SUMS OF SQUARES OVER TOTALLY REAL FIELDS ARE RATIONAL SUMS OF SQUARES

CHRISTOPHER J. HILLAR

(Communicated by Bernd Ulrich)

ABSTRACT. Let K be a totally real number field with Galois closure L. We prove that if $f \in \mathbb{Q}[x_1, \ldots, x_n]$ is a sum of m squares in $K[x_1, \ldots, x_n]$, then f is a sum of

$$4m \cdot 2^{[L:\mathbb{Q}]+1} \binom{[L:\mathbb{Q}]+1}{2}$$

squares in $\mathbb{Q}[x_1, \ldots, x_n]$. Moreover, our argument is constructive and generalizes to the case of commutative K-algebras. This result gives a partial resolution to a question of Sturmfels on the algebraic degree of certain semidefinite programming problems.

1. INTRODUCTION

In recent years, techniques from semidefinite programming have produced numerical algorithms for finding representations of positive semidefinite polynomials as sums of squares. These algorithms have many applications in optimization, control theory, quadratic programming, and matrix analysis [20, 21, 23, 18, 19]. For a noncommutative application of these techniques to a famous trace conjecture, see the papers [1, 8, 13, 16], which continue the work of [9].

One major drawback with these algorithms is that their output is, in general, numerical. For many applications, however, exact polynomial identities are needed. In this regard, Sturmfels has asked whether a representation with real coefficients implies one over the rationals.

Question 1.1 (Sturmfels). If $f \in \mathbb{Q}[x_1, \ldots, x_n]$ is a sum of squares in $\mathbb{R}[x_1, \ldots, x_n]$, then is f also a sum of squares in $\mathbb{Q}[x_1, \ldots, x_n]$?

It is well-known that a polynomial is a sum of real polynomial squares if and only if it can be written in the form

$$(1.1) \qquad f = \mathbf{v}^T B \mathbf{v},$$

in which $\mathbf{v}$ is a column vector of monomials and B is a real positive semidefinite (square) matrix [25]; in this case, the matrix B is called a *Gram matrix* for f. If B happens to have rational entries, then f is a sum of squares in $\mathbb{Q}[x_1, \ldots, x_n]$ (this

follows from a Cholesky factorization argument or from a matrix generalization of Lagrange's four square theorem [10]). Thus, in the language of quadratic forms, Sturmfels is asking whether the existence of a positive semidefinite Gram matrix for $f \in \mathbb{Q}[x_1, \ldots, x_n]$ over the reals implies that one exists over the rationals.

Although a complete answer to Question 1.1 is not known, Parrilo and Peyrl have written an implementation of SOSTOOLS in the algebra package Macaulay 2 that attempts to find rational representations of polynomials that are sums of squares [22]. Their idea is to approximate a real Gram matrix B with rational numbers and then project back to the linear space of solutions governed by equation (1.1).

The following result says that Question 1.1 has a positive answer in some "generic" sense; it also explains the difficulty of finding counterexamples.

Theorem 1.2. *Let $f \in \mathbb{Q}[x_1, \ldots, x_n]$. If there is an invertible Gram matrix B for f, then there is a Gram matrix for f with rational entries.*

Proof. Let B be a real positive semidefinite matrix and $\mathbf{v}$ a vector of monomials such that $f = \mathbf{v}^T B \mathbf{v}$. Consider the set L of real symmetric matrices $S = S^T = (s_{ij})$ of the same size as B for which $f = \mathbf{v}^T S \mathbf{v}$. This space corresponds to the solutions of a set of linear equations in the s_{ij} over $\mathbb{Q}$. From elementary linear algebra (Gaussian elimination), it follows that there is an integer k such that

$$L = \{S_0 + t_1 S_1 + \cdots + t_k S_k : t_1, \ldots, t_k \in \mathbb{R}\}$$

for some rational symmetric matrices $S_1, \ldots, S_k$. The subset of matrices in L that are positive definite is determined by a finite set of strict polynomial inequalities in the $t_1, \ldots, t_k$ produced by setting all the leading principal minors to be positive [11, p. 404]. By continuity, a real positive definite solution B guarantees a rational one, and this completes the proof.   □

Remark 1.3. The argument above shows that we may find a rational Gram matrix of the same size as the original Gram matrix B. Is this true even if B is not invertible? We suspect not.

Although the general case seems difficult, Question 1.1 has a positive answer for univariate polynomials due to results of Landau [15], Pourchet [24], and (algorithmically) Schweighofer [29]. In fact, Pourchet has shown that at most 5 polynomial squares in $\mathbb{Q}[x]$ are needed to represent every positive semidefinite polynomial in $\mathbb{Q}[x]$, and this is best possible.

It follows from Artin's solution to Hilbert's 17th problem [26, Theorem 2.1.12] that if $f \in \mathbb{Q}[x_1, \ldots, x_n]$ is a sum of squares of rational functions in $\mathbb{R}(x_1, \ldots, x_n)$, then it is a sum of squares in $\mathbb{Q}(x_1, \ldots, x_n)$. Moreover, from the work of Voevodsky on the Milnor conjectures, it is known that 2^{n+2} such squares suffice [14, p. 530]. However, the transition from rational functions to polynomials is often a very delicate one. For instance, not every polynomial that is a sum of squares of rational functions is a sum of squares of polynomials [14, p. 398].

More generally, Sturmfels is interested in the algebraic degree [17] of maximizing a linear functional over the space of all sum of squares representations of a given polynomial that is a sum of squares. In the special case of Question 1.1, a positive answer signifies an algebraic degree of 1 for this optimization problem.

General theory (for instance, Tarski's Transfer Principle for real closed fields [26, Theorem 2.1.10]) reduces Question 1.1 to one involving real algebraic numbers. In this paper, we present a positive answer to this question for a special class of fields.

Recall that a *totally real number field* is a finite algebraic extension of $\mathbb{Q}$ all of whose complex embeddings lie entirely in $\mathbb{R}$. For instance, the field $\mathbb{Q}(\sqrt{d})$ is totally real for positive, integral d. Our main theorem is the following.

Theorem 1.4. *Let K be a totally real number field with Galois closure L. If $f \in \mathbb{Q}[x_1, \ldots, x_n]$ is a sum of m squares in $K[x_1, \ldots, x_n]$, then f is a sum of*

$$4m \cdot 2^{[L:\mathbb{Q}]+1} \binom{[L:\mathbb{Q}]+1}{2}$$

squares in $\mathbb{Q}[x_1, \ldots, x_n]$.

Our techniques also generalize naturally to the following situation. Let R be a commutative $\mathbb{Q}$-algebra and let K be a totally real number field. Also, set $S := R \otimes_{\mathbb{Q}} K$, which we naturally identify as a ring extension of $R = R \otimes_{\mathbb{Q}} \mathbb{Q}$. If f is a sum of the form

$$f = \sum_{i=1}^{m} p_i^2, \quad p_i \in R,$$

then we say that f is a *sum of squares* over R. It is a difficult problem to determine those f which are sums of squares over R. In this setting, Theorem 1.4 generalizes in the following way.

Theorem 1.5. *Let K be a totally real number field with Galois closure L. If $f \in R$ is a sum of m squares in $R \otimes_{\mathbb{Q}} K$, then f is a sum of $4m \cdot 2^{[L:\mathbb{Q}]+1} \binom{[L:\mathbb{Q}]+1}{2}$ squares over R.*

Remark 1.6. One can view Theorem 1.5 as a "going-down" theorem [4] for certain quadratic forms over the rings R and $R \otimes_{\mathbb{Q}} K$. We do not know how much the factor $2^{[L:\mathbb{Q}]+1} \binom{[L:\mathbb{Q}]+1}{2}$ can be improved upon, although we suspect that for polynomial rings, it can be improved substantially. We remark that it is known [2] that arbitrarily large numbers of squares are necessary to represent any sum of squares over $\mathbb{R}[x_1, \ldots, x_n]$, $n > 1$, making a fixed bound (for a given n) as in the rational function case impossible.

Our proof of Theorem 1.4 is also constructive.

Example 1.7. Consider the polynomial
$$f = 3 - 12y - 6x^3 + 18y^2 + 3x^6 + 12x^3y - 6xy^3 + 6x^2y^4.$$
This polynomial is a sum of squares over $\mathbb{R}[x, y]$. To see this, let $\alpha, \beta, \gamma \in \mathbb{R}$ be the roots of the polynomial $u(x) = x^3 - 3x + 1$. Then, a computation reveals that
$$f = (x^3 + \alpha^2 y + \beta xy^2 - 1)^2 + (x^3 + \beta^2 y + \gamma xy^2 - 1)^2 + (x^3 + \gamma^2 y + \alpha xy^2 - 1)^2.$$
Using our techniques, we can construct from this representation one over $\mathbb{Q}$:
$$\left(x^3 + xy^2 + 3y/2 - 1\right)^2 + \left(x^3 + 2y - 1\right)^2 + \left(x^3 - xy^2 + 5y/2 - 1\right)^2$$
$$+ \left(2y - xy^2\right)^2 + 3y^2/2 + 3x^2y^4.$$
This example will be revisited many times in the sequel to illustrate our proof. □

We believe that in Theorem 1.4 the field K may be replaced by any real algebraic extension of the rationals (thus giving a positive answer to Sturmfels' question); however, our techniques do not readily generalize to this situation. We shall discuss the obstructions throughout our presentation.

The organization of this paper is as follows. In Section 2, we set up our notation and state a weaker (but still sufficient) version of our main theorem. Section 3 describes a matrix factorization for Vandermonde matrices. This construction is applied in the subsequent section to reduce the problem to the case $K = \mathbb{Q}(\sqrt{l_1}, \ldots, \sqrt{l_r})$ for positive integers l_k. Finally, the proof of Theorem 1.4 is completed in Section 5. For simplicity of exposition, we shall focus on the polynomial version of our main theorem although it is an easy matter to translate the techniques to prove the more general Theorem 1.5.

2. Preliminaries

An equivalent definition of a totally real number field K is that it is a field generated by a root of an irreducible polynomial $u(x) \in \mathbb{Q}[x]$, all of whose zeroes are real. For instance, the field $K = \mathbb{Q}(\alpha, \beta, \gamma) = \mathbb{Q}(\alpha)$ arising in Example 1.7 is a totally real (Galois) extension of $\mathbb{Q}$ in this sense. A splitting field of $u(x)$ (a Galois closure of K) is also totally real, so we lose no generality in assuming that K is a totally real Galois extension of $\mathbb{Q}$. We will therefore assume from now on that $K = \mathbb{Q}(\theta)$ is Galois and that θ is a real algebraic number, all of whose conjugates are also real. We set $r = [K : \mathbb{Q}]$ and let G be the Galois group $\mathrm{Gal}(K/\mathbb{Q})$. For the rest of our discussion, we will fix $K = \mathbb{Q}(\theta)$ with these parameters.

We begin by stating a weaker formulation of Theorem 1.4. For the purposes of this work, a *rational sum of squares* is a linear combination of squares with positive rational coefficients.

Theorem 2.1. *Let K be a totally real number field that is Galois over $\mathbb{Q}$. Then for any $p \in K[x_1, \ldots, x_n]$, the polynomial*

$$f = \sum_{\sigma \in G} (\sigma p)^2$$

can be written as a rational sum of $2^{[K:\mathbb{Q}]+1} \binom{[K:\mathbb{Q}]+1}{2}$ squares in $\mathbb{Q}[x_1, \ldots, x_n]$.

Remark 2.2. Elements of the form $\sum_{\sigma \in G} (\sigma p)^2$ are also sometimes called *trace forms* for the field extension K [14, p. 217].

It is elementary, but important that this result implies Theorem 1.4.

Proof of Theorem 1.4. Let $f = \sum_{i=1}^{m} p_i^2 \in \mathbb{Q}[x_1, \ldots, x_n]$ be a sum of squares with each $p_i \in K[x_1, \ldots, x_n]$. Summing both sides of this equation over all actions of $G = \mathrm{Gal}(K/\mathbb{Q})$, we have

$$f = \frac{1}{|G|} \sum_{i=1}^{m} \sum_{\sigma \in G} (\sigma p_i)^2.$$

The conclusions of Theorem 1.4 now follow immediately from Theorem 2.1 and Lagrange's four square theorem (every positive rational number is the sum of at most four squares). □

Remark 2.3. This averaging argument can also be found in the papers [3, 6].

We will focus our remaining efforts, therefore, on proving Theorem 2.1.

3. Vandermonde factorizations

To prepare for the proof of Theorem 2.1, we describe a useful matrix factorization. It is inspired by Ilyusheckin's recent proof [12] that the discriminant of a symmetric matrix of indeterminates is a sum of squares, although it would not surprise us if the factorization was known much earlier.

Let $A = A^T$ be an $r \times r$ symmetric matrix over a field F of characteristic not equal to 2, and let $y_1, \ldots, y_r$ be the eigenvalues of A in an algebraic closure of F. Also, let V_r be the Vandermonde matrix

$$V_r = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ y_1 & y_2 & \cdots & y_r \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{r-1} & y_2^{r-1} & \cdots & y_r^{r-1} \end{bmatrix}.$$

The matrix $B = V_r V_r^T$ has as its (i,j)th entry the $(i+j-2)$th Newton power sum of the eigenvalues of A:

$$\sum_{k=1}^{r} y_k^{i+j-2}.$$

Since the trace of A^m is also the mth Newton power sum of the y_k, it follows that we may write $B = [\mathrm{tr}(A^{i+j-2})]_{i,j=1}^{r} \in F^{r \times r}$. We next give another factorization of B in the form CC^T.

Let E_{ij} be the $r \times r$ matrix with a 1 in the (i,j) entry and 0's elsewhere. A basis for $r \times r$ symmetric matrices is then given by the following $\binom{r+1}{2}$ matrices:

$$\{E_{ii} : i = 1, \ldots, r\} \ \cup \ \{(E_{ij} + E_{ji})/\sqrt{2} : 1 \leq i < j \leq r\}.$$

For example, the "generic" symmetric 2×2 matrix

(3.1) $$A = \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix},$$

with entries in the field $F = \mathbb{Q}(x_{11}, x_{12}, x_{22})$, is represented in this basis as

$$x_{11} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + x_{22} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \sqrt{2} \cdot x_{12} \begin{bmatrix} 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 \end{bmatrix}.$$

This basis is useful since the inner product of two symmetric matrices P and Q with respect to this basis is simply $\mathrm{tr}(PQ)$, as one can easily check.

Express the powers A^m in terms of this basis and place the vectors of the coefficients as rows of a matrix C. The entries of the $r \times \binom{r+1}{2}$ matrix C will be in $F[\sqrt{2}]$. Our construction proves the formal identity

(3.2) $$V_r V_r^T = [\mathrm{tr}(A^{i+j-2})]_{i,j=1}^{r} = CC^T.$$

Example 3.1. With A given by (3.1), the factorization reads:

$$\begin{bmatrix} 1 & 1 \\ y_1 & y_2 \end{bmatrix} \begin{bmatrix} 1 & y_1 \\ 1 & y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ x_{11} & x_{22} & \sqrt{2} \cdot x_{12} \end{bmatrix} \begin{bmatrix} 1 & x_{11} \\ 1 & x_{22} \\ 0 & \sqrt{2} \cdot x_{12} \end{bmatrix}.$$

Algebraically, this equation reflects the fact that for a 2×2 symmetric matrix A,

$$\mathrm{tr}(A) = x_{11} + x_{22},$$
$$\mathrm{tr}(A^2) = \mathrm{tr}(A)^2 - 2\det(A) = x_{11}^2 + x_{22}^2 + 2x_{12}^2.$$

$\square$

In the next section, we will use the matrix factorization (3.2) to replace a Gram matrix over $\mathbb{Q}(y_1, \ldots, y_r)$ with one over a much smaller field.

4. Symmetric matrices with prescribed characteristic polynomial

Let $K = \mathbb{Q}(\theta)$ be totally real and Galois, and set $\sigma_1, \ldots, \sigma_r$ to be the elements of $\mathrm{Gal}(K/\mathbb{Q})$. Given $p \in K[x_1, \ldots, x_n]$, we may express it in the form

$$p = \sum_{i=0}^{r-1} q_i \theta^i,$$

for elements $q_i \in \mathbb{Q}[x_1, \ldots, x_n]$. With this parameterization, the sum

$$\sum_{\sigma \in G} (\sigma p)^2 = \sum_{j=1}^{r} \left(\sum_{i=0}^{r-1} q_i (\sigma_j \theta)^i \right)^2$$

appearing in the statement of Theorem 2.1 may be written succinctly as
(4.1)
$$\begin{bmatrix} q_0 \\ \vdots \\ q_{r-1} \end{bmatrix}^T \begin{bmatrix} 1 & \cdots & 1 \\ \sigma_1\theta & \cdots & \sigma_r\theta \\ \vdots & \ddots & \vdots \\ (\sigma_1\theta)^{r-1} & \cdots & (\sigma_r\theta)^{r-1} \end{bmatrix} \begin{bmatrix} 1 & \sigma_1\theta & \cdots & (\sigma_1\theta)^{r-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \sigma_r\theta & \cdots & (\sigma_r\theta)^{r-1} \end{bmatrix} \begin{bmatrix} q_0 \\ \vdots \\ q_{r-1} \end{bmatrix}.$$

Let V_r be the Vandermonde matrix appearing in equation (4.1). We would like to construct a factorization as in (3.2) to replace the elements of K with numbers from $\mathbb{Q}(\sqrt{2})$. To apply the techniques of Section 3, however, we must find an $r \times r$ *symmetric* matrix A whose eigenvalues are $\sigma_1\theta, \ldots, \sigma_r\theta$ (the roots of the minimal polynomial for θ over $\mathbb{Q}$). A necessary condition is that these numbers are all real, but we would like a converse. Unfortunately, a converse with matrices over $\mathbb{Q}$ is impossible. For the interested reader, we include a proof of this basic fact.

Proposition 4.1. *There is no rational, symmetric matrix with characteristic polynomial $u(x) = x^2 - 3$.*

Proof. We argue by way of contradiction. Let

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \quad a, b, c \in \mathbb{Q},$$

and suppose that

$$u(x) = \det(xI - A) = x^2 - (a + c)x + (ac - b^2).$$

It follows that there are rational numbers a and b such that $a^2 + b^2 = 3$. Multiplying by a common denominator, one finds that there must be integer solutions u, v, w to the diophantine equation

(4.2) $u^2 + v^2 = 3w^2.$

Recall from elementary number theory that a number n is the sum of two integral squares if and only if every prime $p \equiv 3 \pmod 4$ that appears in the prime factorization of n appears to an even power. This contradiction finishes the proof. □

If we allow A to contain square roots of rational numbers, however, then there is always such a symmetric A. This is the content of a result of Fiedler [5]. We include his proof for completeness.

Theorem 4.2 (Fiedler)**.** *Let $u(x) \in \mathbb{C}[x]$ be monic of degree r and let $b_1, \ldots, b_r$ be distinct complex numbers such that $u(b_k) \neq 0$ for each k. Set $v(x) = \prod_{k=1}^{r}(x - b_k)$ and choose any complex numbers $d_1, \ldots, d_r$ and δ that satisfy*

$$\delta v'(b_k)d_k^2 - u(b_k) = 0, \quad k = 1, \ldots, r.$$

Let $d = [d_1, \ldots, d_r]^T$ and $B = \mathrm{diag}(b_1, \ldots, b_r)$. Then the symmetric matrix

$$A = B - \delta dd^T$$

has characteristic polynomial equal to $u(x)$.

Proof. Applying the Sherman-Morrison formula [7, p. 50] for the determinant of a rank 1 perturbation of a matrix, we have

$$\det(xI - A) = \det(xI - B) + \delta \det(xI - B)d^T(xI - B)^{-1}d$$

(4.3)
$$= \prod_{k=1}^{r}(x - b_k) + \delta \sum_{k=1}^{r} d_k^2 \prod_{i=1, i \neq k}^{r}(x - b_i).$$

Since the monic polynomial $\det(xI - A)$ and $u(x)$ agree for $x = b_1, \ldots, b_r$, it follows that they are equal. □

Remark 4.3. There are simpler, tridiagonal matrices which can replace the matrix A (see [28]); however, square roots are still necessary to construct them.

The following corollary allows us to form a real symmetric matrix with characteristic polynomial equal to the minimal polynomial for θ over $\mathbb{Q}$.

Corollary 4.4. *If $u(x) \in \mathbb{Q}[x]$ is monic of degree r and has r distinct real roots, then there are positive rational numbers $l_1, \ldots, l_r$ and a symmetric matrix A with entries in $\mathbb{Q}(\sqrt{l_1}, \ldots, \sqrt{l_r})$ such that the eigenvalues of A are the roots of $u(x)$.*

Proof. Let $b_1, \ldots, b_{r-1}$ be rational numbers such that exactly one b_i is (strictly) between consecutive roots of $u(x)$, and let b_r be a rational number either smaller than the least root of $u(x)$ or greater than the largest root of $u(x)$. Also, set $\delta \in \{-1, 1\}$ such that $l_k = \delta u(b_k)/v'(b_k)$ is positive for each k. The corollary now follows from Theorem 4.2 by setting $d_k = \sqrt{l_k}$ for each k. □

Example 4.5. Consider the polynomial $u(x) = x^3 - 3x + 1$ from Example 1.7. Choosing $(b_1, b_2, b_3) = (0, 1, 2)$ and $\delta = 1$, we have $d = [\sqrt{2}/2, 1, \sqrt{6}/2]^T$ and

$$A = \begin{bmatrix} -1/2 & -\sqrt{2}/2 & -\sqrt{3}/2 \\ -\sqrt{2}/2 & 0 & -\sqrt{6}/2 \\ -\sqrt{3}/2 & -\sqrt{6}/2 & 1/2 \end{bmatrix}.$$

One can easily verify that the characteristic polynomial of A is $u(x)$. □

Combining Corollary 4.4 and the construction found in Section 3, we have proved the following theorem.

Theorem 4.6. *Let K be a totally real Galois extension of $\mathbb{Q}$ and set $r = [K : \mathbb{Q}]$. Then for any $p \in K[x_1, \ldots, x_n]$, there are positive integers $l_1, \ldots, l_r$ such that*

$$\sum_{\sigma \in G}(\sigma p)^2 = q^T CC^T q,$$

in which q is a vector of polynomials in $\mathbb{Q}[x_1, \ldots, x_n]$ and C is an $r \times \binom{r+1}{2}$ matrix with entries in $F = \mathbb{Q}(\sqrt{l_1}, \ldots, \sqrt{l_r}, \sqrt{2})$.

To illustrate the computations performed in the proof of Theorem 4.6, we present the following.

Example 4.7. We continue with Example 4.5. Let $\alpha \in \mathbb{R}$ be the root of $u(x)$ with $\alpha \in (1, 2)$. Then, setting $\beta = 2 - \alpha - \alpha^2$ and $\gamma = \alpha^2 - 2$, we have $u(x) = (x - \alpha)(x - \beta)(x - \gamma)$. The Galois group of $K = \mathbb{Q}(\alpha)$ is cyclic and is generated by the element $\sigma \in G$ such that $\sigma(\alpha) = \beta$. If we let $\mathbf{v} = [x^3 + 2xy^2 - 1, -xy^2, y - xy^2]^T$, then the factorization obtained by Theorem 4.6 is given by

$$
f = \mathbf{v}^T
\begin{bmatrix}
1 & -1/2 & 3/2 \\
1 & 0 & 2 \\
1 & 1/2 & 5/2 \\
0 & -1 & 2 \\
0 & -\sqrt{6}/2 & \sqrt{6}/2 \\
0 & -\sqrt{3} & 0
\end{bmatrix}^T
\begin{bmatrix}
1 & -1/2 & 3/2 \\
1 & 0 & 2 \\
1 & 1/2 & 5/2 \\
0 & -1 & 2 \\
0 & -\sqrt{6}/2 & \sqrt{6}/2 \\
0 & -\sqrt{3} & 0
\end{bmatrix}
\mathbf{v}.
$$

One can check that this factorization already produces the rational sum of squares representation we encountered in Example 1.7. □

We note that when K is an arbitrary number field, Galois over $\mathbb{Q}$, our approach still produces a result similar in spirit to Theorem 4.6. The only difference is that we must allow negative integers l_k in the statement.

Theorem 4.8. *Let K be a finite Galois extension of $\mathbb{Q}$ and set $r = [K : \mathbb{Q}]$. Then for any $p \in K[x_1, \ldots, x_n]$, there are integers $l_1, \ldots, l_r$ such that*

$$
\sum_{\sigma \in G} (\sigma p)^2 = q^T C C^T q,
$$

in which q is a vector of polynomials in $\mathbb{Q}[x_1, \ldots, x_n]$ and C is an $r \times \binom{r+1}{2}$ matrix with entries in $F = \mathbb{Q}(\sqrt{l_1}, \ldots, \sqrt{l_r}, \sqrt{2})$.

The following corollary is the closest we come to answering Sturmfels' question in the general case. It follows from applying Theorem 4.8 in the same way that Theorem 4.6 will be used below to prove Theorem 5.1.

Corollary 4.9. *Let K be a finite extension of $\mathbb{Q}$. If $f \in \mathbb{Q}[x_1, \ldots, x_n]$ is a sum of squares over $K[x_1, \ldots, x_n]$, then it is a difference of two sums of squares over $\mathbb{Q}[x_1, \ldots, x_n]$.*

Example 4.10. Consider the degree 2 field extension $K = \mathbb{Q}(i\sqrt{2})$, which is the splitting field of $u(x) = x^2 + 2$. One can check that setting $(b_1, b_2) = (0, 1)$, $\delta = -1$, and $d = [\sqrt{2}, i\sqrt{3}]^T$ in Theorem 4.2 produces the symmetric matrix

$$
A = \begin{bmatrix} 2 & i\sqrt{6} \\ i\sqrt{6} & -2 \end{bmatrix}.
$$

It follows that the 2×2 Vandermonde matrix V_2 as in (4.1) satisfies $V_2 V_2^T = C C^T$, in which

$$
C = \begin{bmatrix} 1 & 1 & 0 \\ 2 & -2 & 2i\sqrt{3} \end{bmatrix}.
$$

This calculation expresses the polynomial $f = (x + i\sqrt{2}y)^2 + (x - i\sqrt{2}y)^2$ as the difference

$$
f = (x + 2y)^2 + (x - 2y)^2 - 12y^2.
$$

□

5. Proof of Theorem 2.1

In this section, we complete the proof of our main theorem. The results so far show that if f is a sum of m squares in $K[x_1, \dots, x_n]$ for a totally real field K, Galois over $\mathbb{Q}$, then f is a sum of $m \cdot \binom{[K:\mathbb{Q}]+1}{2}$ squares in $L[x_1, \dots, x_n]$, where $L = \mathbb{Q}(\sqrt{l_1}, \dots, \sqrt{l_r}, \sqrt{2})$ for some positive integers l_k. The proof of Theorem 2.1 is thus complete if we can show the following.

Theorem 5.1. *Let $l_1, \dots, l_{r+1}$ be positive integers and set $L = \mathbb{Q}(\sqrt{l_1}, \dots, \sqrt{l_{r+1}})$. If $f \in \mathbb{Q}[x_1, \dots, x_n]$ is a sum of s squares in $L[x_1, \dots, x_n]$, then f is a rational sum of at most $s \cdot 2^{r+1}$ squares in $\mathbb{Q}[x_1, \dots, x_n]$.*

Proof. Let l be a positive integer and let $L = F(\sqrt{l})$ be a quadratic extension of a field F of characteristic 0. We shall prove: If $f \in F[x_1, \dots, x_n]$ is a rational sum of s squares in $L[x_1, \dots, x_n]$, then f is a rational sum of at most $2s$ squares in $F[x_1, \dots, x_n]$. The theorem then follows by repeated applications of this fact.

If $L = F$, then there is nothing to prove. Otherwise, let $\sigma \in \mathrm{Gal}(L/F)$ be such that $\sigma(\sqrt{l}) = -\sqrt{l}$, and let $f \in F[x_1, \dots, x_n]$ be a sum of s squares in $L[x_1, \dots, x_n]$:

$$f = \sum_{i=1}^{s} p_i^2 = \frac{1}{2} \sum_{i=1}^{s} \left(p_i^2 + (\sigma p_i)^2 \right).$$

It therefore suffices to prove that for fixed $p \in L[x_1, \dots, x_n]$, the element $p^2 + (\sigma p)^2$ is a rational sum of 2 squares. Finally, writing $p = a + b\sqrt{l}$ for $a, b \in F[x_1, \dots, x_n]$, we have that

$$(a + b\sqrt{l})^2 + (a - b\sqrt{l})^2 = 2a^2 + 2lb^2.$$

This completes the proof. $\qquad\square$

References

[1] S. Burgdorf, *Sums of Hermitian Squares as an Approach to the BMV Conjecture*, preprint.
[2] M. D. Choi, Z. D. Dai, T. Y. Lam, B. Reznick, *The Pythagoras number of some affine algebras and local algebras.* J. Reine Angew. Math. **336** (1982), 45–82. MR671321 (84f:12012)
[3] M.D. Choi, T. Y. Lam, B. Reznick, *Even symmetric sextics*, Math. Z. **195** (1987), no. 4, 559–580. MR900345 (88j:11019)
[4] R. Elman, T. Y. Lam, *Quadratic forms under algebraic extensions*, Math. Ann. **219** (1976), 21–42. MR0401649 (53:5476)
[5] M. Fiedler, *Expressing a polynomial as the characteristic polynomial of a symmetric matrix*, Lin. Alg. Appl. **141** (1990), 265–270. MR1076118 (92b:15016)
[6] K. Gatermann, P. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, J. Pure Applied Algebra **192** (2004), 95–128. MR2067190 (2005d:68155)
[7] G. H. Golub, C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 1996. MR1417720 (97g:65006)
[8] D. Hägele, *Proof of the cases $p \le 7$ of the Lieb-Seiringer formulation of the Bessis-Moussa-Villani conjecture*, J. Stat. Phys. **127** (2007), 1167–1171. MR2331034 (2008c:82007)
[9] C. Hillar, *Advances on the Bessis-Moussa-Villani Trace Conjecture*, Lin. Alg. Appl. **426** (2007), 130–142. MR2344564
[10] C. Hillar, J. Nie, *An elementary and constructive solution to Hilbert's 17th problem for matrices*, Proc. Amer. Math. Soc. **136** (2008), 73–76. MR2350390

[11] R. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985. MR832183 (87e:15001)

[12] N. V. Ilyusheckin, *Some identities for elements of a symmetric matrix*, Journal of Mathematical Sciences **129** (2005), 3994–4008. MR2037535 (2005a:15014)

[13] I. Klep, M. Schweighofer, *Sums of hermitian squares and the BMV conjecture*, J. Stat. Phys., to appear.

[14] T. Y. Lam, *Introduction to Quadratic Forms over Fields*, American Mathematical Society, 2004. MR2104929 (2005h:11075)

[15] E. Landau, *Uber die Darstellung definiter Funktionen durch Quadrate*, Math. Ann. **62** (1906), 272–285. MR1511376

[16] P. Landweber, E. Speer, *On D. Hägele's approach to the Bessis-Moussa-Villani conjecture*, preprint.

[17] J. Nie, K. Ranestad, B. Sturmfels, *The algebraic degree of semidefinite programming*, Mathematical Programming, to appear.

[18] A. Papachristodoulou, P. A. Parrilo, S. Prajna, *Introducing SOSTOOLS: A General Purpose Sum of Squares Programming Solver*, Proceedings of the IEEE Conference on Decision and Control (CDC), Las Vegas, NV, 2002.

[19] A. Papachristodoulou, P. A. Parrilo, S. Prajna, *New Developments in Sum of Squares Optimization and SOSTOOLS*, Proceedings of the American Control Conference (ACC), Boston, MA, 2004.

[20] P. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*, Math. Program., Ser. B **96** (2003), 293–320. MR1993050 (2004g:90075)

[21] P. Parrilo, *Exploiting algebraic structure in sum of squares programs*, Positive Polynomials in Control, Lecture Notes in Control and Information Sciences, Vol. 312, pp. 181–194, Springer, 2005. MR2123524 (2005i:93038)

[22] P. Parrilo, H. Peyrl, *A Macaulay 2 package for computing sum of squares decompositions of polynomials with rational coefficients*, SNC '07: Proceedings of the 2007 International Workshop on Symbolic-Numeric Computation, pp. 207–208, ACM, 2007.

[23] P. Parrilo, B. Sturmfels, *Minimizing polynomial functions*, Algorithmic and Quantitative Real Algebraic Geometry, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 60, pp. 83–99, Amer. Math. Soc., Providence, RI, 2003. MR1995016 (2004e:13038)

[24] Y. Pourchet, *Sur la représentation en somme de carrés des polynômes à une indéterminée sur un corps de nombres algébriques*, Acta Arith. **19** (1971), 89–104. MR0289442 (44:6632)

[25] V. Powers, T. Woermann, *An algorithm for sums of squares of real polynomials*, J. Pure and Appl. Alg. **127** (1998), 99–104. MR1609496 (99a:11047)

[26] A. Prestel, C. N. Delzell, *Positive Polynomials: From Hilbert's 17th Problem to Real Algebra*, Springer, 2001. MR1829790 (2002k:13044)

[27] B. Reznick, *Uniform denominators in Hilbert's seventeenth problem*, Math. Z. **220** (1995), 75–97. MR1347159 (96e:11056)

[28] G. Schmeisser, *A real symmetric tridiagonal matrix with a given characteristic polynomial*, Lin. Alg. Appl. **193** (1993), 11–18. MR1240269 (94i:15006)

[29] M. Schweighofer, *Algorithmische Beweise für Nichtnegativ- und Positivstellensätze*, Diplomarbeit an der Universität Passau, 1999.

DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, COLLEGE STATION, TEXAS 77843
E-mail address: chillar@math.tamu.edu