Southern Illinois University Carbondale
OpenSIUC

Articles and Preprints Department of Mathematics

3-2001

Norms of Sums of Squares

Robert W. Fitzgerald
Southern Illinois University Carbondale, rfitzg@math.siu.edu

Follow this and additional works at: http://opensiuc.lib.siu.edu/math_articles

Part of the Number Theory Commons

Published in *Linear Algebra and its Applications, 325, 1-6.*

Recommended Citation
Fitzgerald, Robert W. "Norms of Sums of Squares." (Mar 2001).

This Article is brought to you for free and open access by the Department of Mathematics at OpenSIUC. It has been accepted for inclusion in Articles and Preprints by an authorized administrator of OpenSIUC. For more information, please contact opensiuc@lib.siu.edu.

NORMS OF SUMS OF SQUARES

Robert W. Fitzgerald

Southern Illinois University

Abstract. For a finite separable extension K/F of fields of characteristic not 2, the norm of a sum of 2^n squares in K is a sum of 2^n squares in F. We find explicit identities.

Let F be a field of characteristic not 2 and let K/F be a finite separable extension. A special case of Scharlau's Norm Principle [4, I 8.6] says that if α is a sum of 2^n squares in K, for some n, then $N_{K/F}(\alpha)$ is a sum of 2^n squares in F. The goal here is to find explicit identities expressing $N_{K/F}(\alpha)$ as a sum of 2^n squares in F.

The motivation is a previous result [1, 1.1] for monic, separable $f(x) \in F[x]$: If $f(x)$ is the characteristic polynomial of a symmetric $k \times k$ matrix over F then the discriminant of $f(x)$ is a sum of 2^n squares, where $2^{n-1} < k \leq 2^n$. The converse is true for $n = 2$ and false for $n \geq 4$. To settle the open case $n = 3$ requires an explicit identity for the discriminant as a sum of four squares. Such an identity follows, by the proof of [1, 1.1], from the identity expressing the norm, from a cubic extension, of a sum of three squares in K as a sum of four squares in F. Although we obtain such an identity here, it is too complicated to immediately settle the $n = 3$ case. This application will be postponed.

The proofs depend on a matrix theoretic approach to the transfer of a quadratic form, which seems to be new. Then one imitates the proof of Scharlau's Norm Principle (and Witt's Cancellation Theorem), replacing the usual vector space arguments with explicit matrix computations.

We have presented formulas for several cases other than the one which motivated this work because there is a curious difference between the case of two squares and the case of four or more squares. This is best illustrated by an example. Suppose $[K : F] = 2$ with σ the non-trivial F-automorphism of K. Let $x_1, x_2 \in K$. Then:

$$N_{K/F}(x_1^2 + x_2^2) = (x_1^2 + x_2^2)(\sigma(x_1)^2 + \sigma(x_2)^2).$$

We have the standard identity:

$$(x_1^2 + x_2^2)(y_1^2 + y_2^2) = (x_1y_1 + x_2y_2)^2 + (x_1y_2 - x_2y_1)^2. \tag{0.1}$$

1991 *Mathematics Subject Classification.* 11E25, 15A18,12Y05.

Key words and phrases. Norm, sums of squares, symmetric matrix.

Typeset by $\mathcal{AMS}$-TEX

If we set $y_1 = \sigma(x_1)$ and $y_2 = \sigma(x_2)$ then the second term of (0.1) $p_2 := x_1y_2 - x_2y_1$ becomes $x_1\sigma(x_2) - x_2\sigma(x_1)$ which is not in F (i.e. $\sigma(p_2) = -p_2$). Thus we have not written $N_{K/F}(x_1^2 + x_2^2)$ as a sum of two squares in F. However, if instead we set $y_1 = \sigma(x_2)$ and $y_2 = \sigma(x_1)$ then:

$$\begin{aligned} p_1 &= x_1\sigma(x_2) + x_2\sigma(x_1) = tr_{K/F}(x_1\sigma(x_2)) \in F \\ p_2 &= x_1\sigma(x_1) - x_2\sigma(x_2) = N_{K/F}(x_1) - N_{K/F}(x_2) \in F, \end{aligned}$$

and we have the desired identity. We will show such a substitution is always possible for sums of two squares and any finite K/F.

The situation is different for sums of four squares. Let $x_i \in K$ for $1 \leq i \leq 4$. We have:

$$N_{K/F}(x_1^2 + x_2^2 + x_3^2 + x_4^2) = (x_1^2 + x_2^2 + x_3^2 + x_4^2)(\sigma(x_1)^2 + \sigma(x_2)^2 + \sigma(x_3)^2 + \sigma(x_4)^2),$$

and the Lagrange identity (see e.g. [2, p. 287]):

$$\begin{aligned} (x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) &= p_1^2 + p_2^2 + p_3^2 + p_4^2 \\ \text{where} \qquad p_1 &= x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4 \\ p_2 &= x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3 \\ p_3 &= x_1y_3 - x_3y_1 - x_2y_4 + x_4y_2 \\ p_4 &= x_1y_4 - x_4y_1 + x_2y_3 - x_3y_2. \end{aligned}$$

There are 24 possible substitutions, namely, for each permutation π on four letters set $y_i = \sigma(x_{\pi(i)})$. None of these substitutions yield all four p_i in F. The identity for the norm of a sum of four squares must involve division.

1. Sums of two squares.

We begin by finding an identity for $\prod_{i=1}^m (x_{i1}^2 + x_{i2}^2)$, where x_{i1}, x_{i2} are independent variables over F. Let X denote the set of x_{ij}'s. Let $A_k^m(X)$ denote the sum of all products $x_{1\,i_1} \cdots x_{m\,i_m}$ with exactly k of the i_j's equal to 1. We will frequently use:

$$A_{k+1}^{m+1}(X) = A_k^m(X)x_{m+1\,1} + A_{k+1}^m(X)x_{m+1\,2}. \tag{1.1}$$

Lemma 1.2. *For $m \geq 2$:*

$$\prod_{i=1}^m (x_{i1}^2 + x_{i2}^2) = P_1^m(X)^2 + P_2^m(X)^2,$$

where:

$$P_1^m(X) = \begin{cases} \sum_{i=0}^{2r} (-1)^{i+1} A_{2i}^m(X), & \text{if } m = 4r \\ \sum_{i=0}^{2r} (-1)^{i+1} A_{2i+1}^m(X), & \text{if } m = 4r+1 \\ \sum_{i=0}^{2r} (-1)^{i} A_{2i+1}^m(X), & \text{if } m = 4r+2 \\ \sum_{i=0}^{2r+1} (-1)^{i+1} A_{2i}^m(X), & \text{if } m = 4r+3. \end{cases}$$

and

$$P_2^m(X) = \begin{cases} \sum_{i=0}^{2r-1}(-1)^{i+1}A_{2i+1}^m(X), & \text{if } m = 4r \\ \sum_{i=0}^{2r}(-1)^{i+1}A_{2i}^m(X), & \text{if } m = 4r+1 \\ \sum_{i=0}^{2r+1}(-1)^{i+1}A_{2i}^m(X), & \text{if } m = 4r+2 \\ \sum_{i=0}^{2r+1}(-1)^{i}A_{2i+1}^m(X), & \text{if } m = 4r+3. \end{cases}$$

Proof. This is not quite a straightforward induction argument so we provide most of the details. For $m = 2$ we apply (0.1) to:

$$(x_{11}^2 + x_{12}^2)(x_{22}^2 + x_{21}^2),$$

Note that we have switched the order of x_{21} and x_{22} in the sum. This yields:

$$\begin{aligned} P_1^2(X) &= x_{11}x_{22} + x_{12}x_{21} = A_1^2(X) \\ P_2^2(X) &= x_{11}x_{21} - x_{12}x_{22} = A_2^2(X) - A_0^2(X), \end{aligned}$$

as desired.

When $m > 2$ we have four cases to consider. First suppose $m = 4r+3$. Then:

$$\prod_{i=1}^{m}(x_{i1}^2 + x_{i2}^2) = (P_1^{m-1}(X)^2 + P_2^{m-1}(X)^2)(x_{m1}^2 + x_{m2}^2),$$

Note that this time we did not switch the order of $x_{m\,1}$ and $x_{m\,2}$. With (0.1) this yields:

$$\begin{aligned} P_1^m(X) &= P_1^{m-1}(X)x_{m\,1} + P_2^{m-1}(X)x_{m\,2} \\ &= \sum_{i=0}^{2r}(-1)^i A_{2i+1}^{m-1}(X)x_{m\,1} + \sum_{i=0}^{2r+1}(-1)^{i+1}A_{2i}^{m-1}(X)x_{m\,2} \\ &= \sum_{i=0}^{2r}(-1)^i\left[A_{2i+1}^{m-1}(X)x_{m\,1} + A_{2i+2}^{m-1}(X)x_{m\,2}\right] - A_0^{m-1}(X)x_{m\,2} \\ &= \sum_{i=1}^{2r+1}(-1)^{i+1}A_{2i}^m(X) - A_0^m(X) \\ &= \sum_{i=0}^{2r+1}(-1)^{i+1}A_{2i}^m(X), \end{aligned}$$

where we used (1.1) in the fourth line. Similarly,

$$\begin{aligned} P_2^m(X) &= P_1^{m-1}(X)x_{m\,2} - P_2^{m-1}(X)x_{m\,1} \\ &= \sum_{i=0}^{2r}(-1)^i A_{2i+1}^{m-1}(X)x_{m\,2} - \sum_{i=0}^{2r+1}(-1)^{i+1}A_{2i}^{m-1}(X)x_{m\,1} \\ &= -A_{m-1}^{m-1}(X)x_{m\,1} + \sum_{i=0}^{2r}(-1)^i\left[A_{2i+1}^{m-1}(X)x_{m\,2} + A_{2i}^{m-1}(X)x_{m\,1}\right] \\ &= \sum_{i=0}^{2r+1} A_{2i+1}^m(X), \end{aligned}$$

Next assume $m = 4r$. Again we switch the order of the sum in the last term so that our product is $(P_1^{m-1}(X)^2 + P_2^{m-1}(X)^2)(x_{m\,2}^2 + x_{m\,1}^2)$. The induction proceeds as before. The argument for $m = 4r + 1$ is the same except we do not switch the order of the sum, using $(P_1^{m-1}(X)^2 + P_2^{m-1}(X)^2)(x_{m\,1}^2 + x_{m\,2}^2)$. Lastly, suppose $m = 4r + 2$. We switch and use:

$$(P_1^{m-1}(X)^2 + P_2^{m-1}(X)^2)(x_{m\,2}^2 + x_{m\,2}^2) = (-P_1^m(X))^2 + (-P_2^m(X))^2.$$

We have:

$$\begin{aligned} P_1^m(X) &= -P_1^{m-1}(X)x_{m\,2} - P_2^{m-1}(X)x_{m\,1} \\ &= -\sum_{i=0}^{2r}(-1)^{i+1}A_{2i+1}^{m-1}(X)x_{m\,2} - \sum_{i=0}^{2r}(-1)^{i+1}A_{2i}^{m-1}(X)x_{m\,1} \\ &= \sum_{i=0}(-1)^i\left[A_{2i+1}^{m-1}(X)x_{m\,2} + A_{2i}^{m-1}(X)x_{m\,1}\right] \\ &= \sum_{i=0}^{2r}(-1)^i A_{2i+1}^m(X), \end{aligned}$$

as desired. The derivation for $P_2^m(X) = -P_1^{m-1}(X)x_{m\,1} + P_2^{m-1}(X)x_{m\,2}$ is similar. □

Theorem 1.3. *Let K/F be a separable field extension of degree m. Let $\sigma_1, \sigma_2, \dots, \sigma_m$ be the F-monomorphisms of Kinto a fixed algebraic closure of F. Let $u, v \in K$. Write $P_i^m(u,v)$ for:*

$$P_i^m(\sigma_1(u), \dots, \sigma_m(u), \sigma_1(v), \dots, \sigma_m(v)),$$

with $i = 1, 2$. Then each $P_i^m(u,v) \in F$ and

$$N_{K/F}(u^2 + v^2) = P_1^m(u,v)^2 + P_2^m(u,v)^2.$$

Proof. We have by (1.2):

$$N_{K/F}(u^2 + v^2) = \prod_{i=1}^m(\sigma_i(u)^2 + \sigma_i(v)^2) = P_1^m(u,v)^2 + P_2^m(u,v)^2.$$

Thus it suffices to show each $P_i^m(u,v) \in F$. This is clear if $v = 0$ so suppose $v \neq 0$. Let $s_k(z_1, \dots, z_m)$ denote the sum of all products of k distinct z's. Then:

$$A_k^m(\sigma_1(u), \dots, \sigma_m(u), \sigma_1(v), \dots, \sigma_m(v)) = N_{K/F}(v) \cdot s_k(\sigma_1(u/v), \dots, \sigma_m(u/v)).$$

Write $s_k(u/v)$ for $s_k(\sigma_1(u/v), \dots, \sigma_m(u/v))$. For any i, $\sigma_i(s_k(u/v)) = s_k(u/v)$. Thus $s_k(u/v)$, and so $A_k^m(u,v)$ and $P_i^m(u,v)$, is in F. □

Corollary 1.4. *Let K/F be a finite Galois extension and let $R \subset K$ be a ring invariant under the action of the Galois group. If $u, v \in R$ then $N_{K/F}(u^2 + v^2) = a^2 + b^2$, with $a, b \in R \cap F$.*

Remark. If $\operatorname{char} F > m$ then one can express both $P_i^m(u, v)$, $i = 1, 2$, in terms of norms and traces of elements of K. Define:

$$t_k(z_1, \dots, z_m) = \sum_{i=1}^{m} z_i^k.$$

From the theory of symmetric polynomials we know there are $Q_k \in \mathbb{Z}[z_1, \dots, z_m]$ such that:

$$s_k = \frac{1}{k} Q_k(t_1, \dots t_m).$$

For instance, $s_2 = (t_1^2 - t_2)/2$ and $s_3 = (t_1^3 - 3t_1t_2 + 2t_3)/6$. Now:

$$\begin{aligned} t_k(\sigma_1(u/v), \dots, \sigma_m(u/v)) &= tr_{K/F}((u/v)^k) \\ s_k(u/v) &= \frac{1}{k} Q_k(tr_{K/F}(u/v), \dots, tr_{K/F}((u/v)^m), \end{aligned}$$

which expresses s_k, hence A_k^m and P_i^m, in terms of traces and norms.

2. The transfer of a matrix.

The transfers of isometric quadratic foms are themselves isometric. We present a new matrix-theoretic proof of this basic fact.

We recall the transfer map. Let K/F be a finite field extension and let $t : K \to F$ be a non-trivial linear functional. Let V be an m-dimensional space over K and let $B : V \times V \to K$ be symmetric and bilinear. The transfer of B is $t_*B : V \times V \to F$ defined by $(t_*B)(v, w) = t(B(v, w))$, where we now view V as an F-space.

For the matrix version of the transfer, first fix a basis $\mathcal{D} = \{\gamma_1, \dots, \gamma_n\}$ of K over F. We denote the standard basis of K^m by $\mathcal{E}$. Let Q be a symmetric $m \times m$ matrix over K. Then the transfer of Q, with respect to $\mathcal{D}$, is $t_*Q = (t(w_i^T Q w_j))$, where the w_i range through

$$\mathcal{E}\mathcal{D} = \{\gamma_1 e_1, \gamma_2 e_1, \dots, \gamma_n e_1, \gamma_1 e_2, \dots, \gamma_n e_m\}.$$

If Q_1 and Q_2 are conguent symmetric matrices over K then t_*Q_1 and t_*Q_2 are congruent. This is easily checked using the transfer of spaces as Q_1 and Q_2 are the matrices of the same bilinear space with respect to different bases. However, we need an explicit matrix proof.

We fix the field extension K/F, linear functional t, and basis $\mathcal{D}$ of K/F throughout this section.

Definition. For $x \in K$ define $\alpha(x)$ to be the matrix, with respect to $\mathcal{D}$, representing multiplication by x. Explicitly, $\alpha(x)$ is the $n \times n$ matrix whose i-th column is the $\mathcal{D}$-coordinates of $\gamma_i x$. For a matrix $M = (m_{ij})$ over K we define $\alpha(M)$ to be the block matrix $(\alpha(m_{ij}))$,

Proposition 2.1. *Let Q_1 and Q_2 be symmetric $m \times m$ matrices over K and suppose $P^T Q_1 P = Q_2$, for some invertible matrix P. Then $\alpha(P)^T(t_* Q_1)\alpha(P) = t_* Q_2$.*

Proof. We work with the bilinear space (V, B) where $V = K^m$ and $B(v, w) = v^T Q_1 w$. The matrix of B with respect to $\mathcal{E}$ is Q_1. Let $\mathcal{C}$ be the set of columns of P. Then the matrix of B with respect to $\mathcal{C}$ is Q_2. The matrix of $t_* B$ with respect to $\mathcal{ED}$ is $t_* Q_1$ and the matrix of $t_* B$ with respect to $\mathcal{CD}$ is $t_* Q_2$. Thus we need the transition matrix R for $\mathcal{CD}$ to $\mathcal{ED}$.

Write $P = (p_{ij})$ and $p_{ij} = p_{ij1}\gamma_1 + \cdots p_{ijn}\gamma_n$. Set $\Gamma = (\gamma_1, \ldots \gamma_n)$. By definition, $\alpha(p_{ij})e_s \bullet \Gamma = \gamma_s p_{ij}$. Let c_i denote the i-th column of P. Then the $\mathcal{ED}$-coordinates of $\gamma_s c_i$ are:

$$(\alpha(p_{i1})e_s, \alpha(p_{i2})e_s, \ldots, \alpha(p_{im})e_s),$$

a vector with mn entries. Thus the block of $\mathcal{ED}$-coordinates of $\gamma_1 c_i, \ldots, \gamma_n c_i$ is:

$$(\alpha(c_{i1}), \alpha(c_{i2}), \ldots, \alpha(c_{im}))^T = \alpha(c_i).$$

Hence $R = \alpha(P)$ and $R^T Q_1 R = Q_2$. □

Proposition 2.2. *(1) The map $\alpha : K \to M_n(F)$ is an F-algebra homomorphism.*

(2)Suppose $P^T P = S \cdot I$, where $S \in K$ and I is the $m \times m$ identity matrix. Then $\alpha(P^T)\alpha(P) = \alpha(S) \cdot I$, where the latter denotes the block diagonal matrix with m copies of $\alpha(S)$ along the diagonal.

Proof. (1) is clear since $\alpha(x)$ represents multiplication by x. (2) follows immediately from (1). □

3. Norms of sums of three squares for quadratic extensions.

We begin with an example showing (1.4) fails for sums of four squares.

Example 3.1. Let $F = \mathbb{Q}(r, s, t)$ and $K = F(\sqrt{7})$. Let $R = \mathbb{Z}[r, s, t, \sqrt{7}]$, which is invariant under the action of $\mathrm{Gal}(K/F)$. Let:

$$u = 1 + s\sqrt{7} \qquad \text{and} \qquad v = r + t\sqrt{7}.$$

We show that $N_{K/F}(1 + u^2 + v^2)$ is **not** a sum of four squares in $R \cap F = \mathbb{Z}[r, s, t]$. This implies the remark made in the introduction, namely that there is no substitution $y_i = \sigma(x_{\pi(i)}), \pi \in S_4$, that makes each of the four terms of Lagrange's identity an element of F.

We sketch the proof. First, we have:

$$\begin{aligned} N_{K/F}(S) &= (2 + 7s^2 + r^2 + 7t^2)^2 - 28(s + rt)^2 \\ &= 4 + 4r^2 + 28t^2 + 49s^4 + 14r^2s^2 + 98s^2t^2 + r^4 - 14r^2t^2 + 49t^4 - 56rst. \end{aligned}$$

Suppose this can be written as the sum of four squares of the form:

$$(a_i r^2 + b_i rs + c_i s^2 + d_i rt + e_i st + f_i t^2 + g_i r + h_i s + k_i t + l_i)^2.$$

Set $A = (a_1, a_2, a_3, a_4) \in \mathbb{Z}^4$ and similarly for the other letters. Comparing coefficients yields equations such as $A \bullet A = 1, A \bullet B = 0$, etc, using the standard inner product. We can assume that $A = (1,0,0,0)$. The coefficient of r^2t^2 yields $2A \bullet F + D \bullet D = -14$. Then $f_1 \leq -7$. The t^4 term gives $F \bullet F = 49$ so that $f_1 = -7$ and hence $F = (-7,0,0,0)$ and $D = 0$.

There are now many cases to consider. In each case the goal is to determine C and L and then use the following facts:

(i) A positive integer of the form $4^a(8b-1)$ is not a sum of three squares in $\mathbb{Z}$.
(ii) There do not exist $v, w \in \mathbb{Z}^3$ such that $v \bullet v = 14, v \bullet w = 0$ and $w \bullet w = 49$.
(iii) There do not exist $v, w \in \mathbb{Z}^4$ such that $v = (\pm1,\pm1,\pm1,\pm1), v \bullet w = 0$ and $w \bullet w = 49$.

Statement (i) is standard ([2, p. 174]). For (ii) note that $(v+w) \bullet (v+w) = 63$, contradicting (i). For (iii) one must check each representation of 49 as a sum of four squares.

We will present an outline of the cases and the details of two typical ones. Begin by setting V equal to the span (over $\mathbb{Q}$) of A, B, C, K, L and setting W equal to the span of E, G, H. Then V and W are orthogonal.

In the first case, suppose $\dim V = 1$, so that V consists of multiples of A. Then $A \bullet B = A \bullet K = 0$ implies $B = K = 0$. Then the equation from the r^2s^2 term, $2A \bullet C = B \bullet B = 14$, gives $c_1 = 7$ and so, as $C \bullet C = 49$, we have $C = (7,0,0,0)$. Next, L is a multiple of A so $L = (l_1,0,0,0)$. $L \bullet L = 4$ yields $l_1 = \pm2$. If $L = (2,0,0,0)$ then the coefficient of t^2, namely $2F \bullet L + K \bullet K$ computes to -28, while it should be 28. Hence $L = (-2,0,0,0)$. Then looking at the coefficients of s^2 and r^2s give $H \bullet H = -2C \bullet L = 28$ and $A \bullet H = 0$. Thus $h_1 = 0$ and $h_2^2 + h_3^2 + h_4^2 = 28$, contradicting (i).

The subsequent cases are (2) $\dim V = 2$ and $B \neq 0$, (3) $\dim V = 2$ and $B = 0$, (4) $\dim W = 1$ and $H \neq 0$, (5) $\dim W = 1$ and $E = H = 0$, all of which end in similar contradictions. This leaves $\dim V = 3$, $\dim W = 1$, $H = 0$ and $E \neq 0$. Then $l_1 \neq 0$ else $K \bullet K = 28$ and $A \bullet K = 0$ which implies 28 is a sum of three squares, contrary to (i). Further, $l_1 \neq \pm2$ else $c_1 = 0$ which implies $B \bullet B = 14, B \bullet C = 0$ and $C \bullet C = 49$, contrary to (ii). Thus $L = (\pm1,\pm1,\pm1,\pm1)$. But $L \bullet C = 0$ and $C \bullet C = 49$, contradicting (iii). This completes the example.

Theorem 3.2. *Let $K = F(\sqrt{d})$ be a separable extension. Let $n = 2^k$, with $k \geq 1$ and let $u_1, u_2, \ldots, u_n \in K$. Then:*

$$(\sum_{i=1}^{n} u_{i\,2}^2) N_{K/F}(1 + \sum_{i=1}^{n} u_i^2) = (S_2)^2 + \sum_{i=1}^{n} (u_{i\,1}S_2 - u_{i\,2}S_1)^2$$

where $S = 1 + \sum u_i^2$, $u_i = u_{i\,1} + u_{i\,2}\sqrt{d}$ and $S = S_1 + S_2\sqrt{d}$.

Proof. Of course, the result can be verified by expanding out both sides but we will derive it. We imitate Scharlau's proof of his norm principle, keeping track of the matrices that arise. Let P be a $2n \times 2n$ matrix such that $P^TP = S \cdot I$. Let $t : K \to F$ be the linear functional with $t(1) = 0$ and $t(\sqrt{d}) = 1$. Let t_* be the induced transfer map of quadratic forms. Then :

$$t_*(\langle 1 \rangle) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad t_*(\langle S \rangle) = \begin{pmatrix} S_2 & S_1 \\ S_1 & dS_2 \end{pmatrix}.$$

Call the first matrix q_1 and the second q_2. Set Q_1 equal to the block diagonal matrix with $2n$ copies of q_1; Q_1 is the matrix of $t_*(2n \cdot \langle 1 \rangle)$. Similarly define Q_2, the matrix of $t_*(2n \cdot \langle S \rangle)$ with respect to the standard basis. Let R be the block matrix $\alpha(P)$. Then $R^T Q_1 R = Q_2$ by (2.1). Next diagonalize q_2 as follows:

$$m = \begin{pmatrix} 1 & -S_1 \\ 0 & S_2 \end{pmatrix}$$

$$m^T q_2 m = \begin{pmatrix} S_2 & 0 \\ 0 & -S_2 N_{K/F}(S) \end{pmatrix}$$

Set M equal to the block diagonal matrix with $2n$ copies of m.. Then:

$$(RM)^T Q_1 (RM) = S_2 \mathrm{diag}(1, -N_{K/F}(S), 1, -N_{K/F}(S), \dots).$$

Now $e_1^T Q_1 e_1 = 0$. Set $w = (RM)^{-1} e_1$. Then:

$$\sum_{j=1}^{2n} w_{2j-1}^2 = N_{K/F}(S)(\sum_{j=1}^{2n} w_{2j}^2) \tag{3.3}$$

which will give our identity. $(RM)^{-1}$ can be easily found as follows. Let $R^* = \alpha(P^T)$. Then $R^* R$ is the block diagonal matrix with $2n$ copies of $\alpha(S)$ by (2.2)(2). Let L be the block diagonal matrix with $2n$ copies of:

$$\alpha(S)^{-1} = \frac{1}{N_{K/F}(S)} \begin{pmatrix} S_1 & -dS_2 \\ -S_2 & S_1 \end{pmatrix}.$$

Then $R^{-1} = LR^*$. Write $P = (p_{ij}) = (p_{ij\,1} + p_{ij\,2}\sqrt{d})$. Then for $w = (RM)^{-1} e_1$ we have:

$$w_i = \begin{cases} S_1 p_{1\,j\,2} - S_2 p_{1\,j\,1}, & \text{if } i = 2j \\ N_{K/F}(S) p_{1\,j\,2}, & \text{if } i = 2j-1. \end{cases}$$

Plugging into (3.3) gives:

$$N_{K/F}(S)^2 \sum_{j=1}^{2n} p_{1\,j\,2}^2 = N_{K/F}(S) \sum_{j=1}^{2n} (S_1 p_{1\,j\,2} - S_1 p_{1\,j\,1})^2.$$

Now we can assume the first column of P is $(1, 0, \dots, 0, -u_1, -u_2 \dots, -u_n)$ by[2,XI,1.2]. This gives the desired result. $\square$

One can easily derive the identities for $N_{K/F}(1 + u_1^2 + \dots + u_n^2)$ from (3.2): multiply both sides by $\sum u_{i2}^2$ and apply the $n = 2^k$ identity. For instance, in the case of example (3.1) we obtain:

$$N_{K/F}(S) = \left(\frac{2s(s+rt)}{s^2+t^2}\right)^2 + \left(\frac{2t(s+rt)}{s^2+t^2}\right)^2 + \left(\frac{2(t-rs)(s+rt)}{s^2+t^2}\right)^2$$
$$+ \left(\frac{4rst + r^2t^2 - 7s^4 - r^2s^2 - 14s^2t^2 - 2t^2 - 7t^4}{s^2+t^2}\right)^2$$

One can also derive the identities for $N_{K/F}(u_1^2+\cdots+u_{2n}^2)$ by factoring $u_1^2+\cdots u_{2n}^2 = (u_1^2+\cdots u_n^2)(1+v_1^2+\cdots v_n^2)$, where:

$$\sum_{i=1}^{n} v_i^2 = (\sum_{i=n+1}^{2n} u_i^2)/(\sum_{i=1}^{n} u_i^2),$$

is found with the $n=2^k$ identity. Then take norms of both sides.

4. Norms of sums of three squares for cubic extensions.

We introduce some notation and assumptions. Let K/F be a separable extension of degree 3. We assume -1 is not a sum of two squares in F (otherwise it is trivial to write any element of F as a sum of three squares). Let $u, v \in K$ and set $S = 1+u^2+v^2$. We assume $S \notin F$. Then $K = F(S)$. Suppose $S^3 = aS^2+bS+c$, with $a,b,c \in F$. Note that $N_{K/F}(S) = c$. Write:

$$u = u_0+u_1S+u_2S^2 \qquad \text{and} \qquad v = v_0+v_1S+v_2S^2.$$

Theorem 4.1. *Use the above notation and assumptions. Suppose $\chi := (2c+bc-1)/c$ is not zero. Then:*

$$(w_2^2+w_5^2+w_8^2+w_{11}^2)N_{K/F}(S) = w_3^2+w_6^2+w_9^2+w_{12}^2,$$

where:

$$\begin{aligned}
w_2 &= -2\mu(u)u_2 - 2\mu(v)v_2\\
w_3 &= c - \frac{bc+1}{\chi} - 2\mu(u)\left(\gamma(u) - \frac{(bc+1)\varphi(u)}{c\chi}\right) - 2\mu(v)\left(\gamma(v) - \frac{(bc+1)\varphi(v)}{c\chi}\right)\\
w_5 &= 2\mu(u)v_2 - 2\mu(v)u_2\\
w_6 &= 2\mu(u)\left(\gamma(v) - \frac{(bc+1)\varphi(v)}{c\chi}\right) - 2\mu(v)\left(\gamma(u) - \frac{(bc+1)\varphi(u)}{c\chi}\right)\\
w_8 &= u_1 + au_2 + \frac{2cu_2}{\chi} - \frac{4\mu(u)}{\chi}(u_2\varphi(u)+v_2\varphi(v)) - \frac{4\mu(v)}{\chi}(u_2\varphi(v)-v_2\varphi(u))\\
w_9 &= \frac{bcu_2 - u_2 + 2cu_0}{2} + \frac{2c\gamma(u)}{\chi} - 2\mu(u)\left(\frac{bc+1}{2c} + \frac{2}{\chi}(\varphi(u)\gamma(u)+\varphi(v)\gamma(v))\right)\\
&\quad - \frac{4\mu(v)}{\chi}(\varphi(v)\gamma(u)-\varphi(u)\gamma(v))\\
w_{11} &= v_1 + av_2 + \frac{2cv_2}{\chi} - \frac{4\mu(u)}{\chi}(v_2\varphi(u)-u_2\varphi(v)) - \frac{4\mu(v)}{\chi}(u_2\varphi(u)+v_2\varphi(v))\\
w_{12} &= \frac{bcv_2 - v_2 + 2cv_0}{2} + \frac{2c\gamma(v)}{\chi} - \frac{4\mu(u)}{\chi}(\varphi(u)\gamma(v)-\varphi(v)\gamma(u))\\
&\quad - 2\mu(v)\left(\frac{bc+1}{2c} + \frac{2}{\chi}(\varphi(u)\gamma(u)+\varphi(v)\gamma(v))\right).
\end{aligned}$$

Here we use:

$$\begin{aligned}
\varphi(u) &= \frac{u_0 - bcu_0 + 2c^2u_1}{2c} \\
Q &= \chi(\left(1 + \frac{4\varphi(u)^2}{\chi^2} + \frac{4\varphi(v)^2}{\chi^2}\right)
\end{aligned}$$

$$\begin{aligned}
\psi(u) &= \frac{bcu_2 + u_2 + 2cu_0}{2} \\
\mu(u) &= \frac{1}{Q}\left(\psi(u) + \frac{2c\varphi(u)}{\chi}\right) \\
\gamma(u) &= \frac{-bcu_0 - u_0 + 2c^2u_1}{2c}
\end{aligned}$$

Proof. The reader who wishes to verify, with the aid of a computer, this identity needs to notice that the u_i and v_i are not independent. Comparing coefficients with respect to the basis $\{1, S, S^2\}$ of $S = 1 + u^2 + v^2$ yields:

$$\begin{aligned}
0 &= 1 + u_0^2 + 2cu_1u_2 + acu_2^2 + v_0^2 + 2cv_1v_2 + acv_2^2 \\
1 &= 2u_0u_1 + 2bu_1u_2 + (ab + c)u_2^2 + 2v_0v_1 + 2bv_1v_2 + (ab + c)v_2^2 \\
0 &= 2u_0u_2 + u_1^2 + 2au_1u_2 + (a^2 + b)u_2^2 + 2v_0v_2 + v_1^2 + 2av_1v_2 + (a^2 + b)v_2^2.
\end{aligned}$$

These relations,plus a Grobner basis program, are required for direct verification. We will, instead, derive the identity.

We begin, as in (3.1), by following the proof of Scharlau's Norm Principle. Set:

$$P = \begin{pmatrix} 1 & 0 & u & v \\ 0 & 1 & -v & u \\ -u & v & 1 & 0 \\ -v & -u & 0 & 1 \end{pmatrix}.$$

Then $P^TP = S \cdot I$. Let $t : K \to F$ be the linear functional with $t(1) = 1, t(S) = 0$ and $t(S^2) = 0$ and let t_* be the induced transfer map. The key values are :

$$q_1 := t_*\langle 1\rangle = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & c \\ 0 & c & ac \end{pmatrix} \qquad q_3 := t_*\langle S\rangle = \begin{pmatrix} 0 & 0 & c \\ 0 & c & ac \\ c & ac & bc + a^2c \end{pmatrix}.$$

Let $Q_1 = diag(q_1, q_1, q_1, q_1)$ and similarly for Q_3. Then if $R = \alpha(P)$ we have $R^TQ_1R = Q_3$, by (2.1).

We diagonalize q_3 as follows: set $q_2 = diag(1, c, -1)$. Then for:

$$m = \begin{pmatrix} (1 - bc)/(2c) & 0 & (1 + bc)/(2c) \\ -a & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$$

we have $m^T q_3 m = q_2$. Let M be the block diagonal matrix of four copies of m and similarly for Q_2. Then:

$$(RM)^T Q_1 (RM) = Q_2.$$

The object is to modify RM to N, as in the proof of Witt's Cancellation Theorem, so that $N^T Q_1 N = Q_2$ as before while the $\{1, 4, 7, 10\}$-submatrix of N is the identity. Then if $w = N^{-1} e_2$ we have $0 = Q_1(e_2) = Q_2(w)$ and $w_1 = w_4 = w_7 = w_{10} = 0$. Hence:

$$c(w_2^2 + w_5^2 + w_8^2 + w_{11}^2) = w_3^2 + w_6^2 + w_9^2 + w_{12}^2. \tag{4.2}$$

Since $c = N_{K/F}(S)$, (4.2) is our desired identity.

We will use hyperplane reflections to modify RM. Let B be the symmetric bilinear form associated to Q_1, that is, $B(x, y) = x^T Q_1 y$. If $Q_1(y) \neq 0$ then the hyperplane reflection at y is:

$$T_y : x \mapsto x - \frac{2B(x, y)}{Q_1(y)} y.$$

The matrix of T_y, with respect to the standard basis, is $I - (2/Q_1(y)) y y^T Q_1$.

Set $z_1 = RMe_1 - e_1$, where e_1 is the first vector in the standard basis. Then:

$$\begin{aligned} Q_1(z_1) &= Q_2(e_1) - 2B(e_1, RMe_1) + Q_1(e_1) \\ &= 2 - 2\frac{1 - bc}{2c} = \chi. \end{aligned}$$

We are assuming $\chi \neq 0$. (If $\chi = 0$ then z_1 must be replaced by $RMe_1 + e_1$ and the identity reworked.) Set $T_1 = T_{z_1}$. Then $(T_1 RM)^T Q_1 (T_1 RM) = Q_2$ and:

$$T_1 RM = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{pmatrix}.$$

We continue the process. Set $z_2 = T_1 T RMe_4 - e_4$. Then $z_2 = RMe_4 - e_4$ and $Q_1(z_2) = \chi \neq 0$. Set $T_2 = T_{z_2}$. Next, set $z_3 = T_2 T_1 RMe_7 - e_7$. Then:

$$\begin{aligned} z_3 &= RMe_7 - e_7 - \frac{2\varphi(v)}{\chi} z_2 + \frac{2\varphi(u)}{\chi} z_1 \\ Q_1(z_3) &= \chi \left(1 + \frac{4\varphi(u)^2}{\chi^2} + \frac{4\varphi(v)^2}{\chi^2}\right). \end{aligned}$$

Since we are assuming that -1 is not a sum of two squares in F, we have $Q_1(z_3) \neq 0$. We shorten $Q(z_3)$ to Q. So set $T_3 = T_{z_3}$.

Lastly, set $z_4 = T_3 T_2 T_1 RMe_{10} - e_{10}$. Then:

$$\begin{aligned} z_4 &= RMe_{10} - e_{10} + \frac{2\varphi(u)}{\chi} z_2 + \frac{2\varphi(v)}{\chi} z_1 \\ Q_1(z_4) &= Q. \end{aligned}$$

Set $T_4 = T_{z_4}$. Our modified isometry is $N = T_4T_3T_2T_1RM$, that is, $N^TQ_1N = Q_2$ and the $\{1,4,7,10\}$-submatrix of N is the identity matrix. Our identity arises from the coordinates of $w = N^{-1}e_2$.

We sketch the computation of w. First, every hyperplane reflection is an involution so that $T_i^{-1} = T_i$ for each i. A striaght-forward computation yields:

$$\begin{aligned} T_1T_2T_3T_4(e_2) &= e_2 - \frac{2c}{\chi}z_1 - 2\mu(u)z_3' - 2\mu(v)z_4' \\ \text{where} \quad z_3' &= RMe_7 - e_7 + \frac{2\varphi(v)}{\chi}z_2 - \frac{2\varphi(u)}{\chi}z_1 \\ z_4' &= RMe_{10} - e_{10} - \frac{2\varphi(u)}{\chi}z_2 - \frac{2\varphi(v)}{\chi}z_1. \end{aligned}$$

R^{-1} can be computed using (2.2)(2). Set $R^* = \alpha(P^T)$. Then $R^*R = \alpha(S)\cdot I$. Now:

$$\alpha(S) = \begin{pmatrix} 0 & 0 & c \\ 1 & 0 & b \\ 0 & 1 & a \end{pmatrix} \quad \text{and} \quad \alpha(S)^{-1} = \frac{1}{c}\begin{pmatrix} -b & c & 0 \\ -a & 0 & c \\ 1 & 0 & 0 \end{pmatrix}.$$

So $R^{-1} = diag(\alpha(S)^{-1}, \alpha(S)^{-1}, \alpha(S)^{-1}, \alpha(S)^{-1})R^*$. Lastly,

$$m^{-1} = \begin{pmatrix} c & 0 & (bc+1)/2 \\ 0 & 1 & a \\ c & 0 & (bc-1)/2 \end{pmatrix}.$$

The formulas for the w_i, $i \neq 1,4,7,10$, follow. □

The final remarks of Section 3 hold here as well. To get an identity for $N_{K/F}(1+u^2+v^2)$ multiply both sides of (4.1) by $w_2^2 + w_5^2 + w_8^2 + w_{11}^2$ and apply the four square identity. To get an identity for $N_{K/F}(u_1^2+u_2^2+u_3^2+u_4^2)$ factor:

$$\begin{aligned} u_1^2+u_2^2+u_3^2+u_4^2 &= (u_1^2+u_2^2)\left(1+\frac{u_3^2+u_4^2}{u_1^2+u_2^2}\right) \\ &= (u_1^2+u_2^2)\left(1+\left(\frac{u_1u_3+u_2u_4}{u_1^2+u_2^2}\right)^2+\left(\frac{u_1u_4-u_2u_3}{u_1^2+u_2^2}\right)^2\right). \end{aligned}$$

Apply the identity from (4.1) to the norm of the second term and the identity of (1.3) to the norm of the first term, then apply the two square identity twice.

Lastly, the argument of (4.1) can be used to find other identities. For the norm of $1+u^2+v^2$ from larger extensions one needs more hyperplane reflections fo modify RM. For sums of more squares one requires a different initial matrix P. For instance, for one plus a sum of four squares use the matrix implicit in [3, p. 2] and for one plus the sum of eight squares use [3, p.93].

References

1. R. Fitzgerald, *Characteristic polynomials of symmetric matrices*, Linear and Multilinear Alg. **36** (1994), 233–237.
2. T.Y. Lam, *The Algebraic Theory of Quadratic Forms*, Benjamin, Reading, Mass., 1973.
3. A.R. Rajwade, *Squares*, London Math. Soc. Lecture Note Series, vol. 171, Cambridge University, Cambridge, 1993.
4. W. Scharlau, *Quadratic and Hermitian Forms*, Grundlehren Math. Wissenschaften, vol. 270, Springer-Verlag, Berlin/Heidelberg/New York/Tokyo, 1985.

Carbondale, IL 62901
E-mail address: rfitzg@math.siu.edu