Representations as Sums of Squares

Henri Cohen,
Université Bordeaux I,
Institut de Mathématiques, U.M.R. 5251 du C.N.R.S,
351 Cours de la Libération,
33405 TALENCE Cedex, FRANCE

March 1, 2009

Abstract

This is a translation of a paper [5] I wrote in 1971, and may help for Parimala's course. Evidently completely outdated, but still may be useful. I changed some notation so as to be compatible with the course.

1 Introduction and Notation

Let K be a field, let $S_m(K) = \{x_1^2 + \cdots + x_m^2\,, x_i \in K\}$ be the set of sums of m squares of K, $S(K) = S_\infty(K) = \bigcup_m S_m(K)$ the set of elements of K which are sums of squares. Clearly if $m \le m'$ then $S_m(K) \subset S_{m'}(K)$, $S_m(K) + S_{m'}(K) = S_{m+m'}(K)$, and $S_m(K) \cdot S_{m'}(K) \subset S_{mm'}(K)$. With the usual convention that $\inf(\emptyset) = +\infty$, we will set:

$$s(K) = \inf\{m\,,\ -1 \in S_m(K)\} \quad \text{and} \quad p(K) = \inf\{m\,,\ S_m(K) = S(K)\}$$

($s(K)$ is called the *level* of K, and the letter s is from the German name "Stufe", the letter p is from Pythagoreas). Clearly $S_m(K) = S(K)$ if and only if $S_m(K) = S_{m+1}(K)$. Also, in characteristic 2 we have trivially $s(K) = p(K) = 1$, so if needed we may assume that $\mathrm{char}(K) \neq 2$.

2 Quadratic Forms

Let V be a K-vector space of finite dimension, and let q be a quadratic form defined on V. We say that $a \in K$ is *represented* by q if there exists $X \in V$ with $X \neq 0$ and such that $q(X) = a$ (note that the condition $X \neq 0$ is needed only when $a = 0$). We say that q is *isotropic* if it represents 0, and *universal* over K if any nonzero $a \in K^*$ is represented. Recall the link between these two notions:

Proposition 2.1 *Let q be a nondegenerate quadratic form over K with* $\mathrm{char}(K) \neq 2$*. Then if q is isotropic, it is universal over any K-algebra, and in particular over K.*

(Note: in the course, nondegenerate is called "regular".)

Let q and q' be two quadratic forms on V and V' respectively. We will say that q and q' are *equivalent* (in the course, "isometric"), and write $q \sim q'$, if there exists an isomorphism ϕ from V to V' such that $q'(\phi(X)) = q(X)$ for all $X \in V$, or equivalently $q'(Y) = q'(\phi^{-1}(Y))$. Clearly two equivalent forms represent the same elements. In addition, when $\mathrm{char}(K) \neq 2$ it is easy to show that any quadratic form is equivalent to a *diagonal* form

$$\langle a_1, a_2, \dots, a_n \rangle \ ,$$

which will be our notation for the form $a_1 x_1^2 + a_2 x_2^2 + \cdots + a_n x_n^2$ on $V = K^n$.

Since we study forms up to equivalence, there is little loss of generality in restricting to diagonal forms and to $V = K^n$.

We define the following binary operations $\oplus$ and $\otimes$ on diagonal forms as follows:

$$\langle a_1, \dots, a_m \rangle \oplus \langle b_1, \dots, b_n \rangle = \langle a_1, \dots, a_m, b_1, \dots, b_n \rangle \ ,$$
$$\langle a_1, \dots, a_m \rangle \otimes \langle b_1, \dots, b_n \rangle = \langle a_i b_j \ , \ 1 \leq i \leq m, \ 1 \leq j \leq n \rangle \ .$$

In addition, if $c \in K$ we define

$$c\langle a_1, \dots, a_m \rangle = \langle c a_1, \dots, c a_m \rangle \ .$$

It is not difficult to show that these operations are compatible with equivalence of quadratic forms.

3 Multiplicative Quadratic Forms

The definitions and theorems of this section are due to Pfister; see [10].

Definition 3.1 *Let q be a quadratic form over K. We will say that q is* multiplicative *if for all $d \in K^*$* represented by *q we have $q \sim dq$.*

Fundamental Examples:

• The form $\langle 1 \rangle = x^2$ is trivially multiplicative.

• The form $\langle 1, a \rangle = x^2 + ay^2$ is multiplicative since $(x_1^2 + a x_2^2)(x^2 + ay^2) = (x_1 x - a x_2 y)^2 + a(x_1 y + x_2 x)^2$, and since when $d = x_1^2 + a x_2^2 \neq 0$ the determinant of the transformation matrix is nonzero.

Before stating and proving the first of Pfister's theorems, we prove a lemma:

Lemma 3.2 *Assume that $a \neq 0$ is represented by the form $\langle b, c \rangle$. Then*

$$\langle b, c \rangle \sim \langle a, abc \rangle \ .$$

Proof. Write $a = bx_1^2 + cx_2^2$. Generalizing the above example, we have the identity

$$(bx_1^2 + cx_2^2)x^2 + (b^2cx_1^2 + bc^2x_2^2)y^2 = b(x_1x - cx_2y)^2 + c(bx_1y + x_2x)^2 \ ,$$

and since the determinant of the transformation matrix is $bx_1^2 + cx_2^2 = a \neq 0$, the lemma follows. $\square$

Pfister's first theorem is the following:

Theorem 3.3 *If q is multiplicative, then*

$$q \otimes \langle 1, a\rangle = q \oplus aq$$

is also multiplicative.

Proof. Let $d \neq 0$ be represented by the form $q \oplus aq$. We can thus write $d = b + ac$, where b and c are values of q. We can immediately take care of the cases where either one is 0: if $c = 0$ then $d = b$ and $d(q \oplus aq) = bq \oplus abq \sim q \oplus aq$ since q is multiplicative. Similarly, if $b = 0$ then $d = ac$ and $d(q \oplus aq) = acq \oplus a^2cq \sim aq \oplus q$ since q is multiplicative and trivially $a^2q \sim q$. We may therefore assume that b and c are nonzero, so are really represented by q. Thus, again since q is multiplicative we have

$$d(q \oplus aq) = dq \oplus adq \sim dq \oplus abdq \sim dq \oplus abcdq \ ,$$

so by the above lemma and the fact that $d = b \cdot 1^2 + ac \cdot 1^2$ is represented by $\langle b, ac\rangle$ we have

$$\begin{aligned} d(q \oplus aq) &\sim \langle d, abcd\rangle \otimes q \sim \langle b, ac\rangle \otimes q \\ &\sim bq \oplus acq \sim q \oplus aq \end{aligned}$$

again since q is multiplicative and b and c are represented by q, finishing the proof. $\square$

Corollary 3.4 *The so-called* Pfister forms *in 2^k variables*

$$\langle 1, a_1\rangle \otimes \langle 1, a_2\rangle \otimes \cdots \langle 1, a_k\rangle$$

are multiplicative.

Corollary 3.5 *For any $k \geq 0$ the set $S_{2^k}(K) \setminus \{0\}$ is a multiplicative subgroup of K^*, and in particular $S_{2^k}(K) \cdot S_{2^k}(K) = S_{2^k}(K)$.*

Proof. Immediate and left to the reader. $\square$

Note that before Pfister the above corollary was known (in fact with K replaced by a commutative *ring*) only for $k = 0$, 1, 2, and 3, since it corresponds to the multiplicativity of the norm in $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$ (the noncommutative Hamilton *quaternions*), and $\mathbb{O}$ (the nonassociative Cayley *octonions*). For $k \geq 4$ the result is *not* true for a general commutative ring, even for a domain of characteristic 0. In fact:

Exercise:

(1) Following in detail the above proof and using a computer algebra system, compute explicitly rational functions Z_i in the 16 variables X_i and Y_i, such that

$$\left(\sum_{1\leq i\leq 8} X_i^2\right)\left(\sum_{1\leq i\leq 8} Y_i^2\right) = \left(\sum_{1\leq i\leq 8} Z_i^2\right) .$$

In particular, note that they are *not* polynomials, so do not correspond to octonion multiplication.

(2) If you have the courage, do the same for sums of 16 squares.

4 Computation of $s(K)$ and $p(K)$ for Nonordered Fields

Recall that by a theorem of Artin–Schreier, a field cannot have an order compatible with the field laws if and only if $-1 \in S(K)$, in other words if and only if $s(K)$ is finite (by abuse of language we will say that K is nonordered). For such a field we have trivially $p(K) \geq s(K)$. On the other hand, $-1 \in S_m(K)$ means that the form $x_1^2 + \cdots + x_m^2 + x_{m+1}^2$ is isotropic in K, hence that it is universal if $\mathrm{char}(K) \neq 2$ by Proposition 2.1. We have thus shown:

Proposition 4.1 *If K is a nonordered field then $S(K) = K$ and $s(K) \leq p(K) \leq s(K) + 1$, in other words $p(K) = s(K)$ or $p(K) = s(K) + 1$.*

In [9], Pfister shows that the possible values of $s(K)$, hence of $p(K)$, are extremely restricted in this case; this is Pfister's second main theorem proved in the present paper:

Theorem 4.2 *Let K be a nonordered field. There exists k such that $s(K) = 2^k$.*

Proof. Let k be the unique integer such that $2^k \leq s(K) < 2^{k+1}$. By definition we can thus write $-1 = a+b$ with $a \in S_{2^k}(K)$ and $b \in S_{s(K)-2^k}(K) \subset S_{2^k-1}(K)$. Note that $a \neq 0$ since otherwise $-1 \in S_{2^k-1}(K)$ so $s(K) < 2^k$. We can thus write $-1 = (b+1)/a$. Since by Pfister's theorem $S_{2^k}(K) \setminus \{0\}$ is a subgroup of K^* and since $b+1 \in S_{2^k}(K)$, it follows that $-1 \in S_{2^k}(K)$ so that $s(K) \leq 2^k$, proving the theorem. $\square$

We will see below in Proposition 8.2 that this is best possible, in other words that for every k there exists a field K such that $s(K) = 2^k$.

Thanks to this theorem it is easy to give the values of $s(K)$ and $p(K)$ for some simple nonordered fields. For instance:

Field K	$s(K)$	$p(K)$
$\mathbb{F}_{2^m}$	1	1
$\mathbb{F}_q,\ q \equiv 1 \pmod 4$	1	2
$\mathbb{F}_q,\ q \equiv 3 \pmod 4$	2	2
$\mathbb{F}_{2^m}(X_1, \dots, X_n)$	1	1
$\mathbb{F}_q(X_1, \dots, X_n),\ q \equiv 1 \pmod 4$	1	2
$\mathbb{F}_q(X_1, \dots, X_n),\ q \equiv 3 \pmod 4$	2	3
$\mathbb{Q}_2$	4	4
$\mathbb{Q}_p,\ p \equiv 1 \pmod 4$	1	2
$\mathbb{Q}_p,\ p \equiv 3 \pmod 4$	2	3
$\mathbb{Q}_2(X_1, \dots, X_n)$	4	5
$\mathbb{Q}_p(X_1, \dots, X_n),\ p \equiv 1 \pmod 4$	1	2
$\mathbb{Q}_p(X_1, \dots, X_n),\ p \equiv 3 \pmod 4$	2	3

In the above, n is implicitly assumed to be at least 1. Note also that trivially if K is algebraically closed we have $s(K) = p(K) = 1$.

5 Ordered Fields; Pfister's Third Theorem

The results of this section are also due to Pfister; see [11].

The case of ordered fields is much more difficult, and many conjectures remain. We evidently have $s(K) = +\infty$, but on the other hand we cannot a priori determine whether $p(K)$ is finite or not. It is in fact easy to construct fields with $p(K) = +\infty$, for instance $K = \mathbb{R}(X_i\,;\ i \geq 1)$, see Proposition 8.1 below.

The first fields which are natural to study are the fields $\mathbb{R}(X_1, \dots, X_n)$. We have the beautiful result of Pfister [11]:

Theorem 5.1

$$p(\mathbb{R}(X_1, \dots, X_n)) \leq 2^n \ .$$

Before proving this theorem, a few remarks are in order. First, by a classical theorem of Artin [1], the set $S(\mathbb{R}(X_1, \dots, X_n))$ of sums of squares in $\mathbb{R}(X_1, \dots, X_n)$ is equal to the set of rational functions which are nonnegative for all values of the variables for which they are defined. It follows from Pfister's theorem that such a rational function is in fact the sum of the squares of at most 2^n rational functions.

Second, note that the theorem is trivial for $n = 0$, very easy for $n = 1$, and was proved by Hilbert for $n = 2$; see [7]. For $n = 3$, it was first proved by Ax using cohomological methods; see [2]. As we will see, Pfister's proof is quite elementary.

Since K is ordered, $i = \sqrt{-1} \notin K$, and the idea is to study quadratic forms over $K(i)$. We will in fact prove the following theorem, which as we will see generalizes the above:

Theorem 5.2 *Assume that in $K(i)$ any quadratic form in $m > 2^n$ variables is isotropic. Then $p(K) \leq 2^n$.*

Before proving this theorem, we need some preliminary results. In what follows we assume that the assumption of the theorem is satisfied. In addition, since $p(K) = 1$ when $\mathrm{char}(K) = 2$ we may of course assume that $\mathrm{char}(K) \neq 2$.

We first prove two simple lemmas:

Lemma 5.3 *Let q be a nondegenerate quadratic form in 2^n variables over $K(i)$. Then q is universal on $K(i)$.*

Proof. Indeed, if $a \in K(i)^*$ then by assumption the form $q \oplus \langle -a \rangle$ which has $2^n + 1$ variables is isotropic, hence either a is represented by q, or q itself is isotropic, hence universal by Proposition 2.1. □

Lemma 5.4 *Let q be a nondegenerate* multiplicative *quadratic form over K. If q represents $b + ic \neq 0$ in $K(i)$, then q represents $b^2 + c^2$ in K.*

Proof. By assumption there exists X and Y in K^n such that $q(X + iY) = b + ic$, hence since $i \notin K$ by conjugation $q(X - iY) = b - ic$. We may assume $Y \neq 0$, otherwise $c = 0$ and q represents b, hence also b^2 in K since it is multiplicative. In addition we may assume that $q(Y) \neq 0$, otherwise q would be isotropic, hence universal, so would represent $b^2 + c^2$. Thus, if we denote by $B(X, Y)$ the bilinear form associated with the quadratic form q we have

$$\begin{aligned} b^2 + c^2 &= q(X + iY)q(X - iY) \\ &= (q(X) - q(Y) + 2iB(X, Y))(q(X) - q(Y) - 2iB(X, Y)) \\ &= (q(X) - q(Y))^2 + 4B(X, Y)^2 \ . \end{aligned}$$

On the other hand, note that

$$\begin{aligned} q(2B(X, Y)X + (q(Y) - q(X))Y) &= 4B(X, Y)^2 q(X) + (q(Y) - q(X))^2 q(Y) \\ &\quad + 4B(X, Y)^2 (q(Y) - q(X)) = q(Y)((q(Y) - q(X))^2 + 4B(X, Y)^2) \\ &= q(Y)(b^2 + c^2) \ , \end{aligned}$$

and since $q(Y) \neq 0$ and q is multiplicative we deduce that $b^2 + c^2$ is represented by q. □

We can now begin the proof proper.

Proof of Theorem 5.2. It is of course sufficient to prove that $S_{2^n+1}(K) = S_{2^n}(K)$. Thus, let $a \neq 0$ be such that $a = e_0^2 + e_1^2 + \cdots + e_{2^n}^2$. We will set

$$\begin{aligned} a_i &= \sum_{2^i < j \leq 2^{i+1}} e_j^2 \quad \text{for } 1 \leq i \leq n-1 \ , \quad a_n = -a \ , \quad \text{and} \\ q_i &= \langle 1, a_1 \rangle \otimes \cdots \otimes \langle 1, a_i \rangle \quad \text{for } 0 \leq i \leq n \end{aligned}$$

(where by convention $q_0 = \langle 1 \rangle$). All the partial sums occuring in a can be assumed to be nonzero, otherwise trivially $a \in S_{2^n}(K)$. In particular $e_1^2+e_2^2 \neq 0$ and $a_i \neq 0$ for all i, so the forms q_i are nondegenerate, and by Pfister's first theorem they are all multiplicative. In particular the form q_n is a nondegenerate multiplicative form in 2^n variables, so by Lemma 5.3 it is universal in $K(i)$, so it represents $e_1 + ie_2$, hence by Lemma 5.4 it represents $e_1^2 + e_2^2$ in K. Now note the following immediate additional lemma:

Lemma 5.5 *For all i such that $0 \leq i \leq n$ we have:*

(1)
$$q_i = \langle 1 \rangle \oplus a_1 q_0 \oplus a_2 q_1 \oplus \cdots \oplus a_i q_{i-1} \ .$$

(2) *Any element of K represented by q_i is a sum of $2^{i+1} - 1$ squares.*

Proof. (1) is trivial by induction, and (2) is also immediate by induction by Pfister's first theorem, since $a_i \in S_{2^i}(K)$. ☐

Resuming the proof of the theorem, for $0 \leq i \leq n$ set $q_i' = \langle 1 \rangle \oplus q_i$. By definition of the a_i we have

$$0 = e_0^2 + e_1^2 + e_2^2 + a_1 + \cdots + a_{n-1} + a_n \ ,$$

so by the above lemma and the fact that q_n represents $e_1^2 + e_2^2$, the form

$$\begin{aligned} q_0' \oplus a_1 q_0' \oplus \cdots \oplus a_n q_{n-1}' &= \langle 1 \rangle \oplus \langle 1 \rangle \oplus a_1 q_0 \oplus \cdots \oplus a_n q_{n-1} \oplus \langle a_1, \ldots, a_n \rangle \\ &= \langle 1 \rangle \oplus q_n \oplus \langle a_1, \ldots, a_n \rangle \end{aligned}$$

represents 0 nontrivially. Since $a_n = -a$, this means that there exist X and Y in K^n such that

$$a q_{n-1}'(X) = (q_0' \oplus a_1 q_0' \oplus \cdots a_{n-1} q_{n-2}')(Y) \ .$$

By construction, the first component of X is equal to 1 (it is the coefficient of a_n in $e_0^2+e_1^2+e_2^2+a_1+\cdots+a_{n-1}+a_n$), and in particular $X \neq 0$. If we had $q_{n-1}'(X) = 0$ then by (2) of the above lemma we would have a nontrivial sum of 2^n squares which vanishes, so by Proposition 2.1 $a \in S_{2^n}(K)$. Otherwise, using again (2) of the above lemma we see by induction that $(q_0' \oplus a_1 q_0' \oplus \cdots a_{n-1} q_{n-2}')(Y)$ is a sum of $2^1 + 2^1 + 2^2 + \cdots + 2^{n-1} = 2^n$ squares, and so is $q_{n-1}'(X)$ again by (2) of the lemma, so $a \in S_{2^n}(K)$ by Pfister's first theorem once again, finishing the proof of Theorem 5.2. ☐

Pfister's third theorem now follows from a classical theorem of Tsen–Lang:

Corollary 5.6 $p(\mathbb{R}(X_1, \ldots, X_n)) \leq 2^n$.

Proof. If $K = \mathbb{R}(X_1, \ldots, X_n)$ we have $K(i) = \mathbb{C}(X_1, \ldots, X_n)$, and by a theorem of Tsen–Lang which we will assume (see for instance [6]), the assumption of Theorem 5.2 is satisfied. ☐

6 Ordered Fields; Cassels' Theorem

Pfister's theorem only gives an upper bound on $p(\mathbb{R}(X_1, \ldots, X_n))$. To find a lower bound, we use the following elegant theorem of Cassels [3]:

Theorem 6.1 *Let q be a nondegenerate quadratic form over a field K with $\operatorname{char}(K) \neq 2$. If q represents $f \in K[X]$ in $K(X)$, then q also represents f in $K[X]$.*

Proof. Since q is nondegenerate and $\operatorname{char}(K) \neq 2$, we may assume that $q = \langle a_1, \ldots, a_n \rangle$ with $a_i \neq 0$ for all i. If q is isotropic in K then by Proposition 2.1 it is universal in the algebra $K[X]$ (note that it is *not* a field), so represents f. We may therefore assume that q is not isotropic. By assumption, there exist $f_i \in K[X]$ with $f_0 \neq 0$ such that $ff_0^2 = a_1 f_1^2 + \cdots a_n f_n^2$. Choose f_0 with the lowest possible degree, and assume by contradiction that $\deg(f_0) > 0$. Computing the Euclidean division of f_i by f_0 we obtain polynomials g_i such that $\deg(f_i - g_i f_0) < \deg(f_0)$ and $g_0 = 1$.

Denote by $B(X, Y)$ the bilinear form associated to the quadratic form $\langle -f \rangle \oplus q$, and set $F = (f_0, \ldots, f_n)$ and $G = (g_0, \ldots, g_n)$. The assumption is equivalent to $B(F, F) = 0$. On the other hand, since $g_0 = 1$ and f_0 has the lowest possible degree, but $\deg(f_0) > 0$, we have $B(G, G) \neq 0$. Now set

$$H = B(G, G)F - 2B(F, G)G \; .$$

I claim that $H \neq 0$. Indeed, otherwise we would have $B(F, G) = 0$ (otherwise $G = \lambda F$ for some λ so $B(G, G) = \lambda^2 B(F, F) = 0$, absurd), so $B(G, G)F = 0$, hence $F = 0$, again a contradiction, proving my claim. On the other hand one checks immediately that $B(H, H) = 0$, and that if $H = (h_0, \ldots, h_n)$ we have

$$h_0 = (\sum_i a_i g_i^2 - f) f_0 - 2(\sum_i a_i f_i g_i - f_0 f) = \frac{1}{f_0} \sum_i a_i (f_i - g_i f_0)^2 \; .$$

Thus H is a nonzero vector of polynomials satisfying $B(H, H) = 0$ with $\deg(h_0) < 2\deg(f_0) - \deg(f_0) < \deg(f_0)$, and $h_0 \neq 0$ otherwise q would be isotropic. Thus f_0 is not of minimal degree, a contradiction which proves the theorem. $\square$

Remark. There is an analogous theorem for $\mathbb{Q}$ and $\mathbb{Z}$ due to Davenport and Cassels, as follows:

Theorem 6.2 *Let q be a positive definite quadratic form in n variables with integer matrix coefficients, in other words such that $B(X, Y) \in \mathbb{Z}$ for any X, Y in $\mathbb{Z}^n$. Assume that for any $X \in \mathbb{Q}^n$ there exists $Y \in \mathbb{Z}^n$ such that $q(X - Y) < 1$. Then if $k \in \mathbb{Z}$ is represented by q in $\mathbb{Q}$, it is also represented by q in $\mathbb{Z}$.*

The proof follows similar lines to the above proof, and left as an excellent excercise for the reader.

7 Consequences of Cassels's Theorem

In what follows we always assume implicitly that all quadratic forms are non-degenerate and that $\text{char}(K) \neq 2$.

Lemma 7.1 *Let $q = \langle a_1, \dots, a_n \rangle$ with $a_i \in K^*$. Then q represents the polynomial $a_n X^2 + d$ in $K(X)$ if and only if at least one of the following conditions is satisfied:*

(1) *q is isotropic in K.*

(2) *d is represented by $\langle a_1, \dots, a_{n-1} \rangle$ in K.*

Proof. By Proposition 2.1 it is clear that these conditions are sufficient. Conversely, assume that q is not isotropic in K and represents $a_n X^2 + d$ in $K(X)$. By Cassels's theorem, q also represents $a_n X^2 + d$ in $K[X]$, and since q is not isotropic (exercise: why is this necessary?), there exist b_i and c_i in K such that

$$a_n X^2 + d = \sum_{1 \leq i \leq n} a_i (b_i X + c_i)^2 \ .$$

Since $\text{char}(K) \neq 2$ one of the two elements $1 \pm b_n$ is invertible, so there exist $c \in K$ and a suitable sign such that $c = \pm(b_n c + c_n)$. Thus

$$a_n c^2 + d = \sum_{1 \leq i \leq n} a_i (b_i c + c_i)^2 = a_n c^2 + \sum_{1 \leq i \leq n-1} a_i (b_i c + c_i)^2 \ ,$$

so that $d = \sum_{1 \leq i \leq n-1} a_i (b_i c + c_i)^2$ as claimed. $\square$

Corollary 7.2 *Let K be an ordered field and $L = K(X_1, \dots, X_n)$. Then*

$$X_1^2 + \cdots + X_n^2 \notin S_{n-1}(L) \quad \textit{and}$$
$$1 + X_1^2 + \cdots + X_n^2 \notin S_n(L) \ .$$

In particular, $p(L) \geq n + 1$.

Proof. Immediate from the lemma by induction. $\square$

Remark. Cassels's theorem does not generalize to several variables: for instance q can represent f in $K(X,Y)$ without representing it in $K[X,Y]$ (although by Cassels's theorem it will be representable in $K(X)[Y]$ and in $K(Y)[X]$). For instance, let $f(X,Y) = 1 - 3X^2Y^2 + X^2Y^4 + X^4Y^2$. We check that

$$f(X,Y) = \frac{(1 - X^2Y^2)^2 + X^2(1 - Y^2)^2 + X^2Y^2(1 - X^2)^2}{1 + X^2} \ ,$$

so as a quotient of a sum of 4 squares it is a sum of 4 squares in $K(X,Y)$. On the other hand it is immediate to check that f is not even a sum of squares in $K[X,Y]$.

The above results imply that $n+1 \leq p(\mathbb{R}(X_1,\ldots,X_n)) \leq 2^n$, which leaves a large margin of uncertainty. One conjectures that Pfister's upper bound is in fact the correct value. For $n=0$ and $n=1$ this is clear, and in fact is immediate directly. For $n=2$, in [4], Cassels–Ellison–Pfister have shown that the polynomial $f(X,Y)$ given in the above remark is *not* a sum of 3 squares in $\mathbb{R}(X,Y)$, so that indeed $p(\mathbb{R}(X,Y))=4$, which implies by induction the slightly stronger inequality $n+2 \leq p(\mathbb{R}(X_1,\ldots,X_n)) \leq 2^n$. Nothing better is known for $n \geq 3$, for instance $p(\mathbb{R}(X,Y,Z)) = 5$, 6, 7, or 8.

In the case of other ordered fields such as $K=\mathbb{Q}$, even less is known. For instance Euler–Lagrange's theorem on sums of four squares of integers together with the trivial fact that $7 \notin S_3(\mathbb{Q})$ says that $p(\mathbb{Q})=4$. In [12], Pourchet proved that $p(\mathbb{Q}(X))=5$, and in an unpublished preprint Pop generalized this to number fields with the inequality $5 \leq p(K(X)) \leq 6$ for K a number field. On the other hand, one does not even know whether $p(\mathbb{Q}(X,Y))$ is finite or not.

8 Some Additional Examples

Proposition 8.1 *There exists a field L such that $p(L)=+\infty$.*

Proof. Indeed, choose $L=\mathbb{R}(X_i\,;\ i \geq 1)$. Taking $K=\mathbb{R}(X_i\,;\ i>n)$, Corollary 7.2 tells us that $1+X_1^2+\cdots+X_n^2 \notin S_n(L)$, so that for all n we have $p(L)>n$. ◻

Proposition 8.2 *For any $k \geq 0$ there exists a nonordered field K with $s(K)=2^k$.*

We have seen above that these are the only possible values of $s(K)$.

Proof. Set $n=2^k$. We will choose $K=\mathbb{R}(X_1,\ldots,X_n)(Y)$ where Y is a root of the equation $Y^2+X_1^2+\cdots+X_n^2=0$ (more abstractly $K=\mathbb{R}(X_1,\ldots,X_{n+1})/(X_1^2+\cdots+X_{n+1}^2)$). By construction $-1=\sum_{1\leq i\leq n}(X_i/Y)^2$, so that $s(K)\leq n$. Let us show that we cannot have $s(K)\leq n-1$. Indeed, this would mean that there exist $f_1,\ldots,f_n$ in $\mathbb{R}[X_1,\ldots,X_n][Y]$ not all zero and such that $f_1^2+\cdots+f_n^2=0$. Replacing all Y^2 by $-\sum_{1\leq i\leq n}X_i^2$, we may assume that the degree in Y of all the f_i is at most equal to 1. Thus, taking representatives F_i of the f_i in $\mathbb{R}[X_1,\ldots,X_{n+1}]$ of degree at most 1 in X_{n+1}, the identity $f_1^2+\cdots+f_n^2=0$ means that there exists $P\in\mathbb{R}[X_1,\ldots,X_{n+1}]$ such that

$$F_1^2+\cdots+F_n^2 = P\cdot(X_1^2+\cdots+X_{n+1}^2)$$

in $\mathbb{R}[X_1,\ldots,X_{n+1}]$. Since $\sum_i F_i^2$ has degree at most equal to 2 in X_{n+1} we must have $P\in\mathbb{R}[X_1,\ldots,X_n]$. If we replace X_{n+1} by 0, it follows from the fact that $n=2^k$ and multiplicativity of $S_{2^k}(L)$ for a *field* L that $P\in S_n(\mathbb{R}(X_1,\ldots,X_n))$. Since the F_i are not all equal to 0, applying again multiplicativity we deduce that $X_1^2+\cdots+X_{n+1}^2\in S_n(\mathbb{R}(X_1,\ldots,X_{n+1}))$, which clearly contradicts Corollary 7.2 (2). ◻

References

[1] E. Artin, *Über die Zerlegung definiter Funktionen in Quadrate*, Hamb. Abh. **5** (1927), 100–115; Collected Papers 273–288.

[2] J. Ax, *On ternary definite rational functions*, unpublished.

[3] J.W.S. Cassels, *On the representation of rational functions as sums of squares*, Colloque C.N.R.S. Clermont Ferrand (1964), 55–65; Acta Arith. **9** (1964), 79–82.

[4] J.W.S. Cassels, W.J. Ellison, and A. Pfister, *On sums of squares and on elliptic curves over function fields*, J. Numb. Th. **3** (1971), 125–149.

[5] H. Cohen, *Représentations comme sommes de carrés*, Sém. Th. Nombres Bordeaux (1972), Exposé **21**.

[6] M.J. Greenberg, *Forms in many variables*, Math. Lect. Notes series, Benjamin (1969).

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

[8] S. Lang, *On quasi algebraic closure*, Ann. Math. **55** (1952), 373–390.

[9] A. Pfister, *Darstellung von* -1 *als Summe von Quadraten in einem Körper*, J. London Math. Soc. **40** (1965), 159–165.

[10] A. Pfister, *Multiplikative quadratische Formen*, Arch. Math. **16** (1965), 363–370.

[11] A. Pfister, *Zur Darstellung definiter Funktionen als Summe von Quadraten*, Invent. Math. **4** (1967), 229–237.

[12] 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.