

**Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig**

**Positive Semidefinite Univariate Matrix
Polynomials**

by

Christoph Hanselka and Rainer Sinn

Preprint no.: 43

2017



POSITIVE SEMIDEFINITE UNIVARIATE MATRIX POLYNOMIALS

CHRISTOPH HANSELKA AND RAINER SINN

ABSTRACT. We study sum-of-squares representations of symmetric univariate real matrix polynomials that are positive semidefinite along the real line. We give a new proof of the fact that every positive semidefinite univariate matrix polynomial of size $n \times n$ can be written as a sum of squares $M = Q^T Q$, where Q has size $(n+1) \times n$, which was recently proved by Blekherman-Plaumann-Sinn-Vinzant. Our new approach using the theory of quadratic forms allows us to prove the conjecture made by these authors that these minimal representations $M = Q^T Q$ are generically in one-to-one correspondence with the representations of the nonnegative univariate polynomial $\det(M)$ as sums of two squares.

In parallel, we will use our methods to prove the more elementary hermitian analogue that every hermitian univariate matrix polynomial M that is positive semidefinite along the real line, is a square, which is known as the matrix Fejér-Riesz Theorem.

INTRODUCTION

A symmetric (or hermitian) matrix M whose entries are polynomials with real (or complex) coefficients in s variables x_1, \dots, x_s is said to be positive semidefinite if the constant symmetric (or hermitian) matrix $M(x)$ is positive semidefinite for all $x \in \mathbb{R}^s$.

In this paper, we study sum-of-squares certificates for symmetric matrices whose entries are univariate real polynomials in t to be positive semidefinite, that is factorizations of a univariate matrix polynomial $M(t)$ as a hermitian square, i.e.

$$M(t) = Q(t)^* Q(t),$$

where $Q(t)$ is a univariate $r \times n$ matrix polynomial for some integer r and $Q(t)^*$ is the conjugate transpose of $Q(t)$ (complex conjugation is applied coefficient-wise to the entries of Q). Such a factorization makes it immediate that $M(t)$ is positive semidefinite at each $t \in \mathbb{R}$. This type of sum-of-squares certificates simultaneously generalizes the case of univariate polynomials, which we recover for $n = 1$, and the case of constant matrices, which follows from the spectral theorem for matrices. So it is the simplest generalization of classical results in real algebraic geometry to the setup of matrix polynomials and a first step in gaining a better understanding of techniques that can be used to understand positive semidefinite matrix polynomials.

The existence of sum-of-squares certificates for hermitian univariate matrix polynomials was known since the Fifties at least and is often known as the matrix Fejér-Riesz Theorem, see [WM57]. This theorem has received much attention and has been generalized to various contexts in analysis, see [DR10] for a recent survey. Its mentioned matrix version directly implies the existence of sum-of-squares certificates for real symmetric $n \times n$ univariate matrix polynomials $M = Q^T Q$,

Date: July 27, 2017.

2010 Mathematics Subject Classification. Primary: 14P05; Secondary: 47A68, 11E08, 11E25, 13J30.

Key words and phrases. Matrix factorizations, matrix polynomial, sum of squares, Smith normal form.

where Q has size $2n \times n$, which was proved by Choi-Lam-Reznick [CLR80]. Generalizations of this result to coefficient fields other than \mathbb{R} have been developed in [FRS06].

In the symmetric case, the bound on the size of the matrix $Q(t)$ was later improved to $r = n + 1$ by Blekherman-Plaumann-Sinn-Vinzant using techniques from projective algebraic geometry [BPSV16] (and Leep [Lee06] in an unpublished manuscript using techniques from the theory of quadratic forms). This bound $r = n + 1$ is smallest possible and Blekherman-Plaumann-Sinn-Vinzant further observed that the number of essentially different sum-of-squares certificates $M = Q^T Q$ for a generic matrix polynomial M and $r = n + 1$ is finite and conjectured a count in terms of the degree of the determinant of M as a univariate polynomial, [BPSV16, Introduction].

In this paper, we prove this conjectured count by showing that factorizations $M = Q^T Q$ of a generic positive semidefinite real symmetric univariate $n \times n$ matrix polynomial M , where Q is an $(n + 1) \times n$ matrix polynomial, are in one-to-one correspondence with representations of the nonnegative univariate polynomial $\det(M)$ as a sum of two squares.

Theorem (Corollary 4.7). *Let $M \in \text{Sym}_n(\mathbb{R}[t])$ be positive semidefinite with nonzero and square free determinant $\det(M)$. Then there is a bijection between the sets*

$$\{ Q \in \text{Mat}_{(n+1) \times n}(\mathbb{R}[t]) \mid Q^T Q = M \}$$

and

$$\{ g \in \mathbb{R}[t]^2 \mid g^T g = \det(M) \}$$

modulo the left action of the orthogonal groups $O_n(\mathbb{R})$ and $O_2(\mathbb{R})$, respectively.

Another interesting consequence of our approach is a characterization of all real symmetric matrix polynomials M that are squares, i.e. that can be factored as $M = Q^T Q$, where the matrix polynomial Q is of equal size $n \times n$.

Theorem (Corollary 2.7). *Let $M \in \text{Sym}_n(\mathbb{R}[t])$ be positive semidefinite with nonzero determinant $\det(M)$. Then M admits a square factorization, $M = Q^T Q$ for some $Q \in \text{Mat}_n(\mathbb{R}[t])$, if and only if $\det(M)$ is a square in $\mathbb{R}[t]$.*

We develop our theory mostly in parallel for both the real symmetric and complex hermitian setting. On the one hand, this leads to the following result, which is, as we learned during preparation of this paper, also due to [ESS16, Theorems 2,3].

Theorem (Corollary 4.2). *Let $M \in \text{Her}_n(\mathbb{C}[t])$ be positive semidefinite with nonzero and square free determinant $\det(M)$. Then the determinant map induces a bijection between the sets*

$$\{ Q \in \text{Mat}_n(\mathbb{C}[t]) \mid Q^* Q = M \}$$

and

$$\{ g \in \mathbb{C}[t] \mid g^* g = \det(M) \}$$

modulo the left action of the unitary groups $U_n(\mathbb{C})$ and $U_1(\mathbb{C})$, respectively.

On the other hand, this parallel approach highlights the essential differences between the real and the complex case. While the latter can be treated completely elementary, the former requires a considerable amount of additional work and quadratic forms theory. An indication that the complex hermitian case is indeed simpler is the fact that the determinant induces the bijection between factorization of M and its determinant in Corollary 4.2, whereas the bijection in the real symmetric case in Corollary 4.7 is a lot more subtle and quite surprising (see Remark 4.8).

One of the central results for the proof of the above mentioned theorems is about Smith normal forms over the polynomial ring.

Theorem (Theorem 4.1). *Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and $M \in \text{Her}_n(\mathbb{K}[t])$ be positive semidefinite with nonzero determinant. Then the equivalence classes of $n \times n$ -factorizations $M = Q^*Q$ are in one-to-one correspondence to those of the monic Smith normal form of M .*

Reader's Guide. There are two main technical ingredients for the proofs of our main results (Corollaries 4.2 and 4.7), which are the following. Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and $M \in \text{Her}_n(\mathbb{K}[t])$ be positive semidefinite.

- We first show that factorizations $M = Q^*Q$ over $\mathbb{K}[t]$ are essentially the same as those over the localization \mathcal{O} at zeros of the determinant of M (Theorem 2.10).
- Then we show that M and its monic Smith normal form become congruent over \mathcal{O} (Theorem 3.7).

The presentation is structured as follows. After fixing our conventions and recalling basic definitions in Section 1, we describe a special consequence of Witt's local-global principle to hermitian squares over the rational function field $\mathbb{R}(t)$. In Section 2 we study how hermitian factorizations over the rational function field can be turned into factorizations over the polynomial ring. We prove the existence of factorizations $M = Q^T Q$ of the generically smallest possible size $r = n + 1$ and characterize those M that admit factorizations of square size $r = n$. The main result for the count of the number of smallest hermitian square representations is Theorem 2.10, which establishes a correspondence of factorizations over the polynomial ring $\mathbb{R}[t]$ and the rational function field $\mathbb{R}(t)$. This accomplishes the first of the two main steps described above. The second one is the main result of Section 3 in which we show that a positive semidefinite symmetric (or hermitian) matrix is congruent to its Smith normal form, if we allow certain denominators in the congruence transformations. An essential technical difficulty is to control these denominators. We then combine these two steps, in the following Section 4 to prove the main results. In a short Appendix, Section 5, we discuss applications of the prime avoidance lemma to hermitian forms that we need earlier in the paper.

Acknowledgments. We are grateful to Markus Schweighofer. Our approach extends fruitful discussions with him. The first author is supported by the Faculty Research Development Fund (FRDF) of The University of Auckland (project no. 3709120). The second author would like to thank Bernd Sturmfels and the Max-Planck-Institute in Leipzig for their hospitality and support.

1. PRELIMINARIES

We recall some notions from linear algebra and fix our notation and terminology to avoid confusion. As a general reference for the theory of quadratic forms over rings, we refer to [O'M71].

- The polynomial rings in this paper will be over the field of real numbers or the field of complex numbers. Many statements will be developed in parallel for both cases, so we use \mathbb{K} to mean \mathbb{R} or \mathbb{C} . We want to note that throughout the paper, \mathbb{R} and \mathbb{C} can be replaced by any real closed field and its algebraic closure, respectively.
- The polynomial ring $\mathbb{C}[t]$ in one variable over the complex numbers is equipped with an involution (written as \cdot^*), given by coefficient wise complex conjugation and $t^* = t$.

- In the following, let R be a commutative ring with involution written as \cdot^* . It might be the trivial one, as is the case for $\mathbb{R}[t]$.
- We write $\mathbb{1}_n$ for the $n \times n$ identity matrix.
- We call a square matrix $M \in \text{Mat}_n(R)$ with entries in a ring R *non-degenerate* if $\det(M)$ is nonzero.
- The *adjoint* of a matrix $A \in \text{Mat}_{m \times n}(R)$, denoted by A^* , is the entry-wise conjugate of the transpose A^T .
- We denote the set of *hermitian* $n \times n$ matrices over a ring R , i.e. those $A \in \text{Mat}_n(R)$ such that $A^* = A$, by $\text{Her}_n(R)$.
- We write $\text{U}_n(R) = \{ U \in \text{Mat}_n(R) \mid U^*U = \mathbb{1}_n \}$ for the *unitary group* over R .
- If the involution on the ring R is trivial, then $\text{Her}_n(R) = \text{Sym}_n(R)$ is the set of *symmetric matrices* and $\text{U}_n(R) = \text{O}_n(R)$ is the *orthogonal group*.
- Given two square matrices $A \in \text{Mat}_n(R)$ and $B \in \text{Mat}_m(R)$ we denote by $A \oplus B \in \text{Mat}_{n+m}(R)$ the block diagonal matrix with blocks A and B .
- We write $\langle a_1, \dots, a_n \rangle$ for the diagonal matrix with diagonal entries a_1, \dots, a_n .
- We call $M, N \in \text{Her}_n(R)$ *congruent over the ring* R , written as $M \simeq_R N$, if there exists an invertible matrix $S \in \text{Mat}_n(R)$ over R such that $M = S^*NS$.
- For $M \in \text{Her}_n(R)$ and $k \in \mathbb{N}$, we denote by $\Phi_k(M, R)$ the set of equivalence classes (orbits) of $k \times n$ -factorizations of Q

$$\Phi_k(M, R) := \{ Q \in \text{Mat}_{k \times n}(R) \mid Q^*Q = M \} / \text{U}_k(R)$$

where unitary group $\text{U}_k(R)$ acts on the left.

- For a principal ideal domain R and a non-degenerate matrix $M \in \text{Mat}_n(R)$ with determinant $d := \det(M)$ we define

$$\begin{aligned} \mathcal{O}_M &:= \left\{ \frac{a}{b} \in \text{Quot}(R) \mid b \text{ is coprime to } d \right\} \\ &= \bigcap_{\mathfrak{p} \in Z(d)} R_{\mathfrak{p}} \end{aligned}$$

where $Z(d)$ is the set of prime ideals containing d . \mathcal{O}_M is a semi-local principal ideal domain and will play a central role in what follows.

1.1. Quadratic forms over the rational function field $\mathbb{R}(t)$. One central piece of our argument over the real numbers relies on Witt's local-global principle, which states that every totally indefinite quadratic form of dimension at least three over a function field of transcendence degree one over \mathbb{R} is isotropic, i.e. represents zero non-trivially. A proof can be found in [PD01, Theorem 3.4.11]. Essential for the present paper is the following consequence, which is well-known in the quadratic forms community.

Corollary 1.1. *Let $a, b \in \mathbb{R}(t)$ be nonzero and positive semidefinite. Then $\langle a, b \rangle$ represents 1 over $\mathbb{R}(t)$. In particular,*

$$\langle a, b \rangle \simeq_{\mathbb{R}(t)} \langle 1, ab \rangle.$$

Proof: Consider the totally indefinite form $\langle a, b, -1 \rangle$ and apply Witt's local-global principle to get a representation nontrivial representation

$$ax^2 + by^2 - z^2 = 0.$$

Since a and b are positive semidefinite, z must be nonzero. Dividing by z and adding 1 to the equation we get a representation of 1 by the form $\langle a, b \rangle$. An appropriate base change thus yields $\langle a, b \rangle \simeq \langle 1, c \rangle$ for some $c \in \mathbb{R}(t)$. Comparing

determinants we get that ab and c differ by a square. Rescaling the second basis vector, we may assume that $ab = c$. \square

Applying this corollary inductively one can easily characterize those non-degenerate $M \in \text{Sym}_n(\mathbb{R}(t))$ that admit a square factorization $M = Q^T Q$ over the rational function field.

Corollary 1.2. *Let $M \in \text{Sym}_n(\mathbb{R}(t))$ be non-degenerate and positive semidefinite wherever it is defined. Then there exists $Q \in \text{Mat}_n(\mathbb{R}(t))$ with $M = Q^T Q$ if and only if $\det(M)$ is a square in $\mathbb{R}(t)$.*

Proof: Clearly, if $M = Q^T Q$ then $\det(M) = (\det Q)^2$ is a square. Conversely, assume that $d := \det(M)$ is a square. After diagonalization of M (as a quadratic form) we may assume that $M = \langle a_1, \dots, a_n \rangle$ for some $a_i \in \mathbb{R}(t)$. Applying Corollary 1.1 $n - 1$ times, we get

$$M \simeq_{\mathbb{R}(t)} \langle 1, a_1 a_2, a_3 \dots, a_n \rangle \simeq_{\mathbb{R}(t)} \cdots \simeq_{\mathbb{R}(t)} \langle 1, \dots, 1, \prod_{i=1}^n a_i \rangle.$$

Since $\prod_{i=1}^n a_i = d$ is a square, we therefore have $M \simeq \mathbb{1}_n$. \square

A classical theorem due to Cassels [Cas64] (and shortly after generalized by Pfister [Pfi65]) says that for any field k a polynomial $f \in k[t]$ that is a sum of squares of rational functions is already a sum of squares of polynomials (with the same number of squares). Tignol proved in [Tig96] a version for univariate polynomial rings over central simple algebras, which, applied to the matrix algebra $\text{Mat}_n(\mathbb{K})$, gives that any matrix $M \in \text{Her}_n(\mathbb{K}[t])$ that admits a rational factorization $M = Q^* Q$, where $Q \in \text{Mat}_n(\mathbb{K}(t))$, also admits a polynomial factorization $M = P^* P$, where $P \in \text{Mat}_n(\mathbb{K}[t])$. This statement can also be shown using techniques of Leep's from [Lee06]. We will prove it below, see Corollary 2.7, as a result of a more elementary and more explicit proof of a Cassels-Pfister Theorem for matrices over $\mathbb{K}[t]$, which we give in the following section.

2. A CASSELS-PFISTER THEOREM FOR MATRICES

In this section, we present an elementary and explicit proof of the known fact that any matrix $M \in \text{Her}_n(\mathbb{K}[t])$ that admits a rational factorization $M = Q^* Q$, where $Q \in \text{Mat}_n(\mathbb{K}(t))$, also admits a polynomial factorization $M = P^* P$, where $P \in \text{Mat}_n(\mathbb{K}[t])$. Our approach allows us to investigate the pole behavior more closely. The ring \mathcal{O}_M that we associate to a non-degenerate matrix $M \in \text{Mat}_n(\mathbb{K}[t])$ will play a central role. It consists of those rational functions, that have no poles wherever M becomes degenerate. We prove that, up to equivalence, factorizations of M over \mathcal{O}_M correspond exactly to factorizations over the polynomial ring, see Theorem 2.10.

2.1. Splitting-off Matrix Zeros. To every zero of a (scalar) polynomial corresponds a linear factor that can be split off. Almost the same can be done in the matrix case, if we take care of the order of multiplication in evaluating matrix polynomials: Let $P = \sum_i P_i t^i \in \text{Mat}_n(\mathbb{K}[t])$ be a matrix polynomial and let $A \in \text{Mat}_n(\mathbb{K})$ be a constant matrix. We fix the notation

$$P_A := \sum_i P_i A^i,$$

$${}_A P := \sum_i A^i P_i$$

for the right and left evaluation of P at A , respectively.

Lemma 2.1. For P and A as above, we have

$$P_A = 0 \Leftrightarrow \exists S \in \text{Mat}_n(\mathbb{K}[t]): P = S(t\mathbb{1}_n - A)$$

and

$${}_A P = 0 \Leftrightarrow \exists S \in \text{Mat}_n(\mathbb{K}[t]): P = (t\mathbb{1}_n - A)S.$$

Proof: Just as in the scalar case, we use the identity

$$(t^k\mathbb{1}_n - A^k) = (t^{k-1}\mathbb{1}_n + t^{k-2}A + \cdots + A^k)(t\mathbb{1}_n - A)$$

to split off $(t\mathbb{1}_n - A)$ from the right of $P - P_A$ which coincides with P if $P_A = 0$. The argument obviously can be adapted to the case ${}_A P = 0$. \square

For a scalar polynomial $q \in \mathbb{K}[t]$ and a complex number $z \in \mathbb{C}$ which is a zero of q^*q , we can conclude that z or z^* must be a zero of q . The following proposition shows how this can be generalized to matrix polynomials.

Proposition 2.2. Let $Q \in \text{Mat}_n(\mathbb{K}[t])$ and $z \in \mathbb{C}$ such that $(Q^*Q)(z) = 0$. In case $\mathbb{K} = \mathbb{R}$ we further assume that n is even. Then there exists a constant matrix $A \in \text{Mat}_n(\mathbb{K})$ with the following properties:

- A is normal and its only eigenvalues are z and z^* .
- A is a zero of Q under left evaluation, i.e. ${}_A Q = 0$.

In particular, we can split off a linear factor from Q

$$Q = (t\mathbb{1}_n - A)P \quad \text{for some } P \in \text{Mat}_n(\mathbb{K}[t])$$

and we have $(t\mathbb{1}_n - A)^*(t\mathbb{1}_n - A) = (t - z)^*(t - z)\mathbb{1}_n$.

Before proving this proposition, we give the following simple observation that we need for the case $\mathbb{K} = \mathbb{R}$. Its use in the proof of Proposition 2.2 has been inspired by a similar argument in [FRS06]. In order to distinguish from taking the adjoint, we denote the entry-wise complex conjugation of a vector $v \in \mathbb{C}^n$ or a matrix $M \in \text{Mat}_n(\mathbb{C})$ by \bar{v} and \bar{M} , respectively.

Lemma 2.3. Let n be even and let $U \subseteq \mathbb{C}^n$ be a subspace that is orthogonal to its conjugate \bar{U} and maximal under inclusion with this property. Then $\dim U = \frac{n}{2}$. In particular, \bar{U} is the orthogonal complement of U .

Proof: The condition $\bar{U} \perp U$ (with respect to the standard hermitian inner product) just means that U is totally isotropic with respect to the bilinear form

$$\begin{aligned} \beta: \mathbb{C}^n \times \mathbb{C}^n &\rightarrow \mathbb{C} \\ (v, w) &\mapsto v^T w. \end{aligned}$$

Since -1 is a square in \mathbb{C} and n is even, β is hyperbolic and thus every maximal totally isotropic subspace is of dimension $\frac{n}{2}$ (see for example [Lam05, Corollary I.4.4]). \square

Proof of Proposition 2.2: The condition $0 = Q^*(z)Q(z) = (Q(z^*))^*Q(z)$ can be read as orthogonality of the images of the linear maps $Q(z)$ and $Q(z^*)$ in \mathbb{C}^n , that is,

$$\text{im } Q(z) \perp \text{im } Q(z^*).$$

We are going to choose a subspace $U \subseteq \mathbb{C}^n$ such that

$$(*) \quad \text{im } Q(z) \subseteq U \quad \text{and} \quad U \perp \text{im } Q(z^*).$$

For $\mathbb{K} = \mathbb{C}$, we just take $U = \text{im } Q(z)$. For $\mathbb{K} = \mathbb{R}$, the appropriate choice of U will ensure that the entries of the constructed matrix lie in \mathbb{R} . Rewriting $(*)$ yields

$$(**) \quad U^\perp \subseteq \text{im } Q(z)^\perp = \ker Q^*(z^*) \quad \text{and} \quad U \subseteq \text{im } Q(z^*)^\perp = \ker Q^*(z).$$

We choose A to be the representing matrix of the operator $z^*\pi_U + z\pi_{U^\perp} \in \text{End}(\mathbb{C}^n)$, where π_U and π_{U^\perp} are the orthogonal projections onto U and U^\perp , respectively. In other words A^* acts on U as multiplication by z and on U^\perp as multiplication by z^* . Combining this with (**) we conclude

$$\forall u \in U: Q_{A^*}^* u = Q^*(z)u = 0$$

as well as

$$\forall w \in U^\perp: Q_{A^*}^* w = Q^*(z^*)w = 0.$$

Since U and U^\perp span \mathbb{C}^n , this means $Q_{A^*}^* = 0$, or equivalently ${}_A Q = 0$. Moreover, A clearly is normal and its only eigenvalues are z and z^* , as desired.

If $\mathbb{K} = \mathbb{C}$ we are done at this point. So for the rest of the proof we assume $\mathbb{K} = \mathbb{R}$ and n is even. In this case, $\overline{Q} = Q$. Since $\text{im } Q(z)$ is orthogonal to $\text{im } Q(z^*) = \text{im } \overline{Q}(z^*) = \overline{\text{im } Q(z)}$ we also have

$$\text{im } Q(z) \perp \overline{\text{im } Q(z)}.$$

We choose a subspace $U \subseteq \mathbb{C}^n$ containing $\text{im } Q(z)$ and maximal with $U \perp \overline{U}$. Since $\overline{U} \supseteq \overline{\text{im } Q(z)} = \text{im } Q(z^*)$ we also have $U \perp \text{im } Q(z^*)$ as required in (*). Due to the maximality of U it is the orthogonal complement of its conjugate \overline{U} , as observed in Lemma 2.3. In particular $\pi_{\overline{U}} = \pi_{U^\perp}$.

It is easily seen that the conjugate of the representing matrix of π_U is the representing matrix of $\pi_{\overline{U}} = \pi_{U^\perp}$. Using this, it is clear that the matrix A we constructed with the above choice of U has real entries. \square

2.2. Pole Cancellation. In this subsection, we show how to produce polynomial factorizations from given rational ones using unitary matrices “capturing” the poles of the factors, see Theorem 2.10. In the complex case $\mathbb{K} = \mathbb{C}$, our approach is similar to the approach in [ESS16].

Remark 2.4. In the following, we will often use the simple fact that if $\sum_i a_i^* a_i = 1$ for some polynomials $a_i \in \mathbb{K}[t]$, then all a_i are in fact constant. The reason is that all leading coefficients of the $a_i^* a_i$ are positive and hence the coefficients of the highest degree term cannot cancel each other. In particular, a polynomial unitary matrix $U \in \text{U}_n(\mathbb{K}[t])$ has constant entries, i.e. $U \in \text{U}_n(\mathbb{K})$.

Lemma 2.5. *For any $M \in \text{Her}_n(\mathbb{K}[t])$ and $k \in \mathbb{N}$, the classes of square factorizations of M correspond one-to-one to those of $M \oplus \mathbb{1}_k$. More precisely, the map*

$$\begin{aligned} \Phi_n(M, \mathbb{K}[t]) &\rightarrow \Phi_{n+k}(M \oplus \mathbb{1}_k, \mathbb{K}[t]) \\ [Q] &\mapsto [Q \oplus \mathbb{1}_k] \end{aligned}$$

is a bijection.

Proof: To show injectivity, let $U(Q_1 \oplus \mathbb{1}_k) = Q_2 \oplus \mathbb{1}_k$ for some unitary $(n+k)$ matrix. Then U must be of the form $U_1 \oplus \mathbb{1}_k$ and we have $U_1 Q_1 = Q_2$, that is $[Q_1] = [Q_2]$. To show surjectivity, let $P \in \text{Mat}_{(n+k)}(\mathbb{K}[t])$ such that $P^* P = M \oplus \mathbb{1}_k$. Then the last k columns of P form an orthonormal system and therefore must have constant entries, see Remark 2.4. Extending them to an orthonormal basis of \mathbb{K}^{n+k} shows that there exists a unitary matrix $U \in \text{U}_{n+k}(\mathbb{K})$ which has the same last k columns as P . Then $U^* P = Q \oplus \mathbb{1}_k$ for some $Q \in \text{Mat}_n(\mathbb{K}[t])$ with $Q^* Q = M$. \square

We use the fact that we can split off linear factors coming from matrix zeros in order to show that we can get rid of poles in (rational) factorizations of polynomial matrices.

Theorem 2.6. *Let \mathcal{O} be a $*$ -invariant subring of $\mathbb{K}(t)$ containing $\mathbb{K}[t]$ and let $S \in \text{Mat}_n(\mathcal{O})$. If S^*S has polynomial entries, then there exists a unitary matrix $U \in \text{U}_n(\mathcal{O})$ such that US has polynomial entries. In other words, for positive semidefinite $M \in \text{Her}_n(\mathbb{K}[t])$ the canonical map*

$$\Phi_n(M, \mathbb{K}[t]) \rightarrow \Phi_n(M, \mathcal{O})$$

is surjective.

Proof: Note that for every fixed n , the two formulations of the theorem are in fact equivalent. For the moment we assume that n is even if $\mathbb{K} = \mathbb{R}$, and keep the odd case for later. First we prove the following intermediate claim:

If $a \in \mathbb{K}[t]$ and $Q \in \text{Mat}_n(\mathbb{K}[t])$ such that a^*a divides the entries of Q^*Q in $\mathbb{K}[t]$, then there exist $U \in \text{U}_n(\mathbb{K}[t]_{a^*})$ such that $UQ \in \text{Mat}_n(\mathbb{K}[t])$ and a divides the entries of UQ in $\mathbb{K}[t]$. For this we may assume that a is irreducible, otherwise we repeat the argument for each irreducible factor of a . So let a be monic and irreducible such that a^*a divides the entries of Q^*Q . We consider two cases.

Case 1: a is linear, say $a = t - z$. Since $(Q^*Q)(z) = 0$ we can use Proposition 2.2 to split off a linear factor $T := (t - A)$ with $T^*T = a^*a\mathbb{1}_n$ from the left of Q , i.e. $Q = TP$ for some $P \in \text{Mat}_n(\mathbb{K}[t])$. Then $U := \frac{1}{a^*}T^*$ is unitary and $UQ = \frac{1}{a^*}T^*TP = aP$ is divisible by a .

Case 2: a is quadratic, say $a = (t - z)^*(t - z)$ (in particular, $\mathbb{K} = \mathbb{R}$ and n is even). Then $a^2|Q^*Q$. Just as in the first case, we can split off a linear factor T_0 from Q with $T_0^T T_0 = a\mathbb{1}_n$. So let $Q = T_0 P$. Then $a|P^T P$ and we can split off another linear factor T_1 from P with $T_1^T T_1 = a\mathbb{1}_n$. Then $T := T_0 T_1$ divides Q from the left and $T^T T = a^2\mathbb{1}_n$. In particular, $U := \frac{1}{a}T^T$ is orthogonal and UQ is divisible by a . This proves the intermediate claim.

Now let $S \in \text{Mat}_n(\mathcal{O})$ be a matrix of rational functions such that S^*S has polynomial entries. Denote by $a \in \mathbb{K}[t]$ the smallest common denominator of the entries of S . Using that $\mathbb{K}[t]$ is a principal ideal domain, it is not hard to see, that $\frac{1}{a} \in \mathcal{O}$. Indeed, given $\frac{c_1}{a}, \dots, \frac{c_r}{a} \in \mathcal{O}$ such that a is coprime to c_1, \dots, c_r , then there exists a linear combination $1 = \alpha a + \sum_i \gamma_i c_i$ with $\alpha, \gamma_1, \dots, \gamma_r \in \mathbb{K}[t]$ and therefore

$$\frac{1}{a} = \alpha + \sum_i \gamma_i \frac{c_i}{a} \in \mathcal{O}.$$

Since \mathcal{O} is $*$ -invariant, also $\frac{1}{a^*} \in \mathcal{O}$. We set $Q := aS \in \text{Mat}_n(\mathbb{K}[t])$. Since S^*S has polynomial entries, $Q^*Q = a^*a S^*S$ is divisible by a^*a . Using the claim, there exists $U \in \text{U}_n(\mathbb{K}[t]_{a^*}) \subseteq \text{U}_n(\mathcal{O})$ such that $UQ \in \text{Mat}_n(\mathbb{K}[t])$ and a divides the entries of UQ and hence $US = \frac{1}{a}UQ$ has polynomial entries, as claimed. This proves the case $\mathbb{K} = \mathbb{C}$ or n even.

For the remaining case let $\mathbb{K} = \mathbb{R}$ and n be odd. We prove the second formulation of the theorem. So let $M \in \text{Sym}_n(\mathbb{K}[t])$ be positive semidefinite. We look at the following commutative diagram consisting of the canonical maps.

$$\begin{array}{ccc} \Phi_n(M, \mathbb{R}[t]) & \longrightarrow & \Phi_n(M, \mathcal{O}) \\ \downarrow & & \downarrow \\ \Phi_{n+1}(M \oplus \langle 1 \rangle, \mathbb{R}[t]) & \longrightarrow & \Phi_{n+1}(M \oplus \langle 1 \rangle, \mathcal{O}) \end{array}$$

The left hand map and bottom map are surjective by Lemma 2.5 and the even case, respectively. Since the right hand map is clearly injective (see the proof of Lemma 2.5), this gives the surjectivity of the top map, completing the proof. \square

Corollary 2.7. *Let $M \in \text{Sym}_n(\mathbb{R}[t])$ be positive semidefinite and non-degenerate. Then there exists $Q \in \text{Mat}_n(\mathbb{R}[t])$ with $M = Q^T Q$ if and only if $\det(M)$ is a square in $\mathbb{R}[t]$.*

Proof: Combine Corollary 1.2 with Theorem 2.6. \square

Corollary 2.8. *Let $M \in \text{Sym}_n(\mathbb{R}[t])$ be positive semidefinite. Then M admits a factorization $M = Q^T Q$ for some $Q \in \text{Mat}_{(n+1) \times n}(\mathbb{R}[t])$.*

Proof: First we reduce to the case that M is non-degenerate, i.e. has nonzero determinant. Since $\mathbb{R}[t]$ is a principal ideal domain, we can choose a basis of $\ker M$ and extend it to a basis of $\mathbb{R}[t]^n$. The according congruence transformation on M results in a block matrix of the form $M' \oplus 0$, where M' is non-degenerate and positive semidefinite. From a factorization of M' we get one of M . Replacing M by M' , we may assume that $d := \det(M)$ is nonzero. Then by the previous Corollary 2.7, $M \oplus d$ has a square factorization of size $n + 1$. The first n columns of the latter give the desired $(n + 1) \times n$ -factorization of M . \square

One of the main aims of this paper is not only to prove existence, but to give a precise classification of all such factorizations up to unitary equivalence. For every non-degenerate positive semidefinite matrix $M \in \text{Her}_n(\mathbb{K}[t])$, there exists only one square factorization up to unitary equivalence over $\mathbb{K}(t)$. Namely, let $M = Q^* Q = P^* P$, with $Q, P \in \text{Mat}_n(\mathbb{K}(t))$. Then Q and P only differ by the unitary matrix $U = Q^{-1} P \in \text{U}_n(\mathbb{K}(t))$. That is, over $\mathbb{K}(t)$ all square factorizations of M are equivalent. However, the situation changes if we require the involved unitary matrices to have no poles wherever M is singular, i.e. to have entries in \mathcal{O}_M . The next proposition shows that any two polynomial factorizations over $\mathbb{K}[t]$ that are equivalent over \mathcal{O}_M , are already equivalent over $\mathbb{K}[t]$. Due to Remark 2.4, they are even equivalent over \mathbb{K} .

Proposition 2.9. *Let $M \in \text{Her}_n(\mathbb{K}[t])$ be positive semidefinite and non-degenerate. Let \mathcal{O} be a subring of \mathcal{O}_M containing $\mathbb{K}[t]$. Given $Q_1, Q_2 \in \text{Mat}_n(\mathbb{K}[t])$ and $U \in \text{U}_n(\mathcal{O})$ such that $M = Q_1^* Q_1 = Q_2^* Q_2$ and $U Q_1 = Q_2$, then $U \in \text{U}_n(\mathbb{K}[t])$. In particular, the canonical map*

$$\Phi_n(M, \mathbb{K}[t]) \rightarrow \Phi_n(M, \mathcal{O})$$

is injective.

Proof: By assumption, the entries of $U = Q_2 Q_1^{-1}$ lie in $\mathcal{O} \subseteq \mathcal{O}_M$. By definition of \mathcal{O}_M this means that potential poles of U can only occur wherever $\det(M)$ does not vanish. But in these points M and thus Q_1 are invertible. Therefore, U is defined everywhere and hence polynomial. \square

We have now proved the Cassels-Pfister Theorem for matrices over $\mathbb{K}[t]$ that we need for our purposes.

Theorem 2.10. *Let $M \in \text{Her}_n(\mathbb{K}[t])$ be positive semidefinite and non-degenerate and let $\mathcal{O} = \mathcal{O}_M$. Then the canonical map*

$$\Phi_n(M, \mathbb{K}[t]) \rightarrow \Phi_n(M, \mathcal{O})$$

is a bijection.

Proof: Combine Theorem 2.6 and Proposition 2.9. \square

3. THE SMITH NORMAL FORM OF POSITIVE SEMIDEFINITE MATRICES

We will see that Theorem 2.10 gives us a much stronger result than mere existence of polynomial factorizations as in Corollaries 2.7 and 2.8. Namely it allows us to work over the semi-local ring \mathcal{O}_M instead of $\mathbb{K}[t]$ in order to count the number of square factorizations. The advantage of working over \mathcal{O}_M lies in the main result of this section, Theorem 3.7, which states that a positive semidefinite matrix $M \in \text{Sym}_n(\mathbb{K}[t])$ is congruent to its Smith normal form if we allow congruence transformations over \mathcal{O}_M .

Recall that the *Smith normal form* of a matrix $M \in \text{Mat}_n(R)$ over a principal ideal domain R is a diagonal matrix $D = \langle a_1, \dots, a_k, 0, \dots, 0 \rangle$ (k the rank of M) where $a_1, \dots, a_k \in R$ with $a_i | a_{i+1}$ ($i = 1, \dots, k-1$) such that there exist invertible matrices $S, T \in \text{Mat}_n(R)$ with $SMT = D$. The a_i are called the *invariant factors* of M and are uniquely determined up to units in R . See [Mac33, Chapter IV] for background. Moreover, if $R = \mathbb{K}[t]$, then requiring the invariant factors to be monic makes them unique and in that case we refer to $\langle a_1, \dots, a_k, 0, \dots, 0 \rangle$ as the *monic Smith normal form* of M .

In general the transformation to obtain the Smith normal form of a symmetric matrix M cannot be chosen to be a congruence transformation, i.e. $S = T^*$ for the above transformation matrices. However, if $R = \mathbb{K}[t]$ and M is positive semidefinite, then this is possible locally “around” the roots of $\det(M)$. For a precise statement see Theorem 3.7.

3.1. Diagonalization over semi-local principal ideal domains. First we show that the Gram-Schmidt method for orthogonalization leads to a Smith normal form of a given matrix over many semi-local principal ideal domains.

Lemma 3.1. *Let \mathcal{O} be a semi-local principal ideal domain containing \mathbb{Q} . Moreover, let $A \in \text{Sym}_n(\mathcal{O})$. Then A is congruent to its Smith normal form. More precisely, there exist $a_1, \dots, a_k \in \mathcal{O}$, where k is the rank of A , with $a_1 | a_2 | \dots | a_k$ and $A \simeq_{\mathcal{O}} \langle a_1, \dots, a_k, 0, \dots, 0 \rangle$.*

Proof: We proceed by induction on n . If $n = 0$ there is nothing to prove. So let $n > 0$. Then either $A = 0$ or the entries of A have a greatest common divisor, denoted by a_1 . So we have $A = a_1 B$ for some $B \in \text{Sym}_n(\mathcal{O})$. For each of the finitely many maximal ideals of \mathcal{O} there is at least one entry of B not contained in it. Denote $q_B: (v, w) \mapsto v^T B w$ the bilinear form defined by B . By the prime avoidance Lemma 5.2 (Appendix) for quadratic forms, there exists $v \in \mathcal{O}^n$ such that $q_B(v, v)$ is not contained in any of the maximal ideals and is hence a unit in \mathcal{O} . Since v represents a unit, the submodule $\mathcal{O}v$ of \mathcal{O}^n has an orthogonal complement, as can be shown using the Gram-Schmidt orthogonalization method. Restricting the bilinear form q_B to the orthogonal complement and applying the induction hypothesis to any representing matrix, we get the diagonalization as claimed. \square

Remark 3.2. Under one additional assumption, the same proof also works for a hermitian matrix A over a semi-local principal ideal domain $\mathcal{O} \supset \mathbb{Q}$ with involution, using Lemma 5.3. In the induction step, we need to be able to choose the greatest common divisor a_1 of the entries of A to be hermitian, i.e. $*$ -invariant. This is possible, if (and only if) every $*$ -invariant ideal (in this case, the ideal generated by the entries of A) has a $*$ -invariant generator. If \mathcal{O} is a subring of $\mathbb{C}(t)$ containing $\mathbb{C}[t]$, then this is true, since every ideal is generated by a monic polynomial, which must have real coefficients, if the ideal is $*$ -invariant.

In fact, this condition just means that \mathcal{O} is unramified over its subring of $*$ -invariant elements. To illustrate that this assumption is necessary, we equip $\mathbb{R}[t]$

with the \mathbb{R} -linear involution given by $t \mapsto -t$. The ring of $*$ -invariant elements is $\mathbb{R}[t^2]$. Clearly, the $*$ -invariant ideal (t) has no $*$ -invariant representative and it is obviously not possible to diagonalize the hermitian matrix $\begin{pmatrix} 0 & t \\ -t & 0 \end{pmatrix}$, even over the localization $\mathcal{O} = \mathbb{R}[t]_{(t)}$.

3.2. Avoiding Denominators. The most technical step in the proof of Theorem 3.7 is to keep track of denominators in transforming quadratic forms. Using induction, similarly as in Corollary 1.2, we reduce to the case of two dimensional forms and elements represented by them. As is common in quadratic forms theory, it can be quite useful to consider quadratic forms of the form $\langle 1, -c \rangle$. The essential advantage that we are going to exploit is the additional multiplicative structure that we gain by viewing these as norm forms of quadratic ring extensions.

We first give a variant of a standard exercise in number theory about the ring of integers in quadratic number fields.

Lemma 3.3. *Let A be a principal ideal domain with field of fractions K and let $c \in A$ be square free. If $2 \in A^\times$, then $A[\sqrt{c}]$ is the integral closure of A in $K[\sqrt{c}]$.*

Proof: Clearly, every element of $A[\sqrt{c}]$ is integral over K . Now let $a, b \in K$ and set $x := a + b\sqrt{c}$. Suppose x is integral over A . Then $x^* = a - b\sqrt{c}$ is integral, too. Therefore, $2a = x^* + x \in A$ and thus $a \in A$. Now also $b\sqrt{c}$ is integral. In particular, $b^2c \in A$. Since c is square free, also $b \in A$. \square

Lemma 3.4. *Let B be a factorial ring with involution. Then for every $a \in L := \text{Quot}(B)$ with $a^*a \in B$ there exists $b \in B$ such that $b^*b = a^*a$.*

Proof: Let $a = \frac{c}{d}$ with $c, d \in B$ coprime. Then d divides c^*c since $a^*a \in B$. But since c and d are coprime, already d divides c^* . In other words d^* divides c , i.e. $a = \frac{bd^*}{d}$ for some $b \in B$. Clearly $b^*b = a^*a$. \square

The following somewhat technical lemma is used to avoid denominators in transformation in the aforementioned two-dimensional forms. Recall that we write $Z_B(e)$ for the set of all prime ideals of B containing e .

Lemma 3.5. *Let A be a semi-local principal ideal domain with field of fractions K and $2 \in A^\times$. Moreover, let $c, e \in A$ such that c is square free and denote $B := A[\sqrt{c}]$. Suppose that for all the zeros $\mathfrak{q} \in Z_B(e)$ of e , the residue field $k(\mathfrak{q}) = B/\mathfrak{q}$ is quadratically closed¹. Then for every $\gamma \in L := K[\sqrt{c}]$ such that $N_{L|K}(\gamma) \in A^\times$, there exists $\alpha \in A[e\sqrt{c}]$ such that $N_{L|K}(\alpha) = N_{L|K}(\gamma)$.*

Proof: Note that since 2 is a unit, B is the integral closure of A in L by Lemma 3.3 and hence a Dedekind domain, see [Neu99, Chapter I, Proposition (8.1)]. Also, B is clearly semi-local and therefore, a principal ideal domain [Neu99, Chapter IV, §4, Exercise 4]. Denote $N := N_{L|K}$ the norm form of $L|K$, see [Neu99, Chapter I, §2]. As an involution $*$ on L and B we fix the nontrivial K -automorphism of L . Then $N(x) = x^*x$ for all $x \in L$.

Using Lemma 3.4 we may assume that $\gamma \in B$, since $\gamma^*\gamma \in A \subseteq B$. We are going to construct $\varepsilon \in B^\times$ such that $\varepsilon^2\gamma \in A[e\sqrt{c}]$. Then, for $\alpha := \frac{\varepsilon^2}{N(\varepsilon)}\gamma$, we have $N(\alpha) = N(\gamma)$ and $\alpha \in A[e\sqrt{c}]$, as desired, because $N(\varepsilon) \in A^\times$.

To construct such an ε , we use the fact that for all $\mathfrak{q} \in Z_B(e)$ the residue field $k(\mathfrak{q}) = B/\mathfrak{q}$ is quadratically closed as well as $\gamma^*(\mathfrak{q}) \neq 0$ (because $\gamma \in B^\times$) to conclude by Hensel's Lemma for complete discrete valuations rings (see [Neu99,

¹i.e. every element is a square.

Chapter II, §4, Lemma (4.6)) that γ^* is a square modulo any power of \mathfrak{q} . So we can choose a unit $\varepsilon \in B^\times$ satisfying the following finitely many congruences for $\mathfrak{q} \in Z_B(e)$

$$\varepsilon^2 \equiv \gamma^* \pmod{\mathfrak{q}^{n_{\mathfrak{q}}}}$$

where

- $n_{\mathfrak{q}} = 2v_{\mathfrak{q}}(e) + 1$ if $\mathfrak{q}|A \cap \mathfrak{q}$ is ramified (which is the case if and only if \mathfrak{q} is a zero of \sqrt{c})
- $n_{\mathfrak{q}} = v_{\mathfrak{q}}(e)$ otherwise.

Here, $v_{\mathfrak{q}}$ denotes the discrete valuation corresponding to \mathfrak{q} . Let $a, b \in A$ with $\gamma\varepsilon^2 = a + b\sqrt{c}$. Then we have

$$b = \frac{\gamma\varepsilon^2 - (\gamma\varepsilon^2)^*}{2\sqrt{c}}.$$

By the choice of ε , we get that $v_{\mathfrak{p}}(b) \geq v_{\mathfrak{p}}(e)$ for all $\mathfrak{p} \in Z_A(e)$ and hence e divides b . This means $\gamma\varepsilon^2 \in A[e\sqrt{c}]$ as desired. \square

Combining the previous lemma with Witt's local-global principle leads to the following proposition, which is the main step in the proof of Theorem 3.7.

Proposition 3.6. *Let $a, b, d \in \mathbb{R}[t]$ be nonzero with a and b positive semidefinite and $a|b|d$. Write*

$$\mathcal{O} := \left\{ \frac{c}{e} \mid c, e \in \mathbb{R}[t], e \text{ coprime to } d \right\}$$

and let $u, v \in \mathcal{O}^\times$ be positive semidefinite. Then

$$\langle au, bv \rangle \simeq_{\mathcal{O}} \langle a, buv \rangle.$$

Proof: Since a divides b , we can factor out a and therefore assume that $a = 1$. By Corollary 1.1 to Witt's local-global principle, 1 is represented by $\langle u, bv \rangle$ over the rational function field $\mathbb{R}(t)$. Dividing by u , we get that $\frac{1}{u}$ is represented over $\mathbb{R}(t)$ by

$$\left\langle 1, \frac{bv}{u} \right\rangle \cong_{\mathbb{R}(t)} \langle 1, -c \rangle$$

where $c \in \mathcal{O}$ is the square free part of $-\frac{bv}{u}$, i.e. c is square free and $-\frac{bv}{u} = e^2c$ for some $e \in \mathcal{O}$. Being represented by $\langle 1, -c \rangle$ means being the norm of an element $\mathbb{R}(t)[\sqrt{c}]$. Clearly, c is negative semidefinite. Since c is square free, it cannot have a real zero. In particular, all quotients of $\mathcal{O}[\sqrt{c}]$ modulo its maximal ideals are isomorphic to \mathbb{C} and hence quadratically closed. Now we can use Lemma 3.5 to get a representation of $\frac{1}{u}$ as a norm of an element of $\mathcal{O}[e\sqrt{c}]$. In other words $\frac{1}{u}$ is represented by $\langle 1, -e^2c \rangle$ over \mathcal{O} or (multiplying by u) we get that 1 is represented by

$$\langle u, -ue^2c \rangle = \langle u, bv \rangle.$$

Hence the latter is congruent to $\langle 1, buv \rangle$ over \mathcal{O} . \square

Similarly to Corollary 1.2 we apply this Proposition inductively to obtain the main result of this section.

Theorem 3.7. *Let $M \in \text{Her}_n(\mathbb{K}[t])$ positive semidefinite. Then M and its monic Smith normal form are congruent over \mathcal{O}_M .*

Proof: Let $\mathcal{O} := \mathcal{O}_M$. Using Lemma 3.1 and Remark 3.2 we get $b_1 | \dots | b_k \in \mathcal{O}$ such that

$$M \simeq_{\mathcal{O}} E := \langle b_1, \dots, b_k, 0, \dots, 0 \rangle.$$

We may assume that $k = n$, i.e. M is non-degenerate. Now let $D = \langle a_1, \dots, a_n \rangle$ be the monic Smith normal form of M over $\mathbb{K}[t]$. Then both D and E are Smith normal forms of M over \mathcal{O} . Due to uniqueness there exist units $u_i \in \mathcal{O}^\times$ such that $b_i = a_i u_i$ for all i . We now show that E and D are congruent over \mathcal{O} .

Let $i \in \{1, \dots, n\}$. The determinant of M is divisible by a_i . Since u_i is a unit in \mathcal{O} , the numerator and denominator of u_i (in a representation in lowest terms) are coprime to a_i by the definition of $\mathcal{O} = \mathcal{O}_M$. In particular, the rational functions u_i and a_i cannot have simultaneous sign changes. Since $b_i = a_i u_i$ is positive semidefinite, both a_i and u_i must be positive semidefinite as well.

If $\mathbb{K} = \mathbb{C}$, then u_1, \dots, u_n are hermitian squares of units in \mathcal{O} and we are already done. So now we consider the case $\mathbb{K} = \mathbb{R}$.

We apply Proposition 3.6 to get that the subform $\langle a_1 u_1, a_n u_n \rangle$ of D is congruent to $\langle a_1, a_n u_1 u_n \rangle$ over \mathcal{O} . Replacing u_n by $u_1 u_n$, we can therefore assume that $u_1 = 1$. Repeating this argument we can also assume that $u_2 = \dots = u_{n-1} = 1$. Up to a positive constant, $\prod_{i=1}^n a_i$ is the determinant of M . Since congruent transformations only change the determinant by a square, we conclude that u_n must be a square and hence can also be assumed to be 1, which finishes the proof. \square

4. PROOF OF THE MAIN THEOREMS

As in the previous sections, we let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. We combine our previous work in order to prove the main results.

Theorem 4.1. *Let $M \in \text{Her}_n(\mathbb{K}[t])$ be positive semidefinite and non-degenerate. Then the equivalence classes of $n \times n$ -factorizations $M = Q^*Q$ are in one-to-one correspondence to those of the monic Smith normal form of M .*

Proof: Denote $\mathcal{O} := \mathcal{O}_M$ and let D be the monic Smith normal form of M . We consider the following diagram.

$$\begin{array}{ccc} \Phi_n(M, \mathbb{K}[t]) & \dashrightarrow & \Phi_n(D, \mathbb{K}[t]) \\ \downarrow & & \downarrow \\ \Phi_n(M, \mathcal{O}) & \longrightarrow & \Phi_n(D, \mathcal{O}) \end{array}$$

By Theorem 2.10 we have vertical bijections induced by the inclusion. By Theorem 3.7 we have $M \simeq_{\mathcal{O}} D$, i.e. there exists an invertible matrix $T \in \text{Mat}_n(\mathcal{O})$ with $D = T^*MT$. Right multiplication with T induces a bijection on the bottom. \square

From this we obtain the result for $\mathbb{K} = \mathbb{C}$ as mentioned in the introduction.

Corollary 4.2. *Let $M \in \text{Her}_n(\mathbb{C}[t])$ be positive semidefinite with nonzero and square free determinant. Then the determinant map induces a bijection between the sets*

$$\{ Q \in \text{Mat}_n(\mathbb{C}[t]) \mid Q^*Q = M \}$$

and

$$\{ g \in \mathbb{C}[t] \mid g^*g = \det(M) \}$$

modulo the left action of the unitary groups $U_n(\mathbb{C})$ and $U_1(\mathbb{C})$, respectively.

Proof: Without loss of generality we can assume that $d := \det(M)$ is monic. Since d is square free, the monic Smith normal form of M is given by

$$D = \langle 1, \dots, 1, d \rangle = \mathbb{1}_{n-1} \oplus \langle d \rangle.$$

We combine the two bijections

$$\Phi_n(M, \mathbb{C}[t]) \rightarrow \Phi_n(D, \mathbb{C}[t]) = \Phi_n(\mathbb{1}_{n-1} \oplus \langle d \rangle, \mathbb{C}[t]) \rightarrow \Phi_1(\langle d \rangle, \mathbb{C}[t])$$

from Theorem 4.1 and Lemma 2.5, respectively. By Remark 2.4, $U_n(\mathbb{C}[t]) = U_n(\mathbb{C})$. Following the construction, it is clear that the map is induced by the determinant. \square

4.1. Cauchy-Binet. We want to show that the analogue of Theorem 4.1 holds also for $(n+1) \times n$ instead of $n \times n$ -factorizations. For this we observe the following. For any integral domain R and nonzero $d \in R$, the decompositions

$$d = a^2 + b^2$$

as a sum of two squares are basically the same as 2×2 -factorizations

$$\langle d, d \rangle = Q^T Q$$

since for $d = a^2 + b^2$ we can extend the vector $(a, b)^T$ to the matrix $Q = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. This observation can also be generalized to higher dimensions in the following way.

Lemma 4.3. *Let R be an integral domain, $M \in \text{Sym}_n(R)$ with nonzero determinant $d = \det(M)$ and $Q \in \text{Mat}_{(n+1) \times n}(R)$ such that $Q^T Q = M$. Then Q can be extended to a square factorization of $M \oplus \langle d \rangle$. More precisely, there exists a vector $v \in R^{n+1}$ such that for $P = (Q|v) \in \text{Mat}_{n+1}(R)$ we have $P^T P = M \oplus \langle d \rangle$. Moreover, v is uniquely determined up to a scalar factor ± 1 .*

Proof: Uniqueness is clear, since the vector space (over $\text{Quot } R$) of solutions to $Q^T v = 0$ is one-dimensional. So there are at most 2 solutions with the additional requirement $v^T v = d$. To show existence, we define the i -th component of v to be the i -th maximal minor of Q with sign $(-1)^i$. Then by Cramer's rule $Q^T v = 0$. Moreover, using the Cauchy-Binet formula to compute the determinant of $Q^T Q$, we get that $v^T v = \det(Q^T Q) = d$. \square

Remark 4.4. We want to note that the analogue of the preceding lemma holds for factorizations of hermitian matrices over rings with involutions. The difference is that the extending vector v is determined up to a unit of norm 1, i.e. a factor $u \in R^\times$ with $u^* u = 1$ (instead of $u^2 = 1$).

With R and M as above (using Remark 4.4, M might as well be hermitian, if R is a ring with involution), this lemma can be essentially reformulated as follows.

Corollary 4.5. *Removal of the last column induces a bijection*

$$\Phi_{n+1}(M \oplus \langle \det(M) \rangle, R) \rightarrow \Phi_{n+1}(M, R). \quad \square$$

Using this observation, we get the corresponding result of Theorem 4.1 for $(n+1) \times n$ -factorizations.

Corollary 4.6. *Let $M \in \text{Her}_n(\mathbb{K}[t])$ be positive semidefinite and non-degenerate. Then the equivalence classes of $(n+1) \times n$ -factorizations $M = Q^* Q$ are in one-to-one correspondence to those of the monic Smith normal form of M .*

Proof: We may assume, that $d := \det(M)$ is monic. If D is the monic Smith normal form of M , then $D \oplus \langle d \rangle$ is the monic Smith normal form of $M \oplus \langle d \rangle$. So we can complete the following diagram to make it commute.

$$\begin{array}{ccc} \Phi_{n+1}(M \oplus \langle d \rangle, \mathbb{K}[t]) & \longrightarrow & \Phi_{n+1}(D \oplus \langle d \rangle, \mathbb{K}[t]) \\ \downarrow & & \downarrow \\ \Phi_{n+1}(M, \mathbb{K}[t]) & \dashrightarrow & \Phi_{n+1}(D, \mathbb{K}[t]) \end{array}$$

We have a bijection on top by Theorem 4.1 and vertical bijections by Corollary 4.5. \square

Now we can prove the case of particular interest over the field of real numbers.

Corollary 4.7. *Let $M \in \text{Sym}_n(\mathbb{R}[t])$ be positive semidefinite with nonzero and square free determinant $\det(M)$. Then there is a bijection between the sets*

$$\{ Q \in \text{Mat}_{(n+1) \times n}(\mathbb{R}[t]) \mid Q^T Q = M \}$$

and

$$\{ g \in \mathbb{R}[t]^2 \mid g^T g = \det(M) \}$$

modulo the left action of the orthogonal groups $O_n(\mathbb{R})$ and $O_2(\mathbb{R})$, respectively.

Proof: The proof works just as for Corollary 4.2. Only we use Corollary 4.6 instead of Theorem 4.1 in order to obtain a bijection

$$\Phi_{n+1}(M, \mathbb{R}[t]) \rightarrow \Phi_2(\langle d \rangle, \mathbb{R}[t]).$$

\square

Remark 4.8. While it is clear in the complex case, Corollary 4.2, that the map in question is given by the determinant, we do not have such an obvious description in the real case.

Given a factorization $M = Q^T Q$, where $Q \in \text{Mat}_{n+1}(\mathbb{R}[t])$, we follow the construction steps to get a representation of $d := \det(M)$ as a sum of two squares. By Cauchy-Binet, we get a representation of d as a sum of $n + 1$ -squares, as in Lemma 4.3. More precisely, there exists $v \in \mathbb{R}[t]^{n+1}$ such that for $Q_1 := (Q|v) \in \text{Mat}_{n+1}(\mathbb{R}[t])$ we have $M \oplus \langle d \rangle = Q_1^T Q_1$. By Theorem 3.7 there exists an invertible matrix $T \in \text{Mat}_n(\mathcal{O}_M)$ such that $T^T M T = \mathbb{1}_{n-1} \oplus d$. For $Q_2 := (QT, v)$ we therefore get $\mathbb{1}_{n-1} \oplus \langle d, d \rangle = Q_2^T Q_2$. By Theorem 2.10 there exists $U \in O_{n+1}(\mathcal{O}_M)$ such that UQ_2 is polynomial and by further applying Lemma 2.5 we may assume that $UQ_2 = \mathbb{1}_{n-1} \oplus A$, for some $A \in \text{Mat}_2(\mathbb{R}[t])$ with $\langle d, d \rangle = A^T A$. Therefore, the last column of UQ_2 is given by $Uv = (0, \dots, a, b)^T$, where $g = (a, b)^T \in \mathbb{R}[t]^2$ is the last column of A , which is the desired 2×1 -factorization of d .

In short, for the representation $v \in \mathbb{R}[t]^{n+1}$ of $\det(M)$ as a sum of $n + 1$ -squares coming from the application of Cauchy-Binet to compute $\det(Q^T Q)$, there exists an *appropriate* orthogonal matrix $U \in O_{n+1}(\mathcal{O}_M)$, such that Uv is essentially a vector g of length two.

The first natural question that arises from this observation is whether v can already be compressed over \mathbb{R} to a vector g of length two, that is, can we choose the above matrix U to have entries in \mathbb{R} ? We want to give an example that this is generally not the case, i.e. the denominators in U are really necessary:

Pick any $a, b \in \mathbb{R}[t]$ such that $1, a, b$ are \mathbb{R} -linearly independent (e.g. $a = x$, $b = x^2$). We set

$$Q := \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ a & b \end{pmatrix} \in \text{Mat}_{3 \times 2}(\mathbb{R}[t])$$

which is a factorization of

$$M := Q^T Q = \begin{pmatrix} a^2 + 1 & ab \\ ab & b^2 + 1 \end{pmatrix}.$$

The vector $v = (-a, -b, 1)^T$ of maximal minors of Q gives a representation $a^2 + b^2 + 1 = v^T v$ of the determinant of M as a sum of three squares. By the above, v is equivalent to a vector of length two, over \mathcal{O}_M , not however over \mathbb{R} by the choice of a, b .

The next question now is, whether any $U \in \mathcal{O}_{n+1}(\mathcal{O}_M)$ for which Uv is of length two does the job. More generally, there is the following open problem.

Question. Given any $g, h \in \mathbb{R}[t]^2$ with $g^T g = h^T h =: d$ such that g and h are $\mathcal{O}_n(\mathcal{O}_{\langle d \rangle})$ -equivalent (after appending $n - 2$ zeros), are g and h already $\mathcal{O}_2(\mathbb{R})$ -equivalent?

5. APPENDIX: SOME PRIME AVOIDANCE

The following is a polynomial version of the prime avoidance lemma. Since we could not find a reference, we include it here.

Lemma 5.1. *Let R be a commutative ring and $\mathfrak{p}_1, \dots, \mathfrak{p}_n \in \text{Spec } R$ such that all quotients R/\mathfrak{p}_i are infinite. Moreover let $f \in R[x_1, \dots, x_s]$ be a polynomial that does not lie in $\bigcup_{i=1}^n \mathfrak{p}_i R[x_1, \dots, x_s]$. Then there exists $a \in R^s$ such that $f(a) \notin \bigcup_{i=1}^n \mathfrak{p}_i$.*

Proof: We may assume that the \mathfrak{p}_i are pairwise incomparable with respect to inclusion. Passing over to the quotient modulo $I := \bigcap_i \mathfrak{p}_i$ we may further assume that $I = 0$. We define the multiplicative set $S := R \setminus \bigcup_i \mathfrak{p}_i$, the localization $Q := S^{-1}R$, and the prime ideals $\mathfrak{q}_i := \mathfrak{p}_i Q$. Any element that is not contained in $\bigcup_i \mathfrak{q}_i$ is invertible. Using the Prime Avoidance Lemma [Eis95, Lemma 3.3] we thus conclude, that the \mathfrak{q}_i are maximal ideals. In particular, they are pairwise co-prime. By the Chinese Remainder Theorem, Q is isomorphic to the product of the fields $K_i := Q/\mathfrak{q}_i$. Via this isomorphism we identify elements $q \in Q$ with tuples $(q_1, \dots, q_r) \in \prod_{i=1}^r K_i$ and refer to q_1, \dots, q_r as the components of q . Likewise we identify $f \in Q[\underline{x}]$ with a tuple of polynomials

$$f = (f_1, \dots, f_r) \in \prod_{i=1}^r K_i[\underline{x}] = Q[\underline{x}]$$

and by assumption none of its components are zero. Since the K_i are infinite there exists $b_i \in K_i^s$ with $f_i(b_i) \neq 0$ for every i . Again we identify the tuple $(b_1, \dots, b_r) \in \prod_{i=1}^r K_i^s$ with an element $b \in Q^s$. Denote by $c \in S$ a common denominator of all components of b , i.e. $cb \in R^s$. None of the components $g_i := f_i(tb_i)$ of the univariate polynomial $g := f(tb) \in Q[t]$ is the zero polynomial, since $g_i(1) \neq 0$. By possibly replacing c by a suitable power of c , we may assume that every component c_i of c is either 1 or not a root of unity. Then for each i the set $\{c_i^d \mid d \in \mathbb{N}\}$ is either infinite or $\{1\}$. We can therefore choose $0 < d \in \mathbb{N}$ such that for all i we have $g_i(c_i^d) \neq 0$. Now $a := c^d b$ lies in R^s and has the property that $f(a) = g(c^d)$ does not lie in any of the \mathfrak{p}_i , as desired. \square

Lemma 5.2. *Let R be a commutative ring containing \mathbb{Q} and let $\mathfrak{p}_1, \dots, \mathfrak{p}_r \in \text{Spec } R$. Suppose we have $A \in \text{Sym}_n(R)$ such that $A \notin \bigcup_i \mathfrak{p}_i^{n \times n}$. Then there exists $v \in R^n$ with*

$$q_A(v, v) = v^T A v \notin \bigcup_i \mathfrak{p}_i.$$

Proof: Since $2 \in A^\times$ and $A \notin \bigcup_i \mathfrak{p}_i^{n \times n}$ we conclude that the polynomial $x^T A x \in R[x_1, \dots, x_n]$ does not lie in $\bigcup_i \mathfrak{p}_i R[x_1, \dots, x_n]$ (use the polarization identity for quadratic forms). The claim now follows from Lemma 5.1. \square

For the sake of completeness, we also provide the following generalization for hermitian forms.

Lemma 5.3. *Let $R \supset \mathbb{Q}$ be a commutative ring with involution and let $\mathfrak{p}_1, \dots, \mathfrak{p}_r \in \text{Spec } R$. Suppose we have $M \in \text{Her}_n(R)$ such that $M \notin \bigcup_i \mathfrak{p}_i^{n \times n}$. Then there exists $v \in R^n$ with*

$$h_M(v, v) = v^* M v \notin \bigcup_i \mathfrak{p}_i.$$

Proof: Denote by S the ring of elements which are fixed by the involution on R . In order to make the proof more transparent, we first consider the case that there exists a skew-hermitian unit, i.e. $u \in R^\times$ with $u^* = -u$. In this case $R = S + uS$, since $2 \in R^\times$. We write $M = A + uB$, with $A, B \in \text{Mat}_n(S)$, A symmetric and B skew-symmetric, thinking of A as the real part and B as the imaginary part of M .

Now we fix $i \in \{1, \dots, r\}$ for a moment and consider $\mathfrak{p} = \mathfrak{p}_i$. If $A \notin \mathfrak{p}^{n \times n}$, we find $z \in S^n$ such that $h_M(z, z) = z^* M z = z^T A z \notin \mathfrak{p}$ using the polarization identity. Otherwise, $B \notin \mathfrak{p}^{n \times n}$ and we take $z = e_i + u e_j$ if the ij -th entry of B does not lie in \mathfrak{p} . Then again $h_M(z, z) \notin \mathfrak{p}$. We repeat this for every i and apply Lemma 5.1 to the polynomial map $S^n \times S^n \rightarrow S, (x, y) \rightarrow h_M(x + uy, x + uy)$ to get v such that $h_M(v, v) \notin \mathfrak{p}$.

In case there exists no such skew-hermitian unit $u \in R^\times$, we still can write $M = A + B$, where $A \in \text{Sym}_n(S)$ and $B \in \text{Mat}_n(R)$ is skew-symmetric with skew-hermitian entries. Then for every $i \in \{1, \dots, r\}$, we choose skew-hermitian $u_i \in R$ such that

- $u_i = 0$, if $A \notin \mathfrak{p}_i^{n \times n}$,
- any $u_i \notin \mathfrak{p}_i$, if $A \in \mathfrak{p}_i^{n \times n}$ and hence $B \notin \mathfrak{p}_i^{n \times n}$.

Similarly as above, it is easy to check that for every i there exist $z_i \in R^n$ of the form $z_i = x_i + u_i y_i$ ($x_i, y_i \in S^n$) such that $h_M(z_i, z_i) \notin \mathfrak{p}_i$. Now we consider the map

$$\begin{aligned} \varphi: (S^n)^{r+1} &\rightarrow S^n \\ w = (w_0, \dots, w_r) &\mapsto w_0 + \sum u_i w_i. \end{aligned}$$

Again, applying Lemma 5.1 to the polynomial map $w \mapsto h_M(\varphi(w), \varphi(w))$, proves the existence of $v = \varphi(w)$ such that $h_M(v, v) \notin \bigcup \mathfrak{p}_i$. \square

REFERENCES

- [BPSV16] G. Blekherman, D. Plaumann, R. Sinn, and C. Vinzant. Low-rank sum-of-squares representations on varieties of minimal degree. *to appear in Int. Math. Res. Not. IMRN*, 2016.
- [Cas64] J. W. S. Cassels. On the representation of rational functions as sums of squares. *Acta Arith.*, 9:79–82, 1964.
- [CLR80] M. D. Choi, T.Y. Lam, and B. Reznick. Real zeros of positive semidefinite forms. I. *Math. Z.*, 171(1):1–26, 1980.
- [DR10] M.A. Dritschel and J. Rovnyak. The operator Fejér-Riesz theorem. In *A glimpse at Hilbert space operators*, volume 207 of *Oper. Theory Adv. Appl.*, pages 223–254. Birkhäuser Verlag, Basel, 2010.
- [Eis95] David Eisenbud. *Commutative algebra*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. With a view toward algebraic geometry.
- [ESS16] L. Ephremidze, I. Selesnick, and I. Spitkovsky. On non-optimal spectral factorizations. *ArXiv e-prints*, 1609.02058, 2016.
- [FRS06] José F. Fernando, Jesús M. Ruiz, and Claus Scheiderer. Sums of squares of linear forms. *Math. Res. Lett.*, 13(5-6):947–956, 2006.

- [Lam05] T. Y. Lam. *Introduction to quadratic forms over fields*, volume 67 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2005.
- [Lee06] D.B. Leep. Sums of squares of polynomials and the invariant $g_n(r)$. *unpublished manuscript, personal communication*, 2006.
- [Mac33] C.C. MacDuffee. The theory of matrices. *Zentralblatt MATH*, 59(02), 1933.
- [Neu99] J. Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.
- [O'M71] O. T. O'Meara. *Introduction to quadratic forms*. Springer-Verlag, New York-Heidelberg, 1971. Second printing, corrected, Die Grundlehren der mathematischen Wissenschaften, Band 117.
- [PD01] Alexander Prestel and Charles N. Delzell. *Positive polynomials*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2001. From Hilbert's 17th problem to real algebra.
- [Pfi65] Albrecht Pfister. Multiplikative quadratische Formen. *Arch. Math. (Basel)*, 16:363–370, 1965.
- [Tig96] J.-P. Tignol. A Cassels-Pfister theorem for involutions on central simple algebras. *J. Algebra*, 181(3):857–875, 1996.
- [WM57] N. Wiener and P. Masani. The prediction theory of multivariate stochastic processes. I. The regularity condition. *Acta Math.*, 98:111–150, 1957.

CHRISTOPH HANSELKA, THE UNIVERSITY OF AUCKLAND, DEPARTMENT OF MATHEMATICS, PRIVATE BAG 92019, AUCKLAND 1142, NEW ZEALAND
E-mail address: c.hanselka@auckland.ac.nz

RAINER SINN, MAX-PLANCK-INSTITUT FÜR MATHEMATIK IN DEN NATURWISSENSCHAFTEN, INSELSTRASSE 22, 04103 LEIPZIG, GERMANY
E-mail address: rsinn@mis.mpg.de