universität
innsbruck

Fakultät für Mathematik,
Informatik und Physik

Quantifier Elimination in Matrix Algebras

Master thesis
submitted for the degree of Diplom-Ingenieur

submitted by
Mag. Phil. Clemens Brüser

supervised and assessed by
Univ.-Prof. Dr. Tim Netzer

Institut für Mathematik
Fakultät für Mathematik, Informatik und Physik
Leopold-Franzens-Universität Innsbruck

Freiburg and Innsbruck, January 2022

Abstract

This thesis is based on an article by Igor Klep and Marcus Tressl proving that under certain conditions the matrix algebra $M_n(F)$ admits quantifier elimination if the underlying field F does so too. Namely, they found that if one extends the language in which F admits quantifier elimination by a tuple of constants interpreted as matrix units, then $M_n(F)$ admits quantifier elimination. If we restrict the class of fields considered, it is instead sufficient to add two unary function symbols for the trace and the transposition to get quantifier elimination.

In the first chapter, these findings will be replicated. Proofs that are omitted or kept short in the original paper will be spelled out in full detail whenever reasonably feasible. In a subsequent chapter, we will use the results to find a quantifier-free characterisation of positive (semi-)definite real or complex matrices as well as a quantifier-free criterion for invertibility of real or complex matrices. These considerations make use of the so-called Newton Identities, which will be applied to the characteristic polynomial of a matrix. Two further applications of these results give a description of contractive maps and criteria for the solubility of certain matrix equations, among them the well-known Sylvester Equation.

These examples will also allow us to come to the conclusion that while quantifier elimination exists in both $M_k(F)$ and $M_n(F)$ for $1 \leq k < n$, we cannot generally say that these results are dimensionally compatible, that is: Assuming that for a given formula φ we have found quantifier-free formulae φ^n and φ^k satsifying

$$\begin{aligned} M_n(F) \vDash \varphi[h] &\iff M_n(F) \vDash \varphi^n[h] \\ M_k(F) \vDash \varphi[h] &\iff M_k(F) \vDash \varphi^k[h] \end{aligned}$$

then there will be evaluations h in $M_k(F)$ such that

$$M_k(F) \vDash \varphi^k[h] \iff M_k(F) \vDash \varphi^n[h]$$

does not hold.

Clemens Brüser
January 2022

Dedication

Est enim amicitia nihil aliud nisi omnium divinarum humanarumque rerum cum benevolentia et caritate consensio.
Indeed, friendship is nothing less than sharing a common attitude towards all things divine and human, together with benevolence and kindness.
Cicero: Laelius de Amicitia, 20.

I extend my heartfelt gratitude to several people who have accompanied me on my academic way that has led to this thesis.

First among them is my supervisor, Prof. Dr. Tim Netzer. It was he, who originally nourished my interest for the fascinating field of algebra, and he always lent me an ear whenever I had run into difficulties pertaining to my thesis. Many errors both small and large would have gone unnoticed without his keen eye.

While studying, I have met many interesting people and found good friends, who broadened my horizons and supported me in writing this thesis. They also kept me sane during the Corona-related lockdowns of the years 2020 and 2021. To all of you - and to Julia in particular - I say: Thank you!

Finally, my family - Winfried, Annette, Florian, Johanna and Lukas - deserve unreserved gratitude for their unwavering support, which extends way beyond my academic life. While I wish I could, it is impossible to express these feelings in words confined to this short page.

Contents

1 Introduction

The concept of quantifier elimination is a powerful property that a mathematical theory (in the sense of model theory) can exhibit. Given any formula containing any finite number of quantifiers $\forall, \exists$, it will guarantee that there is a different, yet equivalent, formula that does not make use of these quantifiers. A well-known example for such corresponding formulae over the real numbers is given by

$$(\exists x : x^2 + px + q = 0) \longleftrightarrow (p^2 - 4q \geq 0)$$

where the right hand side obviously is free of quantifiers. If we were to consider the same formula over the complex numbers, we would even have the equivalence

$$(\exists x : x^2 + px + q = 0) \longleftrightarrow 0 = 0$$

by the fundamental theorem of algebra. Again, the right hand side does not use any quantifiers.

It is a standard result of real algebra and geometry that the real numbers - or more general: The theory of real closed fields in the language of rings extended by the ordering on the field - admits quantifier elimination. This is also known as the transfer principle of Tarski-Seidenberg. Similarly, the theory of algebraically closed fields - and hence the complex numbers - admits quantifier elimination.

Naturally, one may ask how these results can be generalised or applied to new settings. This marks the entrance of matrix algebras. We are interested in whether a matrix algebra $M_n(F)$ over a field F admits quantifier elimination provided that F does. The answer is not trivial and we will follow [KT20] in answering it. First, we will prove that we get quantifier elimination in $M_n(F)$ if we equip the language of rings with a tuple of new constant symbols - so-called matrix units. These are best thought of as the standard matrix units $(E_{ij})_{i,j=1}^n$. We will then proceed to give a similar result, replacing the matrix units by two unary functions that will take the place of the trace and the transposition (or involution respectively). This additionally requires us to place further restrictions on the field F that we consider.

In a later chapter, we will prove that these results about quantifier elimination in $M_n(F)$ cannot generally be applied to $M_k(F)$ for $k < n$. Neither is it possible to embed $M_k(F)$ into $M_n(F)$ to save the results derived for the n-dimensional case.

The main part of this thesis comprises a collection of formulae that exemplify the power of quantifier elimination in matrix algebras. We will derive quantifier-free formulae for the following questions commonly found in settings of algebra.

- Is a given matrix $A \in M_n(F)$ positive (semi-)definite?
- Is a given matrix $A \in M_n(F)$ invertible?
- Is the linear mapping defined by a given matrix $A \in M_n(F)$ a contraction?
- Given $A, B \in M_n(F)$, does the linear matrix equation $AX + B = 0$ have a solution?
- Given $A, B \in M_n(F)$, under which circumstances is there a unique solution to Sylvester's Equation $AX - XB = C$ for all $C \in M_n(F)$?

In every case, F may either denote the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. We end with an outlook on what further directions of study may follow from the considerations in this thesis.

2 Preliminaries

This chapter will feature definitions and results that will be used frequently within the thesis. Due to the minor role that it plays in terms of developing and presenting new ideas, not all of the statements presented will be proven rigorously. Whenever that is the case, we will refer to relevant literature. Unless indicated otherwise, all statements are given in the author's own formulations.

As a last remark on this chapter, it should be pointed out that the results below are loosely connected by their appearance later in the thesis alone. Other than that, they are to be regarded as mostly independent and thus they will not allow for coherent reading.

2.1 Algebraic Preliminaries

We begin with a standard result of non-commutative algebra that states that the center of a matrix ring is isomorphic to its underlying ring. This result will be frequently used and we will not reference it at every occurrence.

Lemma 2.1. *Let R be a ring. Then the center C of $M_n(R)$ is isomorphic to R.*

Proof. We show that C equals $R \cdot I_n$ and use the standard identification $R \leftrightarrow R \cdot I_n$. First, we assume that $A \in R \cdot I_n$. It is immediately clear that A lies in C. Now we conversely assume that $A \in C$, yet $A \notin R \cdot I_n$. Then we distinguish two cases.

- Case 1: Not all diagonal elements of A are equal.

  We choose $i \neq j$ such that $A_{ii} \neq A_{jj}$ and observe that $A = T_{ij}^2 A \overset{A \in C}{=} T_{ij} A T_{ij} \neq A$, where T_{ij} shall denote the elementary matrix that switches the i-th and the j-th row (or column respectively). This is a contradiction.

- Case 2: There are indices $i, j \in \{1, ..., n\}$ with $i \neq j$ and $0 \neq A_{ij}$.

We note that $0 = E_{ii}E_{jj}A \overset{A \in C}{=} E_{ii}AE_{jj} = A_{ij}E_{ij} \neq 0$. This, too, is an immediate contradiction.

□

Another well-known set of theorems that is central to the study of matrices and their induced linear maps are the spectral theorems. The version given below will be used frequently and is a rephrased version of [Bos14], Theorem 6, p. 280.

Theorem 2.2 (Spectral theorem for hermitian matrices). *Let $A \in M_n(\mathbb{C})$ be a hermitian matrix, that is $A^* = A$. Then all eigenvalues of A are real and there exists a unitary matrix $P \in M_n(\mathbb{C})$ such that $P^*AP = D$ where D is a real diagonal matrix, the diagonal entries of which are the eigenvalues of A.*

If A is a real matrix, then P may be chosen real as well, such that P is an orthogonal matrix.

Proof. See [Bos14], Theorem 6, p. 280. □

The spectral theorem is one example how a base change might yield a particularly fruitful matrix form. Two other crucial results in this area are the well-known Schur form and the Jordan canonical form. While these matrix representations are usually deduced for complex matrices, there are real versions as well. We cite all statements in shortened versions of the results in [HJ13].

Theorem 2.3 (Schur form, complex version, [HJ13], Theorem 2.3.1.). *Let $A \in M_n(\mathbb{C})$ have eigenvalues $\lambda_1, \dots, \lambda_n$ in any prescribed order and let $x \in \mathbb{C}^n$ be a unit vector such that $Ax = \lambda_1 x$.*

There is a unitary $U = (x, u_2, \dots, u_n) \in M_n(\mathbb{C})$ such that $U^*AU = T = (t_{ij})_{i,j=1}^n$ is upper triangular with diagonal entries $t_{ii} = \lambda_i$, $i = 1, \dots, n$.

Proof. See [HJ13], Theorem 2.3.1. □

Theorem 2.4 (Schur form, real version, [HJ13], Theorem 2.3.4.). *Let $A \in M_n(\mathbb{R})$ be given. There is a real orthogonal $Q \in M_n(\mathbb{R})$ such that Q^TAQ is a real upper quasitri-*

angular matrix

$$\begin{pmatrix} A_1 & & & * \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_m \end{pmatrix}$$

such that each A_i is 1-by-1 or 2-by-2 with the following properties:

1. *Its 1-by-1 diagonal blocks display the real eigenvaules of A.*

2. *Each of its 2-by-2 diagonal blocks has a conjugate pair of non-real eigenvalues.*

3. *The ordering of its diagonal blocks may be prescribed in the following sense: If the real eigenvalues and conjugate pairs of non-real eigenvalues of A are listed in a prescribed order, then the real eigenvalues and conjugate pairs of non-real eigenvalues of the respective diagonal blocks $A_1, \ldots, A_m$ of $Q^T AQ$ are in the same order.*

Proof. See [HJ13], Theorem 2.3.4. □

Theorem 2.5 (Jordan canonical form, complex version, [HJ13], Theorem 3.1.11.)**.** *Let $A \in M_n(\mathbb{C})$ be given. There is a nonsingular $S \in M_n(\mathbb{C})$, positive integers q and $n_1, \ldots, n_q$ with $n_1 + n_2 + \ldots + n_q = n$, and scalars $\lambda_1, \ldots, \lambda_q \in \mathbb{C}$ such that*

$$A = S \begin{pmatrix} J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_q}(\lambda_q) \end{pmatrix} S^{-1}$$

where $J_{n_k}(\lambda_k)$ denotes the Jordan block of size n_k corresponding to the eigenvalue λ_k of A, that is:

$$J_{n_k}(\lambda_k) = \begin{pmatrix} \lambda_k & 1 & \ldots & 0 \\ 0 & \lambda_k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \ldots & 0 & \lambda_k \end{pmatrix} \in M_{n_k}(\mathbb{C})$$

Proof. See [HJ13], Theorem 3.1.11. □

Theorem 2.6 (Jordan canonical form, real version, [HJ13], Theorem 3.4.15.)**.** *Each $A \in M_n(\mathbb{R})$ is similar via a real similarity to a real block diagonal matrix of the form*

$$C_{n_1}(a_1, b_1) oplus \ldots \oplus C_{n_p}(a_p, b_p) \oplus J_{m_1}(\mu_1) oplus \ldots \oplus J_{m_r}(\mu_r)$$

in which $\lambda_k = a_k + ib_k$, $k = 1, 2, \ldots, p$ are non-real eigenvalues of A, each a_k and b_k is real and $b_k > 0$, and $\mu_1, \ldots, \mu_r$ are real eigenvalues of A. Each real block triangluar matrix $C_{n_k}(a_k, b_k) \in M_{2n_k}$ is of the form

$$C_k(a,b) = \begin{pmatrix} C(a,b) & I_2 & & & \\ & C(a,b) & I_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & I_2 \\ & & & & C(a,b) \end{pmatrix}$$

with

$$C(a,b) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

and corresponds to a pair of conjugate Jordan blocks $J_{n_k}(\lambda_k), J_{n_k}(\overline{\lambda_k}) \in M_{n_k}$ with non-real λ_k in the Jordan canonical form of A. The real Jordan blocks $J_{m_r}(\mu_r)$ are the Jordan blocks in the Jordan canonical form that have real eigenvalues.

Proof. See [HJ13], Theorem 3.4.15. □

The following theorem will occur only once. The tools required to prove it go well beyond what is reasonable within the context of this thesis, which is why we omit the proof.

Definition 2.7. Let A be an F-algebra and let C be its center. A is called a central simple algebra, if $C = F$ and there are no non-trivial ideals of A.

Theorem 2.8 (Skolem-Noether, [Bre14], Theorem 4.46.)**.** *Let A be a finite dimensional central simple algebra. If S is a simple sub-algebra of A that contains the unity 1 of A, then every homomorphism φ from S into A that maps 1 into 1 can be extended to an inner automorphism of A, that is there is $x \in A$ satisfying*

$$\varphi(a) = xax^{-1}$$

for all $a \in A$.

Proof. See [Bre14], Theorem 4.46. □

The most interesting example of a central simple algebra in this thesis is the matrix algebra $M_n(F)$. The following lemma proves that it is in fact a central simple algebra.

Lemma 2.9. *Let F be a field. Then $M_n(F)$ is a central simple algebra.*

Proof. By Lemma 2.1, $M_n(F)$ is central. In order to show that $M_n(F)$ is simple, we follow [Bre14], Example 1.10. For that, let I be a non-zero ideal in $M_n(F)$ and let $0 \neq A \in I$. Then there exist $j, k \in \{1, ..., n\}$ such that $A_{jk} \neq 0$. In particular,

$$E_{ij} A E_{kl} = A_{jk} E_{il} \in I$$

for all $i, l \in \{1, ..., n\}$. Hence,

$$(x A_{ij}^{-1} E_{ij}) A E_{kl} = x E_{il} \in I$$

for all $x \in F$ and thus, $I = M_n(F)$. □

2.1.1 Real Closed Fields

The following results aim to provide the reader with some basic ideas of the theory of real closed fields. The statements should be regarded as a generalisation of the real numbers. For an extended discussion, see [PD01].

Definition 2.10 ([PD01], Definition 1.1.1)**.** Let F be a field. A relation $\leq$ on F^2 is called an ordering on F, if $\leq$ is a linear ordering on F and additionally it satisfies

$$\begin{aligned} a \leq b &\Longrightarrow a + c \leq b + c \\ 0 \leq a, 0 \leq b &\Longrightarrow 0 \leq ab \end{aligned}$$

for all $a, b, c \in F$. If $\leq$ is an ordering on F, then $(F, \leq)$ is called an ordered field.

The following lemma will not explicitly be used later. However, it is useful in motivating the definition of real fields.

Lemma 2.11 ([PD01], Definition 1.1.6 and Lemma 1.1.7)**.** *Let F be a field and let $T \subseteq F$ be a subset satisfying*

$$-1 \notin P \qquad P + P \subseteq P \qquad P \cdot P \subseteq P \qquad \sum F^2 \subseteq P \qquad P \cup -P = F$$

Then P defines an ordering by the relation $a \leq b :\Longleftrightarrow b - a \in P$. Every ordering on F arises this way and there is a bijection between these two notions of ordering.

Proof. We roughly follow [PD01], Lemma 1.1.7. If $P \subseteq F$ satisfies the conditions above, then we immediately get the following for any $a, b, c \in F$:

- Reflexivity: $a \leq a$, since $a - a = 0 \in P$ by $P \cup -P = F$.

- Anti-Symmetry: Assume that $a \leq b$ and $b \leq a$. Then $a - b, b - a \in P$. If now $a-b \neq 0$, then without loss of generality $\frac{1}{a-b} \in P$ (since otherwise $-\frac{1}{a-b} = \frac{1}{b-a} \in P$) and thus $-1 = \frac{b-a}{a-b} \in P$, a contradiction. Hence, $a = b$ must hold.

- Transitivity: Assume that $a \leq b$ and $b \leq c$. Then $c - b, b - a \in P$ and hence also $(c - b) + (b - a) = c - a \in P$, so $a \leq c$.

- Linearity: By assumption, $a - b \in P$ or $b - a \in P$. This immediately implies linearity.

- Assume that $a \leq b$. Then $(b + c) - (a + c) = b - a \in P$, which implies that $a + c \leq b + c$.

- Assume that $0 \leq a, b$. Then $a, b \in P$ and as an immediate consequence $ab \in P$, which yields $0 \leq ab$.

Now let $\leq$ be a given ordering on F. Then it is clear that $P := \{a \in F | 0 \leq a\}$ satisfies all of the properties in the lemma. Bijectivity of the two constructions is also clear. □

Definition 2.12 ([PD01], Corollary 1.1.12)**.** A field F is called real, if it can be equipped with an ordering.

Lemma 2.13 ([PD01], Definition 1.1.11)**.** *A field F is real if and only if $-1 \notin \sum F^2$.*

Proof. First not that $T := \sum F^2$ satisfies all conditions of Lemma 2.11 but $T \cup -T = F$. However, any field with such a subset T may be equipped with an ordering by [PD01], Theorem 1.1.9.

If conversely, $-1 \in \sum F^2$, then we cannot find an ordering on F by Lemma 2.11. □

We are particularly interested in so-called real closed fields as they admit quantifier elimination. We first give a definition and then introduce a nice characterisation.

Definition 2.14 ([PD01], Definition 1.2.8)**.** A field R is called real closed, if R is real and there is no proper algebraic field extension F of R that is real.

Theorem 2.15 ([PD01], Theorem 1.2.10)**.** *The following statements are equivalent:*

1. *R is a real closed field.*
2. *K^2 defines an ordering in the sense of Lemma 2.11 and every polynomial $p \in K[x]$ of odd degree has a root in K.*
3. *$R \neq R[\sqrt{-1}]$ and $R[\sqrt{-1}]$ is an algebraically closed field.*

Proof. See [PD01], Theorem 1.2.10. This draws on results of Galois theory and generalises the proof of the fundamental theorem of algebra. □

2.2 Model Theoretic Preliminaries

For the discussions in the next chapters, it is relevant to give some model theoretic preliminaries. It is assumed that the reader is acquainted with basic notations and results of mathematical logic and model theory. As a general reference, we refer to [PD11]. In the following, we will only present statements that are not contained there or which are of great importance.

The first result is a simple statement that gives a handy characterisation of all models of a theory.

Lemma 2.16. *Let A be a structure and let T be the theory of A. Then the structure B is a model of T if and only if $A \equiv B$.*

Proof. $\Longrightarrow$: First note that T is complete by [PD11], p. 48, meaning that for all sentences α, we either have $\alpha \in T$ or $\neg\alpha \in T$. If B is a model of T, then from $A \vdash \alpha$, we conclude that $\alpha \in T$ and thus $B \vdash \alpha$. If $A \not\vdash \alpha$, then $A \vdash \neg\alpha$ and in an analogous fashion as before, we conclude $B \vdash \neg\alpha$, which is equivalent to $B \not\vdash \alpha$. Hence, $A \equiv B$.

$\Longleftarrow$: This direction is trivial. □

In the following discussions, we will often restrict the underlying language of a structure. The next lemma proves that these restrictions do not affect certain relation between two structures in the initial language.

Lemma 2.17. *Let $\mathcal{L}$ be a language and let $\tilde{\mathcal{L}}$ extend $\mathcal{L}$. Further let $\tilde{A}, \tilde{B}$ be two $\tilde{\mathcal{L}}$-structures and let A, B denote the restrictions of $\tilde{A}, \tilde{B}$ to $\mathcal{L}$. Then the following statements hold.*

1. *If $\tilde{A} \equiv \tilde{B}$, then $A \equiv B$.*
2. *If $\tilde{A} \cong \tilde{B}$, then $A \cong B$.*
3. *If $\tilde{A} \preceq \tilde{B}$, then $A \preceq B$.*

Proof. Let φ be an $\mathcal{L}$-formula. Then we may in particular interpret φ as an $\tilde{\mathcal{L}}$-formula. We now turn to our statements.

1. For elementary equivalence, we only need to consider the case that φ is a sentence and we note that clearly, $\tilde{A} \vDash \varphi$ if and only if $A \vDash \varphi$, since A and $\tilde{A}$ have the same universe. The same holds for B. Using elementary equivalence in $\tilde{\mathcal{L}}$, we now note that

   $$A \vDash \varphi \iff \tilde{A} \vDash \varphi \overset{\tilde{A} \equiv \tilde{B}}{\iff} \tilde{B} \vDash \varphi \iff B \vDash \varphi$$

   This proves the claim.

2. The argument works by the observation that $\tilde{A} \vDash \varphi[h]$ if and only if $A \vDash \varphi[h]$ for any evaluation h in A, since A and $\tilde{A}$ have the same universe. Using the same argument for B instead of A, we may then continue in an analogous fashion to the previous statement.

3. This works by analogous arguments as in the previous statements.

□

As our main results will focus on quantifier elimination and model completeness, we require criteria for these properties. They are captured in the following central theorems.

Theorem 2.18 ([PD11], Lemma 3.3.1.)**.** *$\Sigma \subseteq Sent(L)$ is model complete if and only if, for any two models A, B of Σ with $A \subseteq B$, we even have $A \preceq B$.*

Proof. See [PD11], Lemma 3.3.1. □

Theorem 2.19 ([PD11], Theorem 3.4.3.)**.** *Let T be an $\mathcal{L}$-theory. Then T admits quantifier elimination if and only if T is model complete and T has the amalgamation property over substructures.*

Proof. See [PD11], Theorem 3.4.3. □

Remark 2.20. In the above setting, we may work with finitely generated substructures instead, that is: T admits quantifier elimination if and only if T is model complete and T has the amalgamation property over finitely generated substructures. To see this, it is sufficient to note that in the proof of theorem 3.4.3. in [PD11], we may always work with a finitely generated substructure without having to change anything else.

Throughout this thesis, we will use that both real closed fields (in particular, $\mathbb{R}$) and algebraically closed fields (in particular, $\mathbb{C}$) admit quantifier elimination. For reference purposes, we write down this result in the theorem below.

Theorem 2.21. *The theory of real closed fields and the theory of algebraically closed fields admit quantifier elimination.*

Proof. See [PD01], Theorem 2.1.6. for real closed fields and [PD11], Theorem 3.4.4. for algebraically closed fields. □

We have now laid out all necessary preliminaries in order to present the first main result of this thesis - quantifier elimination in matrix algebras.

3 Quantifier Elimination in Matrix Algebras

In this chapter, we will present the key findings of [KT20]. All statements that are labelled to be taken from them and the corresponding proofs will rely on their work, though we will fill in some details. Note that contrary to their approach, we will always consider unital rings, that is: The language of rings is defined as $\mathcal{L}_{Ri} := \{+, \cdot, -, 0, 1\}$. Ring homomorphisms will not be required to be unital.

3.1 Quantifier Elimination using Matrix Units

In the first lemma, we start by collecting some properties that are inspired by the standard matrix units $(E_{ij})_{i,j=1}^n$.

Definition 3.1. Let A be a (not necessarily commutative) ring and let C_A denote its center. A subfield $F \subseteq A$ is called a central subfield, if $F \subseteq C_A$.

Lemma 3.2 ([KT20], 2.1.1.)**.** *Let A be a ring, $n \in \mathbb{N}$ and for $i, j \in \{1, ..., n\}$ let $a_{ij} \in A$. Suppose that for all $i, j, s, t \in \{1, ..., n\}$, we have*

$$a_{ij}a_{st} = \left\{ \begin{array}{ll} a_{it} & j = s \\ 0 & else \end{array} \right\} = \delta_{js}a_{it}$$

Then the following hold:

1. *For all $i, j, s, t \in \{1, ..., n\}$, we have $a_{ss}a_{ij}a_{tt} = \delta_{is}\delta_{jt}a_{ij}$.*

2. *If $a_{ij} = 0$ for some $i, j \in \{1, ..., n\}$, then $a_{st} = 0$ for all $s, t \in \{1, ..., n\}$. If, however, $a_{st} \neq 0$ for all $s, t \in \{1, ..., n\}$, then the family $(a_{ij})_{i,j=1}^n$ is F-linearly independent for any central subfield F of A.*

3. *Let F be a central subfield of A. If $(x_{ij})_{i,j=1}^n, (y_{ij})_{i,j=1}^n \in M_n(F)$, then we have*

$$\left(\sum_{i,j=1}^n x_{ij}a_{ij}\right) \cdot \left(\sum_{i,j=1}^n y_{ij}a_{ij}\right) = \sum_{i,j=1}^n \left(\sum_{k=1}^n x_{ik}y_{kj}\right) a_{ij}$$

4. *If F is a central subfield of A, then the map*

$$\begin{aligned} \varphi : M_n(F) &\to A \\ (x_{ij})_{i,j=1}^n &\mapsto \sum_{i,j=1}^n x_{ij}a_{ij} \end{aligned}$$

is an F-algebra-homomorphism. If $a_{ij} \neq 0$ for all $i,j \in \{1,...,n\}$, then φ is injective.

Proof. 1. Note that by assumption, we have

$$a_{ss}a_{ij}a_{tt} = \delta_{si}a_{sj}a_{tt} = \delta_{is}\delta_{jt}a_{st} = \delta_{is}\delta_{jt}a_{ij}$$

for all $i,j,s,t \in \{1,...,n\}$, from which the statement follows immediately.

2. First, we note that if $a_{ij} = 0$ for some choice of $i,j \in \{1,...,n\}$, then for all $s,t \in \{1,...,n\}$, we get

$$a_{st} = a_{si}a_{ij}a_{jt} = 0$$

Now we assume that $a_{ij} \neq 0$ for all $i,j \in \{1,...,n\}$ and that $F \subseteq A$ is a central subfield. We further assume that we have a linear combination adding up to 0, that is

$$\sum_{i,j=1}^n f_{ij}a_{ij} = 0$$

for certain $f_{ij} \in F$. Multiplying with a_{ss} and a_{tt} and using that F is central, we observe that

$$0 = a_{ss}\left(\sum_{i,j=1}^n f_{ij}a_{ij}\right) a_{tt} = \sum_{i,j=1}^n f_{ij}a_{ss}a_{ij}a_{tt} = f_{st}a_{st}$$

for any $s,t \in \{1,...,n\}$, where we used 1 in the last step. Assuming that $f_{s,t} \neq 0$, we get an immediate contradiction, since $a_{st} \neq 0$ and $f_{st} \in F$ is invertible. We conclude that $f_{st} = 0$, which implies F-linear independence.

3. The statement follows from the commuting properties of x_{ij} and y_{ij} with $i,j \in \{1,...,n\}$ and the following observation:

$$\begin{aligned}\left(\sum_{i,j=1}^{n} x_{ij}a_{ij}\right)\cdot\left(\sum_{i,j=1}^{n} y_{ij}a_{ij}\right) &= \sum_{i,j,k,l=1}^{n} x_{ij}a_{ij}y_{kl}a_{kl} = \\ \overset{F\text{ central}}{=} \sum_{i,j,k,l=1}^{n} x_{ij}y_{kl}\underbrace{a_{ij}a_{kl}}_{=\delta_{jk}a_{il}} &= \sum_{i,k,l=1}^{n} x_{ik}y_{kl}a_{il} = \\ = \sum_{i,j,k=1}^{n} x_{ik}y_{kj}a_{ij} &= \sum_{i,j=1}^{n}\left(\sum_{k=1}^{n} x_{ik}y_{kj}\right)a_{ij}\end{aligned}$$

4. F-linearity is clear, multiplicativity follows from part 3. Assuming that all a_{ij} are non-zero, part 2 yields that $\varphi(x) = 0$ if and only if $x = 0$ and thus, φ is injective.

□

Remark 3.3. If in the above setting, $n \geq 2$ and $a_{ij} \in C_A$ for some $i,j \in \{1,...,n\}$, then $a_{st} = 0$ for all $s,t \in \{1,...,n\}$. If indeed $a_{st} \neq 0$ for some choice of $s,t \in \{1,...,n\}$, then $a_{st} \neq 0$ for all $s,t \in \{1,...,n\}$ by 2. We may then conclude

$$0 \neq a_{it} = a_{ij}a_{jt} = a_{jt}a_{ij} = \delta_{it}a_{jj} \implies i = t$$

for arbitrary t. This is a clear contradiction.

Example 3.4 ([KT20], 2.1.1.)**.** φ as above need not be unital. Take, for example, any field F, set $n = 1, m \geq 2$ and $A := M_m(F)$. Now choose $a = a_{11} \in A\backslash\{0, I_m\}$ such that $a^2 = a$, e.g. $a = E_{11}$.

Now note that F naturally embeds into A by the identification $[c \mapsto c \cdot I_m]$. Obviously, $F \subseteq C_A$. Setting $\varphi(x) := xa$ defines an F-algebra-homomorphism since $\varphi(xy) = xya = xya^2 = xaya = \varphi(x)\varphi(y)$. However, since $\varphi(1) = a \in A\backslash\{0, I_m\}$, φ is not unital.

Example 3.5. Choosing $A = M_n(F)$ and $a_{ij} = E_{ij}$ satisfies all four properties of Lemma 3.2. Also note that by the Lemma 2.1, $C_A \cong F$. This example may be regarded as prototypical. If indeed, $a_{ij} \neq 0$ for $i,j \in \{1,...,n\}$ are given and F is a given central subfield of A, then we define $A' := span_F(a_{ij}|i,j \in \{1,...,n\})$. By Lemma 3.2 4 , we get that

$$\begin{aligned} M_n(F) &\to A' \\ E_{ij} &\mapsto a_{ij}\end{aligned}$$

defines an F-algebra isomorphism. Then $A' \subseteq A$ is itself a ring. In many cases, this lets us restrict our considerations to $A = M_n(F)$. Note that the 1 in A' need not coincide with the 1 in A.

Our goal is to formalise the results of Lemma 3.2, which motivates the formulae of the following definition. If given as below, ε defines a tuple $(a_{ij})_{i,j}$ of matrix units, δ defines the elements commuting with all a_{ij}, giving a notion of center elements. λ defines the map previously denoted by φ (not necessarily as an algebra homomorphism) and γ describes the whole set of properties from Lemma 3.2. In particular, φ will once again be a homomorphism.

Definition 3.6 ([KT20], 2.1.2.)**.** Let F be a field and define $M := M_n(F)$ for some $n \in \mathbb{N}$. Then the center $C := C_n$ of M_n is isomorphic to F by Lemma 2.1. We may embed $F \cong C \hookrightarrow M_N(C)$ for all $N \in \mathbb{N}$ and thus interpret $M_N(C)$ both as a subset of M^{N^2} and as an F-algebra.

We now take $2N^2 + 2$ variables

$$\overline{u} := (u_{ij} | i, j \in \{1, \dots, N\}), \overline{x} := (x_{ij} | i, j \in \{1, \dots, N\}), v, y$$

and define the following $\mathcal{L}_{Ri}$-formulae:

$$\varepsilon := \varepsilon_N(\overline{u}) \qquad \text{by} \qquad \bigwedge_{i,j,t=1}^{N} u_{ij}u_{jt} = u_{it} \neq 0 \wedge \bigwedge_{\substack{i,j,s,t=1 \\ j \neq s}}^{N} u_{ij}u_{st} = 0$$

$$\delta := \delta_N(v, \overline{u}) \qquad \text{by} \qquad \bigwedge_{s,t=1}^{N} vu_{st} = u_{st}v$$

$$\lambda := \lambda_N(\overline{x}, y, \overline{u}) \qquad \text{by} \qquad y = \sum_{i,j=1}^{N} x_{ij}u_{ij}$$

$$\gamma := \gamma_N(\overline{x}, y, \overline{u}) \qquad \text{by} \qquad \lambda_N(\overline{x}, y, \overline{u}) \wedge \varepsilon_N(\overline{u}) \wedge \bigwedge_{i,j=1}^{N} \delta_N(x_{ij}, \overline{u})$$

A realization $\overline{a}$ of ε is called a tuple of matrix units.

Lemma 3.7 ([KT20], 2.1.3.)**.** *Let F be a field and let C be the center of $M := M_n(F)$. For $N, n \in \mathbb{N}$ with $N \leq n$ and $i, j \in \{1, \dots, N\}$, let $E_{ij} \in M_N(C)$ be the $N \times N$-matrix that has exactly one non-zero entry, namely $1 \in C$, at position (i, j).*

1. *If $\Theta : M_N(C) \to M$ is an embedding of F-algebras (which need not be unital), then the N^2-tuple $\overline{a} := (a_{ij})_{i,j=1}^N = (\Theta(E_{ij}))_{i,j=1}^N \in M^{N^2}$ is a realisation of $\varepsilon_N(\overline{u})$ and $\gamma_N(\overline{x}, y, \overline{a})$ defines the graph of Θ.*

2. *Let $\overline{a} := (a_{ji})_{i,j=1}^N \in M^{N^2}$ be a realization of $\varepsilon_N(\overline{u})$ in M. Then there is a unique F-algebra embedding $\Theta_{\overline{a}} : M_N(C) \to M_n(F)$ such that $\Theta(E_{ij}) = a_{ij}$ for all $i, j \in \{1, \ldots, N\}$. Explicitly, the graph of $\Theta_{\overline{a}}$ is defined by $\gamma_N(\overline{x}, y, \overline{a})$.*

Proof. 1. It can be verified by simple caclucations that $\overline{a}$ is a realization of ε as all caclulations are done in the matrix ring $M_N(C) \cong M_N(F)$. Now $(E_{ij}|i, j = 1, \ldots, N)$ is a basis of $M_N(C)$ and thus the graph of Θ is defined by γ as all $X \in M_N(C)$ have a representation

$$X = \sum_{i,j=1}^N x_{ij} E_{ij}$$

2. Since $\overline{a}$ realizes ε, the a_{ij} are C-linearly independent by Lemma 3.2. Thus, $\Theta_{\overline{a}} : M_N(C) \to M_n(F)$ satisfying $\Theta(E_{ij}) = a_{ij}$ defines an embedding after unique linear extension to $M_N(C)$. Now $\Theta_{\overline{a}}$ is an F-algebra embedding, since $\overline{a}$ realizes ε. Indeed, observe that for $A = (c_{ij})_{i,j=1}^N, B = (d_{ij})_{i,j=1}^N \in M_N(C)$

$$\Theta_{\overline{a}}(AB) = \Theta_{\overline{a}}\left(\left(\sum_{i,j=1}^N c_{ij}E_{ij}\right)\left(\sum_{k,l=1}^N d_{kl}E_{kl}\right)\right) = \Theta_{\overline{a}}\left(\sum_{i,j,k=1}^N c_{ij}d_{jk}E_{ik}\right) = \\ \sum_{i,j,k=1}^N c_{ij}d_{jk}a_{ik} = \left(\sum_{i,j=1}^N c_{ij}a_{ij}\right)\left(\sum_{k,l=1}^N d_{kl}a_{kl}\right) = \Theta_{\overline{a}}(A)\Theta_{\overline{a}}(B)$$

The statement about the graph of $\Theta_{\overline{a}}$ follows from part 1.

□

Lemma 3.7 establishes a bijection between the two sets

$$\Phi := \{\Theta : M_N(C) \to M_n(F) | \Theta \text{ is an } F\text{-algebra embedding}\}$$

of embeddings and

$$E := \{(\overline{a} := (a_{ij})_{i,j=1}^N \in (M_n(F))^{N^2} | \overline{a} \text{ realizes } \varepsilon_N(\overline{u})\}$$

of tuples of matrix units. This bijection is given by

$$\begin{aligned} \Phi &\longleftrightarrow E \\ \Theta &\longmapsto (\Theta(E_{ij}))_{i,j=1}^N \\ \Theta_{\overline{a}} &\longleftarrow\!\shortmid \overline{a} \end{aligned}$$

There are several immediate corollaries, which we will state below.

Corollary 3.8 ([KT20], 2.1.3.)**.** *The family of all embeddings $\Theta : M_N(C) \to M_n(F)$ of F-algebras is definable without quantifiers in $M = M_n(F)$ by $\gamma_N(\overline{x}, y, \overline{u})$ and the family's parameter set is defined without quantifiers by $\varepsilon_N(\overline{u})$.*

Proof. This is an immediate consequence of Lemma 3.7. □

Corollary 3.9 ([KT20], 2.1.4.)**.** *For any field F, the center of $M_n(F)$ is existentially defined by*

$$\exists \overline{u} : (\varepsilon_n(\overline{u}) \wedge \delta_n(v, \overline{u}))$$

Proof. This uses the linear independence of $(E_{ij} | i, j = 1, \ldots, N)$ and that for $n = N$, Θ is an isomorphism. □

Before proceeding with another corollary that will allow us to establish some important relations between different models of the theory of $M_n(F)$ for a field F, we need the following lemma. This will provide the necessary tools to transfer model theoretic statements about fields to their associated matrix algebras.

Lemma 3.10. *Let F be a field and let C be another $\mathcal{L}_{Ri}$-structure. Then the following hold:*

1. $C \equiv F \iff M_n(C) \equiv M_n(F)$
2. $C \cong F \iff M_n(C) \cong M_n(F)$
3. $C \preceq F \iff M_n(C) \preceq M_n(F)$

Proof. Let φ^n be an $\mathcal{L}_{Ri}$-formula. We first want to show that for any evaluation h^n in $M_n(F)$, we have

$$M_n(F) \vDash \varphi^n[h^n] \iff F \vDash \varphi[h]$$

for a suitable $\mathcal{L}_{Ri}$-formula φ and a suitable evaluation h in F. The choice of these objects, however, is clear as we just access the single entries of matrices by replacing each variable occurring in φ^n by n^2 new variables, each representing one entry of a matrix. Resolving these replacements by the standard matrix addition and multiplication rules, we find φ as desired. h simply gives an entrywise evaluation of h^n.

Conversely, we may start with any formula φ and any evaluation h in F, and define a formula φ_n and an evaluation h_n satisfying

$$M_n(F) \vDash \varphi_n[h_n] \iff F \vDash \varphi[h]$$

by restricting all quantifiers in φ to centers and defining h_n via the natural embedding $F \mapsto F \cdot I_n$.

We may apply the same ideas to C as to F, since by elementary equivalence, C will always be a ring in the scenarios described above. With this established, we deal with the different claims separately:

1. $\Longrightarrow$: We observe that for any $\mathcal{L}_{Ri}$-sentence φ^n, we get
$$M_n(F) \vDash \varphi^n \iff F \vDash \varphi \overset{F\equiv C}{\iff} C \vDash \varphi \iff M_n(C) \vDash \varphi^n$$
This proves our claim.

   $\Longleftarrow$: We contrary assume that $M_n(F) \equiv M_n(C)$. Then an analogous line of argument yields the result.

2. $\Longrightarrow$: If $\tau : C \to F$ is an isomorphism, we define $\tau^n : M_n(C) \to M_n(F)$ by $\tau^n(A) := (\tau(A_{ij}))_{i,j=1}^n$. τ^n is clearly bijective and the properties of an $\mathcal{L}_{Ri}$-isomorphism are easily checked.

   $\Longleftarrow$: We observe that restricting any isomorphism between $M_n(F)$ and $M_n(C)$ to centers again yields an isomorphism.

3. $\Longrightarrow$: We note that for arbitrary φ^n and h^n, we have
$$\begin{aligned} M_n(F) \vDash \varphi^n[h^n] &\iff F \vDash \varphi[h] \\ &\overset{C\preceq F}{\iff} C \vDash \varphi[h] \\ &\iff M_n(C) \vDash \varphi^n[h^n] \end{aligned}$$

   $\Longleftarrow$: We observe that we may argue in an analogous fashion as before.

□

Example 3.11. We illustrate the construction in the previous proof by the following examples.

1. We denote the formula $x + y = 0$ by φ^2 and want to decide, whether it holds in $M_2(F)$ for some evaluation h^2 with $h^2(x) = A, h^2(y) = B$. In order to do so, we define φ as
$$a_{11} + b_{11} = 0 \ \wedge \ a_{12} + b_{12} = 0 \ \wedge \ a_{21} + b_{21} = 0 \ \wedge \ a_{22} + b_{22} = 0$$

which we then decide over F. The definition of the evaluation h is self-evident.

2. We denote the formula $xy = 1$ by φ^2 and want to decide, whether it holds in $M_2(F)$ for some evaluation h with $h^2(x) = A, h^2(y) = B$. In order to do so, we define φ as

$$a_{11}b_{11} + a_{12}b_{21} = 1 \ \wedge \ a_{11}b_{12} + a_{12}b_{22} = 0 \wedge$$
$$a_{21}b_{11} + a_{22}b_{21} = 0 \ \wedge \ a_{22}b_{12} + a_{22}b_{22} = 1$$

which we then decide over F. The definition of the evaluation h is again self-evident.

3. We denote the formula $\exists y : xy = 1$ by φ and want to decide whether it holds in F for some evaluation h with $h(x) = a$. In order to do so, we define φ_n as

$$\exists y : (\forall z : zy = yz) \wedge xy = 1$$

which we then decide over $M_n(F)$. The evaluation h_n satisfies $h(x) = a \cdot I_n$.

Corollary 3.12 ([KT20], 2.1.5.)**.** *Let F be a field.*

1. *The theory of $M_n(F)$ is axiomatized by saying the following about a model A:*

   a) *A is a ring whose center C is elementarily equivalent to F, that is: A satisfies the ring axioms and additionally, whenever $F \vDash \varphi$ for some $\mathcal{L}_{Ri}$-sentence, we add some sentence φ' as an axiom. φ' is defined in the same recursive way as φ, except that whenever in the construction of φ we encounter a quantifier construction $\forall x \psi$, we write*

   $$\forall x : (\forall y : xy = yx) \to \psi$$

   instead. In particular, the axiom set will not be finite.

   b) *There exists $\overline{a} := (a_{ij})_{i,j=1}^n \in A^{n^2}$ realizing ε_n and each such realization $\overline{a}$ defines an isomorphism $\Theta_{\overline{a}} : M_n(C) \to A$ via $\gamma_n(\overline{x}, y, \overline{a})$.*

2. *Let A, B be rings that are elementarily equivalent to $M_n(F)$. If $A \subseteq B$, then $C_A \subseteq C_B$ where C_A is the center of A and C_B that of B. Furthermore, if $\overline{a}$ is a realization of ε_n in A^{n^2}, then the following diagram commutes:*

$$\begin{tikzcd}
A \arrow[r, hook] & B \\
M_n(C_A) \arrow[u, "\Theta_{\overline{a}}", "\cong"'] \arrow[r, hook] & M_n(C_B) \arrow[u, "\Theta_{\overline{a}}", "\cong"']
\end{tikzcd}$$

Proof. 1. Let A be a model satisfying 1a and 1b. By 1b, we get $M_n(C) \cong A$ and by 1a, $C \equiv F$. Thus, $M_n(C) \equiv M_n(F)$ by Lemma 3.10 1. This gives $M_n(F) \equiv A$ by transitivity of elementary equivalence, and we observe that consequently, A is a model of the theory of $M_n(F)$.

Now we contrarily assume that $A \equiv M_n(F)$. Then for 1a, we note that for any $\mathcal{L}_{Ri}$-sentence φ we have

$$C \vDash \varphi \iff A \vDash \varphi'$$

where we define φ' by replacing any string $\forall x \psi$ in the construction of φ as described above. As this is independent of the ring considered, the same holds for $M_n(F)$ and its center F. Using that $A \equiv M_n(F)$, we immediately get $C \equiv F$. This proves 1a.

For 1b, note that the existence of a tuple $(a_{ij})_{i,j=1}^n$ satisfying ε_n that additionally induces the desired isomorphism can be described by the $\mathcal{L}_{Ri}$-sentence γ_n. This sentence trivially holds in $M_n(F)$ and thus, by assumption, also in A.

2. Note that by isomorphism and Corollary 3.9, C_A and C_B are both existentially defined by the formula $\exists \overline{u}(\varepsilon_n(\overline{u}) \wedge \delta_n(v, \overline{u}))$, where n does not depend on whether we work over the ring C_A or C_B. Any $a \in C_A$ will therefore also lie in C_B, since $A \subseteq B$ and any tuple $\overline{a}$ satisfying ε_n in A also satisfies ε_n in B.

For commutativity of the diagram, note that our embeddings are given by the identity. Thus, the statement is trivial.

□

Remark 3.13. In the above setting, we have tacitly used that A and B have the same dimension, if both are elementarily equivalent to $M_n(F)$. This holds as we may describe the existence of a linearly independent tuple $(b_1, \ldots, b_n)$ (over F and C respectively) via the $\mathcal{L}_{Ri}$-sentence

$$\exists b_1, \ldots b_n : \forall c_1, \ldots c_n : \left(\forall x : \bigwedge_{i=1}^{n} c_i x = x c_i \right) \rightarrow \left(\sum_{i=1}^{n} c_i b_i = 0 \rightarrow \bigwedge_{i=1}^{n} c_i = 0 \right)$$

Since the sentence

$$\exists b_1, \ldots, b_n : \forall x : \exists c_1, \ldots c_n : x = \sum_{i=1}^{n} c_i b_i$$

defines $b_1, \ldots, b_n$ as generators of a vector space, this gives us dimensional equality between $M_n(F)$, A, and B over their respective centers. Also note that the existence of a basis of n^2 elements can thus be derived within $M_n(F)$ and then transferred to A. As a consequence, in the above diagram, $\Theta_{\overline{a}} : M_n(C_B) \rightarrow B$ still is an isomorphism.

Definition 3.14. Let U be an $\mathcal{L}$-structure. We call U model complete if the theory $Th(U)$ of U is model complete.

Definition 3.15 ([KT20], 2.1.6.)**.** Let F be a field and let $\tilde{F}$ be an expansion of F in some language $\mathcal{L}$ extending $\mathcal{L}_{Ri}$, that is: F and $\tilde{F}$ have the same universe, but $\tilde{F}$ interprets additional symbols. Then we define $M_n(\tilde{F})$ as an $\mathcal{L}$-structure in the following way:

- $M_n(\tilde{F})$ expands $M_n(F)$, that is: They have the same universe, and symbols in $\mathcal{L}_{Ri}$ are not newly interpreted.

- If R is a new relation symbol, then we give an interpretation of R only on the center C of $M_n(F)$, that is: For arbitrary $\mathcal{L}$-terms $t_1, t_2, \ldots$, we have $M_n(\tilde{F}) \vDash R^{M_n(\tilde{F})}(t_1^{M_n(\tilde{F})}, \ldots)$ if and only if $t_1^{M_n(\tilde{F})}, \ldots \in C$ and $\tilde{F} \vDash R^{\tilde{F}}(t_1^{\tilde{F}}, \ldots)$.

- If f is a new m-place function symbol, then for given $\mathcal{L}$-terms $t_1, t_2, \ldots$, we define $f^{M_n(\tilde{F})}$ on C^m by $f^{M_n(\tilde{F})}(t_1^{M_n(\tilde{F})}, \ldots) := f^{\tilde{F}}(t_1^{\tilde{F}}, \ldots)$. Otherwise, we set the function 0.

- The interpretation of new constant symbols is already given by their interpretation on $\tilde{F}$.

Before putting this expansion to use to prove model completeness of matrix algebras in Lemma 3.18, we make two observations about how a field relates to its associated matrix algebra in an extended language.

Lemma 3.16. *Let C and F be fields in a language $\mathcal{L}$ extending $\mathcal{L}_{Ri}$. We interpret $M_n(F)$ and $M_n(C)$ as $\mathcal{L}$-structures as described in Definition 3.15. If $f : C \to F$ is an $\mathcal{L}$-embedding, then f may be extended to an $\mathcal{L}$-embedding $f : M_n(C) \to M_n(F)$ by applying f entry-wise.*

Proof. Well-definedness follows from linear independence of the family $(E_{ij})_{i,j=1}^n$. The rest is achieved by linearly extending the map. □

Lemma 3.17. *Let F be a field and let $\tilde{F}$ be an expansion of F in a language $\mathcal{L}$ extending $\mathcal{L}_{Ri}$. Now let $\tilde{C}$ be another $\mathcal{L}$-structure. Then the following hold:*

1. $\tilde{C} \equiv \tilde{F} \iff M_n(\tilde{C}) \equiv M_n(\tilde{F})$

2. $\tilde{C} \preceq \tilde{F} \iff M_n(\tilde{C}) \preceq M_n(\tilde{F})$

Proof. 1. By Lemma 2.17 1 and Lemma 3.10 1, $\tilde{C} \equiv \tilde{F}$ implies $C \equiv F$ and $M_n(C) \equiv M_n(F)$. For any $\mathcal{L}$-sentence, we also note that since containment in the center of a matrix ring is expressible as an $\mathcal{L}_{Ri}$-formula, by the definition of new symbols, we immediately get that

$$M_n(\tilde{C}) \vDash \varphi \iff M_n(\tilde{F}) \vDash \varphi$$

for any $\mathcal{L}$-sentence φ and thus $M_n(\tilde{C}) \equiv M_n(\tilde{F})$.

Contrary, if $M_n(\tilde{C}) \equiv M_n(\tilde{F})$, we immediately get $C \equiv F$ by Corollary 3.12 1a (or equivalently Lemma 3.10 1). Again, since containment in the center is expressible by an $\mathcal{L}_{Ri}$-formula, we get the statement.

2. Note that by Lemma 2.17 3 and Lemma 3.10 3, we have $C \preceq F$ and $M_n(C) \preceq M_n(F)$. For any $\mathcal{L}$-formula φ and any evaluation h in $M_n(\tilde{C})$, we note that since containment in the center of a matrix ring is expressible as an $\mathcal{L}_{Ri}$-formula, we get by the embedding property that for any $X \in M_n(\tilde{C})$, $X \in C \cdot I_n$ if and only if $X \in F \cdot I_n$. Thus, by the definition of new symbols, we immediately get that

$$M_n(\tilde{C}) \vDash \varphi[h] \iff M_n(\tilde{F}) \vDash \varphi[h]$$

for any $\mathcal{L}$-sentence φ and any evaluation h in $M_n(C)$. Hence, $M_n(\tilde{C}) \preceq M_n(\tilde{F})$.

Again, if we assume that $M_n(\tilde{C}) \preceq M_n(\tilde{F})$, then we get the converse implication by restricting to centers.

□

Lemma 3.18 ([KT20], 2.1.7.)**.** *Let F be a field and let $\tilde{F}$ be a model complete expansion of F in some language $\mathcal{L}$ extending $\mathcal{L}_{Ri}$. Then $M_n(\tilde{F})$ is also model complete.*

Proof. Let $\tilde{A}, \tilde{B}$ be $\mathcal{L}$-structures that are elementarily equivalent to $M_n(\tilde{F})$. Further assume that $\tilde{A} \subseteq \tilde{B}$. By Theorem 2.18, it is sufficient to prove $\tilde{A} \preceq \tilde{B}$.

In order to do so, first note that since $\tilde{A}, \tilde{B} \equiv M_n(\tilde{F})$ and $\mathcal{L}$ extends $\mathcal{L}_{Ri}$, both $A := \tilde{A}|_{\mathcal{L}_{Ri}}$ and $B := \tilde{B}|_{\mathcal{L}_{Ri}}$ are rings. Recall from Lemma 2.17 1 that elementary equivalence is invariant under restrictions, that is: We get $A \equiv M_n(F) \equiv B$ as $\mathcal{L}_{Ri}$-structures.

By Corollary 3.12 1b, we may chose $\overline{a} \in A^{n^2}$ satisfying ε_n, and by Corollary 3.12 2, we get $C_A \subseteq C_B$ for the centers of A and B respectively. We also get the commutative diagram

$$\begin{tikzcd}
A \arrow[r, hook] & B \\
M_n(C_A) \arrow[u, "\Theta_{\overline{a}}", "\cong"'] \arrow[r, hook] & M_n(C_B) \arrow[u, "\Theta_{\overline{a}}", "\cong"']
\end{tikzcd}$$

Here, we use that $A \subseteq B$ is trivially invariant under the restriction of languages. We may then let $\tilde{A}$ induce an $\mathcal{L}$-structure $\tilde{C_A}$ on C_A by interpreting symbols of $\mathcal{L}$ according to their definition on $\tilde{A}$. Analogously, we define $\tilde{C_B}$. We immediately get the following properties:

- $\tilde{C_A} \subseteq \tilde{C_B}$ as $C_A \subseteq C_B$ and $\tilde{A} \subseteq \tilde{B}$.
- $\tilde{C_A} \equiv \tilde{F} \equiv \tilde{C_B}$, since by assumption, $\tilde{A} \equiv M_n(\tilde{F}) \equiv \tilde{B}$

Thus, we may conclude $\tilde{C_A} \preceq \tilde{C_B}$ by model completeness of $\tilde{F}$.

Now let $\Theta_{\overline{a}} : M_n(C_A) \to A$ denote the algebra isomorphism induced by $\overline{a}$. We claim that this naturally extends to an isomorphism from $M_n(\tilde{C_A})$ to $\tilde{A}$ and we will denote this extended isomorphism again by $\Theta_{\overline{a}}$.

Indeed, if X lies in the center of $M_n(C_A)$, then $X = \sum_{i=1}^n \lambda E_{ii}$ for some $\lambda \in C_A$. Hence, $\Theta_{\overline{a}}(X) = \lambda$ because of

$$1 = \Theta_{\overline{a}}(I_n) = \Theta_{\overline{a}}\left(\sum_{i=1}^n E_{ii}\right) = \sum_{i=1}^n a_{ii}$$

Since new symbols on $M_n(\tilde{C_A})$ are defined only on the center, this proves the claim.

Analogous ideas apply to B. Finally, $\tilde{C_A} \preceq \tilde{C_B}$ translates to $M_n(\tilde{C_A}) \preceq M_n(\tilde{C_B})$ by Lemma 3.17 2. Thus, we get the commutative diagram

$$\begin{tikzcd}
\tilde{A} \arrow[r, hook] & \tilde{B} \\
M_n(\tilde{C_A}) \arrow[u, "\Theta_{\overline{a}}", "\cong"'] \arrow[r, hook] & M_n(\tilde{C_B}) \arrow[u, "\Theta_{\overline{a}}", "\cong"']
\end{tikzcd}$$

and conclude

$$\begin{aligned}
\tilde{A} \models \varphi[h] &\iff M_n(\tilde{C_A}) \models \varphi[\Theta_{\overline{a}}^{-1} \circ h] \\
&\iff M_n(\tilde{C_B}) \models \varphi[\Theta_{\overline{a}}^{-1} \circ h] \\
&\iff \tilde{B} \models \varphi[\Theta_{\overline{a}} \circ \Theta_{\overline{a}}^{-1} \circ h] = \varphi[h]
\end{aligned}$$

□

Corollary 3.19 ([KT20], 2.1.7.)**.** *Both $M_n(\mathbb{C})$ and $M_n(\mathbb{R})$ (the latter expanded by $<$, which is interpreted on its center) are model complete.*

Proof. This is a direct application of Lemma 3.18 to the model complete structures $\langle \mathbb{C}, +, -, \cdot, 0, 1\rangle$ and $\langle \mathbb{R}, <, +, -, \cdot, 0, 1\rangle$. The model completeness of these structures follows from Theorem 2.21. □

Remark 3.20 ([KT20], 2.1.8.)**.** There is no analogous result to that of the Lemma 3.18 that lets us lift quantifier elimination from $\tilde{F}$ to $M_N(\tilde{F})$. For example, $M_n(\mathbb{C})$ does not admit quantifier elimination for $n \geq 2$ (see [Ros78], Theorem 3.2.), yet $\mathbb{C}$ does (see Theorem 2.21).

Still, we may save the result of the Lemma 3.18 for quantifier elimination if we allow matrix units as parameters, which will be done in the next two lemmata. This should come as no surprise. Indeed, let us assume that we have a field $\tilde{F}$ that admits quantifier elimination. If we now admit (standard) matrix units as constant symbols, we may regard them as a way to access single entries of a matrix and consequently to reduce any question regarding quantifier elimination to an equivalent question in the underlying field.

Lemma 3.21 ([KT20], 2.1.9.)**.** *Let F be a field and let $U \subseteq M_n(F)$ be a subring. Further let $E_{ij} \in U$ for all $i, j \in \{1, ..., n\}$. Then*

$$R_U := \{a \in F | a \text{ is the } (1,1)\text{-entry for some } Y \in U\}$$

is a subring of F and $U = M_n(R_U)$.

Proof. If $a, b \in U$, then $X_{11} = a$ and $Y_{11} = b$ for some $X, Y \in U$. As U is a ring, $X + Y \in U$ and thus $a + b = (X + Y)_{11} \in R_U$. Similarly, $-X \in U$ and thus $-a \in R_U$. For multiplication, note that since $E_{11} \in U$, we have $E_{11}XE_{11}, E_{11}YE_{11} \in U$ and thus $ab = (E_{11}XE_{11}E_{11}YE_{11})_{11} \in R_U$. The remaining ring properties are trivial.

If now $Y \in U$, then also $E_{1i}YE_{j1} \in U$, which has only 0-entries except for the entry Y_{ij} at the $(1,1)$-place. Therefore, $Y \in M_n(R_U)$. If contrary $Y \in M_n(R_U)$, then for all $i, j \in \{1, ..., n\}$, we find some $X^{(i,j)} \in U$ such that $x_{11}^{(i,j)} = Y_{ij}$. We then note that

$$Y_{ij}E_{ij} = E_{i1}X^{(i,j)}E_{1j} \in U$$

and thus

$$Y = \sum_{i,j=1}^{n} E_{i1}X^{(i,j)}E_{1j} \in U$$

□

Lemma 3.22 ([KT20], 2.1.11.)**.** *Let F be a field and let $\tilde{F}$ be an expansion of F in some language $\mathcal{L}$ extending $\mathcal{L}_{Ri}$. Let $\tilde{F}$ admit quantifier elimination. Let further $\overline{c} = (c_{ij})_{i,j=1}^{n}$ denote new constant symbols. By $\mathcal{L}(\overline{c})$, we denote the language $\mathcal{L}$ expanded by the constants $\overline{c}$. Then the $\mathcal{L}(\overline{c})$-structure $(M_n(\tilde{F}), \overline{e})$, where $\overline{c}$ is interpreted as a tuple of matrix units $\overline{e}$, also admits quantifier elimination.*

Proof. By Lemma 3.18 and Theorem 2.19, $M_n(\tilde{F})$ is model complete. This immediately transfers to $(M_n(\tilde{F}), \overline{e})$. Indeed if both $A' := (\tilde{A}, \overline{a})$ and $B' := (\tilde{B}, \overline{b})$ are models of the theory $Th(M_n(\tilde{F}), \overline{e})$ satisfying $A' \subseteq B'$, then we observe the following:

- $\overline{a} = \overline{b}$ as $A' \subseteq B'$.
- Both $\tilde{A}$ and $\tilde{B}$ are models of $M_n(\tilde{F})$, where $\tilde{A}, \tilde{B}$ are the restrictions of A', B' to $\mathcal{L}$. This follows from Lemma 2.17 1.
- $\tilde{A} \subseteq \tilde{B}$, and thus already $\tilde{A} \preceq \tilde{B}$ by model completeness.

Now observe that

$$\begin{aligned} A' \vDash \varphi[h] &\iff \tilde{A} \vDash \varphi(\overline{c}/\overline{a})[h] \\ &\iff \tilde{B} \vDash \varphi(\overline{c}/\overline{a})[h] \\ &\iff B' \vDash \varphi[h] \end{aligned}$$

for any evaluation h in A' (i.e. $\tilde{A}$, since the universes coincide). Thus, $A' \preceq B'$.

Now it is sufficient to prove the amalgamation property. For that, let A', B' be elementarily equivalent to $(M_n(\tilde{F}), \overline{e})$ and let $U' := (\tilde{U}, \overline{u})$ be a common substructure of A' and B'. This implies $\overline{u} = \overline{a} = \overline{b}$. Now define A, B and U as the restrictions of $\tilde{A}, \tilde{B}$ and $\tilde{U}$ respectively to $\mathcal{L}_{Ri}$. Then U is a common subring of A and B. Further define the fields C_A and C_B as the centers of A and B respectively.

By Lemma 3.12 1b, we may choose isomorphisms $\varphi : A \to M_n(C_A)$ and $\psi : B \to M_n(C_B)$ that map the matrix units $\overline{u}$ to the matrix units $(E_{ij})_{i,j=1}^{n}$ (note that these are formally different in $M_n(C_A)$ and $M_n(C_B)$). We may extend these isomorphisms to $\varphi : \tilde{A} \to M_n(\tilde{C_A})$ and $\psi : \tilde{B} \to M_n(\tilde{C_B})$ as in the proof of Lemma 3.18.

Now we apply Lemma 3.21 to obtain two subrings $R \subseteq C_A$ and $S \subseteq C_B$ that satisfy $M_n(R) \cong U \cong M_n(S)$. In particular, the isomorphisms are given by $\varphi|_U$ and $\psi|_U$ respectively. We may induce an $\mathcal{L}$-structure on $M_n(R)$ and $M_n(S)$ through these isomorphisms and thus extend the isomorphisms to $\mathcal{L}(\overline{c})$. This gives us the following commutative diagram.

$$\begin{tikzcd}
(M_n(\tilde{C}_A), \overline{E}) & (\tilde{A}, \overline{a}) \arrow[l, "\cong", "\varphi"'] & & (\tilde{B}, \overline{b}) \arrow[r, "\cong", "\psi"'] & (M_n(\tilde{C}_B), \overline{E}) \\
(M_n(\tilde{R}), \overline{E}) \arrow[u, hook] & & (\tilde{U}, \overline{u}), \arrow[ll, "\cong", "\varphi|_U"'] \arrow[rr, "\cong", "\psi|_U"'] \arrow[ul, hook] \arrow[ur, hook] & & (M_n(\tilde{S}), \overline{E}) \arrow[u, hook]
\end{tikzcd}$$

To prove the amalgamation property, we first restrict all maps in that diagram to centers interpreted in $\mathcal{L}$. This leaves isomorphisms invariant and by Lemma 2.1 and the choice of R and S, the outer subring relations will be preserved as well. This gives us the commutative diagram

$$\begin{tikzcd}
\tilde{C}_A & \tilde{C}_A \arrow[l, "\cong"'] & & \tilde{C}_B \arrow[r, "\cong"] & \tilde{C}_B \\
\tilde{R} \arrow[u, hook] & & \tilde{C}_U, \arrow[ll, "\cong"'] \arrow[rr, "\cong"] \arrow[ul] \arrow[ur] & & \tilde{S} \arrow[u, hook]
\end{tikzcd}$$

$\tilde{C}_U$ may now be embedded into $\tilde{C}_A$ via $\tilde{R}$ and similar arguments work for $\tilde{C}_B$. Note that we may not embed via $\tilde{C}_A$ since *a priori*, $\tilde{C}_U$ need not be a subset of $\tilde{C}_A$. Finally, this lets us amalgamate using quantifier elimination in $\tilde{F} \equiv \tilde{C}_A \equiv \tilde{C}_B$, yielding two embeddings

$$\varepsilon : \tilde{C}_A \to \tilde{\Omega}$$
$$\delta : \tilde{C}_B \to \tilde{\Omega}$$

with $\tilde{\Omega} \equiv \tilde{F}$ satisfying $\varepsilon(\varphi(u)) = \delta(\psi(u))$ for all u in $\tilde{C}_U$.

We may extend ε and δ to embeddings

$$\overline{\varepsilon} : M_n(\tilde{C}_A) \to M_n(\tilde{\Omega})$$
$$\overline{\delta} : M_n(\tilde{C}_B) \to M_n(\tilde{\Omega})$$

by applying the embeddings entry-wise using Lemma 3.16. The desired amalgamation is then given by the mappings $\overline{\varepsilon} \circ \varphi$ and $\overline{\delta} \circ \psi$ after noting that ε and δ are in fact $\mathcal{L}(\overline{c})$-homomorphisms - here we use that matrix units are mapped to matrix units. □

3.2 Quantifier Elimination using Trace and Transposition

A question that naturally arises is whether we actually need to access all matrix entries to get quantifier elimination. Ideally, we may work with weaker properties of a matrix that still allow for quantifier elimination. For Pythagorean fields, this question may be answered in the positive, provided they possess the so-called Specht property for the transpose. Indeed, we will only need to consider the trace and the transpose as additional unary functions to get that result. What is more, we will even get an equivalence between quantifier elimination and the Specht property. All terms that were just introduced informally will be properly defined later.

We prepare the main result in Theorem 3.30 by the following two lemmata and one theorem. The first lemma tells us that adding two new functions for the trace and the transposition will not affect model completeness as derived in Lemma 3.18. The second lemma provides us with a criterion that lets us compare the traces of two words. The theorem directly pertains to the Specht property as defined below for certain choices of our base field F.

Lemma 3.23. *Let F be a field and let $\tilde{F}$ be a model complete expansion of F in a language $\mathcal{L}$ extending $\mathcal{L}_{Ri}$. Let $\mathcal{L}'$ be the language $\mathcal{L}$ extended by two new unary function symbolds tr and $invo$. Then the $\mathcal{L}$-structure $(M_n(\tilde{F}), tr_F, [X \mapsto X^T])$ is model complete.*

Proof. By Lemma 3.18, $M_n(\tilde{F})$ is model complete as an $\mathcal{L}$-structure. We now set $M' := (M_n(\tilde{F}), tr_F, [X \mapsto X^T])$ and assume that we are given $A', B' \equiv M'$ satisfying $A' \subseteq B'$. We show that $A' \preceq B'$ and first observe both $A := A'|_{\mathcal{L}Ri}$ and $B := B'|_{\mathcal{L}_{Ri}}$ are elementarily equivalent to $M_n(F)$. Clearly, $\tilde{A} \subseteq \tilde{B}$. Now let φ' be an $\mathcal{L}'$-formula. Then φ' can be translated into an $\mathcal{L}$-formula φ satisfying

$$A' \vDash \varphi'[h'] \iff \tilde{C_A} \vDash \varphi[h]$$

for any evaluation h' in A' and accordingly chosen evaluation h in the $\mathcal{L}$-center $\tilde{C_A}$ of A' (after potentially applying the isomorphism $\Theta_{\overline{a}}$ from Corollary 3.12 1b; compare this to the proof of Lemma 3.10). This translates to $\tilde{C_B} \vDash \varphi[h]$ by model completeness and $\tilde{C_A} \subseteq \tilde{C_B}$ (see Corollary 3.12 2), which in turn is equivalent to $B' \vDash \varphi'[h']$. □

Lemma 3.24 ([KT20], 2.2.1.). *Let K, L be fields and let $\mathcal{L}$ be an extension of $\mathcal{L}_{Ri}$ by a one-place function symbol F. We now consider the $\mathcal{L}$-structures $(M_n(K), tr_K)$ and $(M_n(L), tr_L)$, where tr_K and tr_L are interpretations of F. Let (U, f) be another $\mathcal{L}$-structure (where f interprets F) and suppose that we are given $\mathcal{L}$-embeddings*

$$\varphi : (U, f) \hookrightarrow (M_n(K), tr_K)$$
$$\psi : (U, f) \hookrightarrow (M_n(L), tr_L)$$

Then the following hold.

1. *The subring R of U genereated by the image of f is commutative and $\varphi(R) \subseteq K \cdot I_n, \psi(R) \subseteq L \cdot I_n$.*

2. Assume that $K \cdot I_n$ and $L \cdot I_n$ can be amalgamated over $\varphi|_R$ and $\psi|_R$ into some field Ω by the maps

$$\varepsilon : K \cdot I_n \to \Omega \cdot I_n$$
$$\delta : L \cdot I_n \to \Omega \cdot I_n$$

$$\begin{tikzcd}
 & \Omega \cdot I_n & \\
K \cdot I_n \arrow[ur, hook, "\varepsilon"] & & L \cdot I_n \arrow[ul, hook', "\delta"'] \\
 & R \arrow[ul, hook] \arrow[ur, hook'] &
\end{tikzcd}$$

Then we may naturally (i.e. by mapping the standard matrix units to the corresponding standard matrix units) induce maps

$$\overline{\varepsilon} : M_n(K) \to M_n(\Omega)$$
$$\overline{\delta} : M_n(L) \to M_n(\Omega)$$

and for every $X \in U$ we have

$$tr_\Omega(\overline{\varepsilon}(\varphi(X))) = tr_\Omega(\overline{\delta}(\psi(X)))$$

The following (not necessarily commutative) diagram shows all maps.

$$\begin{tikzcd}
 & & (M_n(\Omega), tr_\Omega) & & \\
(M_n(K), tr_K) \arrow[urr, hook, "\overline{\varepsilon}"] & & & & (M_n(L), tr_L) \arrow[ull, hook', "\overline{\delta}"'] \\
 & & (U, f) \arrow[ull, hook, "\varphi"'] \arrow[urr, hook', "\psi"] & & \\
 & & (R, f_R) \arrow[u, hook] \arrow[uull, "\varphi|_R"] \arrow[uurr, "\psi|_R"'] & &
\end{tikzcd}$$

Proof. 1. First note that U is a ring since as an $\mathcal{L}$-structure it is closed under addition and multiplication. The ring axioms follow from the fact that U embeds into the rings (even fields) K and L. We also get $f(U) \subseteq U$ by closedness of operations. Thus, let $X \in U$. Then by the properties of φ as an $\mathcal{L}$-homomophism, we get $\varphi(f(X)) = tr_K(\varphi(X))$. Now observe that $tr_K(\varphi(X)) \in K \cdot I_n$ and thus $\varphi(f(X)) \in K \cdot I_n$. Hence, $\varphi(f(U)) \subseteq K \cdot I_n$. As φ is an embedding, $f(U) \subseteq R$ inherits commutativity from $K \cdot I_n$. An analogous argument works for L.

2. Let $X \in U$. Then note that

$$\begin{aligned} tr_\Omega(\overline{\varepsilon}(\varphi(X))) &= \varepsilon(tr_K(\varphi(X))) = \\ &= \varepsilon(\varphi(f(X))) = \\ &\overset{\varepsilon\circ\varphi=\delta\circ\psi}{=} \delta(\psi(f(X))) \end{aligned}$$

A similar calculation gives $tr_\Omega(\overline{\delta}(\psi(X))) = \delta(\psi(f(X)))$. Thus, we are done.

□

Theorem 3.25 ([KT20], 2.2.2.)**.** *Let Ω be a real closed field or the algebraic closure of a real closed field. Let further $X_1, \dots, X_d, Y_1, \dots Y_d \in M_n(\Omega)$. The following are equivalent:*

1. *There is some unitary $O \in M_n(\Omega)$ satisfying $OX_iO^* = Y_i$ for all $i \in \{1, ..., n\}$.*

2. *For every word ω in the letters $x_1, \dots x_d, x_1^*, \dots, x_d^*$ we have*

$$tr_\Omega(\omega(X_1, \dots X_d, X_1^*, \dots, X_d^*)) = tr_\Omega(\omega(Y_1, \dots Y_d, Y_1^*, \dots, Y_d^*))$$

3. *For every word ω in the letters $x_1, \dots x_d, x_1^*, \dots, x_d^*$ that has at most length n^2 we have*

$$tr_\Omega(\omega(X_1, \dots X_d, X_1^*, \dots, X_d^*)) = tr_\Omega(\omega(Y_1, \dots Y_d, Y_1^*, \dots, Y_d^*))$$

Proof. 1) $\implies$ 2): Choose O such that 1 holds. Recall that the (orthogonal) conjugation of a matrix leaves its trace unchanged. Thus, $tr_\Omega(X_i) = tr_\Omega(OX_iO^*) = tr_\Omega(Y_i)$ for all $i \in \{1, ..., n\}$. Also note that $OX_i^*O^* = (OX_iO^*)^* = Y_i^*$. If now ω is a word given as specified by 2, we use $O^*O = I_n$ to obtain

$$\begin{aligned} \omega(X_1, \dots X_d, X_1^*, \dots, X_d^*) &= O^*\omega(OX_1O^*, \dots OX_dO^*, OX_1^*O^*, \dots, OX_d^*O^*)O = \\ &= O^*\omega(Y_1, \dots Y_d, Y_1^*, \dots, Y_d^*)O \end{aligned}$$

This immediately yields the desired result about traces.

2) $\implies$ 3): This is trivial.

3) $\implies$ 2): This implication relies on invariant theory and can be found in [Pro76], theorem 7.3., or [Raz74]. Note that dependent on the source, different degree bounds may be given.

2) $\implies$ 1): See [Sib68], Corollary 1 and Lemma 2, which also relies on invariant theory. □

Lemma 3.26 ([KT20], 2.2.3.). *Let F be a real field and $X \in M_n(F)$. Then*

$$X = 0 \iff tr(X^T X) = 0$$

Proof. The statement follows from the identity

$$tr(X^T X) = \sum_{i=1}^{n} \sum_{k=1}^{n} x_{ki}^2$$

and the fact that, in real fields, squares are positive. □

We are now ready to prove the main result of this section. This requires us to introduce the following two definitions.

Definition 3.27 ([Har00], Proposition 16.1). A field F is called Pythagorean, if $1 + a^2$ is a square in F for all $a \in F$.

Lemma 3.28. *A field F is Pythagorean if and only if any sum of squares is again a square.*

Proof. As $1 + a^2$ is a sum of squares for all $a \in F$, one direction is trivial. For the other direction, it is sufficient to consider the case $a^2 + b^2$ for $a, b \in F$ and $a, b \neq 0$. Then $1 + (\frac{b}{a})^2 = c^2$ for some $c \in F$ assumption. Hence, $a^2 + b^2 = (ac)^2$, which proves the claim. □

It is obvious from the definition that every real closed field is Pythagorean. Contrary, a Pythagorean field need not even be real. If, however, we postulate that $1 + a^2 \neq 0$, then it is. If, in fact, -1 were a sum of squares, then $c^2 = -1$ for some $c \in F$. This implies $0 = 1 + c^2$ and thus we have a contradiction.

Definition 3.29 ([KT20], 2.2.4.). Let F be a field. We say that F has the Specht Property for the transpose (SPT), if for any $d, n \in \mathbb{N}$ there is some $D = D(d, n) \in \mathbb{N}$ such that the following holds:

For all d-tuples $X = (X_1, \ldots, X_d), Y = (Y_1, \ldots, Y_d) \in (M_n(F))^d$ with

$$tr(\omega(X, X^T)) = tr(\omega(Y, Y^T))$$

for all words ω in $2d$ variables x, x^T of degree at most D, we find some $O \in M_n(F)$ satisfying $OO^T = I_n$ and $O^T X_i O = Y_i$ for all $i \in \{1, ..., n\}$.

In the case that F is the algebraic closure of a real closed field, we may rephrase the (SPT) using the matrix involution and will refer to it as the Specht Property for the involution (SPI).

Theorem 3.30 ([KT20], 2.2.4.)**.** *Let F be a real, Pythagorean field, let $\tilde{F}$ be an expansion of F in a language $\mathcal{L}$ extending $\mathcal{L}_{Ri}$ and suppose that $\tilde{F}$ admits quantifier elimination in $\mathcal{L}$. We expand $\mathcal{L}$ by two new one-place function symbols to $\mathcal{L}(tr, invo)$. Then the following are equivalent:*

1. *The $\mathcal{L}(tr, invo)$-structure $(M_n(\tilde{F}), tr_F, [X \mapsto X^T])$ admits quantifier elimination.*

2. *F has the (SPT).*

3. *Let $\tilde{K} \equiv \tilde{F}$ and let U be a substructure of $(M_n(\tilde{K}), tr_F, [X \mapsto X^T])$. Let further $\psi : U \to (M_n(\tilde{K}), tr_F, [X \mapsto X^T])$ be an embedding. Then there is an elementary extension $\tilde{K} \preceq \tilde{\Omega}$ and an extension of ψ to an embedding from $(M_n(\tilde{K}), tr_F, [X \mapsto X^T])$ to $(M_n(\tilde{\Omega}), tr_F, [X \mapsto X^T])$. In particular, the following diagram commutes:*

$$\begin{tikzcd}
 & (M_n(\tilde{\Omega}), tr_\Omega, X \mapsto X^T) \\
(M_n(\tilde{K}), tr_K, X \mapsto X^T) \arrow[ur, hook] & (M_n(\tilde{K}), tr_K, X \mapsto X^T) \arrow[u, hook, "\preceq"'] \\
U \arrow[u, hook, "id"'] \arrow[ur, hook, "\psi"'] &
\end{tikzcd}$$

Proof. 2) $\implies$ 1): $\tilde{F}$ is model complete and thus, by Lemma 3.18, so is the $\mathcal{L}$-structure $M_n(\tilde{F})$. This also holds for $(M_n(\tilde{F}), tr_F, [X \mapsto X^T])$ by Lemma 3.23. Hence, it is sufficient to prove the amalgamation property for the theory T of $(M_n(\tilde{F}), tr_F, [X \mapsto X^T])$ over finitely generated substructures as derived in Remark 2.20.

Thus, let $M, N \vDash T$ and let U be a common finitely generated $\mathcal{L}(tr, invo)$-substructure of M and N. We may choose an isomorphism $\overline{\varphi} : M \to (M_n(\tilde{K}), tr_K, [X \mapsto X^T])$ following Corollary 3.12 1b, where $\tilde{K} \equiv \tilde{F}$ denotes the center K of M, lifted to the language $\mathcal{L}$. Note that compatibility with trace and involution follows from the fact that in $M_n(K)$, we may choose a tuple of matrix units $(a_{ij})_{i,j=1}^n$ satisfying $tr_K(a_{ij}) = \delta_{ij}$ and $a_{ij}^T = a_{ji}$ for all $i, j \in \{1, ..., n\}$. Such units must therefore also exist in M by elementary equivalence and the induced isomorphism trivially respects trace and involution. By φ, we denote the restriction of $\overline{\varphi}$ to U. An analogous construction can be done for N, such that we get the following diagram.

$$\begin{tikzcd}
(M_n(\tilde{K}), tr_K, X \mapsto X^T) & M \arrow[l, "\overline{\varphi}"', "\cong"] & & N \arrow[r, "\cong", "\overline{\psi}"'] & (M_n(\tilde{L}), tr_L, X \mapsto X^T) \\
& & U \arrow[ull, "\varphi"] \arrow[ul, hook] \arrow[ur, hook'] \arrow[urr, "\psi"'] & &
\end{tikzcd}$$

By identification via isomorphisms and by writing $U = (\tilde{U}, f, h)$, where f and h are the interpretations of tr and *invo* in U, we get the diagram

$$\begin{tikzcd}
(M_n(\tilde{K}), tr_K, X \mapsto X^T) & & (M_n(\tilde{L}), tr_L, X \mapsto X^T) \\
& (\tilde{U}, f, h) \arrow[ul, "\varphi"] \arrow[ur, "\psi"'] &
\end{tikzcd}$$

By R we denote the subring of U that is generated by the image of f, that is: $R = (f(U))$. By Lemma 3.24 1, R is commutative and $\varphi(R) \subseteq K \cdot I_n$ and $\psi(R) \subseteq L \cdot I_n$. In the future, we will not strictly distinguish between K and $K \cdot I_n$.

Since φ and ψ are $\mathcal{L}$-embeddings, $M_n(\tilde{K})$ and $M_n(\tilde{L})$ induce the same $\mathcal{L}$-structure $\tilde{R}$ on R. Also note that consequently, $\varphi|_R : \tilde{R} \to \tilde{K}$ and $\psi|_R : \tilde{R} \to \tilde{L}$ may be regarded as $\mathcal{L}$-embeddings. As in Lemma 3.24 2, we now amalgamate over the field $\tilde{F}$ - using that $R \subseteq K, L$ and that $\tilde{F}$ admits quantifier elimination - and get the diagram

$$\begin{tikzcd}
& (M_n(\tilde{\Omega}), tr_{\Omega}, X \mapsto X^T) & \\
(M_n(\tilde{K}), tr_K, X \mapsto X^T) \arrow[ur, hook, "\overline{\varepsilon}"] & & (M_n(\tilde{L}), tr_L, X \mapsto X^T) \arrow[ul, hook', "\overline{\delta}"'] \\
& (\tilde{U}, f, h) \arrow[ul, "\varphi"] \arrow[ur, "\psi"'] & \\
& (\tilde{R}, f_R, h_R) \arrow[uul, "\varphi|_R"] \arrow[uur, "\psi|_R"'] \arrow[u, hook] &
\end{tikzcd}$$

There are two things to note. First, $h|_R$ is the identity, since R - being generated by center elements alone (since $f = tr_K|_U = tr_L|_U$) - only consists of center elements, too. Second, only the outer diagram (excluding $(\tilde{U}, f, h)$) may generally be assumed to be commutative (for that, recall that $\varphi(R) \subseteq K$).

Now we recall that U is finitely generated, that is: There are $X_1, \dots, X_d \in U$ that generate U as a ring. We claim that there is an orthogonal matrix $O \in M_n(\Omega)$ that satisfies

$$O\overline{\varepsilon}(\varphi(X_i))O^T = \overline{\delta}(\psi(X_i))$$

for all $i = 1, \dots, d$. Once this claim is established, we define

$$\begin{aligned} \gamma : M_n(\Omega) &\to M_n(\Omega) \\ X &\mapsto OXO^T \end{aligned}$$

where O is chosen as described above. Clearly, γ preserves the trace and transposition. Thus, γ is an $\mathcal{L}(tr, invo)$-automorphism of $(M_n(\tilde{\Omega}), tr_\Omega, [X \mapsto X^T])$. Furthermore, we have $\gamma \circ \overline{\varepsilon} \circ \varphi = \overline{\delta} \circ \psi$ and these mappings define an amalgamation of M and N over φ and ψ. It is therefore left to show that such an O exists.

For that, we write $Y_i := \overline{\varepsilon}(\varphi(X_i))$ and $Z_i := \overline{\delta}(\psi(X_i))$. We let ω be an arbitrary word in the variables x, x^T and define $X := \omega(X_1, \dots, X_d, h(X_1), \dots, h(X_d)) \in U$.

By Lemma 3.24 2, we immediately get that $tr_\Omega(\overline{\varepsilon}(\varphi(X))) = tr_\Omega(\overline{\delta}(\psi(X)))$. Since both $\overline{\varepsilon}$ and φ respect the interpretations of tr and $invo$, we get

$$\begin{aligned} \overline{\varepsilon}(\varphi(X)) &= \overline{\varepsilon}(\varphi(\omega(X_1, \dots, X_d, h(X_1), \dots, h(X_d)))) = \\ &= \omega(\overline{\varepsilon}(\varphi(X_1)), \overline{\varepsilon}(\varphi(X_d)), \overline{\varepsilon}(\varphi(X_1)^T), \overline{\varepsilon}(\varphi(X_d)^T)) = \\ &= \omega(Y_1, \dots, Y_d, Y_1^T, \dots, Y_d^T) \end{aligned}$$

Analogously,

$$\overline{\delta}(\psi(X)) = \omega(Z_1, \dots, Z_d, Z_1^T, \dots, Z_d^T)$$

Thus, $tr_\Omega(\omega(Y, Y^T) = tr_\Omega(\omega(Z, Z^T)$ This proves the existence of O by the (SPT) in Ω. Indeed, by assumption, F has the (SPT) and since $(M_n(\tilde{\Omega}), tr_\Omega, [X \mapsto X^T])$ is elementarily equivalent to $(M_n(\tilde{F}), tr_\Omega, [X \mapsto X^T])$, so does Ω.

This uses the degree bound D in the (SPT) as this allows us to rewrite the (SPT) following Theorem 3.25 as the $\mathcal{L}(tr, invo)$-sentence

$$\begin{aligned} \forall X_1, \dots, X_D, Y_1, \dots, Y_D : &\left(\bigwedge_{\substack{\omega \\ \deg(\omega) \leq D}} tr(\omega(X, X^T)) = tr(\omega(Y, Y^T)) \right) \\ &\to \left(\exists O : O^T O = 1 \wedge \bigwedge_{i=1}^{d} O^T X_i O = Y_i \right) \end{aligned}$$

1) $\implies$ 3): This follows from quantifier elimination. Note that in the given scenario - U being a two-fold substructure of $(M_n(\tilde{K}), tr_K, [X \mapsto X^T])$ - we may first choose an arbitrary amalgamation $\overline{\varepsilon}, \overline{\delta}$ into $(M_n(\tilde{\Omega}), tr_\Omega, [X \mapsto X^T])$ as given in the diagram below.

$$\begin{tikzcd}
 & (M_n(\tilde{\Omega}), tr_{\Omega}, X \mapsto X^T) \\
(M_n(\tilde{K}), tr_K, X \mapsto X^T) \arrow[ur, hook, "\overline{\varepsilon}"] & (M_n(\tilde{K}), tr_K, X \mapsto X^T) \arrow[u, hook, "\overline{\delta}"'] \\
U \arrow[u, hook, "id"'] \arrow[ur, hook, "\psi"'] &
\end{tikzcd}$$

Shuffling around elements (using that $\overline{\varepsilon}$ and $\overline{\delta}$ as embeddings are injective) lets us choose $\overline{\varepsilon}$ as an extension of ψ and $\overline{\delta}$ as the identity. By model completeness of $(M_n(\tilde{F}), tr_F, [X \mapsto X^T])$, we see that the extension given by $\overline{\delta}$ must be elementary. As restricting our considerations to $\mathcal{L}$ and to centers does not change that the embedding is elementary, we immediately get that $\tilde{K} \preceq \tilde{\Omega}$

3) $\implies$ 2): We first prove the claim without a degree bound for all $\tilde{K} \equiv \tilde{F}$. Then we argue by compactness to establish the bound. For that, let $d \in \mathbb{N}, X_1, \ldots, X_d, Y_1, \ldots, Y_d \in M_n(F)$ be given, such that $tr(\omega(X, X^T)) = tr(\omega(Y, Y^T))$ for all words in the $2d$ variables x, x^T.

Let U' be the $\mathcal{L}(tr, invo)$-substructure of $(M_n(\tilde{K}), tr_K, [X \mapsto X^T])$ that is generated by $K \cdot I_n$ and $X_1, \ldots, X_d$ for some $d \in \mathbb{N}$. We define $U := U'|_{\mathcal{L}_{Ri}}$ and interpret U as the K-algebra that is generated by the evaluations in $X_1, \ldots, X_d$ of all words in the $2d$ variables x, x^T. Now let $\varphi : U \to M_n(K)$ be the identity and let $\psi : U \to M_n(K)$ be the K-algebra homomorphism that is defined by setting $\psi(X_i) := Y_i, \psi(X_i^T) := Y_i^T$. This gives us the diagram

$$\begin{tikzcd}
M_n(K) & M_n(K) \\
U \arrow[u, hook, "id"'] \arrow[ur, hook, "\psi"'] &
\end{tikzcd}$$

We now claim the following:

- ψ is well-defined.
- ψ is an $\mathcal{L}(tr, invo)$-homomorphism.

For well-definedness, it is sufficient to prove that for any non-commutative polynomial $p(x, x^T)$, we have the implication that if $p(X, X^T) = 0$, then $p(Y, Y^T) = 0$. By Lemma 3.26, we know that assuming this premise implies $tr(p(X, X^T)^T p(X, X^T)) = 0$. This trace, however, is just a linear combination of words in X and X^T. By the condition of the (SPT), we thus conclude $tr(p(Y,Y^T)^T p(Y,Y^T)) = 0$ and again by Lemma 3.26, we get $p(Y,Y^T) = 0$. It is clear that ψ respects both trace and involution.

This lets us find an amalgamation as defined in 3. Namely, we find $\tilde{\Omega}$ such that $\tilde{K} \preceq \tilde{\Omega}$ and an $\mathcal{L}$-embedding $\overline{\varepsilon} : M_n(\tilde{K}) \to M_n(\tilde{\Omega})$ preserving tr and $invo$ such that $\psi(u) = \overline{\varepsilon}(u)$ for all $u \in U$. Since ψ is a K-algebra homomorphism and $K \subseteq R$, $\overline{\varepsilon}$, too, is a K-algebra homomorphism. We now want to apply the theorem of Skolem-Noether (Theorem 2.8). For that, we first replace $M_n(\tilde{K})$ by $M_n(\tilde{K}) \otimes_K \tilde{\Omega} \cong M_n(\tilde{\Omega})$. We will denote the resulting Ω-algebra homomorphism that is given by $M \otimes \omega \mapsto \omega \cdot \overline{\varepsilon}(M)$ still by $\overline{\varepsilon}$.

We may now apply the theorem of Skolem-Noether working over the algebra $M_n(\Omega)$ (also recall Lemma 2.9) and find some $Z \in M_n(\Omega)^\times$ satisfying $\overline{\varepsilon}(X \otimes \omega) = Z^{-1}(X \otimes \omega)Z$ for all $X \in M_n(K)$. Restricting to elementary tensors of the form $M \otimes \omega \cong \omega M \in M_n(\tilde{K})$, we conclude that

$$Z^{-1}X^T Z = \overline{\varepsilon}(X^T) = \overline{\varepsilon}(X)^T = (Z^{-1}XZ)^T = Z^T X^T (Z^{-1})^T$$

Thus, $ZZ^T X^T = X^T ZZ^T$ for all $X \in M_n(\Omega)$ and ZZ^T lies in the center of $M_n(\Omega)$. Further, we know that there is some $\lambda \in \sum \Omega^2$ with $ZZ^T = Z^T Z = \lambda I_n$ by positivity of ZZ^T and Lemma 2.1.

As the amalgamation diagram commutes, we have

$$Z^{-1} X_i Z = Y_i$$

for all $i = 1, \dots, d$. Since Ω is Pythagorean, λ is a square itself and we replace Z by $\frac{Z}{\sqrt{\lambda}}$. This lets us choose $\lambda = 1$ without loss of generality.

Finally, we define $O := Z$ and observe that O is orthogonal with coefficients in Ω satisfying $O^T X_i O = Y_i$ for all $i = 1, \dots, d$. Since $K \preceq \Omega$, we may choose O with coefficients in K and we are done.

It is only left to prove the claim with the degree bound. For that, we first let a non-principal ultrafilter on $\mathbb{N}$ be given and we assume that $\tilde{K}$ is an ultrapower of $\tilde{F}$ with respect to that ultrafilter. $\tilde{K} \equiv \tilde{F}$ follows from the fact that $\tilde{F} \preceq \tilde{K}$ by [PD11], Corollary 2.6.3.

We now use the abbreviations $X = (X_1, \dots, X_d), Y = (Y_1, \dots, Y_d)$ and for every $k \in \mathbb{N}$, we define $\mathcal{W}_k$ as the set of all words of degree at most k. Then, for every $k \in \mathbb{N}$, we define the set

$$S_k := \left\{ (X,Y) \in M_n(\tilde{K})^{2d} | \forall \omega \in \mathcal{W}_k : tr(\omega(X,X^T)) = tr(\omega(Y,Y^T)) \right\}$$

For every fixed k, S_k is clearly definable in $\mathcal{L}(tr, invo)$, since there are only finitely many words. The intersection $\bigcap_{k,d \in \mathbb{N}} S_k$ is also definable by the (SPT) without the degree bound as derived above. Namely, we define it by the formula

$$\exists O : OO^T = I_n \wedge \bigwedge_{i=1}^{d} O^T X_i O = Y_i$$

Finally, since $\tilde{K}$ is $\aleph_1$-saturated by [PD11], Theorem 2.6.5., we get that the intersection is finite and thus finish the proof (for the argument in the last step - in the formulation for real closed fields - see [PD01], Theorem 2.2.11.). □

It is natural to ask for fields that are both Pythagorean and possess the (SPT). One whole class of such fields - intersections of real closed fields - will be presented below.

Lemma 3.31. *Let F be a field. If F is the intersection of Pythagorean fields, then F is Pythagorean.*

Proof. We write

$$F = \bigcap_{\lambda \in \Lambda} P_\lambda$$

where each P_λ is Pythagorean and embeddable into a large field K. Assume that s is a sum of squares in F. Then s is a sum of squares in each P_λ and we see that s is in fact a square in every such field. Since the square root of an element is unique (up to multiplication with -1) if it exists, s is the same square in every P_λ and thus it lies in F. □

Definition 3.32 ([MSV93], p. 749)**.** Let F be a field. We say that F satisfies the principal axis property, if every symmetric matrix over F is orthogonally similar to a diagonal matrix over F.

Lemma 3.33 ([MSV93], Corollary to Theorem 2)**.** *Let F be a field. If F is the intersection of real closed fields, then F satisfies the principal axis property.*

Proof. [MSV93], Theorem 2 and the subsequent Corollary. □

Theorem 3.34 ([KT20], 2.2.5.)**.** *Let F be a field and let F be the intersection of real closed fields. Then F has the (SPT) and its constant D can be chosen as n^2.*

Proof. Let $d \in \mathbb{N}, X_1, \ldots, X_d, Y_1, \ldots, Y_d \in M_n(F)$ be given such that $tr(\omega(X, X^T)) = tr(\omega(Y, Y^T))$ for all words of length at most n^2 in the $2d$ variables x, x^T.

By Theorem 3.25, for any real closed field $F \subseteq R$, we find $U \in M_n(R)$ such that U is orthogonal and $U^T X_i U = Y_i$ for all $i = 1, \ldots, d$. Now we consider the system of linear equations

$$X_i P = P Y_i$$
$$X_i^T P = P Y_i^T$$

for $i = 1, \ldots, d$. Applying the Gauß Algorithm, we may parametrise all solutions over F as an F-vector space. Hence, they have the form $b_1 P_1 + \ldots + b_r P_r$ for some $r \leq n$ and $b_k \in F, P_k \in M_n(F)$ for all $k = 1, \ldots, r$. Also note that $(P_1, \ldots, P_k)$ is a basis of the space. If we consider the system of equations over R instead, then we will get the same parametrisation except that $b_k \in R$ for $k = 1, \ldots, r$. Additionally, we know that there exists an invertible solution to the system over R. Its determinant is given as the evaluation of a polynomial $p \in F[x_1, \ldots, x_r]$ in the coefficients $b_1, \ldots, b_r$. Hence, p is not the zero polynomial and it must necessarily be non-zero in some evaluation in F. Thus, we find a matrix $P \in M_n(F)$ that is invertible and solves the above system of equations.

This implies $P^{-1} X_i P = Y_i, P^{-1} X_i^T P = Y_i^T$ for all $i = 1, \ldots, d$. In particular,

$$P^{-1} X_i^T P = Y_i^T = (P^{-1} X_i P)^T = P^T X_i^T (P^T)^{-1}$$

and hence PP^T commutes with all X_i and X_i^T.

By Lemma 3.33, F has the principal axis property and we can therefore diagonalise PP^T by an orthogonal matrix $V \in M_n(F)$, that is: There is a diagonal matrix $D \in M_n(F)$ satisfying $V^T P P^T V = D$. By construction, each entry of D is a sum of squares and thus a square itself as F is Pythagorean by Lemma 3.31. This lets us define a squareroot $\sqrt{D} \in M_n(F)$.

Defining $H := V \sqrt{D} V^T \in M_n(F)$, we observe that $H = H^T$ and

$$H^2 = V \sqrt{D} V^T V \sqrt{D} V^T = V \sqrt{D}^2 V^T = P P^T$$

That means that H is a symmetric square root of PP^T and by Lemma 3.35 below, H commutes with everything that commutes with PP^T.

Finally, we set $O := H^{-1} P$ and observe that

$$O^T O = P^T H^{-1} H^{-1} P = P^T H^{-2} P = P^T (P P^T)^{-1} P = I_n$$

and

$$O^T X_i O = O^{-1} X_i O = P^{-1} H X_i H^{-1} P = P^{-1} X_i H H^{-1} P = Y_i$$

Thus, we have found O as desired. □

The argument that proves the existence of $P \in M_n(F)$ stems from a kind personal conversation with Igor Klep, one of the authors of [KT20].

There is one step in the above proof left to be shown. This is done in the next lemma.

Lemma 3.35. *In the above proof, H commutes with every matrix that commutes with H^2.*

Proof. Let V and D be given as in the proof above. We note that $H^2X = XH^2$ if and only if $VDV^TX = XVDV^T$ or equivalently $DV^TXV = V^TXVD$. We write $Y = V^TXV$ and denote the diagonal elements of D by $(d_{ii})_{i=1}^n$.

Now we observe that $d_iY_{ij} = (DY)_{ij} = (YD)_{ij} = d_jY_{ij}$ for all $i, j = 1, \ldots, n$. Therefore, if $Y_{ij} \neq 0$, then $d_i = d_j$ and in particular, $\sqrt{d_i} = \sqrt{d_j}$. This lets us deduce $(\sqrt{D}Y)_{ij} = \sqrt{d_i}Y_{ij} = \sqrt{d_j}Y_{ij} = (Y\sqrt{D})_{ij}$, which proves the claim by $\sqrt{D} = V^THV$ and since X and Y stand in bijection. □

Corollary 3.36 ([KT20], 2.2.6.)**.** *Let F be an intersection of real closed fields and let $\tilde{F}$ be an expanson of F in a language $\mathcal{L}$ extending $\mathcal{L}_{Ri}$. We suppose that $\tilde{F}$ has quantifier elimination in $\mathcal{L}$ and extend $\mathcal{L}$ to $\mathcal{L}(tr, invo)$ by two new unary function symbols. Then $(M_n(\tilde{F}), tr_F, [X \mapsto X^T])$ admits quantifier elimination in $\mathcal{L}(tr, invo)$.*

Proof. This is a direct consequence of Theorem 3.30 and Lemma 3.34. □

Remark 3.37 ([KT20], 2.4.1.)**.** The proof of Theorem 3.30 does not work for complex matrices with trace and transposition, which is shown in [KT20], Example 2.3.3. Similarly, replacing the transposition with the involution will not allow us to give a proof along the lines of Theorem 3.30 as there is no canonical way to interpret an involution on an arbitrary given matrix algebra $M_n(\Omega)$. Instead, we are only able to prove quantifier elimination for the structure $(M_n(\mathbb{C}), \leq, tr_{\mathbb{C}}, [X \mapsto X^*])$ using the (SPI). Here, we interpret $\leq$ on $\mathbb{R} \cdot I_n$. Note that A^* becomes meaningful for the algebraic closure F of any real closed field R, since we always have $F = R[\sqrt{-1}]$ by Theorem 2.15. As our base field in Theorem 3.30, we then consider $\tilde{\mathbb{C}} := (\mathbb{C}, [z \mapsto \overline{z}], \leq)$, where the complex conjugation lets us define $\mathbb{R}$ as a subset of $\mathbb{C}$.

4 Applications

The previous chapter guarantees that in $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$, the following formulae can be expressed without quantifiers using only the trace and transposition (or involution reespectively).

Description	Formula
Invertibile matrices	$\exists B : (AB = 1 \wedge BA = 1)$
Positive semi-definite matrices	$\exists B : A = B^*B$
$\|\cdot\|_2$-contractions	$\forall \Lambda : (\Lambda \in F \cdot I_n \to (\Lambda - A^*A \notin GL_n(F) \to \Lambda < 1))$
Existence of roots of a matrix polynomial P	$\exists X_1, \ldots, X_d : P(X_1, \ldots, X_d) = 0$

Table 4.1: Formulae that may be expressed without quantifiers.

Note that in the second and the third example, the involution coincides with the transposition in the case of real matrices. In the third example, we have additionally used the obvious abbreviations for the formulae describing that a matrix lies in the center of $M_n(F)$ or that it is invertible over F.

For the last example, note that uniqueness may additionally be postulated by the formula $\forall X, Y : P(X) = 0 = P(Y) \to X = Y$. Further note that we may also phrase the formula for inequalities, replacing the symbol $=$ by the property that the evaluation of the polynomial shall be positive (semi-)definite. Another immediate generalisation is to ask for solutions of systems of polynomial equations.

For the above examples, we will present equivalent quantifier-free formulae using only involution and trace. The results will mainly rely on linear algebra. We begin with a characterisation of positive semi-definite matrices as this result will occur repeatedly in the subsequent considerations. In the following, F always denotes either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$.

4.1 Positive Semi-Definite Matrices

Lemma 4.1. *If $A \in M_n(F)$ is a positive semi-definite real- or complex-valued matrix and if $tr(A) = 0$, then $A = 0$.*

Proof. Since all eigenvalues of A are non-negative by assumption and their sum is 0, A only has the eigenvalue 0. Therefore, A is also negative semi-definite. We now observe that consequently,

$$x^* A x = 0$$

for all column vectors x. Hence, by decomposing $A = P^* P$, we conclude

$$0 = x^* A x = (Px)^* P x = \langle Px, Px \rangle$$

for all x and thus $Px = 0$ for all x. This implies $P = 0$ and consequently $A = 0$. □

The following lemma might appear obsolete in light of the more general results below. It is good, though, for giving an idea of the behaviour of the eigenvalues of a positive semi-definite matrix.

Lemma 4.2. *A hermitian matrix $A \in M_2(F)$ is positive semi-definite if and only if*

$$tr(A) + \sum_{i=1}^{k} \left(tr(A)^i - tr(A^i) \right) \geq 0$$

for all $k \in \mathbb{N}$.

Proof. $\Longleftarrow$: Considering $k = 1$ gives us $tr(A) \geq 0$. Hence, we distinguish two cases.

- Case 1: $tr(A) = 0$. Then $tr(A)^2 = 0$ as well. Since $A^2 = A^* A$ is positive semi-definite, $tr(A^2) \geq 0$. If $tr(A^2) > 0$, we get an immediate contradiction by observing that for $k = 2$, the sum above simplifies to $-tr(A^2)$. If $tr(A^2) = 0$, then $A^2 = 0$ by Lemma 4.1. This implies $A = 0$, since otherwise, $A^2 \neq 0$ as well by symmetry.

- Case 2: $tr(A) > 0$. By scaling, we may assume $tr(A) = 1$. Assume that A is not positive semi-definite. Then A has one positive non-zero eigenvalue $\lambda_+ > 1$ and one negative non-zero eigenvalue $\lambda_- = 1 - \lambda_+$. In particular, $tr(A) < \lambda_+$ and $1 = tr(A)^k < \lambda_+^k + \lambda_-^k = tr(A^k)$ for all $k \geq 2$, where we deduce the inequality from

$$\lambda_+^k - |\lambda_-|^k = (\lambda_+ - |\lambda_-|) \cdot (\lambda_+^{k-1} + |\lambda_-| \lambda_+^{k-2} + \ldots + |\lambda_-|^{k-1}) \geq 1$$

Thus, each term in the above sum must be negative. Since A^{2k} is positive semi-definite and $\lambda_+^{2k} > 1 + \varepsilon$ for all $k \in \mathbb{N}$ and some fixed $\varepsilon > 0$, it is not summable and thus

$$tr(A) + \sum_{i=1}^{k} \left(tr(A)^k - tr(A^k)\right) < 0$$

for sufficiently large $k \in \mathbb{N}$.

$\Longrightarrow$: If A is positive semi-definite, then its trace must be positive. It is thus sufficient to prove that $tr(A)^k \geq tr(A^k)$ for all $k \in \mathbb{N}$. Using that both eigenvalues λ_1, λ_2 of A are non-negative by assumption, this follows from

$$tr(A)^k = (\lambda_1 + \lambda_2)^k = \lambda_1^k + \lambda_2^k + d \geq \lambda_1^k + \lambda_2^k = tr(A^k)$$

where d denotes all mixed terms after reducing the binomial. □

Lemma 4.2 is not useful for checking whether a matrix is positive semi-definite. In fact, we cannot bound the range of the sum. That is, we will always find a matrix A such that the sum is positive for all $k \leq K$, where K is arbitrary but fixed. However, the same matrix fails to be positive semi-definite, which can only be established by checking the sum for some $n > K$. Consider, for example, the matrix

$$\begin{pmatrix} 1+\varepsilon & 0 \\ 0 & -\varepsilon \end{pmatrix}$$

For any given $K \in \mathbb{N}$, we may obviously choose $\varepsilon > 0$ small enough such that

$$tr(A) + \sum_{i=1}^{k} \left(tr(A)^i - tr(A^i)\right) \geq 0$$

for all $k \leq K$. Since, however, $(1+\varepsilon)^k$ diverges for $k \to \infty$, we get that ultimately, the condition will not hold for all natural numbers.

Instead, we will now give a handy - and crucially quantifier-free - criterion that will work for all real and complex 2×2-matrices.

Theorem 4.3. *The following statements about a hermitian matrix $A \in M_2(F)$ are equivalent:*

1. *A is positive semi-definite.*

2. *Both of the following hold:*

- $tr(A) \geq 0$
- $tr(A)^2 - tr(A^2) \geq 0$

Proof. $\Longrightarrow$: We first note that if A is positive semi-definite and has trace 0, then $A = 0$ by Lemma 4.1. If its trace is non-zero, then it must be strictly positive and $tr(A)^2 - tr(A^2) \geq 0$ is a special case of the proof of Lemma 4.2.

$\Longleftarrow$: If $tr(A) = 0$, then by assumption $tr(A^2) = 0$ as well. Consequently the two eigenvalues λ_1, λ_2 of A satisfy $\lambda_1 + \lambda_2 = \lambda_1^2 + \lambda_2^2 = 0$. Since A is hermitian, both λ_1 and λ_2 are real and they must consequently be zero. Hence, A is positive semi-definite.

Now assume that $tr(A) > 0$. Without loss of generality, we may assume that $tr(A) = 1$. Based on the assumption that A is not positive semi-definite, we now deduce a contradiction. For that, we note that both eigenvalues of A are real numbers by Lemma 2.2. Denoting them λ_+, λ_- we observe that they necessarily satisfy $1 < \lambda_+$ and $\lambda_- < 0$. In particular, we get

$$tr(A)^2 = 1 < \lambda_+^2 + \lambda_-^2 = tr(A^2)$$

in contradiction to our assumption. This finishes the proof. □

The result of Theorem 4.3 can be generalised for arbitrary matrix sizes, though we will have to fundamentally adapt our methods and adapt to considering the eigenvalues of a matrix wholistically - instead of one by one. The characteristic polynomial of a matrix will allow us to do just that. We carry out the generalisation in the case of 3×3-matrices, which will give a motivation for the full characterisation of positive semi-definite matrices of arbitrary size $n \times n$ in terms of the trace and the transposition. However, we must first derive some preliminaries. The first such result - Descartes' rule of signs and some of its immediate corollaries - will allow us to determine the number of positive roots of a polynomial given only certain information about its coefficients. We will later apply this rule to characteristic polynomials.

Definition 4.4. Let $a = (a_i)_{i=0}^n$ be a real sequence. Then we define $SC(a)$ to be the number of sign changes in a. More precisely, it is the number of indices $i \in \{0, \ldots, n\}$ such that for the first index $j > i$ with $a_j \neq 0$, we get $sign(a_j) = -sign(a_i)$.

If a is the coefficient sequence of a polynomial p, then we define $SC(p) := SC(a)$.

Lemma 4.5 (Descartes' rule of signs, [Wan04b])**.** *Let $p = \sum_{i=0}^n a_i t^{b_i}$ be a polynomial with natural exponents $0 \leq b_0 < \ldots < b_n$ and the real coefficient sequence $(a_i)_{i=0}^n$ satisfying $a_i \neq 0$ for all i. Then p has $SC(p) - 2l = SC(a) - 2l$ positive (non-zero) roots for some $l \in \mathbb{N}$.*

Proof. We follow the proof given in [Wan04b]. As a first simplification, we may factor out the monomial t^{b_0} and then assume without loss of generality that $b_0 = 0$. We now claim that if $a_0 a_n > 0$, then the number of positive roots of p is even, and that if $a_0 a_n < 0$, then the number of positive roots of p is odd. We argue by a case distinction and consider 4 cases:

- Case 1: $a_0 > 0, a_n > 0$: This implies that $p(0) > 0$ and $p(x) \overset{x\to\infty}{\longrightarrow} \infty$. Hence, the positive x-axis will be crossed an even number of times. If the graph of p touches the positive x-axis without truly crossing it, we have a root of p with an even multiplicity. A true crossing will give rise to a root with an odd multiplicity. Therefore, we have an even number of roots with an odd multiplicity and some other roots with an even multiplicity, which taken together proves the claim.

- Case 2: $a_0 < 0, a_n < 0$: We replace p by $-p$ and repeat the proof of the first case.

- Case 3: $a_0 > 0, a_n < 0$: A similar argument as above gives us an uneven number of (true) crossings of the graph of p with the positive x-axis. This proves the claim.

- Case 4: $a_0 < 0, a_n > 0$: Again, we look at $-p$ instead of p and are in the previous case.

In the following, let $z(p)$ denote the number of positive roots of p. Now we deduce our initial claim by induction on the number n of non-zero coefficients of p. If $n = 1$, there is nothing to prove.

Now assume that the claim holds up to $n-1$ for some $n \in \mathbb{N}$. Then we may prove the induction step by distinguishing two cases. We define p' as the formal derivative of p.

- Case 1: $a_0 a_1 > 0$: Then $SC(p) = SC(p')$. By the above considerations, we then know that both $z(p)$ and $z(p')$ have the same parity, so $z(p) \equiv z(p') \mod 2$. The induction hypothesis yields $z(p') \leq SC(p')$ and $z(p') \equiv SC(p') \mod 2$. Taken together, we conclude $z(p) \equiv SC(p) \mod 2$. We further know by Rolle's Theorem that $z(p') \geq z(p) - 1$. Thus,
$$SC(p) = SC(p') \geq z(p') \geq z(p) - 1 > z(p) - 2$$
and we conclude $z(p) < SC(p) + 2$ and hence $z(p) \leq SC(p)$.

- Case 2: $a_0 a_1 < 0$: Then $SC(p') = SC(p) - 1$. We immediately get that $z(p) - z(p') \equiv 1 \mod 2$ by the above considerations about the parity of roots. Again, we use the induction hypothesis to get $z(p') \leq SC(p')$ and $z(p') \equiv SC(p') \mod 2$ and infer $z(p) \equiv SC(p) \mod 2$. Rolle's theorem now states that $z(p') \geq z(p) - 1$ and we conclude
$$SC(p) = SC(p') + 1 \geq z(p') + 1 \geq z(p)$$

thus finishing the second case and the proof.

□

Corollary 4.6. *Let p be a polynomial with alternating non-zero coefficients. Then all real roots of p are positive.*

Proof. Note that the coefficients of the polynomial $q := p(-t)$ all have the same sign and thus, q has no positive real roots by Lemma 4.5. The claim now follows from the definition of q. □

Corollary 4.7. *Let $p = \sum_{i=0}^{n} a_i t^i$ be a polynomial with real coefficients satisfying $a_{i+1} \geq 0 \iff a_i \leq 0$ for all $i = 0, \ldots, n-1$. Further assume that p has no roots in $\mathbb{C} \backslash \mathbb{R}$. Then all roots of p are non-negative. Furthermore, all zero coefficients of p are grouped together at the beginning of its coefficient sequence.*

Proof. Note that $p(-t)$ has no sign changes in its coefficient sequence. Thus, by Theorem 4.5, p has no strictly negative roots. Since all roots are real by assumption, they must be non-negative.

For the second claim, we factor out t^k with maximal k and call the resulting polynomial q. Then q is a polynomial that has only strictly positive roots and thus, by Theorem 4.5, all coefficients of q must be alternating in signs. In particular, they are non-zero. □

A second preliminary result of interest are the Newton Identities. They allow for the calculation of the sums of the powers of a polynomial's roots. If we assume that polynomial to be the characteristic polynomial of a matrix A, this immediately gives us the traces of powers of A. The following results hold for more general fields that just $\mathbb{R}$ and $\mathbb{C}$ and the proofs remain the same. However, this generality is not necessary here. Our formulations closely follow [Net18].

Definition 4.8 (Newton Sums)**.** Let p be a polynomial over F and let $\alpha_1, \ldots, \alpha_n$ denote its roots over the algebraic closure of F. We define the k-th Newton Sum of p by

$$\nu_k(p) := \alpha_1^k + \ldots + \alpha_n^k$$

Theorem 4.9 (Newton Identities)**.** *Let $p = t^n + a_1 t^{n-1} + \ldots + a_{n-1} t + a_n$ be a monic polynomial over F. For all $k \geq 1$ we have the following identity:*

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

Proof. We follow the combinatorial proof given in [Zei84]. We let $\alpha_1, \ldots, \alpha_n$ denote the roots of p over the algebraic closure of F. Using the elementary symmetric polynomials $s_{r,n}$ in $\alpha_1, \ldots, \alpha_n$, we then observe that the coefficients of p may also be written as

$$a_r = (-1)^r s_{r,n} = (-1)^r \sum_{1 \leq i_1, < \ldots < i_r \leq n} \alpha_{i_1} \cdots \alpha_{i_r}$$

Thus, we may write the Newton Identities as

$$\sum_{r=0}^{k-1} (-1)^r \left(\sum_{1 \leq i_1, < \ldots < i_r \leq n} \alpha_{i_1} \cdots \alpha_{i_r} \right) \left(\sum_{j=1}^{n} \alpha_j^{k-r} \right) + (-1)^k \left(\sum_{1 \leq i_1, < \ldots < i_k \leq n} \alpha_{i_1} \cdots \alpha_{i_r} \right) k = 0$$

We now consider the set $\mathcal{A}$ dependent on n and k consisting of all pairs (A, j^l) that satisfy the following:

1. $A \subseteq \{1, ..., n\}$
2. $j \in \{1, ..., n\}$
3. $|A| + l = k$
4. $l \geq 0$ and $l = 0$ implies $j \in A$

We equip each such pair with a weight given by $\omega(A, j^l) := (-1)^{|A|} \left(\prod_{i \in A} \alpha_i \right) \alpha_j^l$. Adding the weights of all pairs in $\mathcal{A}$, we get the left hand side of the Newton Identities. It is left to show that it equals zero. For that, we define a map $T : \mathcal{A} \to \mathcal{A}$ by

$$T(A, j^l) = \begin{cases} (A \backslash \{j\}, j^{l+1}) & , j \in A \\ (A \cup \{j\}, j^{l-1}) & , j \notin A \end{cases}$$

It is obvious that the codomain of T is $\mathcal{A}$ and that it is bijective (by surjectivity on finite sets of equal size) without having a fix point. Furthermore $T^2 = id_{\mathcal{A}}$. Since we also have $\omega(T(A, j^l)) = -\omega(A, j^l)$, we may arrange the weights in mutually cancelling pairs, which proves the claim. □

Corollary 4.10. *If p is a polynomial over F, then $\nu_k(p) \in F$ for all $k \in \mathbb{N}$.*

Proof. Note that $-\nu_1$ appears as a coefficient of p and hence lies in F. Thus, by the recursion given in Theorem 4.9, we derive the claim. □

Corollary 4.11. *If $A \in M_n(F)$ is a polynomial over a field F, and χ_A denotes its characteristic polynomial, then $\nu_k(\chi_A) = tr(A^k)$ for all $k \in \mathbb{N}$.*

Proof. This follows from the fact that if λ is an eigenvalue of A, then λ^k is an eigenvalue of A^k. □

The following theorem and its proof contain all essential arguments that we need to find a characterisation of all positive semi-definite matrices. It is the result of combining both Descartes' rule of signs and the Newton Idendities.

Theorem 4.12. *The following statements about a hermitian matrix $A \in M_3(F)$ are equivalent:*

1. *A is positive semi-definite.*
2. *All of the following hold:*
   - $tr(A) \geq 0$
   - $tr(A)^2 - tr(A^2) \geq 0$
   - $2tr(A^3) - 3tr(A^2)tr(A) + tr(A)^3 \geq 0$

Proof. We begin by giving an idea on how the result may be derived. The case that $A = 0$ is trivial and should be regarded as a fringe case. By scaling we assume that $tr(A) = 1$ whenever $tr(A) > 0$. We first note that by (repeatedly) applying the Newton Identities in Lemma 4.9 to the characteristic polynomial $\chi_A = \sum_{i=0}^{n} a_{n-i}t^i$ of A we get the identities

$$\begin{aligned} tr(A^2) &= 1 - 2(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3) \\ tr(A^3) &= 1 - 3(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3) + 3\lambda_1\lambda_2\lambda_3 \end{aligned}$$

where $\lambda_i, i = 1, 2, 3$ denote the eigenvalues of A. Note that if A is positive semi-definite, all eigenvalues are non-negative and a simple observation lets us conclude

$$\begin{aligned} 1 - tr(A^2) &= 2(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3) \geq 0 \\ 2tr(A^3) - 3tr(A^2) + 1 &= 6\lambda_1\lambda_2\lambda_3 \geq 0 \end{aligned}$$

Here, we also see that the coefficients of χ_A must be alternating (with the last few coefficients potentially being 0, depending on the multiplicity of the eigenvalue 0). This motivates the following converse implication.

Assume that A is hermitian. Then χ_A is a polynomial with only real roots and of the form

$$\chi_A = \sum_{i=0}^{n} a_{n-i}t^i$$

with $a_0 = 1$. Further assume that $a_{n-i} \neq 0$ for all $i = 0, \dots, k$ for some $k \leq n$ and assume that the coefficients are alternating. We note that by Lemma 4.7, χ_A has no strictly negative roots. In other words, χ_A has only non-negative roots and A is positive semi-definite. Since for $n = 3$, the coefficients of χ_A are just $-(\lambda_! + \lambda_2 + \lambda_3), \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3$ and $-\lambda_1\lambda_2\lambda_3$, this proves the claim. □

Remark 4.13. We recall that for a given positive semi-definite matrix, as an immediate consequence of Corollary 4.7, any 0 coefficients of its characteristic polynomial must be grouped at the beginning of its coefficient sequence. This is especially relevant once we generalise to larger matrix sizes (see Corollary 4.16). Thus, if during testing for positive semi-definiteness a 0 coefficient occurs, all consecutive coefficients must be 0 as well.

We are now ready to generalise our result to arbitrary matrix sizes. The majority of work has already been done. Still, we need the following lemma as a preparation to do so. The lemma should be read as an inverse to Theorem 4.9.

Lemma 4.14. *Let $p = t^n + a_1 t^{n-1} + \ldots + a_{n-1}t + a_n$ be a monic polynomial of degree n and let the Newton Sums $\nu_k(p)$ be given for $k = 1, \dots, n$. Then we may calculate the coefficient sequence $(a_i)_{i=0}^n$ of p from the following recursion:*

$$\begin{aligned} a_1 &= -\nu_1(p) \\ \nu_k(p) + a_1 v_{k-1}(p) + \ldots a_{k-1}(p) + k a_k &= 0 \end{aligned}$$

Proof. This is an immediate result of the Newton Identities (Theorem 4.9). □

The explicit calculation of the coefficients of the characteristic polynomial usually follows a version of the Faddeev-LeVerrier-algorithm, see also Algorithm 8.17. and Algorithm 8.11. in [BPR06]. In the given setting, however, it is possible to omit all divisions as we are only interested in whether the calculated coefficients are positive, negative or equal to 0. This leads us to the following important remark.

Remark 4.15. In the above lemma, the Newton Identities can be multiplied by a natural number $N(k) > 0$ such that $N(k)a_k$ can be calculated as a $\mathbb{Z}$-linear combination of products of Newton Sums. This lets us use these calculations (up to integer scaling) in logical formulae.

Corollary 4.16. *The following statements about a hermitian matrix $A \in M_n(F)$ are equivalent:*

1. *A is positive semi-definite.*

2. All of the formulae

$$(-1)^k N(k) a_k \geq 0$$

for $k = 1, \ldots, n$ hold. Here, a_k shall be defined by the characteristic polynomial χ_A of A through $\chi_A = t^n + a_1 t^{n-1} + \ldots + a_{n-1} t + a_n$ and $N(k)$ is chosen in the sense of Remark 4.15.

Example 4.17. We want to give a complete characterisation of positive semi-definite matrices in $M_n(F)$ for $n \leq 5$. Note that the characterisation of 5×5-matrices naturally gives a characterisation of all smaller matrices by removing the conditions of higher degrees. For that, we postulate that a hermitian matrix $A \in M_5(F)$ satisfies the following inequalities:

- $tr(A) \geq 0$
- $tr(A)^2 - tr(A^2) \geq 0$
- $tr(A)^3 - 3tr(A)tr(A^2) + 2tr(A^3) \geq 0$
- $tr(A)^4 - 6tr(A)^2 tr(A^2) + 3tr(A^2)^2 + 8tr(A)tr(A^3) - 6tr(A^4) \geq 0$
- $tr(A)^5 + 10tr(A)^3 tr(A^2) - 15tr(A)tr(A^2)^2 - 20tr(A)^2 tr(A^3) + 20tr(A^2)tr(A^3) + 30tr(A)tr(A^4) - 24tr(A^5) \geq 0$

The calculations can be replicated using the following sample code in the Mathematica language, which does not claim to be in any way efficient. Instead, it aims at explaining the results of the previous section in as clear terms as possible. The comments (marked by the # sign) do not appear in the code used by the author and only serve explanatory purposes.

```
In[1]:= NSums = {v1, v2, v3, v4, v5};
        #defining Newton Sums of the zeros of chi_A
In[2]:= a=Table[0,5]; a[[1]]=-NSums[[1]];
        #initialising table for calculating coefficients of chi_A
        #a_1 = -v1
In[3]:= For[i=2,i<6,i++,
            S=0;
            For[j=1,j<i,j++,
                S = S+a[[i-j]]*NSums[[j]]];
            S=S+NSums[[i]];
            a[[i]]=-S/i;];
        #iteratively calculating the coefficients according to the
        #Newton Identities
In[4]:= a=Expand[a];
        #writing the coefficients of chi_A as linear combinations
        #of monomials in the Newton Sums
In[5]:= F=Table[Factorial[i],{i,5}];
        #F = (1,2,6,24,120)
In[6]:= Expand[a*F]
Out[7]= {-v1,
        v1^2 - v2,
        -v1^3 + 3v1v2- 2v3,
        v1^4 - 6v1^2v2 + 3v2^2 + 8v1v3 - 6v4,
        -v1^5 + 10v1^3v2 - 15v1v2^2 - 20v1^2v3 + 20v2v3 + 30v1v4 - 24v5}
        #writing suitable multiples of the coefficients of chi_A as
        #Z-linear combinations of monomials in the Newton Sums
```

In a final remark of this section, we want to note that the close relationship between the Newton Sums, traces of matrix powers and the characteristic polynomial of that matrix has already been highlighted in both [Bre14], pp. 154–157 and [Whe19] before. However, the considerations there are not primarily concerned with questions of quantifier elimination.

4.1.1 Positive Definite Matrices

It is not difficult to refine our result about positive semi-definiteness to get the following result about positive definite matrices.

Corollary 4.18. *The following statements about a hermitian matrix $A \in M_n(F)$ are equivalent.*

1. *A is positive definite*

2. *All of the formulae*

$$(-1)^k N(k) a_k > 0$$

for $k = 1, \ldots, n$ hold. Again, a_k shall be defined by the characteristic polynomial χ_A of A through $\chi_A = t^n + a_1 t^{n-1} + \ldots + a_{n-1} t + a_n$ and $N(k)$ is chosen in the sense of Remark 4.15.

Proof. The formula $(-1)^n N(k)a_k > 0$ guarantees that no eigenvalue of A is 0. All other formulae give us non-negativity of the eigenvalues of A. Taken together, we immediately deduce the claim. □

Remark 4.19. In the above criterion, the only formula that truly needs the strict inequality is the one resulting from $k = n$. This follows from Remark 4.13.

An alternative characterisation of positive definite matrices is given by combining the result on positive semi-definite matrices with that on invertibility below.

4.1.2 Invertible Matrices

This subsection, too, is an immediate application of the previous results. It provides us with a characterisation of invertibility that also relies on the Newton Sums.

Corollary 4.20. *The following statements about a matrix $A \in M_n(F)$ are equivalent:*

1. *A is invertible.*

2. *$N(n)a_n \neq 0$ where a_n denotes the constant term in the characteristic polynomial χ_A of A and $N(n)$ is chosen as in Remark 4.15.*

Proof. A is invertible if and only if all eigenvalues of A are non-zero. Since by definition, a_n is the product of all eigenvalues of A, this implies the claim. □

4.2 $\|\cdot\|_2$-Contractive Maps

We first prove that the characterisation of contractions given in Table 4.1 is valid.

Lemma 4.21. *A matrix $A \in M_n(F)$ is a $\|\cdot\|_2$-contraction if and only if each eigenvalue λ of A^*A satisfies $|\lambda| < 1$.*

Proof. If there were an eigenvalue λ of A^*A with $|\lambda| > 1$, then we could choose an according eigenvector x and get

$$\|A^*Ax\| = \|\lambda x\| = |\lambda|\|x\| > 1 \cdot \|x\|$$

Hence, $||A^*A|| > 0$, and using $||A^*A|| = ||A||^2$ since $M_n(\mathbb{C})$ is a C^*-algebra, we get a contradiction. If conversely, all eigenvalues λ of A^*A satisfy $|\lambda| < 1$, then by definition of $\| \cdot \|_{op,2}$, we get that A is a contraction. □

Fascinatingly, both the eigenvalues of A and its explicit norm are not necessary for us to decide whether A is a contraction. Instead, it is sufficient to work with the trace and the involution on $M_n(\mathbb{C})$. Now, we will concern ourselves with deriving an explicit formula for such a characterisation. For that, we have the following preparatory lemma that lets us restrict the matrices in question to positive semi-definite matrices.

Lemma 4.22. *Let $A \in M_n(F)$. Then A is a contraction if and only if A^*A is a contraction.*

Proof. If A is a contraction, then by submultiplicativity of the norm, so is A^*A. If A^*A is contractive, then we may use that $\|A^*A\| = \|A\|^2$ as $M_n(\mathbb{C})$ is a C^*-algera and conclude the converse. □

To test whether an arbitrary matrix A is a contraction, it is thus sufficient to decide the claim for A^*A. This can be done by the following equivalence.

Corollary 4.23. *The following statements about a matrix $A \in M_n(F)$ are equivalent:*

1. *A is a contraction.*

2. *$I_n - A^*A$ is positive semi-definite.*

In particular, by Corollary 4.16, we have found a quantifier-free criterion to determine whether A is a contraction.

Proof. Note that since A^*A is hermitian, we may choose a unitary matrix U such that $U^*A^*AU = D$ for a diagonal matrix D, the entries of which are the non-negative eigenvalues of A^*A. Now obviously, D is a contraction if and only if all its diagonal entries are smaller than 1. In other words: D is a contraction if and only if $I_n - D$ is positive semi-definite.

Since definiteness is preserved under similarity conjugation, we immediately get that $I_n - A^*A = U(I_n - D)U^*$ is positive semi-definite as well. This readily implies our claim. □

Remark 4.24. If we replace semi-definiteness by definiteness in Corollary 4.23, then we get a criterion for strict contractions, that is $\|A\| < 1$.

4.3 Solubility of Matrix Equations

We begin by stating the most basic (non-trivial) examples for quantifier elimination regarding the existence of roots of a polynomial. Still, F may denote either $\mathbb{R}$ or $\mathbb{C}$.

Example 4.25. 1. We consider the real- or complex-valued polynomial $p = ax + b$ with $a, b \in F$ and we assume that $p \neq 0$. It is well-known that p has a root in F if and only if $a \neq 0$.

2. We consider the polynomial $q = ax^2 + bx + c$ with $a, b, c \in F, a \neq 0$. It is well-known that over the real numbers, q has a root if and only if $b^2 - 4ac \geq 0$. Over the complex numbers, q will always have a root.

Our goal in this section is to give results in the spirit of the above examples. However, we will work over matrix algebras, which means that we can neither rely on commutativity, nor on the existence of multiplicative inverses, nor on the existence of square roots. It is thus considerably harder to develop quantifier-free formulae describing the existence of roots of matrix polynomials. Our starting point will be the seemingly simple examples AX and $AX + B$.

Lemma 4.26. *The polynomial $P = AX$ with $A \in M_n(F)$ has a non-trivial root $X \in M_n(F)$ if and only if A is a zero divisor. In other words: A must not be invertible. This may be expressed without quantifiers.*

Proof. If A is invertible, then $AX = 0$ implies $X = 0$ by multiplication with A^{-1}. If A is not invertible, then we may choose a non-zero column vector x such that $Ax = 0$. Defining $X = (x, 0, \dots, 0)$, we get a non-zero matrix X satisfying $AX = 0$.

The quantifier-free formulation is directly obtained by Corollary 4.20. □

The next theorem provides the algebraical groundwork that will allow us to find a quantifier-free characterisation of the solubility of the equation $AX + B = 0$. First, however, we prove two preparatory lemmata.

Lemma 4.27. *Let $A \in M_n(F)$ be a real- or complex-valued square matrix. Then $rk(A^*A) = rk(A) = rk(AA^*)$.*

Proof. We follow the proof provided online by [htt13] and observe that the real case is a special case of the complex case. Note that it is sufficient to prove that the kernels of the induced linear maps of A and A^*A are identical. We start by assuming $Ax = 0$ and by abuse of notation, we write $x \in ker(A)$. This implies $A^*Ax = 0$ and therefore $x \in ker(A^*A)$. Now, we let $x \in ker(A^*A)$. Thus, $A^*Ax = 0$ and in particular $(Ax)^*Ax = x^*A^*Ax = 0$. By the properties of the scalar product, we get $Ax = 0$ and thus $x \in ker(A)$. Hence, $ker(A) = ker(A^*A)$.

Finally, we conclude that $rk(A) = rk(A^*A)$ since $ker(A) = ker(A^*A)$ implies dimensional equality of the kernels and thus the dimensions of the images of the maps induced by A and AA^* are equal as well. The other equality follows by applying the same arguments to the matrix A^* which is known to satisfy $rk(A^*) = rk(A)$. □

Lemma 4.28. *Let $A, B \in M_n(F)$ be two positive semi-definite matrices. Then the rank of $A + B$ is greater or equal than the rank of both A and B.*

Proof. We note that since $A + B$ is positive semi-definite, any 0-eigenvector of $A + B$ must be a 0-eigenvector of both A and B. Again by positive semi-definiteness A, B and $A + B$ are diagonalisable and their eigenvalues' algebraic and geometric multiplicities coincide. Thus, the 0-eigenspace of A has a higher (or equal) dimension than that of $A + B$ and the claim follows for A. An analogous argument works for B. □

Theorem 4.29. *Let $A, B \in M_n(F)$. Then the polynomial $AX + B$ has a root if and only if $AA^* + BB^*$ has the same rank as AA^*.*

Proof. According to Theorem 2.3 and Theorem 2.4, we may transform A into Schur form $S_A = P^{-1}AP$, where P is unitary. We note that X is a solution to $AX + B = 0$ if and only if $Y := P^{-1}XP$ is a solution to $S_AY = -P^{-1}BP =: C$. By cancelling all zero-rows at the bottom of S_A, we see that the resulting map induced by the reduced form of S_A is surjective. Thus, we have a solution if and only if C has the same (and possibly more) zero-rows as S_A.

We observe that the rank of $S_AS_A^*$ is the same as that of S_A by Lemma 4.27. Further, we observe that S_A has the block form

$$S_A := \begin{pmatrix} M & 0 \\ 0 & 0 \end{pmatrix}$$

In conclusion, we get that as a necessary condition, $S_AS_A^* + CC^*$ has the same rank as $S_AS_A^*$. If this were not the case, then the rank would be higher by Lemma 4.28. This, however is impossible by the choice of zero-rows in S_A and C. Multiplying by P and P^{-1} from left and right respectively now yields the statement. □

Corollary 4.30. *The existence of a matrix X satisfying $AX + B = 0$ is given without quantifiers by the formula*

$$\bigwedge_{k=0}^{n} (N(k)a_k^Q = 0 \to N(k)a_k^P = 0)$$

We use the abbreviations $P := AA^* + BB^*$ and $Q := BB^*$ and let a_k^P and a_k^Q denote the coefficients of the characteristic polynomials of P and Q respectively. $N(k)$ is chosen according to Remark 4.15.

Proof. This is an immediate consequence of the fact that both P and Q are positive semi-definite. The rank of both matrices is determined by the algebraic multiplicity of the eigenvalue 0, which may be determined by the number of zero-coefficients in the characteristic polynomial. The comparison between the matrices is then possible according to Remark 4.13. □

It is now time to make this criterion explicit in the case that $n = 2$ and $n = 3$. The formulae for larger matrices can be derived in an analogous fashion. For the specific steps taken to derive these formulae, compare the previous results on positive semi-definite matrices, mainly Corollary 4.16 and Exampe 4.17.

Example 4.31. If $n = 2$, our formula is given by

$$\begin{gathered} tr(BB^*) = 0 \to tr(AA^* + BB^*) = 0 \land \\ tr(BB^*)^2 - tr((BB^*)^2) = 0 \to tr(AA^* + BB^*)^2 - tr((AA^* + BB^*)^2) = 0 \end{gathered}$$

If $n = 3$, the calculations are in no way more difficult, though it is ostensibly more tedious to spell out the formula in detail. It is given by

$$\begin{aligned} tr(BB^*) = 0 \to tr(AA^* + BB^*) = 0 \land \\ tr(BB^*)^2 - tr((BB^*)^2) = 0 \to tr(AA^* + BB^*)^2 - tr((AA^* + BB^*)^2) = 0 \land \\ tr(BB^*)^3 - 3tr(BB^*)tr((BB^*)2) + 2tr((BB^*)^3) = 0 \to \\ tr(AA^* + BB^*)^3 - 3tr(AA^* + BB^*)tr((AA^* + BB^*)^2) + 2tr((AA^* + BB^*)^3) = 0 \end{aligned}$$

For the next slight generalisation to the equation $AXB = C$, we need a few preliminaries developed in [Wan04a]. Note that in comparison, our setting is significantly less general.

Definition 4.32 ([Wan04a], p. 44)**.** Let $A \in M_n(F)$. We call $A^+ \in M_n(F)$ a reflexive inverse of A, if it satisfies both $A^+AA^+ = A^+$ and $AA^+A = A$.

Lemma 4.33 ([Wan04a], Proposition 1.1.). *Every $A \in M_n(F)$ possesses a reflexive inverse.*

Proof. We start by solving the equation $AXA = A$. First, we assume that $F = \mathbb{C}$ and choose $P \in Gl_n(\mathbb{C})$ such that $P^{-1}AP =: J$ has Jordan canonical form according to Theorem 2.5. We now solve the equation $JYJ = J$. This is straightforward as without loss of generality we may restrict our considerations to single Jordan blocks $J_m(\lambda)$, where $m \leq n$ and λ denotes an arbitrary eigenvalue of A. Hence, Y will be a block diagonal matrix as well. There are three cases:

- Case 1: $\lambda \neq 0$: Then $J_m(\lambda)$ is invertible and there is a unique solution for Y restricted to that block.
- Case 2: $\lambda = 0$ and $J_m(\lambda) = 0$. Then we simply define $Y = I$ on that block.
- Case 3: $\lambda = 0$ and $J_m(\lambda) \neq 0$: Then we consider a matrix Y that satisfies

$$YJ_\lambda = \begin{pmatrix} 0_F & 0 \\ 0 & I \end{pmatrix}$$

  and observe that this yields a solution to $J_\lambda Y J_\lambda = J_\lambda$. Y exists by observing that it simply encodes repeated changes of rows.

By defining $X := PYP^{-1}$, we get a solution to the initial equation. Also, XAX will now be the desired reflexive inverse since

$$\begin{aligned} A(XAX)A &= (AXA)XA = AXA = A \\ (XAX)A(XAX) &= X(AXA)(XAX) = XA(XAX) = X(AXA)X = XAX \end{aligned}$$

If $F = \mathbb{R}$ and there are non-real eigenvalues, then we may still transform A into a block diagonal form by Theorem 2.6 and repeat the above arguments. □

The following lemma lets us reduce the question of solubility of the equation $AXB = C$ to that of the two equations $AX = C$ and $XB = C$, which we know how to deal with following Corollary 4.30.

Lemma 4.34 ([Wan04a], Lemma 2.2.). *Let $A, B, C \in M_n(F)$. The following statements are equivalent.*

1. *The equation $AXB = C$ has a solution $X \in M_n(F)$.*
2. *$AA^+CB^+B = C$.*

3. $AA^+C = C$ and $CB^+B = C$.

Proof. 1) $\implies$ 3): Let X be a solution of $AXB = C$. Then $AA^+C = AA^+AXB = AXB = C$. Similarly, $CB^+B = AXBB^+B = AXB = C$.

3) $\implies$ 2): Note that $AA^+CB^+B = CB^+B = C$.

2) $\implies$ 1): We define $X = A^+CB^+$. Then $AXB = AA^+CB^+B = C$. □

Corollary 4.35. *Let $A, B, C \in M_n(\mathbb{C})$. Then the polynomial $AXB - C$ has a root if and only if $AA^* + CC^*$ has the same rank as AA^* and $B^*B + C^*C$ has the same rank as B^*B.*

Proof. We make use of Lemma 4.34 3. By Theorem 4.30, $AA^+C = C$ translates to the solubility of $AX = C$ and $CB^+B = C$ translates to the solubility of $XB = C$, or equivalently $B^*X^* = C^*$. Both these statements may be expressed without quantifiers through the above criteria following Theorem 4.29 and Corollary 4.30. □

4.3.1 Unique Solubility of Sylvester's Equation

We now turn to a question raised, but not answered, in [KT20], 2.2.7. We want to give a characterisation of the solubility of Sylvester's Equation only in terms of the trace and the transpose. A thorough analysis of characteristic polynomials will once again prove to be the crucial step in finding such a description.

Definition 4.36 (Sylvester's Equation for square matrices)**.** Let $A, B, C \in M_n(F)$. The matrix equation

$$AX - XB = C \tag{4.1}$$

is called Sylvester's Equation.

Theorem 4.37 (Sylvester-Rosenblum, [KT20], 2.2.7.)**.** *For $A, B \in M_n(F)$, Sylvester's Equation $AX - XB = C$ has a unique solution X for any $C \in M_n(F)$ if and only if the (complex) spectra of A and B are disjoint.*

Proof. $\implies$: This proof follows the argument given in [con21]. Assume that the spectra of A and B are not disjoint and choose vectors $u, v \in \mathbb{C}^n$ satisfying $Au = \lambda u$ and

$v^*B = \lambda v^*$ for some shared eigenvalue λ. Then we note that $uv^* \neq 0$ since both u and v are non-zero vectors. Consequently, $Auv^* - uv^*B = \lambda uv^* - u\lambda v^* = 0$ is a non-trivial solution to Sylvester's Equation in contradiction to our assumption.

If we consider the real case, then we may consider the matrices $Re(uv^*)$ and $Im(uv^*)$ to get the same result.

$\Longleftarrow$: We follow the proof given in [BR97], Theorem VII.2.1. For that, we define the linear operator $\tau : M_n(F) \to M_n(F)$ by $\tau(X) = AX - XB$ and prove that τ is invertible if the spectra of A and B are disjoint. This will imply our statement.

We first define two further operators $\mathcal{A}$ and $\mathcal{B}$ on $M_n(F)$ by $\mathcal{A}(X) := AX$ and $\mathcal{B}(X) = XB$. Using that $\mathcal{A}(\mathcal{B}(X)) = AXB = \mathcal{B}(\mathcal{A}(X))$, we follow [HJ13], Theorem 2.3.3., and choose a basis such that both $\mathcal{A}$ and $\mathcal{B}$ are upper triangular. Their eigenvalues now lie on the diagonal and hence, the spectrum $\sigma(\mathcal{A} - \mathcal{B})$ is contained in the difference $\sigma(\mathcal{A}) - \sigma(\mathcal{B})$.

It is now easy to observe that if λ is an eigenvalue of $\mathcal{A}$, then it is one of A as well. For that assume that $AX = \mathcal{A}(X) = \lambda X$ for some $0 \neq X \in M_n(F)$. If we restrict to single columns of X, we immediately get λ-eigenvectors of the matrix A and thus $\sigma(\mathcal{A}) \subseteq \sigma(A)$. Repeating this argument for B and $\mathcal{B}$, this yields that if the spectra of A and B are disjoint, then 0 is not an eigenvalue of $\tau = \mathcal{A} - \mathcal{B}$ and thus, τ is invertible. □

Definition 4.38. Let p be a real or complex polynomial. By $Z(p)$, we denote the number of different zeros of p in $\mathbb{C}$.

Corollary 4.39. *For $A, B \in M_n(F)$, Sylvester's Equation $AX - XB = C$ has a unique solution X for any $C \in M_n(F)$ if and only if $Z(\chi_A\chi_B) = Z(\chi_A) + Z(\chi_B)$.*

Proof. This is an immediate consequence of Theorem 4.37. □

Definition 4.40 ([NPT13], p. 410)**.** Let p be a monic polynomial of degree d over F. We define the Hermite-Matrix of p via its Newton Sums by

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

The Hermite-Matrix has some really nice properties that we will not talk about here. Instead, we will just derive that its rank equals the number of solutions of p in the algebraic closure of F (here, $\mathbb{C}$). This will be of great use in deciding whether Sylvester's Equation has a unique solution or not. The theorem in question is a weaker version of Theorem 6.2.6. in [BCR98], which we will cite in the version of [Net18], Theorem 1.3.4. We start with a small lemma first.

Lemma 4.41. *Let K be a field, $n \in \mathbb{N}$ and let $s \leq n$. Further, let $v_1, \ldots, v_s$ be linearly independent column vectors in K^n. Then*

$$\sum_{i=1}^{n} v_i v_i^T$$

is a matrix of rank s.

Proof. By assumption, the matrix $A := (v_1, \ldots, v_s)$ has full column rank (that is s). Thus, AA^T will also have column rank s. Now we observe that $\sum_{i=1}^{n} v_i v_i^T = AA^T$ and finish the proof. □

Theorem 4.42. *Let p be a real or a complex polynomial of degree $d \geq 1$. Then the number of different roots of p in $\mathbb{C}$ is equal to the rank of $\mathcal{H}(p)$.*

Proof. We follow the proofs given in [Net18], Theorem 1.3.4., and [BCR98], Theorem 6.2.6. Let $\alpha_1, \ldots, \alpha_d$ denote all roots of p. We write

$$\omega_i := (1, \alpha_i, \ldots, \alpha_i^d)$$

It is then clear that

$$\mathcal{H}(p) = \sum_{i=1}^{d} \omega_i \omega_i^T$$

We now assume without loss of generality that $\alpha_1, \ldots, \alpha_s$ are the different roots of p and we denote the algebraic multiplicity of α_i by n_i. Then consequently

$$\mathcal{H}(p) = \sum_{i=1}^{s} n_i \omega_i \omega_i^T$$

By our assumption and the properties of the Vandermonde matrix, $(\omega_i | i = 1, \ldots, s)$ is a linearly independent tuple. By Lemma 4.41, this proves that $\mathcal{H}(p)$ has rank s and thus the statement follows. □

Corollary 4.43. *A quantifier-free criterion for the unique solubility of Sylvester's Equation $AX + XB = C$ dependent on A and B is given by the following formula.*

$$rk(\mathcal{H}(\chi_A)) + rk(\mathcal{H}(\chi_B)) = rk(\mathcal{H}(\chi_A \chi_B))$$

Proof. It is a direct consequence of Corollary 4.39 that the criterion fully describes the solubility of Silvester's equation . It can be realised without quantifiers following Lemma 4.27 and then applying Remark 4.13 as in the proof of Corollary 4.30. Note that for comparability of dimensions, prior to applying our results it is necessary to embed the Hermite-Matrices $\mathcal{H}(\chi_A), \mathcal{H}(\chi_B)$ into a larger matrix that is then filled with zero-entries, that is: We map

$$(\mathcal{H}(\chi_A), \mathcal{H}(\chi_B)) \mapsto \begin{pmatrix} \mathcal{H}(\chi_A) & 0 \\ 0 & \mathcal{H}(\chi_B) \end{pmatrix}$$

such that

$$rk(\mathcal{H}(\chi_A)) + rk(\mathcal{H}(\chi_B)) = rk\left(\begin{pmatrix} \mathcal{H}(\chi_A) & 0 \\ 0 & \mathcal{H}(\chi_B) \end{pmatrix}\right)$$

□

Note that in the above setting of characteristic polynomials, we always get $\mathcal{H}(\chi_A) = (tr(A^{i+j})_{i,j=0}^{n-1}$.

The criterion derived is not yet very tangible and can only be spelled out if we refer back to our previous results step by step. We exemplify an idea of the specific formula in the simplest case, $n = 2$.

Example 4.44. We first spell out the relevant Hermite-Matrices in detail:

$$\begin{aligned}
\mathcal{H}(\chi_A) &= \begin{pmatrix} 2 & tr(A) \\ tr(A) & tr(A^2) \end{pmatrix} \\
\mathcal{H}(\chi_B) &= \begin{pmatrix} 2 & tr(B) \\ tr(B) & tr(B^2) \end{pmatrix} \\
\mathcal{H}(\chi_A\chi_B) &= \begin{pmatrix} 4 & tr(A)+tr(B) & tr(A^2)+tr(B^2) & tr(A^3)+tr(B^3) \\ tr(A)+tr(B) & tr(A^2)+tr(B^2) & tr(A^3)+tr(B^3) & tr(A^4)+tr(B^4) \\ tr(A^2)+tr(B^2) & tr(A^3)+tr(B^3) & tr(A^4)+tr(B^4) & tr(A^5)+tr(B^5) \\ tr(A^3)+tr(B^3) & tr(A^4)+tr(B^4) & tr(A^5)+tr(B^5) & tr(A^6)+tr(B^6) \end{pmatrix}
\end{aligned}$$

We now embed $\mathcal{H}(\chi_A)$ and $\mathcal{H}(\chi_B)$ as in the proof of 4.43 and thus we need to compare the rank of

$$\mathcal{H}(\chi_A\chi_B) \quad \text{and} \quad \begin{pmatrix} 2 & tr(A) & 0 & 0 \\ tr(A) & tr(A^2) & 0 & 0 \\ 0 & 0 & 2 & tr(B) \\ 0 & 0 & tr(B) & tr(B^2) \end{pmatrix}$$

or equivalently that of the positive semi-definite matrices

$$\mathcal{H}(\chi_A\chi_B)^*\mathcal{H}(\chi_A\chi_B) \quad \text{and} \begin{pmatrix} \mathcal{H}(\chi_A)^*\mathcal{H}(\chi_A) & 0 \\ 0 & \mathcal{H}(\chi_B)^*\mathcal{H}(\chi_B) \end{pmatrix}$$

where

$$\mathcal{H}(\chi_A)^*\mathcal{H}(\chi_A) = \begin{pmatrix} 4 + |tr(A)|^2 & 2tr(A) + \overline{tr(A)}tr(A^2) \\ \overline{tr(A)} + tr(A)\overline{tr(A^2)} & |tr(A)|^2 + |tr(A^2)|^2 \end{pmatrix}$$

and similarly for B. This can be done following Corollary 4.30.

For $n = 3$, the procedure to derive the formula is the same, though the formula itself is much longer. We refrain from presenting it here.

5 Dimensional Compatibility of Quantifier Elimination

For this chapter, we return to our general model theoretic setting. We have already seen that we have quantifier elimintation in suitable matrix algebras $M_n(\tilde{F})$ expanded by trace and transposition for arbitrary but fixed n. We are now interested in whether some sort of homogeneity between dimensions of matrix algebras exists, that is: We assume that $k < n$ and we want to discuss an arbitrary formula (potentially containing quantifiers) in $M_k(F)$. Is it possible to embed $M_k(F)$ into $M_n(F)$ and then work in the larger algebra instead? Or is it possible to simply use a quantifier-free formula derived for the n-dimensional case $M_n(F)$ and apply it to $M_k(F)$?

Although desirable, in both cases the answer is no. For $n > k$ and standard identifications $\iota : M_k(F) \to M_n(F)$, it is not possible to work in $M_n(F)$ instead. This shall be exemplified below and then generalised to show that fixing this problem by a better choice of identification is hopeless. Similarly, we may not recycle quantifier-free formulae in order to use them in lower dimensions. This will also be exemplified. Our results from the previous sections prove to be tailor-made for highlighting the problems that may occur.

Example 5.1. For $k < n$, we consider the identification of $M_k(\mathbb{R})$ in $M_n(\mathbb{R})$ via

$$A \mapsto \iota(A) := \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$$

We hope to show that

$$M_n(\mathbb{R}) \vDash \varphi \left[\begin{pmatrix} h & 0 \\ 0 & 0 \end{pmatrix} \right] \iff M_k(\mathbb{R}) \vDash \varphi[h]$$

for all quantifier-free formulae φ and all evaluations h in $M_k(\mathbb{R})$. By quantifier elimination, we may look at arbitrary formulae φ.

We want to decide whether A is positive definite (not semi-definite) and wish to apply the quantifier-free formula derived in Corollary 4.18. This formula, however, will not hold for $\iota(A)$ as $\iota(A)$ always has the eigenvalue 0 with multiplicity $n - k$. Similarly, invertibility does not carry over. Note that this may be attributed to the fact that I_k keeps its rank under ι.

Example 5.2. For $k < n$, we consider the identification of $M_k(F)$ in $M_n(F)$ via

$$A \mapsto \iota(A) := \begin{pmatrix} A & 0 \\ 0 & I_{n-k} \end{pmatrix}$$

We hope to show that

$$M_n(F) \vDash \varphi \left[\begin{pmatrix} h & 0 \\ 0 & I_{n-k} \end{pmatrix} \right] \iff M_k(F) \vDash \varphi[h]$$

for all quantifier-free formulae φ and all evaluations h in $M_k(F)$. By quantifier elimination, we may look at arbitrary formulae φ.

We want to decide whether the difference of matrices $B - A$ is invertible in $M_k(F)$ and use the criterion established in Corollary 4.20. For our matrices, we choose $A = I_k$ and $B = 2I_k$. Now $B - A$ clearly is invertible in $M_k(F)$. However, $\iota(B) - \iota(A)$ is not invertible in $M_n(F)$. Note that this may be attributed to the fact that ι is not an embedding.

Lemma 5.3. *Let f be an injective function from $M_k(F)$ into $M_n(F)$, where $k < n$. Then we will always find a formula φ violating*

$$M_n(F) \vDash \varphi[f \circ h] \iff M_k(F) \vDash \varphi[h]$$

for some evaluation h in $M_k(F)$. In particular, in the above examples we cannot replace ι by another function that yields the desired result.

Proof. Necessarily, f must map matrix units to matrix units. Otherwise, we may choose the formula ε_k (see Definition 3.6) as a counterexample. By Lemma 3.7, we conclude that f must be an F-algebra homomorphism and that the rank of a matrix is preserved under f. That, however, means that $f(I_k)$ will never be invertible in $M_n(F)$, which is a contradiction. □

Remark 5.4. In absence of a possibility to embed $M_k(F)$ into $M_n(F)$ and then work with quantifier-free formulae in $M_n(F)$, we might hope to instead take a formula φ, find equivalent quantifier-free formulae φ^n in $M_n(F)$ and φ^k in $M_k(F)$ and hope that

$$M_k(F) \vDash \varphi^n[h] \iff M_k(F) \vDash \varphi^k[h]$$

for all evaluations h in $M_k(F)$. This, however, will also prove to be not the case.

Example 5.5. We want to decide, whether an element $a \in \mathbb{R} = M_1(\mathbb{R})$ is invertible and use the criterion established in Corollary 4.20 for $M_2(\mathbb{R})$, that is: a is invertible if and only if $tr(a)^2 - tr(a^2) \neq 0$. In our case, however, this will never be satisfied for $a \neq 0$, such that we may not use the criterion.

6 Conclusion and Outlook

In this thesis, we have established that under certain circumstances, the matrix algebra $M_n(F)$ over a field F admits quantifier elimination. These results have been exemplified by a series of examples, most notably by giving a characterisation of positive semi-definite matrices, from which other examples could be derived. We have also seen that switching between dimensions (that is: Considering n in $M_n(F)$ as variable) while keeping results on quantifier elimination is generally not possible.

At the same time, many things could not be done in this thesis. Indeed, there are several starting points for further development of the ideas presented in the previous chapters. We want to sketch some of them without giving an assessment of whether they merit further study in the formulations below. These considerations will mark the end of this thesis.

6.1 Systems of Linear Equations

It is a small step to generalise the setting of section 4.3 to allow for systems of linear matrix equations. The question we ask is: How far can we stretch the results derived so far? [Wan04a] derives solutions to the system of equations

$$\begin{aligned} A_1XB_1 &= C_1 \\ A_2XB_2 &= C_2 \end{aligned}$$

which, while still looking deceptively simple, already requires much more elaborate arguments. Also, this work is not done with quantifier elimination in mind.

While desirable, it seems out of reach to derive a general criterion for the solubility of linear matrix equations in one variable. If we allow equations in more than one variable, this goal seems even more distant.

6.2 Non-Linear Matrix Equations

Another very natural questions to ask is whether we can generalise the matrix polynomials that we want to find roots of. In the simplest terms possible: How can we decide whether the polynomial $AX^2 + BX + C$ has a root? What does a tractable formula to decide the (unique?) solubility of such an equation look like?

There are ideas for the one-variable case in [Wil], though it would still be necessary to derive a quantifier-free criterion in trace and transposition from his ideas.

Further complications are to be expected when generalising to more than one variable or to systems of equations.

6.3 Quantifier Elimination in Other Languages

A less obvious point of further study concerns other or further extensions of the language $\mathcal{L}$ by new functions giving us information about a matrix A without accessing its entries. For example, take $\|\cdot\|$ as an additional function and interpret it as some operator norm. Does $M_n(F)$ still admit quantifier elimination?

Alternatively, we could extend $\mathcal{L}(tr, invo)$ by some unary function symbol f where f is interpreted as a linear map $f^{M_n(F)} : M_n(F) \to M_n(F)$. Is it possible to still get quantifier elimination? What does that tell us about certain properties of f? Such results would immediately lead to the discussion of the C^*-algebra $M_n(\mathbb{C})$ and questions associated to operators on it, for example boundedness or positivity.

We are only what we know, and I wished to be much more than I was, sorely.
David Mitchell: Cloud Atlas, 2008, p. 208.

Bibliography

[BCR98] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy. *Real Algebraic Geometry.* Berlin et al.: Springer, 1998.

[Bos14] Siegried Bosch. *Lineare Algebra.* Berlin et al.: Springer, 2014.

[BPR06] Saugata Basu, Richard Pollack, and Marie-Françoise Roy. *Algorithms in Real Algebraic Geometry.* Berlin et al.: Springer, 2006.

[BR97] Rajendra Bhatia and Peter Rosenthal. "How and Why to solve the Operator Equation $AX - XB = Y$". In: *Bulletin of the London Mathematical Society* 29 (1997), pp. 1–21.

[Bre14] Matej Brešar. *Introduction to Noncommutative Algebra.* Cham et al.: Springer, 2014.

[con21] Wikipedia contributors. *Sylvester equation. Wikipedia, The Free Encyclopedia.* 2021. URL: `https://en.wikipedia.org/w/index.php?title=Sylvester_equation&oldid=1059591226`. (accessed: 11.12.2021).

[Har00] Robin Hartshorne. *Geometry. Euclid and Beyond.* New York: Springer, 2000.

[HJ13] Roger A. Horn and Charles R. Johnson. *Matrix Analysis.* Cambridge et al.: Cambridge University Press, 2013.

[htt13] A.D (https://math.stackexchange.com/users/37459/a-d). *Prove* rank $A^TA =$ rank A *for any* $A \in M_{m\times n}$. 2013. Mathematics Stack Exchange: `349966`. URL: `https://math.stackexchange.com/q/349966`. accessed: 05.12.2021.

[KT20] Igor Klep and Marcus Tressl. *A Model Theoretic Perspective on Matrix Rings.* 2020. arXiv: `1810.09024`. URL: `https://arxiv.org/abs/1810.09024v3`. (accessed: 01.11.2021).

[MSV93] David Mornhinweg, Daniel B. Shapiro, and Kenneth G. Valente. "The Principal Axis Theorem over Arbitrary Fields". In: *The American Mathematical Monthly* 100.8 (1993), pp. 749–754.

[Net18] Tim Netzer. *Reelle Algebra und Geometrie.* Innsbruck: Lecture Notes, 2018. URL: `https://arxiv.org/abs/1810.09024v3`.

[NPT13] Tim Netzer, Daniel Plaumann, and Andreas Thom. "Determinantal Representations and the Hermite Matrix". In: *Michigan Mathematical Journal* 62.2 (2013), pp. 407–420.

[PD01] Alexander Prestel and Charles N. Delzell. *Positive Polynomials. From Hilbert's 17th Problem to Real Algebra.* Berlin et al.: Springer, 2001.

[PD11] Alexander Prestel and Charles N. Delzell. *Mathematical Logic and Model Theory. A Brief Introduction.* London et al.: Springer, 2011.

[Pro76] Claudio Procesi. “The Invariant Theory of $n \times n$ Matrices”. In: *Advances in Mathematics* 19 (1976), pp. 306–381.

[Raz74] Ju. P. Razmyslov. “Trace Identities of Full Matrix Algebras over a Field of Characteristic Zero”. In: *Mathematics of the USSR-Izvestiya* 8.4 (1974), pp. 727–760.

[Ros78] Bruce I. Rose. “Rings which Admit Elimination of Quantifiers”. In: *The Journal of Symbolic Logic* 43.1 (1978), pp. 92–112.

[Sib68] Konstantin Sergeevich Sibirskii. “Algebraic Invariants of a System of Matrices”. In: *Siberian Mathematical Journal* 9.1 (1968), pp. 115–124.

[Wan04a] Quin-Wen Wang. “A System of Matrix Equations and a Linear Matrix Equation over Arbitrary Regular Rings with Identity”. In: *Linear Algebra and its Applications* 384 (2004), pp. 43–54.

[Wan04b] Xiaoshen Wang. “A Simple Proof of Descartes’s Rule of Signs”. In: *The American Mathematical Monthly* 111.6 (2004), pp. 525–526.

[Whe19] Nicholas Wheeler. *Newton and the Characteristic Polynomial of a Matrix.* 2019. URL: `https://www.reed.edu/physics/faculty/wheeler/documents/Miscellaneous%20Math/Novel%20Derivation%20of%20Newton's%20Identities.pdf`. (accessed: 30.11.2021).

[Wil] Robert Lee Wilson. *Polynomial Equations over Matrices.* URL: `https://sites.math.rutgers.edu/~rwilson/polynomial_equations.pdf`. (accessed: 14.12.2021).

[Zei84] Doron Zeilberger. “A Combinatorial Proof of Newton’s Identities”. In: *Discrete Mathematics* 49 (1984), p. 319.

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht.

Die vorliegende Arbeit wurde bisher in gleicher oder ähnlicher Form noch nicht als Magister-/Master-/Diplomarbeit/Dissertation eingereicht.

Datum

Unterschrift