$ for $p \in \mathbb{R}\langle \underline{X} \rangle$. Let $\hat{L} : \mathcal{A} \to \mathbb{R}$ be the induced linear functional given by $\hat{L}(\hat{p}) = L(p)$. By construction, $\hat{L}$ is a tracial state and hence by Proposition 4.18 is given by a conic combination of normalized traces on the Artin-Wedderburn blocks of $\mathcal{A}$. More precisely, there exist unital $*$-subalgebras $\mathcal{A}^{(i)}$ of $\mathbb{R}^{s\times s}$, each isomorphic to a full matrix algebra over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$, a $*$-isomorphism

$$\mathcal{A} \to \bigoplus_{i=1}^{N} \mathcal{A}^{(i)},$$

and $\lambda_1, \dots, \lambda_N \in \mathbb{R}_{>0}$ with $\sum_i \lambda_i = 1$, such that for all $A \in \mathcal{A}$,

$$\hat{L}(A) = \sum_{i=1}^{N} \lambda_i \operatorname{Tr}(A^{(i)}).$$

In particular,

$$L(p) = \hat{L}(\hat{p}) = \sum_{i=1}^{N} \lambda_i \operatorname{Tr}(p(A_1^{(i)}, \dots, A_n^{(i)})) \quad \text{for all } p \in \mathbb{R}\langle \underline{X} \rangle. \tag{6.2}$$

As $\operatorname{Tr}(f(A_1^{(i)}, \dots, A_n^{(i)})) \geq f_{\text{inf}} \geq L(f)$ for all $i = 1 \dots, N$, (6.2) implies

$$L(f) = \operatorname{Tr}(f(A_1^{(i)}, \dots, A_n^{(i)})).$$

That is, each of the tuples $(A_1^{(i)}, \dots, A_n^{(i)})$ is a trace-minimizer for f.

6.8 Remark. In the commutative case one can use exactly the same procedure to obtain minimizers from an optimal solution of the SOS relaxation which satisfies the flatness condition, see for instance [La2, Section 4.3].

Implementation

The first step to extract trace-optimizers is straight-forward and can easily be implemented. In the second step one has to implement the decomposition of $\mathcal{A}$ into simple components. The first efficient algorithm to decompose a semisimple algebra over a number field into simple components goes back to Friedl and Rónyai [FR]. Eberly and Giesbrecht [EG] modified their method to obtain an efficient algorithm to find the simple components of a separable algebra over an infinite field by decomposing its centre. In particular, their algorithm works for semisimple algebras over a field of characteristic 0. One can also employ the Murota, Kanno, Kojima, Kojima, and Maehara probabilistic method [MKKK, MM] which produces a unitary change of basis U for $\mathbb{R}^s$ so that the matrix $*$-algebra $\mathcal{A} \subseteq \mathbb{R}^{s\times s}$ decomposes into a direct sum of simple matrix algebras $\mathcal{A}^{(i)}$ which cannot be further decomposed.

The following example, which can also be found in [BCKP, Example 3.13], has been calculated with `NCSOStools`.

6.9 Example. Let

$$\begin{aligned} f = 3 + X^2 + 2X^3 + 2X^4 + X^6 - 4X^4Y + X^4Y^2 + 4X^3Y + 2X^3Y^2 - 2X^3Y^3 \\ + 2X^2Y - X^2Y^2 + 8XYXY + 2X^2Y^3 - 4XY + 4XY^2 + 6XY^4 - 2Y \\ + Y^2 - 4Y^3 + 2Y^4 + 2Y^6. \end{aligned}$$

The minimum of the commutative collapse $\check{f}$ of f is bounded from below by 1.0797. Using the Θ^2-relaxation one obtains the floating-point lower bound 0.2842 for the trace-infimum of f which is different from the bound for $\check{f}$. In particular, the minimizers should not be scalar matrices. The tracial Hankel matrix $M_{\text{sos}} = M_3(y)$ of the optimizer L^{sos} in ($\text{DSDP}_{\min'}$) is of rank 4 and flat over $M_2(y)$. Thus the matrix representation of the multiplication operators $\hat{X}_i$ is given by 4×4 matrices. In fact,

$$\begin{aligned}
\hat{X}_1 &= \begin{bmatrix} -1.0761 & 0.1802 & 0.5107 & 0.2590 \\ 0.1802 & -0.3393 & -0.1920 & 0.9428 \\ 0.5107 & -0.1920 & 0.5094 & 0.0600 \\ 0.2590 & 0.9428 & 0.0600 & -0.3020 \end{bmatrix}, \\
\hat{X}_2 &= \begin{bmatrix} 0.7108 & 0.7328 & 0.1043 & 0.4415 \\ 0.7328 & -0.3706 & 0.4757 & -0.2147 \\ 0.1043 & 0.4757 & 0.0776 & -0.9102 \\ 0.4415 & -0.2147 & -0.9102 & 0.1393 \end{bmatrix}.
\end{aligned}$$

The Artin-Wedderburn decomposition for the matrix $*$-algebra generated by $\hat{X}_1$ and $\hat{X}_2$ gives in this case only one block:

$$\begin{aligned}
A_1 &= \begin{bmatrix} -1.1843 & 0 & -0.2095 & 0.3705 \\ 0 & -1.1843 & 0.3705 & 0.2095 \\ -0.2095 & 0.3705 & 0.5803 & 0 \\ 0.3705 & 0.2095 & 0 & 0.5803 \end{bmatrix}, \\
A_2 &= \begin{bmatrix} -0.1743 & 0 & 0.4851 & -0.8577 \\ 0 & -0.1743 & -0.8577 & -0.4851 \\ 0.4851 & -0.8577 & 0.4529 & 0 \\ -0.8577 & -0.4851 & 0 & 0.4529 \end{bmatrix}.
\end{aligned}$$

One can easily verify that $\operatorname{Tr}(f(A_1, A_2)) = 0.2842$. Hence the solution 0.2842 is in fact the trace-minimum of f.

Note that $\mathcal{A}$ is (as a real $*$-algebra) isomorphic to $M_2(\mathbb{C})$. For instance, one can replace A_1 and A_2 with Remark 4.17 by the complex-valued matrices

$$A_1' = \begin{bmatrix} -1.1843 & 0.3705 - 0.2095\mathrm{i} \\ 0.3705 + 0.2095\mathrm{i} & 0.5803 \end{bmatrix},$$

$$A_2' = \begin{bmatrix} -0.1743 & -0.8577 + 0.4851\mathrm{i} \\ -0.8577 - 0.4851\mathrm{i} & 0.4529 \end{bmatrix}.$$

In this case it is possible to find a unitary matrix $U \in \mathbb{C}^{2\times 2}$ such that

$$A_j'' = U^T A_j' U \in \mathbb{R}^{2\times 2},$$

e.g.

$$U = \begin{bmatrix} 0.180122 - 0.0473861\mathrm{i} & 0.950143 - 0.250076\mathrm{i} \\ 0.950143 + 0.250076\mathrm{i} & -0.180122 - 0.0473861\mathrm{i} \end{bmatrix},$$

gives the real matrices

$$A_1'' = \begin{bmatrix} 0.674861 & 0.0731923 \\ 0.0731923 & -1.27886 \end{bmatrix}, \quad A_2'' = \begin{bmatrix} 0.0705101 & -1.03179 \\ -1.03179 & 0.20809 \end{bmatrix}.$$

Then $(A_1'', A_2'') \in \left(\mathcal{S}\mathbb{R}^{2\times 2}\right)^2$ is also a trace-minimizer for f.

7 Conclusion

The recent interest in positivity questions involving polynomials in non-commuting variables is predicated on the articles of McCullough [McC] and Helton [Hel] in which they proved independently that a polynomial in non-commuting variables is a sum of hermitian squares if and only if it is matrix-positive, i.e. its values in symmetric matrices of any size are positive semidefinite. The investigation of trace-positive polynomials in non-commuting variables has been established in the early 2000s due to its connection to two famous conjectures. First, Lieb and Seiringer [LiS] connected trace-positive polynomials with the BMV conjecture [BMV] from quantum statistical mechanics. Later, Klep and Schweighofer [KS1] observed that Connes' embedding conjecture on type II_1 von Neumann algebras [Con] is equivalent to a problem of describing polynomials who are trace-positive on all symmetric matrices of norm at most 1.

The approach to find representations as a sum of hermitian squares and commutators of the BMV polynomials has been investigated completely, see [Häg, KS2, Bur, LS, CDT, CKP2]. The results in Section 3.4 are part of this. The other results in Chapter 3 show some analogies of classical results from the beginning of Real Algebra in the context of trace-positive polynomials.

The investigation has mostly been reduced to polynomials in two variables. A natural extension of this would be to consider polynomials in three or more variables. Indeed, there are several open cases where it is unknown if an analogy between the classical commutative case and the tracial case holds true, cf. Section 5.1. Further, representations involving tracial quadratic modules instead of sums of hermitian squares and commutators are of interest in order to investigate polynomials which are trace-positive on a given semialgebraic set K.

The theory of positive polynomials in commuting variables is intimately connected by duality with the classical moment problem [Hav] and has played a prominent role in Real Algebra [KM, PS, Put, PV]. Motivated by this duality we introduced the tracial moment problem as dual counterpart of the investigation of trace-positive polynomials and representations as sums of hermitian squares and commutators. The same duality holds true in the free non-commutative context and it was the key ingredient to prove the famous characterization of matrix-positive polynomials [McC, Hel]. To account Connes' embedding conjecture, where one is interested in polynomials being trace-positive on a given set K, the tracial K-moment problem was considered. All results of Chapter 4 have their counterpart in the commutative context. The same holds true for the duality properties shown in Chapter 5 and the proposed application for global trace-minimization in Chapter 6.

Since there are far more results for the classical moment problem than results for the tracial moment problem presented in this work, a natural way to continue this investigation is to look for further analogies or differences between these two contexts. In particular, the tracial analog for Putinar's theorem, which was proved originally with a solution to the classical K-moment problem, would imply Connes' embedding conjecture.

Deutsche Kurzversion

Ein reelles Polynom f in nicht-kommutierenden Variablen heißt *spurpositiv*, falls alle Matrix-Auswertungen in symmetrischen Matrizen gleicher Größe positive Spur haben.

Die Theorie der spurpositiven Polynome ist eng mit tiefgehenden offenen Problemen in der Operatortheorie oder der mathematischen Physik verknüpft. Zum Beispiel ist Connes' Einbettungsvermutung über II_1 Von-Neumann-Algebren äquivalent zu dem Problem, ob man alle Polynome, dessen Spur positiv auf allen Tupeln von symmetrischen Matrizen mit Norm höchtens 1 ist, in bestimmter Weise darstellen kann. Ähnliches gilt für die Vermutung von Bessis, Moussa und Villani aus der statistischen Quantenmechanik. Diese behauptet in einer algebraischen Formulierung von Lieb und Seiringer, dass für alle $m \in \mathbb{N}_0$ und alle positiv semidefiniten Matrizen A, B gleicher Größe das Polynom

$$p(t) := \mathrm{Tr}((A + tB)^m) \in \mathbb{R}[t]$$

nur positive Koeffizienten besitzt. Mit anderen Worten, das Polynom $S_{m,k}(X^2, Y^2)$, welches den Koeffizienten von t^k in $(X^2 + tY^2)^m$ beschreibt ist spurpositiv. Diese Verbindungen motivieren die Untersuchungen der vorliegenden Arbeit und werden in Kapitel 2 näher behandelt.

Darüber hinaus implizieren spurpositive Polynome Spurungleichungen symmetrischer Matrizen, welche dann unabhängig von der Matrizengröße gelten. Um eine Spurungleichung zu verifizieren, nutzen wir die Tatsache, dass eine symmetrische Matrix genau dann eine positive Spur besitzt, wenn sie Summe einer positiv semidefiniten Matrix (d.h., ein hermitesches Quadrat von Matrizen) und einer Matrix von Spur 0 (d.h., ein Kommutator von Matrizen) ist. Die Idee, um einen systematischen Beweis von Spurungleichungen zu erhalten, besteht nun darin, Zertifikate zu finden, die eine Darstellung als Summe hermitescher Quadrate und Kommutatoren auf der Polynomebene ermöglichen. Der Einfachheit halber betrachten wir hier nur Polynome in zwei Variablen. Sei hierfür $\mathbb{R}\langle X, Y\rangle$ der Ring der reellen Polynome in den nicht-kommutierenden Variablen X, Y, versehen mit der Involution $p \mapsto p^*$, welche $X^* = X, Y^* = Y$ und $a^* = a$ für alle $a \in \mathbb{R}$ erfüllt. Diese Involution modelliert die Transposition von Matrizen auf der Polynomebene. Elemente der Form g^*g mit $g \in \mathbb{R}\langle X, Y\rangle$ heißen *hermitesche Quadrate*. Wir interessieren uns für Polynome, welche sich als Summe hermitescher Quadrate und Kommutatoren von Polynomen schreiben lassen. Anders ausgedrückt: Für welche $f \in \mathbb{R}\langle X, Y\rangle$ existieren Polynome $g_i, p_j, q_j \in \mathbb{R}\langle X, Y\rangle$, so dass $f = \sum_i g_i{}^* g_i + \sum_j (p_j q_j - q_j p_j)$ ist? Die Menge dieser Polynome bezeichnen wir mit Θ^2. Offensichtlich ist jedes Element in Θ^2 spurpositiv und induziert dadurch eine Spurungleichung. Dieses soll am folgenden Beispiel erläutert werden.

Beispiel. Für alle symmetrische Matrizen A, B gleicher Größe gilt

$$\mathrm{Tr}(A^2B^2 - ABAB) \geq 0.$$

Um dieses zu zeigen, betrachten wir das Polynom $f = X^2Y^2 - XYXY$. Da

$$\begin{aligned} f &= \frac{1}{2}\left(XY^2X + YX^2Y + XYXY + YXYX\right) \\ &\quad + \frac{1}{2}\left(XYX \cdot Y - Y \cdot XYX + X \cdot XY^2 - XY^2 \cdot X + X^2Y \cdot Y - Y \cdot X^2Y\right) \\ &= \frac{1}{2}(XY - YX)^*(XY - YX) + \text{(Summe von Kommutatoren)} \end{aligned}$$

ist, ist $f(A, B)$ für alle symmetrischen Matrizen A, B gleicher Größe stets eine Summe hermitescher Quadrate und Kommutatoren von Matrizen. Somit hat $f(A, B) = A^2B^2 - ABAB$ stets eine positive Spur.

Spurpositive Polynome liegen zwischen zwei schon gut verstandenen Polynomklassen: Einerseits Polynome in kommutierenden Variablen, die positiv auf einer semialgebraischen Menge des $\mathbb{R}^n$ sind; andererseits Polynome in nicht-kommutierenden Variablen, dessen Matrix-Auswertungen alle positiv semidefinit sind. Daher stellt man sich die folgende Frage: Welche Resultate dieser bekannten Polynomklassen gelten auch in entsprechender Weise für die Klasse der spurpositiven Polynome?

Helton und McCullough zeigten, dass Polynome, dessen Auswertungen in symmetrischen Matrizen stets positiv semidefinit sind, genau die Polynome sind, die sich als Summe hermitescher Quadrate (ohne Kommutatoren) schreiben lassen. Allerdings ist nicht jedes spurpositive Polynom ein Element von Θ^2. Beispielsweise ist folgende Variante des Motzkin-Polynoms

$$M = X^2Y^4 + X^4Y^2 - 3X^2Y^2 + 1 \in \mathbb{R}\langle X, Y\rangle$$

spurpositiv, kann aber nicht als Summe hermitescher Quadrate und Kommutatoren geschrieben werden. Dieses verhält sich analog zum kommutativen Fall: Nicht jedes positive Polynom in kommutierenden Variablen ist eine Quadratsumme von Polynomen. Daher behandeln wir spurige Versionen von Resultaten der Reellen Algebra aus dem kommutativen Kontext. Für Polynome kleinen Grades zeigen wir ein spuriges Analogon des klassischen Resultats von Hilbert über positive bivariate Quartiken.

Theorem. *Für $f \in \mathbb{R}\langle X, Y\rangle$ vom Grad höchstens vier sind äquivalent:*

(i) *f ist spurpositiv;*

(ii) *f ist spurpositiv auf symmetrischen 2×2-Matrizen;*

(iii) *f ist Summe von vier hermiteschen Quadraten und diverser Kommutatoren;*

(iv) *$f \in \Theta^2$.*

Dieses impliziert, dass eine Spurungleichung vom Grad höchstens vier in zwei symmetrischen Matrizen, welche für alle Paare symmetrischer 2×2-Matrizen gilt, stets auch für alle Paare symmetrischer $s \times s$-Matrizen mit beliebigem $s \in \mathbb{N}$ gilt. Dieses wird in Kapitel 3 behandelt. Daneben präsentieren wir Darstellungen der Polynome $S_{m,4}(X^2, Y^2)$ als Summe hermitescher Quadrate und Kommutatoren.

Theorem. *Für alle $m, r \in \mathbb{N}$ gilt $S_{m,4}(X^2, Y^2) \in \Theta^2$ und $S_{4r+2,4}(X, Y) \in \Theta^2$.*

Hieraus folgt, dass die Koeffizienten von t^k in $p(t) = \mathrm{Tr}((A+tB)^m)$, unabhängig von der Matrizengröße der positiv semidefiniten Matrizen A, B und unabhängig von der Potenz m, für $k \leq 4$ stets positiv sind. Darüber hinaus ist der Koeffizient von t^4 in $p(t)$ sogar positiv für alle symmetrischen Matrizen A, B gleicher Größe, wenn die Potenz m die Gestalt $m = 4r+2$ für ein $r \in \mathbb{N}$ hat.

Ein weiteres Thema der Arbeit ist das *spurige Momentenproblem*, welches man als duales Problem der obigen Fragestellung auffassen kann. Das Momentenproblem ist ein klassisches Problem aus der Funktionalanalysis, das wegen seiner Bedeutung und der Vielfalt seiner Anwendungen untersucht wird. Ein einfaches Beispiel ist das (univariate) Hamburger Momentenproblem: Welche Linearformen L auf den univariaten reellen Polynomen sind durch die Momente eines positiven Borelmaßes auf $\mathbb{R}$ gegeben? Ein Satz von Haviland besagt, dass dies genau dann der Fall ist, wenn L auf allen positiven Polynomen positive Werte annimmt. D.h., der Satz von Haviland verknüpft das Momentenproblem mit positiven Polynomen. Dieses gilt entsprechend für das Momentenproblem in mehreren Variablen als auch für das K-Momentenproblem, bei dem wir den Träger des Borelmaßes einschränken. Diese Dualität zwischen dem Momentenproblem und positiven Polynomen liefert beispielsweise Schmüdgens Positivstellensatz basierend auf seiner Lösung des Momentenproblems auf einer basisch abgeschlossenen kompakten semialgebraischen Menge.

In Kapitel 4 definieren wir das spurige Momentenproblem einschließlich spuriger Riesz-Funktionale und spuriger Hankel-Matrizen, die in gleicher Weise wie im klassischen Fall mit einer Folge reeller Zahlen korreliert sind. Ebenso wird das trunkierte spurige Momentenproblem sowie des spurige K-Momentenproblem behandelt. Wir zeigen einige Analogien zwischen dem klassischen Problem und seiner spurigen Version. Es gilt beispielsweise folgendes spuriges Analogon des Satzes von Haviland.

Theorem. *Sei L eine spurige Linearform auf $\mathbb{R}\langle X, Y\rangle$. Es gibt genau dann ein positives Borelmaß μ auf den symmetrischen $s \times s$-Matrizen, so dass*

$$L(w) = \int \operatorname{Tr}(w)\, d\mu$$

für alle Monome w gilt, wenn L auf allen Polynomen, die spurpositiv auf allen $s \times s$-Matrizen sind, nur positive Werte annimmt.

Eine spurige Linearform ist eine Linearform, welche zusätzlich Kommutatoren auf 0 abbildet. Diese Linearformen entsprechen spurigen Folgen, also Folgen y reeller Zahlen, indiziert durch Monome, für die zwei Werte y_u, y_v gleich sind, wenn $u - v$ eine Summe von Kommutatoren oder $u = v^*$ ist. Das spurige Momentenproblem fragt nach einer Charakterisierung der spurigen Folgen y, für die ein $s \in \mathbb{N}$ und ein Wahrscheinlichkeitsmaß μ auf den symmetrischen $s \times s$-Matrizen existiert, so dass alle Werte y_w von y durch

$$y_w = \int \operatorname{Tr}(w)\, d\mu \tag{R}$$

gegeben sind. In diesem Fall nennen wir y eine spurige Momentenfolge. Zunächst zeigen wir diverse Ergebnisse über die allgemeine Struktur der spurigen Momentenfolgen. Beispielsweise ist das trunkierte Momentenproblem allgemeiner als das (unendliche) Momentenproblem, in Analogie zum Satz von Stochel.

Theorem. *Sei y eine spurige Folge. Falls ein $s \in \mathbb{N}$ existiert, so dass für alle $k \in \mathbb{N}$ ein Wahrscheinlichkeitsmaß μ_k auf den symmetrischen $s \times s$-Matrizen existiert mit $y_w = \int \operatorname{Tr}(w)\, d\mu_k$ für alle Monome w vom Grad höchstens k, dann ist y eine spurige Momentenfolge.*

Eine andere Analogie besteht zwischen dem Satz von Bayer und Teichmann aus dem klassischen Kontext und der folgenden spurigen Version. Diese liefert ebenso wie das vorangegangene Theorem im Fall $s = 1$ direkt das kommutative Analogon.

Theorem. *Sei y eine trunkierte spurige Momentenfolge vom Grad k, die durch ein Wahrscheinlichkeitsmaß μ auf den symmetrischen $s \times s$-Matrizen, für ein $s \in \mathbb{N}$, bestimmt ist. Dann hat μ eine Kubaturformel vom Grad k, d.h., y kann durch ein Wahrscheinlichkeitsmaß mit endlichem Träger via (R) dargestellt werden.*

Eine spurige Momentenfolge erfüllt gewisse notwendige Bedingungen, welche häufig mit der zugehörigen spurigen Hankelmatrix, das spurige Analogon einer Hankelmatrix des klassischen Falls, verknüpft ist. Die spurige Hankelmatrix $M(y)$ einer spurigen Folge y ist die durch Monome indizierte Matrix

$$M(y) = [y_{u^*v}]_{u,v}.$$

Die spurige Momentenfolge $M_k(y)$ vom Grad k ist ebenso definiert, außer, dass die Indizes u, v lediglich durch Monome vom Grad höchstens k gegeben sind. Eine spurige Momentenfolge y, welche durch die spurigen Momente eines Wahrscheinkeitsmaßes μ auf den symmetrischen $s \times s$-Matrizen bestimmt ist, erfüllt zwingend folgende Eigenschaften:

(i) $M(y) \succeq 0$,

(ii) $\operatorname{rank} M(y) \leq |\operatorname{supp} \mu| s^2$.

Diese notwendigen Bedingungen sind im Allgemeinen jedoch nicht hinreichend. Deshalb präsentieren wir einige Bedingungen, welche eine Darstellung (R) implizieren. Konkret zeigen wir spurigen Versionen der Resultate von Curto und Fialkow. Für das (unendliche) Momentenproblem ergibt sich folgendes Theorem.

Theorem. *Sei y eine spurige Folge. Hat y eine positiv semidefinite spurige Hankelmatrix $M(y)$ von endlichem Rang, so ist y eine spurige Momentenfolge.*

Im Falle des trunkierten spurigen Momentenproblems ähnelt die hinreichende Bedingung für eine Darstellung (R) ebenfalls einer bekannten Bedingung im klassischen Fall.

Theorem. *Sei y eine trunkierte spurige Folge vom Grad $2k$. Falls $M_k(y)$ positiv semidefinit ist und* $\operatorname{rank} M_k(y) = \operatorname{rank} M_{k-1}(y)$ *gilt, so ist y eine trunkierte spurige Momentenfolge.*

Alternativ können die spurigen Riesz-Funktionale verwendet werden, um hinreichende Bedingungen zu erhalten. Wir zeigen, dass eine trunkierte spurige Folge eine trunkierte spurige Momentenfolge ist, wenn das zugehörige spurige Riesz-Funktional eine positive spurige Erweiterung besitzt. Dieses verallgemeinert ein Ergebnis von Curto und Fialkow aus dem klassischen Fall. In Analogie zu Ergebnissen von Fialkow und Nie gilt außerdem:

Theorem. *Sei y eine trunkierte spurige Folge vom Grad k. Wenn das zugehörige spurige Riesz-Funktional auf allen spurpositiven Polynomen vom Grad höchstens k, die sich nicht als Summe von Kommutatoren schreiben lassen, nur strikt positive Werte annimmt, dann ist y eine trunkierte spurige Momentenfolge.*

Fialkow und Nie nutzten den klassischen Satz von Hilbert und die Dualität zwischen positiven Polynomen und dem Momentenproblem, um das bivariate trunkierte Momentenproblem vom Grad vier partiell zu lösen. Diese Dualität erstreckt sich auch auf den spurigen Fall, welche dann folgendes Resultat impliziert.

Theorem. *Für jedes $k \in \mathbb{N}_0$ sind folgende Aussagen äquivalent:*

(i) *Jedes spurpositive Polynom vom Grad $2k$ liegt in Θ^2;*

(ii) *Jede trunkierte spurige Folge vom Grad $2k$ mit positiv definiter spuriger Hankelmatrix ist eine trunkierte spurige Momentenfolge.*

Dieses wird in Kapitel 5 behandelt, wo außerdem vorangegangene Ergebnisse mittels konvexer Kegel ausgedrückt und besagte Dualität bewiesen wird.

In Kapitel 6 kombinieren wir verschiedene Ergebnisse aus den vorherigen Kapiteln, um eine Anwendung unserer Theorie vorzustellen. Die Frage, ob ein gegebenes Polynom als Summe hermitescher Quadrate und Kommutatoren geschrieben werden kann, kann numerisch durch ein semidefinites Programm beantwortet werden. Dieses basiert auf einem Analogon der Gram-Matrix-Methode. Ein Polynom f ist genau dann eine Summe hermitescher Quadrate und Kommutatoren, wenn es einen Vektor $\mathbf{v}$ von Monomen und eine positiv semidefinite Matrix G gibt, so dass

$$f = \mathbf{v}^* G \mathbf{v} + \text{Summe von Kommutatoren}$$

ist. Dieses lässt sich leicht als semidefinites Programm schreiben. Daher ersetzen wir das Optimierungsproblem

$$f_{\inf} := \inf\{\operatorname{Tr}\big(f(\underline{A})\big) \mid \underline{A} \text{ Tupel symmetrischer Matrizen}\},$$

welches das Spur-Infimum von f sucht, durch folgendes, einfach zu berechnendes Optimierungsproblem
$$f_{\text{sos}} := \sup\{a \in \mathbb{R} \mid f - a \in \Theta^2\}.$$
Dieses liefert eine Relaxierung des ursprünglichen Problems, welche im Allgemeinen zwar eine untere Schranke jedoch nicht die exakte Lösung liefert. Um nun herauszufinden, ob $f_{\text{sos}} = f_{\text{inf}}$ gilt, wird das entsprechende duale Problem betrachtet. Falls f_{sos} angenommen wird und die duale Lösung eine gewisse Bedingung, welche aus der Theorie des spurigen Momentenproblems kommt, erfüllt, so ist die Relaxierung exakt. Darüber hinaus wird gezeigt, wie in diesem Fall aus der dualen Lösung konkrete globale Spur-Minimierer mit Methoden aus Kapitel 4 extrahiert werden können.

Version abrégée en français

Un polynôme réel en des variables non commutatives a une *trace positive* si toutes ses évaluations en des matrices symétriques ont une trace positive.

La théorie des polynômes à trace positive est profondément liée à des problèmes ouverts d'algèbres d'opérateurs et de physique mathématique. En fait, la conjecture de plongement de Connes sur les algèbres de von Neumann de type II_1 est équivalente la description de l'ensemble des polynômes qui ont une trace positive sur des matrices de norme au plus 1. En outre, la conjecture de Bessis, Moussa et Villani, dans une formulation algébrique de Lieb et Seiringer, dit que pour tout $m \in \mathbb{N}_0$ et toutes matrices A, B semi-définies positives, le polynôme

$$p(t) := \mathrm{Tr}((A + tB)^m) \in \mathbb{R}[t]$$

n'a que des coefficients positifs. En d'autres termes, les polynômes $S_{m,k}(X^2, Y^2)$ en des variables non commutatives, qui décrivent les coefficient de t^k en $(X^2 + tY^2)^m$, ont une trace positive. Ces connexions sont la motivation principale de ce travail et seront expliquées en plus détail au chapitre 2.

Un autre objectif de l'étude des polynômes à trace positive est de donner des inégalités de trace concernant les matrices symétriques. Nous proposons une approche autonome de la dimension pour atteindre des inégalités de trace, c'est-à-dire fournissant des certificats convenant indépendamment de la taille des matrices.

Pour vérifier des inégalités de trace, nous utilisons le fait qu'une matrice a une trace positive si et seulement si elle est une somme d'une matrice semi-définie positive (soit un carré hermitien de matrices) et d'une matrice de trace nulle (soit un commutateur de matrices). L'idée principale pour systématiser la vérification des inégalités de trace est de chercher des certificats portant sur des sommes de carrés hermitiens et de commutateurs au niveau des polynômes. Pour simplifier, on ne considère ici que deux variables. Soit $\mathbb{R}\langle X, Y\rangle$ l'anneau des polynômes en deux variables X, Y non commutatives à coefficients réels muni de l'involution $p \mapsto p^*$ avec $X^* = X, Y^* = Y$ et $a^* = a$ pour tous $a \in \mathbb{R}$, modelant la transposée des matrices. Les élements de la forme g^*g avec $g \in \mathbb{R}\langle X, Y\rangle$ sont des *carrés hermitiens*. Nous nous intéressons aux polynômes qui peuvent être écrits comme une somme de carrés hermitiens et de commutateurs de polynômes. Autrement dit, pour quels $f \in \mathbb{R}\langle X, Y\rangle$ existe-t-il des polynômes $g_i, p_j, q_j \in \mathbb{R}\langle X, Y\rangle$ tels que $f = \sum_i g_i{}^* g_i + \sum_j (p_j q_j - q_j p_j)$? Soit Θ^2 l'ensemble de ces polynômes. Évidemment, tout polynôme en Θ^2 a une trace positive, donc il induit une inégalité de trace. Explicitons cette idée sur exemple simple.

Exemple. Pour toutes matrices symétriques A, B de même taille, nous avons

$$\mathrm{Tr}(A^2B^2 - ABAB) \geq 0.$$

Pour le montrer, considérons le polynôme $f = X^2Y^2 - XYXY$. Comme f peut être écrit

$$\begin{aligned} f &= \frac{1}{2}\big(XY^2X + YX^2Y + XYXY + YXYX\big) \\ &\quad + \frac{1}{2}\big(XYX \cdot Y - Y \cdot XYX + X \cdot XY^2 - XY^2 \cdot X + X^2Y \cdot Y - Y \cdot X^2Y\big) \\ &= \frac{1}{2}(XY - YX)^*(XY - YX) + (\text{somme de commutateurs}), \end{aligned}$$

$f(A, B)$ est une somme de carrés hermitiens et de commutateurs des matrices, pour toutes les matrices symétriques A, B de même taille. Ainsi la trace de $f(A, B) = A^2B^2 - ABAB$ est toujours positive.

Les polynômes à trace positive se situent entre deux classes de polynômes bien étudiées. D'un côté, les polynômes en des variables commutatives qui sont positifs sur un ensemble semi-algébrique de $\mathbb{R}^n$. De l'autre côté, les polynômes en des variables non commutatives dont toutes les évaluations en des matrices symétriques sont semi-définies positives. Par conséquent, la question qui se pose naturellement est la suivante : quels résultats sur ces deux classes des polynômes sont aussi valables pour les polynômes à trace positive ?

Montré par Helton et McCullough, les polynômes dont toutes les évaluations en des matrices symétriques sont semi-définies positives sont exactement les sommes de carrés hermitiens (sans commutateurs). D'autre part, les polynômes à trace positive ne sont pas tous des somme de carrés hermitiens et de commutateurs. Par exemple, la version suivante du polynôme de Motzkin

$$M = X^4Y^2 + X^2Y^4 - 3X^2Y^2 + 1 \in \mathbb{R}\langle X, Y\rangle$$

a une trace positive, mais il ne peut pas être écrit comme la somme de carrés hermitiens et de commutateurs. C'est analogue à ce qui se passe dans le cas commutatif : les polynômes positifs en des variables commutatives ne sont pas tous somme de carrés. Par conséquent, nous chercherons des analogues pour les polynômes à trace positive des résultats classiques en algèbre réelle pour les polynômes positifs. Pour les polynômes de petit degré nous établissons un résultat tracial, analogue du résultat classique de Hilbert sur les polynômes quartiques binaires qui sont positifs.

Théorème. *Soit $f \in \mathbb{R}\langle X, Y\rangle$ de degré 4 . Alors, les propositions suivantes sont équivalentes :*

(i) *f a une trace positive ;*

(ii) $\mathrm{Tr}(f(A, B)) \geq 0$ *pour toutes matrices A, B symétriques de taille 2×2 ;*

(iii) *f est une somme de quatre carrés hermitiens et de certains commutateurs ;*

(iv) $f \in \Theta^2$.

En outre, cela implique que toutes les inégalités de trace de degré quatre en deux matrices symétriques valant pour toutes les matrices symétriques de taille 2×2 sont également valables pour n'importe quelle paire de matrices symétriques de taille $s \times s$ avec un $s \in \mathbb{N}$ arbitraire. Ce sera traité dans le chapitre 3. En plus, nous présentons des représentations des polynômes $S_{m,4}(X^2, Y^2)$ comme une somme de carrés hermitiens et de commutateurs.

Théorème. *Pour tout $m, r \in \mathbb{N}$, il est $S_{m,4}(X^2, Y^2) \in \Theta^2$ et $S_{4r+2,4}(X, Y) \in \Theta^2$.*

Ce qui induit que, indépendamment de la taille de matrices A, B semi-définies positives et indépendamment de la puissance m, les coefficients de t^k dans $p(t) = \mathrm{Tr}((A + tB)^m)$ sont positifs pour tout $k \leq 4$. En particulier, nous en déduisons que les coefficients de t^4 dans $p(t)$ sont positifs pour tout choix de matrices A, B symétriques de même taille si la puissance m est de la forme $m = 4r + 2$.

Par dualité on obtient le problème des moments traciaux, un autre sujet principal de cette thèse. Ce problème est une question classique en analyse fonctionnelle, qui est très étudiée en raison de son importance et de la variété de ses applications. Un exemple simple est le problème des moments (en une variable) de Hamburger : quelles formes linéaires L sur les polynômes univariés réels sont les moments d'une mesure de Borel μ positive ? Par le théorème de Haviland, on sait que c'est le cas si et seulement si L est positive sur tous les polynômes positifs sur $\mathbb{R}$. Ainsi, le théorème de Haviland concerne le problème des moments et les polynômes positifs. Il est aussi valable en plusieurs variables et lorsque le support de μ est réduit à un ensemble fermé. La dualité entre le problème des moments et les polynômes positifs a été appliquée, par exemple, à la solution de Schmüdgen du problème des moments sur un ensemble semi-algébrique compact, impliquant le Positivstellensatz de Schmüdgen.

Dans le chapitre 4 nous définissons le problème des moments traciaux, y compris des formes de Riesz traciales et des matrices traciales de Hankel, qui sont reliées à une suite réelle de la même manière que dans le cas classique. Le problème tronqué des moments traciaux, où l'on ne considère que les suites finies, ainsi que l'analogue tracial du problème des K-moments sont également étudiés. Nous montrons plusieurs analogies qui existent entre le problème des moments classique et sa version traciale. Par exemple, un analogue tracial du théorème de Haviland est vrai.

Théorème. *Soit L une forme linéaire traciale sur $\mathbb{R}\langle X, Y\rangle$. Alors il existe une mesure de Borel μ positive sur les matrices symétriques de taille $s \times s$ telle que pour tous les monômes w on ait*

$$L(w) = \int \mathrm{Tr}(w)\, d\mu$$

si et seulement si L ne prend que des valeurs positives sur tous les polynômes qui ont une trace positive sur tous les vecteurs de matrices symétriques de taille $s \times s$.

Une forme linéaire traciale est une forme linéaire qui envoie des commutateurs sur nulle. Elle correspond à une suite traciale, une suite de nombres réels, indexée par des monômes, invariante par permutation cyclique des indices. Le *problème des moments traciaux* cherche une caractérisation des suites traciales y pour lesquelles il existe un entier $s \in \mathbb{N}$ et une mesure de probabilité μ sur les matrices symétriques de taille $s \times s$ tels que toute valeur y_w de y peut s'écrire comme

$$y_w = \int \mathrm{Tr}(w)\, d\mu. \tag{R}$$

Ces suites sont appelées suites de moments traciaux. Nous présentons des résultats sur la structure générale des suites de moments traciaux. Par exemple, nous montrons que le problème tronqué est plus général que le problème (infini) des moments traciaux par analogie avec le théorème de Stochel.

Théorème. *Soit y une suite traciale. S'il existe $s \in \mathbb{N}$ tel que pour tout $k \in \mathbb{N}$ il existe une mesure μ_k sur les matrices symétriques de taille $s \times s$ telle que $y_w = \int \mathrm{Tr}(w)\, d\mu_k$ pour tout monôme w de degré au plus k, alors y est une suite de moments traciaux.*

En outre on a aussi un analogue tracial du théorème classique de Bayer et Teichmann, qui est, comme la théorème précédente, pour $s = 1$ exactement la version classique.

Théorème. *Soit y une suite tronquée de degré k de moments traciaux avec une mesure μ de probabilité sur les matrices symétriques de taille $s \times s$, pour un certain $s \in \mathbb{N}$. Alors, la mesure μ a une formule de cubature de degré k, c'est-à-dire que y peut être représentée par* (R) *avec une mesure à support fini.*

Une suite de moments traciaux satisfait certaines conditions nécessaires, analogues au cas classique, souvent en rapport avec sa matrice traciale de Hankel. La matrice traciale de Hankel $M(y)$ d'une suite traciale y est la matrice

$$M(y) = [y_{u^*v}]_{u,v},$$

indexée par les monômes u, v. La matrice $M_k(y)$ des moments traciaux de degré k est définie de manière similaire, mais elle est indexée par les monômes u, v de degrés au plus k. Une suite y des moments traciaux avec une représentation utilisant une mesure de probabilité sur les matrices symétriques de la taille $s \times s$ satisfait

(i) $M(y) \succeq 0$,

(ii) $\operatorname{rank} M(y) \leq |\operatorname{supp} \mu| s^2$.

Ces conditions nécessaires ne suffisent pas en général. C'est pourquoi nous présentons également quelques conditions qui impliquent une représentation (R). Nous présentons les analogues traciaux des résultats classiques de Curto et Fialkow sur les matrices de Hankel. Pour le problème (infini) des moments traciaux nous avons le théorème suivant.

Théorème. *Soit y une suite traciale. Alors y est une suite de moments traciaux si sa matrice traciale de Hankel est semi-définie positive et de rang fini.*

Pour le problème tronqué, l'existence d'un représentation (R) ressemble aussi à la situation classique.

Théorème. *Soit y une suite traciale tronquée de degree $2k$. Si $M_k(y)$ est semi-définie positive et $\operatorname{rank} M_k(y) = \operatorname{rank} M_{k-1}(y)$, alors y et une suite tronquée de moments traciaux.*

En outre, les formes traciales de Riesz peuvent être utilisées, comme dans le cas commutatif, pour obtenir des conditions suffisantes à une suite traciale y pour avoir une représentation (R). En effet, si la forme L_y de Riesz d'une suite traciale tronquée admet une extension traciale positive, alors la suite traciale tronquée a une telle représentation. Ceci est aussi un analogue tracial d'un résultat de Curto et Fialkow. En plus, si la forme traciale de Riesz est strictement positive, la suite traciale tronquée est une suite tronquée de moments traciaux, par analogie avec le théorème de Fialkow et Nie.

Théorème. *Soit y une suite traciale tronquée de degré k. Si sa forme traciale de Riesz L_y ne prend que des valeurs strictement positives sur tous les polynômes qui ont une trace positive et qui ne sont pas des sommes de commutateurs, alors y est une suite tronquée de moments traciaux.*

Dans un autre contexte, le théorème classique de Hilbert a été utilisé par Fialkow et Nie pour résoudre dans une certaine mesure le problème des moments en deux variables de degré au plus quatre. La dualité entre les polynômes positifs et le problème des moments s'étend au cas tracial. En particulier, l'équivalence suivante est vérifiée.

Théorème. *Pour tout $k \in \mathbb{N}_0$ les assertions suivantes sont équivalentes :*

(i) *Tous les polynômes de degré $2k$ à trace positive sont des éléments de Θ^2 ;*

(ii) *Toutes les suites traciales tronquées de degré $2k$ avec une matrice traciale de Hankel définie positive sont des suites tronqués de moment traciaux.*

Ceci est géré dans le chapitre 5. Ce chapitre résume également les précédents résultats en termes de cônes convexes et montre la dualité des sommes de carrés hermitiens et de commutateurs et du problème des moment traciaux dans le cadre de la dualité conique.

Dans le chapitre 6, nous combinons plusieurs résultats des chapitres précédents afin de donner une application de notre théorie. La question est de savoir si un polynôme donné peut être écrit comme une somme de carrés hermitiens et de commutateurs ; on peut y répondre numériquement par un algorithme utilisant la programmation semi-définie. Ceci est basé sur un analogue de la méthode de Gram. Un polynôme f est une somme de carrés hermitiens et de commutateurs si et seulement s'il y a un vecteur $\mathbf{v}$ de monômes et une matrice G semi-définie positive tels que

$$f = \mathbf{v}^* G \mathbf{v} + \text{une somme de commutateurs.}$$

Nous appliquons cette méthode et son dual. En effet, le problème d'optimisation

$$f_{\text{inf}} := \inf\{\operatorname{Tr}\big(f(\underline{A})\big) \mid \underline{A} \text{ vecteur de matrices symétriques}\},$$

qui cherche l'infimum de la trace d'un polynôme donné sur tous les vecteurs de matrices symétriques de même taille pourra être affaibli en le problème d'optimisation

$$f_{\text{sos}} := \sup\{a \in \mathbb{R} \mid f - a \in \Theta^2\}.$$

Bien que cet affaiblissement ne soit pas toujours exact, il est facile à calculer et fournit une borne. Pour tester si $f_{\text{sos}} = f_{\text{inf}}$, on étudie le programme dual semi-défini. S'il satisfait une certaine condition, qui est directement liée au problème des moment traciaux, alors l'affaiblissement est vrai. Dans ce cas, nous montrons comment on peut extraire des optimiseurs globaux de la trace du polynôme donné, via une procédure fondée sur les méthodes du chapitre 4.

References

[Akh] N.I. Akhiezer, The classical moment problem and some related questions in analysis, Hafner Publishing Co., 1965

[Ali] F. Alizadeh, *Interior Point Methods in Semidefinite Programming with Applications to Combinatorial Optimization*, SIAM Journal on Optimization, vol. 5 (1993) 13–51

[Bar] A. Barvinok, A course in convexity, Graduate Studies in Mathematics, vol. 54 (2002), Amer. Math. Soc.

[BT] C. Bayer and J. Teichmann, *The proof of Tchakaloff's theorem*, Proc. Amer. Math. Soc. 134, no. 10 (2006) 3035–3040

[Ber] S. Berberian, Lectures in Functional Analysis and Operator Theory, Springer, 1973

[BMV] D. Bessis, P. Moussa and M. Villani, *Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics*, J. Math. Phys. 16, no. 11 (1975) 2318–2325

[BCR] J. Bochnak, M. Coste and M.-F. Roy, Real algebraic geometry, Ergebnisse der Mathematik und ihre Grenzgebiete, Bd. 36. (1998), Springer

[Boz] M. Bożejko, *Bessis-Moussa-Villani conjecture and generalized Gaussian random variables*, Infinite Dim. Analysis, Quantum Probability and Related Topics, 11 (2008), 313–321

[Bur] S. Burgdorf, *Sums of Hermitian squares as an approach to the BMV conjecture*, Linear and Multilinear Algebra vol.59 no. 1 (2011), 1–9

[BCKP] S. Burgdorf, K. Cafuta, I. Klep and J. Povh, *Semidefinite programming certificates for tracial matrix inequalities*, preprint, 2010, available from http://www.optimization-online.org/DB_HTML/2010/04/2595.html

[BK1] S. Burgdorf and I. Klep, *The truncated tracial moment problem*, to appear in J. Operator Theory, 2010, available from http://arxiv.org/abs/1001.3679

[BK2] S. Burgdorf and I. Klep, *Trace-positive polynomials and the quartic tracial moment problem*, C. R. Math. Acad. Sci. Paris, vol. 348, no. 13-14 (2010) 721–726

[CKP1] K. Cafuta, I. Klep and J. Povh, *NCSOStools: a computer algebra system for symbolic and numerical computation with noncommutative polynomials*, available from http://ncsostools.fis.unm.si/

[CKP2] K. Cafuta, I. Klep and J. Povh, *On the nonexistence of sum of squares certificates for the BMV conjecture*, J. Math. Phys., vol. 51, 083521 (2010)

[CD] B. Collins and K.J. Dykema, *A linearization of Connes' embedding problem*, New York J. Math. vol. 14 (2008) 617–641

[CDT] B. Collins, K.J. Dykema and F. Torres-Ayala, *Sum-of-squares results for polynomials related to the Bessis–Moussa–Villani conjecture*, J. Stat. Phys., 139, no. 5 (2010) 779–799

[Con] A. Connes, *Classification of injective factors. Cases* II_1, II_∞, III_λ, $\lambda \neq 1$, Ann. Math. 104, no. 1 (1976) 73–115

[CF1] R.E. Curto and L.A. Fialkow, *Solution of the truncated complex moment problem for flat data*, Mem. Amer. Math. Soc. 119, no. 568 (1996)

[CF2] R.E. Curto and L.A. Fialkow, *Flat extensions of positive moment matrices: recursively generated relations*, Mem. Amer. Math. Soc. 136, no. 648 (1998)

[CF3] R.E. Curto and L.A. Fialkow, *An analogue of the Riesz-Haviland theorem for the truncated moment problem*, J. Funct. An., vol. 255, no. 10 (2008) 2709–2731

[CL] M. D. Choi and T. Y. Lam, *Extremal positive semidefinite forms*, Math. Ann. 231, no. 1 (1977/78) 1–18

[CLR] M. D. Choi, T. Y. Lam and B. Reznick, *Sums of squares of real polynomials*, K-theory and algebraic geometry: connections with quadratic forms and division algebras, Proc. Sympos. Pure Math., vol. 58 (1995) 103–126, Amer. Math. Soc.

[DST] M. Drmota, W. Schachermayer and J. Teichmann, *A hyper-geometric approach to the BMV-conjecture*, Monatshefte für Mathematik, vol. 146 (2005) 179–201

[EG] W. Eberly and M. Giesbrecht, *Efficient decomposition of separable algebras*, Journal of Symbolic Computation, vol. 37 (2004) 35–81

[FP] M. Fannes and D. Petz, *Perturbation of Wigner matrices and a conjecture*, Proc. Amer. Math. Soc. 131, no. 7 (2003) 1981–1988

[FN] L. Fialkow and J. Nie, *Positivity of Riesz functionals and solutions of quadratic and quartic moment problems*, J. Funct. An., vol. 258, no. 1 (2010) 328–356

[FF] C. Fleischhack and S. Friedland, *Asymptotic Positivity of Hurwitz Product Traces: Two Proofs*, Linear Algebra Appl., vol. 432, no. 6 (2010) 1363–1383

[FR] K. Friedl and L. Rónyai, *Polynomial time solutions of some problems in computational algebra*, Symp. on Theory of Computing, vol. 17 (1985), 153–162, Amer. Math. Soc.

[Had] D. Hadwin, *A noncommutative moment problem*, Proc. Amer. Math. Soc. 129, no. 6 (2001) 1785–1791

[Häg] D. Hägele, *Proof of the cases $p \leq 7$ of the Lieb-Seiringer formulation of the Bessis-Moussa-Villani conjecture*, J. Stat. Phys. 127, no. 6 (2007) 1167–1171

[Hav] E.K. Haviland, *On the momentum problem for distribution functions in more than one dimension II*, Amer. J. Math. 58, no. 1 (1936) 164–168

[Hel] J.W. Helton, *"Positive" non-commutative polynomials are sums of squares*, Ann. of Math. Second Series 156, no. 2 (2002) 675–694

[HM] J. Helton and S. McCullough, *A positivstellensatz for non-commutative polynomials*, Trans. Amer. Math. Soc. 356 (2004) 3721–3737

[HdOMS] J.W. Helton, M.C. de Oliveira, R.L. Miller and M. Stankus, NCAlgebra, 2010 release edition, available from `http://www.math.ucsd.edu/~ncalg/`

[HL] D. Henrion and J.B. Lasserre, *Detecting global optimality and extracting solutions in GloptiPoly*, Positive polynomials in control, vol. 312 (2005), Lecture Notes in Control and Inform. Sci., 293–310, Springer

[HLL] D. Henrion, J.-B. Lasserre and J. Löfberg, *GloptiPoly 3: moments, optimization and semidefinite programming*, Optimization Methods and Software, vol. 24, nos. 4–5 (2009), 761–779, available from `http://www.laas.fr/~henrion/software/gloptipoly3/`

[Hilb] D. Hilbert, *Über die Darstellung definiter Formen als Summe von Formenquadraten*, Math. Ann. 32 (1888) 342–50

[Hil] C.J. Hillar, *Advances on the The Bessis-Moussa-Villani Trace Conjecture*, Linear Algebra Appl., vol. 426, no. 1 (2007) 130–142

[HJ1] C.J. Hillar and C.R. Johnson, *Eigenvalues of Words in Two Positive Definite Letters*, SIAM. J. Matrix Anal. & Appl. vol. 23, no. 4 (2002) 916–928

[HJ2] C.J. Hillar and C.R. Johnson, *On the positivity of the coefficients of a certain polynomial defined by two positive definite matrices*, J. Stat. Phys. 118, no. 3–4 (2005) 781–789

[Kir] E. Kirchberg, *On non-semisplit extensions, tensor products and exactness of group C^*-algebras*, Invent. Math. 112 (1993) 449–489

[KP] I. Klep and J. Povh, *Semidefinite programming and sums of hermitian squares of noncommutative polynomials*, J. Pure Appl. Algebra 214 (2010) 740–749

[KS1] I. Klep and M. Schweighofer, *Connes' embedding conjecture and sums of hermitian squares*, Adv. Math. 217, no. 4 (2008) 1816–1837

[KS2] I. Klep and M. Schweighofer, *Sums of hermitian squares and the BMV conjecture*, J. Stat. Phys. 133, no. 4 (2008) 739–760

[KMRT] M.-A. Knus, A.S. Merkurjev, M. Rost and J.-P. Tignol, The Book of Involutions, Coll. Pub., vol. 44 (1998), Amer. Math. Soc.

[KN] M.G. Kreĭn and A.A. Nudel′man, The Markov moment problem and extremal problems, Translations of Mathematical Monographs, vol. 50 (1977), Amer. Math. Soc.

[KM] S. Kuhlmann and M. Marshall, *Positivity, sums of squares and the multi-dimensional moment problem*, Trans. Amer. Math. Soc. 354, no. 11 (2002) 4285–4301

[Lam] T.Y. Lam, A first course in noncommutative rings, Graduate Texts in Mathematics, vol. 131 (1991), Springer

[LS] P.S. Landweber and E.R. Speer, *On D. Hägele's approach to the Bessis-Moussa-Villani conjecture*, Linear Algebra Appl. 431, no. 8 (2009) 1317–1324

[La1] J.B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim. vol. 11, no. 3 (2000/2001) 796–817

[La2] J.B. Lasserre, Moments, Positive Polynomials and Their Applications, Imperial College Press Optimization Series, vol. 1 (2010)

[Lau1] M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, Emerging Applications of Algebraic Geometry, vol. 149 (2009) 157–270, IMA Volumes in Mathematics and its Applications, Springer

[Lau2] M. Laurent, *Revisiting two theorems of Curto and Fialkow on moment matrices*, Proc. Amer. Math. Soc. 133, no. 10 (2005) 2965–2976

[LiS] E.H. Lieb and R. Seiringer, *Equivalent forms of the Bessis-Moussa-Villani conjecture*, J. Stat. Phys. 115, no. 1–2 (2004) 185–190

[Lof] J. Löfberg, *YALMIP: A Toolbox for Modeling and Optimization in MATLAB*, Proceedings of the CACSD Conference (2004), Taipei, Taiwan, available from http://users.isy.liu.se/johanl/yalmip/

[MM] T. Maehara and K. Murota, *A Numerical Algorithm for Block-Diagonal Decomposition of Matrix $*$-Algebras, Part II: General Algorithm*, Japan J. Indust. Appl. Math., vol. 27 no. 2 (2010)263–293

[Mar] M. Marshall, Positive polynomials and sums of squares, Mathematical Surveys and Monographs, vol. 146 (2008), Amer. Math. Soc.

[McC] S. McCullough, *Factorization of operator-valued polynomials in several non-commuting variables*, Linear Algebra Appl., vol. 326, no. 1-3 (2001) 193–203

[MP] S. McCullough and M. Putinar, *Noncommutative sums of squares*, Pacific J. Math. 218, no. 1, (2005) 167–171

[Mot] T.S. Motzkin, *The arithmetic–geometric inequality*, Inequalities (1967) 205–224, Academic Press

[Mou] P. Moussa, *On the representation of Tr$(e^{A-\lambda B})$ as a Laplace transform*, Rev. Math. Phys. 12 (2000) 621–655

[MKKK] K. Murota, Y. Kanno, M. Kojima and S. Kojima, *A numerical algorithm for block-diagonal decomposition of matrix $*$-algebras with application to semidefinite programming*, Japan J. Indust. Appl. Math., vol. 27, no.1 (2010) 125–160

[MvN1] F.J. Murray and J. von Neumann, *On Rings of Operators*, Ann. of Math. Second Series, vol. 37, no. 1 (1936) 116–229

[MvN2] F.J. Murray and J. von Neumann, *On Rings of Operators IV*, Ann. of Math. Second Series, vol. 44, no. 4 (1943) 716–808

[Nes] Y.E. Nesterov, *Squared functional systems and optimization problems*, High Performance Optimization (2000) 405–440, Kluwer Academic Publishers

[NN] Y.E. Nesterov and A. Nemirovski, Interior Point Polynomial Methods in Convex Programming, Studies in Applied Mathematics, vol. 13 (1994), SIAM

[Par] P.A. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*, Math. Program. Series B, vol. 96, no. 2 (2003) 293–320

[PaS] P.A. Parrilo and B. Sturmfels, *Minimizing polynomial functions*, Algorithmic and quantitative real algebraic geometry, DIMACS Ser. Discrete Math. Theor. Comput. Sci., vol. 60 (2003) 83–99, Amer. Math. Soc.

[PNA] S. Pironio, M. Navascues and A. Acin, *Convergent relaxations of polynomial optimization problems with non-commuting variables*, SIAM Journal on Optimization, vol. 20, no. 5 (2010) 2157–2180

[Pól] G. Pólya, *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen*, Acta Mathematica 68, no. 1 (1937) 145–254

[PRSS] V. Powers, B. Reznick, C. Scheiderer and F. Sottile, *A new approach to Hilbert's theorem on ternary quartics*, C. R. Math. Acad. Sci. Paris vol. 339, no. 9 (2004) 617–620

[PS] V. Powers and C. Scheiderer, *The moment problem for non-compact semialgebraic sets*, Adv. Geom. 1, no. 1 (2001) 71–88

[PD] A. Prestel and C.N. Delzell, Positive polynomials. From Hilbert's 17th problem to real algebra, Springer Monogr. Math., 2001

[Put] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Indiana Univ. Math. J. 42, no.3 (1993) 969–984

[PV] M. Putinar and F.-H. Vasilescu, *Solving moment problems by dimensional extension*, Ann. of Math. Second Series, vol. 149, no. 3 (1999) 1087–1107

[R1] F. Rădulescu, *Convex sets associated with von Neumann algebras and Connes' approximate embedding problem*, Math. Res. Lett. 6 (1999) 229–236

[R2] F. Rădulescu, *The von Neumann algebra of the non-residually finite Baumslag group $\langle a, b | ab^3a^{-1} = b^2\rangle$ embeds into R^ω*, Hot topics in operator theory, Theta Ser. Adv. Math. 9 (2008) 173–185.

[R3] F. Rădulescu, *A non-commutative, analytic version of Hilbert's 17th problem in type II_1 von Neumann algebras*, Von Neumann algebras in Sibiu, Theta Ser. Adv. Math. 10 (2008) 93–101

[Ram] M.V. Ramana, *An exact duality theory for semidefinite programming and its complexity implications*, Math. Program. Series B, vol. 77, no. 2 (1997) 129–162

[Rez] B. Reznick, *Some concrete aspects of Hilbert's* 17*th Problem*, Contemp. Math., 253 (2000) 251–272

[Rie] M. Riesz, *Sur le problème des moments*, 3éme note, Ark. Mat. 17 (1923) 1–52

[Rud] W. Rudin, Real and Complex Analysis, McGraw-Hill, 1987

[Sch] K. Schmüdgen, *The K-moment problem for compact semi-algebraic sets*, Math. Ann. 289, no. 2 (1991) 203–206

[Sw] M. Schweighofer, *Optimization of polynomials on compact semialgebraic sets*, SIAM Journal on Optimization vol. 15, no. 3 (2005) 805-825

[ST] J.A. Shohat and J.D. Tamarkin, The problem of moments, Amer. Math. Soc. Surveys II, 1943

[Sho] N.Z. Shor, *An approach to obtaining global extremums in polynomial mathematical programming problems*, Kibernetika vol. 5 (1987) 102–106

[Sto] J. Stochel, *Solving the truncated moment problem solves the full moment problem*, Glasg. Math. J., vol. 43, no. 3 (2001) 335–341

[Stu] J. Sturm, *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones*, Optim. Meth. Softw. vol. 11/12, no. 1–4 (1999) 625–653

[Tak] M. Takesaki, Theory of Operator Algebras I, Springer, 1979

[TTT] K.-C. Toh, M.J. Todd and R.-H. Tütüncü, *SDPT3 – a Matlab software package for semidefinite programming*, Optimization Methods and Software, 11 (1999) 545–581 available from http://www.math.nus.edu.sg/~mattohkc/sdpt3.html

[Voi] D. Voiculescu, *The analogues of entropy and of Fisher's information measure in free probability theory II*, Invent. Math. 118 (1994) 411-440

[vN1] J. von Neumann, *Zur Algebra der Funktionaloperationen und Theorie der normalen Operatoren*, Math. Ann. vol. 102, no. 1 (1930) 370-427

[vN2] J. von Neumann, *On Rings of Operators. Reduction Theory*, Ann. of Math. Second Series, vol. 50, no. 2 (1949) 401–485

[YFK] M. Yamashita, K. Fujisawa and M. Kojima, *Implementation and evaluation of SDPA 6.0 (semidefinite programming algorithm 6.0)*, Optim. Methods Softw., vol. 18, no. 4 (2003) 491–505, availabe from http://sdpa.indsys.chuo-u.ac.jp/sdpa/