SOME ASPECTS OF THE ALGEBRAIC THEORY OF QUADRATIC FORMS

R. PARIMALA

(Preliminary notes for lectures at AWS 2009)

There are many good references for this material including [EKM], [L], [Pf], and [S].

1. Quadratic forms

Let k be a field, char$k \neq 2$

Definition 1.1. A **quadratic form** $q : V \to k$ on a finite dimensional vector space V over k is a map satisfying:

(1) $q(\lambda v) = \lambda^2 q(v)$, $v \in V$, $\lambda \in k$.
(2) The map $b_q : V \times V \to k$, defined by

$$b_q(v, w) = \frac{1}{2}[q(v + w) - q(v) - q(w)]$$

is bilinear.

The bilinear form b_q is symmetric; q determines b_q and for all $v \in V$, $q(v) = b_q(v, v)$.

For a choice of basis $\{e_1, \dots, e_n\}$ of V, b_q is represented by a symmetric matrix $A(q) = (a_{ij})$, $a_{ij} = b_q(e_i, e_j)$. If $v = \sum_{1 \leq i \leq n} X_i e_i \in V$, $X_i \in k$, then

$$q(v) = \sum_{1 \leq i,j \leq n} a_{ij} X_i X_j = \sum_{1 \leq i \leq n} a_{ii} X_i^2 + 2 \sum_{i<j} a_{ij} X_i X_j.$$

Thus q is represented by a homogeneous polynomial of degree 2, Clearly, every homogeneous polynomial of degree 2 corresponds to a quadratic form on V with respect to the chosen basis. We define the *dimension of* q to be the dimension of the underyling vector space V and denote it by $\dim(q)$.

Date: March 14 - March 18, 2009.

Definition 1.2. Two quadratic forms (V_1, q_1), (V_2, q_2) are **isometric** if there is an isomorphism $\phi : V_1 \xrightarrow{\sim} V_2$ such that $q_2(\phi(v)) = q_1(v)$, $\forall v \in V_1$.

If $A(q_1)$, $A(q_2)$ are the matrices representing q_1 and q_2 with respect to bases B_1 and B_2 of V_1 and V_2 respectively, ϕ yields a matrix $T \in M_n(k)$, $n = \dim V$ such that

$$TA(q_2)T^t = A(q_1).$$

In other words, the symmetric matrices $A(q_1)$ and $A(q_2)$ are orthogonally equivalent. Thus isometry of quadratic forms yields orthogonal equivalence on symmetric matrices.

Definition 1.3. The form $q : V \to k$ is said to be **regular** if $b_q : V \times V \to k$ is nondegenerate.

Thus q is regular if and only if the map $V \to V^* = Hom(V, k)$, defined by $v \to (w \mapsto b_q(v, w))$, is an isomorphism. This is the case if $A(q)$ is invertible.

Definition 1.4. Let W be a subspace of V and $q : V \to k$ be a quadratic form. The **orthogonal complement** of W denoted $W^\perp$ is the subspace

$$W^\perp = \{v \in V \ : \ b_q(v, w) \ = \ 0 \ \forall \ w \in W\}.$$

Exercise 1.5. Let (V, q) be a regular quadratic form and W a subspace of V. Then

(1) $\dim(W) + \dim(W^\perp) = \dim(V)$
(2) $(W^\perp)^\perp = W$

1.1. **Orthogonal sums.** Let (V_1, q_1), (V_2, q_2) be quadratic spaces. The space

$$(V_1, q_1) \perp (V_2, q_2) = (V_1 \oplus V_2, q_1 \perp q_2),$$

with $q_1 \perp q_2$ defined by

$$(q_1 \perp q_2)(v_1, v_2) = q_1(v_1) + q_2(v_2), \ v_1 \in V_1 \ v_2 \in V_2.$$

is called the *orthogonal sum* of (V_1, q_1) and (V_2, q_2).

1.2. **Diagonalisation.** Let (V, q) be a quadratic space. There exists a basis $\{e_1, \ldots, e_n\}$ of V such that $b_q(e_i, e_j) = 0$ for $i \neq j$. Such a basis is called an *orthogonal basis* for q and, with respect to an orthogonal basis, b_q is represented by a diagonal matrix. If $\{e_1, \ldots, e_n\}$ is an orthogonal basis of q and $q(e_i) = d_i$, $V = ke_1 \oplus \cdots \oplus ke_n$ is an orthogonal sum and $q|ke_i$ is represented by $\langle d_i \rangle$.

Thus every quadratic space is diagonalisable. We simply write $q = \langle d_1, \ldots, d_n \rangle$ if q admits an orthogonal basis $\{e_1, \ldots, e_n\}$ with $q(e_i) = d_i$.

1.3. **Hyperbolic spaces.**

Definition 1.6. A quadratic space (V, q) is said to be **isotropic** if there is a nonzero $v \in V$ such that $q(v) = 0$. It is **anisotropic** if q is not isotropic. A quadratic space (V, q) is said to be **universal** if it represents every nonzero element of F.

Example 1.7. The quadratic form $X^2 - Y^2$ is isotropic over k. Suppose (V, q) is a regular space which is isotropic. Let $v \in V$ be such that $q(v) = 0$, $v \neq 0$. Since q is regular, there exists $w \in V$ such that $b_q(v, w) \neq 0$. After scaling we may assume $b_q(v, w) = 1$. If $q(w) \neq 0$, we may replace w by $w + \lambda v$, $\lambda = -\frac{1}{2} q(w) b_q(v, w)^{-1}$, and assume that $q(w) = 0$. This $W = kv \oplus kw$ is a 2 dimensional subspace of V and $q|W$ is represented by $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ with respect to $\{v, w\}$.

Definition 1.8. A binary quadratic space isometric to $(k^2, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix})$ is called a **hyperbolic plane**. A quadratic space (V, q) is **hyperbolic** if it is isometric to an orthogonal sum of hyperbolic planes. A subspace W of V such that q restricts to zero on W and $\dim W = \frac{1}{2} \dim V$ is called a **Lagrangian**.

Every regular quadratic form which admits a Lagrangian can easily be seen to be hyperbolic.

Exercise 1.9. Let (V, q) be a regular quadratic space and $(W, q|W)$ a regular subspace. Then $(V, q) \xrightarrow{\sim} (W, q|W) \perp (W^\perp, q|W^\perp)$.

Let (V, q) be a quadratic space. Then

$$V_0 = \{v \in V \ : \ b_q(v, w) = 0 \ \forall \ w \in V\}$$

is called the **radical** of V. If V_1 is any complementary subspace of V_0 in V, then $q|V_1$ is regular and $(V, q) = (V_0, 0) \perp (V_1, q|V_1)$. Note that V is regular if and only if the radical of V is zero. If (V, q) is any quadratic space, we define the **rank** of q to be the dimension of $V/V^{\perp}$. Of course if (V, q) is regular, then $\text{rank}(q) = \dim(q)$.

Theorem 1.10. *(Witt's Cancellation Theorem) Let (V_1, q_1), (V_2, q_2), (V, q) be quadratic spaces over k. Suppose*

$$(V_1, q_1) \perp (V, q) \cong (V_2, q_2) \perp (V, q).$$

Then $(V_1, q_1) \cong (V_2, q_2)$.

The key ingredient of Witt's cancellation theorem is the following.

Proposition 1.11. *Let (V, q) be a quadratic space and $v, w \in V$ with $q(v) = q(w) \neq 0$. Then there is an isometry $\tau : (V, q) \xrightarrow{\sim} (V, q)$ such that $\tau(v) = w$.*

Proof. Let $q(v) = q(w) = d \neq 0$. Then

$$q(v + w) + q(v - w) = 2q(v) + 2q(w) = 4d \neq 0.$$

Thus $q(v+w) \neq 0$ or $q(v-w) \neq 0$. For any vector $u \in V$ with $q(u) \neq 0$, define $\tau_u : V \to V$ by

$$\tau_u(z) = z - \frac{2b_q(z, u)u}{q(u)}.$$

Then τ_u is an isometry called the *reflection with respect to u*.

Suppose $q(v - w) \neq 0$. Then $\tau_{v-w} : V \to V$ is an isometry of V which sends v to w. Suppose $q(v + w) \neq 0$. Then $\tau_w \circ \tau_{v+w}$ sends v to w. □

Remark 1.12. The orthogonal group of (V, q) denoted by $O(q)$ is the set of isometries of V onto itself. This group is generated by reflections. This is seen by an inductive argument on $\dim(q)$, using the above lemma.

Theorem 1.13. *(Witt's decomposition) Let (V, q) be a quadratic space. Then there is a decomposition*

$$(V, q) = (V_0, 0) \perp (V_1, q_1) \perp (V_2, q_2)$$

where V_0 is the radical of q, $q_1 = q|V_1$ is anisotropic and $q_2 = q|V_2$ is hyperbolic. If $(V, q) = (V_0, 0) \perp (W_1, f_1) \perp (W_2, f_2)$ with f_1 anisotropic and f_2 hyperbolic, then

$$(V_1, q_1) \cong (W_1, f_1), \ (V_2, q_2) \cong (W_2, f_2).$$

Remark 1.14. A hyperbolic space (W, f) is determined by $\dim(W)$; if $\dim(W) = 2n$, $(W, f) \cong nH$, where $H = (k^2, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix})$ is the hyperbolic plane.

From now on, we shall assume (V, q) is a regular quadratic space. We denote by q_{an} the quadratic space (V_1, q_1) in Witt's decomposition which is determined by q up to isometry. We call $\frac{1}{2}\dim(V_2)$ the *Witt index* of q. Thus any regular quadratic space q admits a decomposition $q \cong q_{an} \perp (nH)$, with q_{an} anisotropic and H denoting the hyperbolic plane. We also sometime denote by H^n the sum of n hyperbolic planes.

2. Witt group of forms

2.1. **Witt groups.** We set

$$W(k) = \{\text{isomorphism classes of regular quadratic spaces over } k\}/\sim$$

where the Witt equivalence $\sim$ is given by:

$$(V_1, q_1) \sim (V_2, q_2) \quad \Longleftrightarrow \quad \begin{array}{l} \text{there exist } r,\ s \in \mathbb{Z} \text{ such that} \\ (V_1, q_1) \perp H^r \cong (V_2, q_2) \perp H^s \end{array}.$$

$W(k)$ is a group under orthogonal sum:

$$[(V_1, q_1)] \perp [(V_2, q_2)] = [(V_1, q_1) \perp (V_2, q_2)].$$

The zero element in $W(k)$ is represented by the class of hyperbolic spaces. For a regular quadratic space (V, q), $(V, q) \perp (V, -q)$ has Lagrangian

$$W = \{(v, v) \ : \ v \in V\}$$

so that $(V, q) \perp (V, -q) \cong H^n$, $n = \dim(V)$. Thus, $[(V, -q)] = -[(V, q)]$ in $W(k)$.

It follows from Witt's decomposition theorem that every element in $W(k)$ is represented by a unique anisotropic quadratic space up to isometry. Thus $W(k)$ may be thought of as a group made out of isometry classes of anisotropic quadratic spaces over k.

The abelian group $W(k)$ admits a ring structure induced by tensor product on the associated bilinear forms. For example, if $q_1 \cong \langle a_1, \ldots, a_n\rangle$ and q_2 is a quadratic form, then $q_1 \otimes q_2 \cong a_1 q_2 \perp a_2 q_2 \perp \cdots \perp a_n q_2$.

Definition 2.1. Let $I(k)$ denote the ideal of classes q of even dimensional quadratic forms in $W(k)$. The ideal $I(k)$ is called the **fundamental ideal**. $I^n(k)$ stands for the n-th power of the ideal $I(k)$.

Definition 2.2. Let $P_n(k)$ denote the set of isomorphism classes of forms of the type

$$\langle\langle a_1, \dots, a_n\rangle\rangle := \langle 1, a_1\rangle \otimes \cdots \otimes \langle 1, a_n\rangle.$$

Elements in $P_n(k)$ are called **n-fold Pfister forms**.

The ideal $I(k)$ is generated by $\langle 1, a\rangle$, $a \in k^*$. Moreover, the ideal $I^n(k)$ is generated additively by n-fold Pfister forms. For instance, for $n = 2$, the generators of $I^2(k)$ are of the form

$$\langle a, b\rangle \otimes \langle c, d\rangle \cong \langle 1, ac, ad, cd\rangle - \langle 1, cd, -bc, -bd\rangle = \langle\langle ac, ad\rangle\rangle - \langle\langle cd, -bc\rangle\rangle$$

Example 2.3. $k = \mathbb{C}$, every 2 dimensional quadratic space over k is isotropic.

$$W(k) \cong \mathbb{Z}/2\mathbb{Z}$$

$$[(V, q)] \mapsto \dim(V) \pmod 2$$

is an isomorphism.

Example 2.4. $k = \mathbb{R}$. Every quadratic form q is represented by

$$\langle 1, \dots, 1, -1, \dots, -1\rangle$$

with respect to an orthogonal basis. The number r of $+1$'s and the number s of -1's in the diagonalisation above are uniquely determined by the isomorphism class of q. The *signature* of q is defined as $r - s$. The signature yields an homomorphism $sgn : W(\mathbb{R}) \to \mathbb{Z}$ which is an isomorphism.

2.2. **Quadratic forms over p-adic fields.** Let k be a finite extension of the field $\mathbb{Q}_p$ of p-adic numbers. We call k a non-dyadic p-adic field if $p \neq 2$. The field k has a discrete valuation v extending the p-adic valuation on $\mathbb{Q}_p$. Let π be a uniformizing parameter for v and κ the residue field for v. The field κ is a finite field of characteristic $p \neq 2$. Let u be a unit in k^* such that $\overline{u} \in \kappa$ is not a square. Then

$$k^*/k^{*2} = \{1, u, \pi, u\pi\}.$$

Since κ is finite, every 3-dimensional quadratic form over κ is isotropic. By Hensel's lemma, every 3-dimensional form $\langle u_1, u_2, u_3\rangle$ over k, with u_i units in k is isotropic. Since every form in k has a diagonal representation

$$\langle u_1, \dots, u_r\rangle \perp \pi\langle v_1, \dots, v_s\rangle,$$

if r or s exceeds 3, q is isotropic. In particular every 5-dimensional quadratic form over k is isotropic. Further, up to isometry, there is a unique quadratic form in dimension 4 which is anisotropic, namely,

$$\langle 1, -u, -\pi, u\pi \rangle.$$

This is the norm form of the unique quaternion division algebra $H(u,\pi)$ over k (cf., section 2.3).

2.3. **Central simple algebras and the Brauer group.** Recall that a finite dimensional algebra A over a field k is a *central simple algebra* over k if A is simple (has no 2-sided ideals) and the center of A is k. Recall also that for a field k,

$$Br(k) = \{\text{Isomorphism classes of central simple algebras over } k\} \,/ \sim$$

where the Brauer equivalence $\sim$ is given by: $A \sim B$ if and only if $M_n(A) \cong M_m(B)$ for some integers m, n. The pair $(Br(k), \otimes)$ is a group. The inverse of $[A]$ is $[A^{\mathrm{op}}]$ where A^{op} is the *opposite algebra* of A: that is, the multiplication structure, $*$, on A^{op} is given by $a * b = ba$. We have a k-algebra isomorphism $\phi : A \otimes A^{\mathrm{op}} \xrightarrow{\sim} \mathrm{End}_k(A)$ induced by $\phi(a \otimes b)(c) = acb$. The identity element in $\mathrm{Br}(k)$ is given by $[k]$. By Wedderburn's theorem on central simple algebras, the elements of $Br(k)$ parametrize the isomorphism classes of finite dimensional central division algebra over k.

For elements $a, b \in k^*$, we define the **quaternion algebra** $H(a,b)$ to be the 4-dimensional central simple algebra over k generated by $\{i, j\}$ with the relations $i^2 = a$, $j^2 = b$, $ij = -ji$. This is a generalization of the standard Hamiltonian quaternion algebra $H(-1,-1)$. The algebra $H(a,b)$ admits a canonical involution $^-$: $H(a,b) \to H(a,b)$ given by

$$\overline{\alpha + i\beta + j\gamma + ij\delta} = \alpha - i\beta - j\gamma - ij\delta$$

This involution gives an isomorphism $H(a,b) \cong H(a,b)^{\mathrm{op}}$; in particular, $H(a,b)$ has order 2 in ${}_2\mathrm{Br}(k)$. The norm form for this algebra is given by $N(x) = x\overline{x}$ which is a quadratic form on $H(a,b)$ represented with respect to the orthogonal basis $\{1, i, j, ij\}$ by $\langle 1, -a, -b, ab \rangle = \langle\langle -a, -b \rangle\rangle$.

2.4. **Classical invariants for quadratic forms.** Let (V, q) be a regular quadratic space. We define $\dim(q) = \dim(V)$ and $\dim_2(q) = \dim(V)$ modulo 2. We have a ring homomorphism $\dim_2 : W(k) \to \mathbb{Z}/2\mathbb{Z}$. We note that $I(k)$ is the kernel of $\dim_2$. This gives an isomorphism

$$\dim_2 : W(k)/I(k) \xrightarrow{\sim} \mathbb{Z}/2\mathbb{Z}.$$

Let $\mathrm{disc}(q) = (-1)^{n(n-1)/2}[\det(A(q))] \in k^*/k^{*2}$. Since $A(q)$ is determined up to orthogonal equivalence, $\det(A(q))$ is determined modulo squares. We have $\mathrm{disc}(H) = 1$ and $\mathrm{disc}(q)$ induces a group homomorphism

$$\mathrm{disc} : I(k) \to k^*/k^{*2}$$

which is clearly onto. It is easy to verify that $ker(\mathrm{disc}) = I^2(k)$. Thus the discriminant homomorphism induces an isomorphism $I(k)/I^2(k) \to k^*/k^{*2}$.

The next invariant for quadratic forms is the Clifford invariant. To each quadratic form (V, q) we wish to determine a central simple algebra containing V whose multiplication on elements of V satisfies $v{\cdot}v = q(v)$. The smallest such algebra (i.e., defining a universal property) will be the Clifford algebra.

Definition 2.5. The **Clifford algebra** $C(q)$, of quadratic space (V, q) is $T(V)/I_q$, where I_q is the two sided ideal in the tensor algebra $T(V)$ generated by $\{v \otimes v - q(v), v \in V\}$.

The algebra $C(q)$ has a $\mathbb{Z}/2\mathbb{Z}$ gradation $C(q) = C_0(q) \oplus C_1(q)$ induced by the gradation $T(V) = T_0(V) \oplus T_1(V)$, where

$$T_0(V) = \bigoplus_{i \geq 0,\ i \text{ even}} V^{\otimes i} \qquad \text{and} \qquad T_1(V) = \bigoplus_{i \geq 1,\ i \text{ odd}} V^{\otimes i}.$$

If $\dim(q)$ is even, then $C(q)$ is a central simple algebra over k. If $\dim(q)$ is odd, $C_0(q)$ is a central simple algebra over k. The Clifford algebra $C(q)$ comes equipped with an involution τ defined by $\tau(v) = -v$, $v \in V$. Thus, if $\dim(q)$ is even, $C(q)$ determines a 2 torsion element in $Br(k)$.

Definition 2.6. The **Clifford invariant** $c(q)$ of (V, q) in $Br(k)$ is defined as

$$c(q) = \begin{cases} [C(q)], & \text{if } \dim(q) \text{ is even} \\ [C_0(q)], & \text{if } \dim(q) \text{ is odd} \end{cases}$$

The Clifford invariant induces a homomorphism $c : I^2(k) \to {}_2Br(k)$, ${}_2Br(k)$ denoting the 2 torsion in the Brauer group of k. The very first case of the Milnor conjecture: c is surjective and $\ker(c) = I^3(k)$.

Theorem 2.7. *(Merkurjev [M1]) The map c induces an isomorphism*

$$I^2(k)/I^3(k) \cong {}_2Br(k)$$

Example 2.8. Let $q \cong \otimes_{i=1}^{n} \langle\langle -a_i, -b_i \rangle\rangle \in I^2(k)$. Then

$$c(q) \cong \otimes_{1 \leq i \leq n} H_i$$

where $H_i = H(a_i, b_i)$.

Exercise 2.9. Given $\bigotimes_{1 \leq i \leq n} H_i$ a tensor product of n quaternion algebras over k, show that there is a quadratic form q over k of dimension $2n+2$ such that $c(q) \cong \bigotimes_{1 \leq i \leq n} H_i$.

Thus the image of $I^2(q)$ in ${}_2Br(k)$ is spanned by quaternion algebras. It was a long standing question whether ${}_2Br(k)$ is spanned by quaternion algebras. Merkurjev's theorem answers this question in the affirmative; further, it gives precise relations between these algebras in ${}_2Br(k)$.

3. Galois cohomology and the Milnor conjecture

Let $\Gamma_k = Gal(\bar{k}|k)$, $\bar{k}$ denoting the separable closure of k, be the absolute Galois group of k. The group

$$\Gamma_k = \varprojlim_{L \subset \bar{k},\, L|k \text{ finite Galois}} Gal(L|k)$$

is a profinite group. A *discrete* Γ_k *module* M is a continuous Γ_k module for the discrete topology on M and the profinite topology on Γ_k. For a discrete Γ_k module M, we define $H^n(k, M)$ as the direct limit of the cohomology of the finite quotients

$$H^n(k, M) = \varinjlim_{L \subset \bar{k},\, L|k \text{ finite Galois}} H^n(Gal(L|k), M^{\Gamma_L}).$$

Suppose $\mathrm{char}(k) \neq 2$ and $M = \mu_2$. The module μ_2 has trivial Γ_k action. We denote this module by $\mathbb{Z}/2\mathbb{Z}$. We have

$$H^0(k, \mathbb{Z}/2\mathbb{Z}) = \mathbb{Z}/2\mathbb{Z}$$
$$H^1(k, \mathbb{Z}/2\mathbb{Z}) \cong k^*/k^{*2}$$
$$H^2(k, \mathbb{Z}/2\mathbb{Z}) \cong_2 Br(k) \text{ (2 torsion in Br(k))}$$

These can be seen from the Kummer short exact sequence of Γ_k-modules:

$$0 \longrightarrow \mu_2 \longrightarrow \bar{k}^* \xrightarrow{\cdot 2} \bar{k}^* \longrightarrow 0$$

and noting that $H^1(\Gamma_k, \bar{k}^*) = 0$ (Hilbert's Theorem 90) and $H^2(\Gamma_k, \bar{k}^*) = Br(k)$.

For an element $a \in k^*$, we denote by (a) its class in $H^1(k, \mathbb{Z}/2\mathbb{Z})$ and for $a_1, \ldots, a_n \in k^*$, the cup product $(a_1) \cup \cdots \cup (a_n) \in H^n(k, \mathbb{Z}/2\mathbb{Z})$ is denoted by $(a_1). \cdots ..(a_n)$.

For $a, b \in k^*$, the element $(a).(b)$ represents the class of $H(a, b)$ in ${}_2Br(k)$. The map

$$c : I^2(k) \to H^2(k, \mathbb{Z}/2\mathbb{Z})$$

sends $\langle 1, -a, -b, ab\rangle$ to the class of $H(a, b)$ in $H^2(k, \mathbb{Z}/2\mathbb{Z})$. The forms $\langle 1, -a, -b, ab\rangle$ additively generate $I^2(k)$. Merkurjev's theorem asserts that $H^2(k, \mathbb{Z}/2\mathbb{Z})$ is generated by $(a).(b)$, $a, b \in k^*$. The Milnor conjecture (quadratic form version) proposes higher invariants $I^n(k) \to H^n(k, \mathbb{Z}/2\mathbb{Z})$ extending the classical invariants.

Milnor conjecture: The assignment

$$\langle 1, a_1\rangle \otimes \cdots \otimes \langle 1, a_n\rangle \mapsto (a_1). \cdots .(a_n)$$

yields a map $e_n : P_n(k) \to H^n(k, \mathbb{Z}/2\mathbb{Z})$. This map extends to a homomorphism $e_n : I^n(k) \to H^n(k, \mathbb{Z}/2\mathbb{Z})$ which is onto and $\ker(e_n) = I^{n+1}(k)$.

The maps *dimension mod* 2, *discriminant* and *Clifford invariant* coincide with e_0, e_1 and e_2. Unlike these classical invariants, which are defined on all quadratic forms, conjecturally e_n, $n \geq 3$ are defined only on elements in $I^n(k)$ on which the invariants e_i, $i \leq n-1$ vanish. Arason [Ar], (1975) proved that $e_3 : I^3(k) \to H^3(k, \mathbb{Z}/2\mathbb{Z})$ is well defined and is one-one on $P_3(k)$. As we mentioned earlier, the first nontrivial case of the Milnor conjecture was proved by Merkurjev for $n = 2$. The Milnor conjecture (quadratic form version) is now a theorem due to Orlov-Vishik-Voevodsky [OVV].

Milnor conjecture gives a classification of quadratic forms by their Galois cohomology invariants: Given anisotropic quadratic forms q_1 and q_2, suppose $e_i(q_1 \perp -q_2) = 0$ for $i \geq 0$. Then $q_1 = q_2$ in $W(k)$. We need only to verify $e_i(q_1 \perp -q_2) = 0$ for $i \leq N$ where $N \leq 2^n$ and $\dim(q_1 \perp -q_2) \leq 2^n$, by the following theorem of Arason and Pfister.

Theorem 3.1 (Arason-Pfister Hauptsatz)**.** *Let k be a field. The dimension of an anisotropic quadratic form in $I^n(k)$ is at least 2^n.*

3.1. **Pfister forms.** The theory of Pfister forms (or multiplicative forms as Pfister called them) evolved from questions on classification of quadratic forms whose nonzero values form a group (hereditarily).

Definition 3.2. A regular quadratic form q over k is called **multiplicative** if the nonzero values of q over any extension field L over k form a group.

We have the following examples of quadratic forms which are multiplicative.

Example 3.3. $\langle 1\rangle$: nonzero squares are multiplicatively closed in k^*.

Example 3.4. $\langle 1, -a\rangle : x^2 - ay^2$, $a \in k^*$ is the norm from the quadratic algebra $k[t]/(t^2 - a)$ over k and the norm is multiplicative.

Example 3.5. $\langle 1, -a\rangle \otimes \langle 1, -b\rangle : x^2 - ay^2 - bz^2 + abt^2$ is a norm form from the quaternion algebra $H(a, b) : N(\alpha + i\beta + j\gamma + ij\delta) = \alpha^2 - a\beta^2 - b\gamma^2 + ab\delta^2$. The norm once again is multiplicative.

Example 3.6. $\langle 1, -a\rangle \otimes \langle 1, -b\rangle \otimes \langle 1, -c\rangle : (x^2 - ay^2 - bz^2 + abt^2) - c(u^2 - av^2 - bw^2 + abs^2)$ is the norm form from an octonian algebra associated to the triple (a, b, c); it is a non-associative algebra obtained from the quaternion algebra $H(a, b)$ by a doubling process. The norm is once again multiplicative.

Theorem 3.7. *(Pfister) An anisotropic quadratic form q over k is multiplicative if and only if q is isomorphic to a Pfister form.*

We shall sketch a proof of this theorem. The main ingredients are

Theorem 3.8. *(Cassels-Pfister) Let $q = \langle a_1, \ldots, a_n\rangle$ be a regular quadratic form over k and $f(X) \in k[X]$, a polynomial over k which is a value of q over $k(X)$. Then there exist polynomials $g_1, \ldots, g_n \in k[X]$ such that $f(X) = a_1 g_1(X)^2 + \cdots + a_n g_n(X)^2$.*

Corollary 3.9. *(Specialisation Lemma) Let $q = \langle a_1, \ldots, a_n\rangle$ be a quadratic form over k, $X = \{X_1, \ldots, X_n\}$, $p(X) \in k(X)$ a rational function represented by q over $k(X)$. Then for any $v \in k^n$ where $p(v)$ is defined, $p(v)$ is represented by q over k.*

Proof. We may assume, by multiplying $p(X)$ by a square that $p(X) \in k[X]$. Let $p(X) = p_1(X_n)$, $p_1 \in k[X_1, \ldots, X_{n-1}]$. By Cassels-Pfister theorem, $p_1(X_n)$ is represented by q over $k(X_1, \ldots, X_{n-1})[X_n]$. Let $v = (v_1, \ldots, v_n)$. Then specialising X_n to v_n, we have $p_1(v_n) \in k[X_1, \ldots, X_{n-1}]$ is represented by q over $k(X_1, \ldots, X_{n-1})$. By an induction argument, one concludes that $p(v_1, \ldots, v_n)$ is a value of q over k. $\square$

Theorem 3.10. *(Subform theorem) Let $q = \langle a_1, \dots, a_n\rangle$, $\gamma = \langle b_1, \dots, b_m\rangle$ be anisotropic quadratic forms over k. Then γ is a subform of q (i.e., $q \cong \gamma \perp \gamma'$ for some form γ' over k) if and only if $b_1X_1^2 + \cdots + b_mX_m^2$ is a value of q over $k(X_1, \dots, X_m)$.*

Lemma 3.11. *Let q be an anisotropic quadratic form over k of dimension n. Then q is multiplicative if and only if for indeterminates $X = (X_1, \dots, X_n)$, $Y = (Y_1, \dots, Y_n)$, $q(X).q(Y)$ is a value of q over $k(X_1, \dots, X_n, Y_1, \dots, Y_n)$.*

Proof. The only non-obvious part is "only if". Suppose $L|k$ is a field extension and $v, w \in L^n$. Let $q(v) = c$ and $q(w) = d$. Since $q(X).q(Y)$ is a value of q over $k(X, Y)$, by specialisation lemma, $q(X).q(w)$ is a value of q over $k(X)$ and by the same lemma, $q(v).q(w)$ is a value of q over k. □

Lemma 3.12. *Let q be an anisotropic quadratic form over k of dimension n. Let $X = \{X_1, \dots, X_n\}$ be a set of n indeterminates. Then q is multiplicative if and only if $q \cong q(X).q$ over $k(X)$.*

Proof. Suppose $q \cong q(X).q$ over $k(X)$. Let A be the matrix representing q over k. There exists $W \in Gl_n(k(X))$ such that $q(X)A = WAW^t$. Let $Y = \{Y_1, \dots, Y_n\}$ be a set of n indeterminates. Over $k(X, Y)$,

$$q(X).q(Y) = Y(q(X)A)Y^t = (YW)A(YW)^t = q(Z)$$

where $Z = YW$. Thus $q(X).q(Y)$ is a value of q over $k(X, Y)$ and by 3.11, q is multiplicative. Suppose conversely that q is multiplicative. Then $q(X).q(Y)$ is a value of q over $k(X, Y)$. By the subform theorem, $q(X).q$ is a subform of q. A dimension count yields $q \cong q(X).q$. □

Proof of Pfister's theorem 3.7. Let $q = \langle 1, a_1\rangle \otimes \cdots \otimes \langle 1, a_n\rangle$ be an anisotropic quadratic form over k. Over any extension $L|k$, either q is anisotropic or isotropic in which case it is universal. Thus it suffices to show that the nonzero values of q form a subgroup of k^*. The proof is by induction on n; for $n = 1$, q is the norm form from a quadratic extension of k (see Example 3.4). Let $n \geq 2$. We have $q \cong q_1 \perp a_nq_1$, where $q_1 = \langle 1, a_1\rangle \otimes \cdots \otimes \langle 1, a_{n-1}\rangle$ is an anisotropic $(n-1)$ fold Pfister form. Let $X = \{X_1, \dots, X_{2^{n-1}}\}$, $Y = \{Y_1, \dots, Y_{2^{n-1}}\}$ be two sets of 2^{n-1} indeterminates. Since q_1 is multiplicative, by Lemma 3.12, $q_1(X).q_1 \cong q_1$ over $k(X)$ and $q_1(Y).q_1 \cong q_1$ over $k(Y)$. We have over $k(X, Y)$

$$q \cong q_1(X)q_1 \perp a_nq_1(Y)q_1 \cong \langle q_1(X), a_nq_1(Y)\rangle \otimes q_1.$$

Since $q(X,Y) = q_1(X) + a_n q_1(Y)$, $\langle q_1(X), a_n q_q(Y)\rangle$ represents $q(X,Y)$. Therefore, by a comparison of discriminants,

$$\begin{aligned} \langle q_1(X), a_n q_1(Y)\rangle &\cong \langle q(X,Y), a_n q(X,Y) q_1(X) q_1(Y)\rangle \\ &\cong q(X,Y)\langle 1 \perp a_n q_1(X) q_1(Y)\rangle \end{aligned}$$

In particular,

$$\begin{aligned} q &\cong q(X,Y)\langle 1, a_n q_1(X) q_1(Y)\rangle \otimes q_1 \\ &\cong q(X,Y)(q_1 \perp a_n q_1) \\ &\cong q(X,Y) q \end{aligned}$$

Thus by 3.12 q is multiplicative.

Conversely, let q be an anisotropic quadratic form over k which is multiplicative. Let n be the largest such that q contains an n fold Pfister form $q_1 = \langle 1, a_1\rangle \otimes \cdots \otimes \langle 1, a_n\rangle$ as a subform. Suppose $q \cong q_1 \perp \gamma$, $\gamma = \langle b_1, \ldots, b_m\rangle$, with $m \geq 1$. Let $Z = \{Z_1, \ldots, Z_{2^n}\}$. Over $k(Z)$,

$$q \cong q(Z,0)q \cong q_1(Z)(q_1 \perp \gamma) \cong q_1(Z).q_1 \perp q_1(Z).\gamma \cong q_1 \perp q_1(Z).\gamma.$$

By Witt's cancellation, $\gamma \cong q_1(Z).\gamma$ over $k(Z)$. Thus γ represents $b_1 q_1(Z)$ over $k(Z)$ and by the subform theorem, $\gamma \cong b_1 q_1 \perp \gamma_1$. Then $q \cong q_1 \perp b_1 q_1 \perp \gamma_1 \cong \langle 1, b_1\rangle \otimes q_1 \perp \gamma_1$ contains a $n+1$ Pfister form $\langle 1, b_1\rangle \otimes q_1$ leading to a contradiction to maximality of n. Thus $q \cong q_1$. □

An important property of Pfister forms in stated in the following.

Proposition 3.13. *Let ϕ be an n fold Pfister form. If ϕ is isotropic then ϕ is hyperbolic.*

Proof. Let $\phi = r.\langle 1, -1\rangle \perp \phi_0$, with ϕ_0 anisotropic, $\dim(\phi_0) \geq 1$ and $r \geq 1$. Let $\dim(\phi) = m$ and $X = \{X_1, \ldots, X_m\}$ be a set of m indeterminates. Over $k(X_1, \ldots, X_m)$

$$\phi = r.\langle 1, -1\rangle \perp \phi_0 \cong \phi(X_1, \ldots, X_m).\phi \cong r\langle 1, -1\rangle \perp \phi(X_1, \ldots, X_m).\phi_0.$$

By Witt's cancellation theorem

$$\phi_0 \cong \phi(X_1, \ldots, X_m).\phi_0$$

and if b is a value of ϕ_0, $b.\phi(X_1, \ldots, X_m)$ is a value of ϕ_0 and by the subform theorem, $b.\phi$ is a subform of ϕ_0 contradicting $\dim(\phi_0) < \dim(\phi)$. Thus $\phi \cong r.\langle 1, -1\rangle$ is hyperbolic. □

Corollary 3.14. *The only integers n such that a product of sums of n squares is again a sum of n squares over any field of characteristic zero are $n = 2^m$ for all $m \geq 0$.*

3.2. **Level of a field.**

Definition 3.15. The **level** of a field k is the least positive integer n such that -1 is a sum of n squares in k. We denote the level of k by $s(k)$.

If the field is formally real (i.e., -1 is not a sum of squares), then the level is infinite. It was a long standing open question whether the level of a field, if finite, is always a power of 2. Pfister's theory of quadratic forms leads to an affirmative answer to this question.

Theorem 3.16. *(Pfister) The level of a field is a power of* 2 *if it is finite.*

Proof. Let $n = s(k)$. We choose an integer m such that $2^m \leq n < 2^{m+1}$. Suppose

$$-1 = (u_1^2 + u_2^2 + \cdots + u_{2^m}^2) + (u_{2^m+1}^2 + \cdots + u_n^2) \tag{3.17}$$

The element $u_1^2 + u_2^2 + \cdots + u_{2^m}^2 \neq 0$ since $s(k) \geq 2^m$. Every ratio of sums of 2^m squares is again a sum of 2^m squares since $\langle 1, 1\rangle^{\otimes m}$ is a multiplicative form. Thus, from (3.17) we see that

$$\begin{aligned} 0 &= 1 + \frac{u_{2^m+1}^2 + \cdots + u_n^2 + 1}{u_1^2 + \cdots + u_{2^m}^2} \\ &= 1 + (v_1^2 + \cdots + v_{2^m}^2) \end{aligned}$$

Therefore, $-1 = v_1^2 + \cdots + v_{2^m}^2$ and $s(k) = 2^m$. $\square$

Remark 3.18. There exist fields with level 2^n for any $n \geq 1$. For instance, $\mathbb{R}(X_1, \ldots, X_{2^n})(\sqrt{-(X_1^2 + \cdots + X_{2^n}^2)})$ is a field of level 2^n.

Exercise 3.19. Let k be a p-adic field with residue field $\mathbb{F}_q$. Show,

(1) $s(k) = 1$ if q is odd and $q \equiv 1$ *mod* 4.
(2) $s(k) = 2$ if q is odd and $q \equiv -1$ *mod* 4.

4. The u-invariant

Definition 4.1. The **u-invariant** of a field k, denoted by $u(k)$, is defined to be the largest integer n such that every $n+1$ dimensional quadratic form over k is isotropic and there is an anisotropic form in dimension n over k.

$$u(k) = \max \{\dim(q) \ : \ q \text{ anisotropic form over } k\}.$$

If k admits an ordering then sums of nonzero squares is never zero and there is a refined u-invariant for fields with orderings, due to Elman-Lam [EL].

Example 4.2. (1) $u(\mathbb{F}_q) = 2$.
(2) $u(k(X)) = 2$, if k is algebraically closed and X is a curve over k (Tsen's theorem).
(3) $u(k) = 4$ for k a p-adic field
(4) $u(k) = 4$ for k a totally imaginary number field. This follows from the Hasse-Minkowski Theorem.
(5) Suppose $u(k) = n < \infty$. Let $k((t))$ denote the field of Laurent series over k. Then $u(k((t))) = 2n$. In fact, the square classes in $k((t))^*$ are $\{u_\alpha, tu_\alpha\}_{\alpha \in I}$ where $\{u_\alpha\}_{\alpha \in I}$ are the square classes in k^*. As in the p-adic field case, every form over $k((t))$ is isometric to $\langle u_1, \ldots, u_r\rangle \perp t\langle v_1, \ldots, v_s\rangle$, $u_i, v_i \in k^*$ and this form is aniostropic if and only if $\langle u_1, \ldots, u_r\rangle$ and $\langle v_1, \ldots, v_s\rangle$ are anisotropic (cf., project notes)
(6) More generally, if K is a complete discrete valuated field with residue field κ of u-invariant n, then $u(K) = 2n$.

Definition 4.3. A field k is $\boldsymbol{C_i}$ if every homogeneous polynomial in N variables of degree d with $N > d^i$ has a nontrivial zero.

Example 4.4. Finite fields and function fields in one variable over algebraically closed fields are C_1.

If k is a C_i field, $u(k) \leq 2^i$. Further, the property C_i behaves well with respect to function field extensions. If $l|k$ is finite and k is C_i then l is C_i; further, if $t_1, \ldots, t_n$ are indeterminates, $k(t_1, \ldots, t_n)$ is C_{i+n}.

Example 4.5. The u-invariant of transcendental extensions:
(1) $u(k(t_1, \ldots, t_n)) = 2^n$ if k is algebraically closed. In fact,
$$u(k(t_1, \ldots, t_n)) \leq 2^n$$
since $k(t_1, \ldots, t_n)$ is a C_n field. Further, the form
$$\langle\langle t_1, \ldots, t_n\rangle\rangle = \langle 1, t_1\rangle \otimes \cdots \otimes \langle 1, t_n\rangle$$
is anisotropic over $k((t_1))((t_2))\ldots((t_n))$ and hence also over $k(t_1, \ldots, t_n)$ (cf., Springer's Theorem, project notes)
(2) $u(\mathbb{F}_q(t_1, \ldots, t_n)) = 2^{n+1}$.

All fields of known u-invariant in the 1950's happened to have u-invariant a power of 2. Kaplansky raised the question whether the u-invariant of a field is always a power of 2.

Proposition 4.6. *The u-invariant does not take the values* $3, 5, 7$.

Proof. Let q be an anisotropic form of dimension 3. By scaling, we may assume that $q \cong \langle 1, a, b\rangle$. Then the form $\langle 1, a, b, ab\rangle$ is anisotropic; if $\langle 1, a, b, ab\rangle$ is isotropic, since discriminant is one, it is hyperbolic and Witt's cancellation yields $\langle a, b, ab\rangle \cong \langle 1, -1, -1\rangle$ is isotropic and $q \cong a\langle a, b, ab\rangle$ is isotropic leading to a contradiction. Thus $u(k) \neq 3$.

Let $u(k) < 8$. Let $q \in I^2(k)$ be any quadratic form. Every 3 fold Pfister form (which has dimension 8) is isotropic and hence hyperbolic. Thus $I^3(k)$ which is generated by 3 fold Pfister forms is zero. For any $c \in k^*$, $\langle 1, -c\rangle.q \in I^3(k)$ is zero and $c.q$ is Witt equivalent to q, hence isometric to q by Witt's cancellation. We conclude that every quadratic form whose class is in $I^2(k)$ is universal.

Suppose $u(k) = 5$ or 7. Let q be an anisotropic form of dimension $u(k)$. Since every form in dimension $u(k)+1$ is isotropic, if $\mathrm{disc}(q) = d$, $q \perp -d$ is isotropic and therefore q represents d. We may write $q \cong q_1 \perp \langle d\rangle$ where q_1 is even dimensional with trivial discriminant. Hence $[q_1] \in I^2(k)$ so that q_1 is universal. This in turn implies that $q_1 \perp \langle d\rangle \cong q$ is isotropic, leading to a contradiction. □

In the 1990's Merkurjev [M2] constructed examples of fields k with $u(k) = 2n$ for any $n \geq 1$, $n = 6$ being the first open case. Since then, it has been shown that the u-invariant does not need to be even. In [I], Izhboldin proves there exist fields k with $u(k) = 9$ and in [V] Vishik has shown that there exist fields k with $u(k) = 2^r + 1$ for all $r \geq 3$.

Merkurjev's construction yields fields k which are not of arithmetic type, i.e., not finitely generated over a number field or a p-adic field. It is still an interesting question whether $u(k)$ is a power of 2 if k is of arithmetic type.

The behaviour of the u-invariant is very little understood under rational function field extensions. For instance, it is an open question if $u(k) < \infty$ implies $u(k(t)) < \infty$ for the rational function field in one variable over k. This is not known even for $k = \mathbb{Q}_2$ and was unknown for $\mathbb{Q}_p$, p odd until late 90's. Conjecturally, $u(\mathbb{Q}_p(t)) = 8$, in analogy with the positive characteristic local field case, $u(\mathbb{F}_p((X))(t)) = 8$.

We indicate some ways of bounding the u-invariant of a field k once we know how efficiently the Galois cohomology groups $H^n(k, \mathbb{Z}/2\mathbb{Z})$ are generated by symbols for all n.

We set

$$H^n_{dec}(k, \mathbb{Z}/2\mathbb{Z}) = \{(a_1).\cdots.(a_n),\ a_i \in k^*\}$$

and call elements in this set symbols. By Voevodsky's theorem on Milnor conjecture, $H^n(k, \mathbb{Z}/2\mathbb{Z})$ is additively generated by $H^n_{dec}(k, \mathbb{Z}/2\mathbb{Z})$.

Proposition 4.7. *Let k be a field such that $H^{n+1}(k, \mathbb{Z}/2\mathbb{Z}) = 0$ and for $2 \leq i \leq n$, there exist integers N_i such that every element in $H^i(k, \mathbb{Z}/2\mathbb{Z})$ is a sum of N_i symbols. Then $u(k)$ is finite.*

Proof. Let q be a quadratic form over k of dimension m and discriminant d. Let $q_1 = \langle d \rangle$ if m is odd and $\langle 1, -d \rangle$ if m is even. Then $q \perp -q_1$ has even dimension and trivial discriminant. Hence $q \perp -q_1 \in I^2(k)$. Let $e_2(q \perp -q_1) = \sum_{j \leq N_2} \xi_{2j}$ where $\xi_{2j} \in H^2_{dec}(k, \mathbb{Z}/2\mathbb{Z})$. Let ϕ_{2j} be 2 fold Pfister forms such that $e_2(\phi_{2j}) = \xi_{2j}$. Then $q_2 = \sum_{j \leq N_2} \phi_{2j}$ has dimension at most $4N_2$ and $e_2(q \perp -q_1 \perp -q_2) = 0$ and $q \perp -q_1 \perp -q_2 \in I^3(k)$, by Merkurjev's Theorem. Repeating this process and using Milnor Conjecture, we get $q_i \in I^i(k)$ which is a sum of N_i i-fold Pfister forms and $q - \sum_{1 \leq i \leq n} q_i \in I^{n+1}(k) = 0$, since $H^{n+1}(k, \mathbb{Z}/2\mathbb{Z}) = 0$. Thus $[q] = \sum_{1 \leq i \leq n} q_i$ and $\dim(q_{an}) \leq \sum_{1 \leq i \leq n} 2^i N_i$. Thus $u(k) \leq \sum_{1 \leq i \leq n} 2^i N_i$. $\square$

Definition 4.8. A field k is said to have **cohomological dimension at most $\boldsymbol{n}$** if $H^i(k, M) = 0$ for $i \geq n+1$ for all finite discrete Γ_k modules M (cf. [Se] §3).

Example 4.9. Finite fields and function fields in one variable over algebraically closed fields have cohomological dimension 1. Totally imaginary number fields and p-adic fields are of cohomological dimension 2. Thus if k is a p-adic field, and $k(X)$ a function field in one variable over k, $cd(k(X)) \leq 3$. In particular, $H^4(k(X), \mathbb{Z}/2\mathbb{Z}) = 0$.

Theorem 4.10. *(Saltman) Let k be a non-dyadic p-adic field and $k(X)$ a function field in one variable over k. Every element in $H^2(k(X), \mathbb{Z}/2\mathbb{Z})$ is a sum of two symbols.*

Theorem 4.11. *(Parimala-Suresh) Let $k(X)$ be as in the previous theorem. Then every element in $H^3(k(X), \mathbb{Z}/2\mathbb{Z})$ is a symbol.*

Corollary 4.12. *For $k(X)$ as above, $u(k(X)) \leq 2 + 8 + 8 = 18$.*

It is not hard to show from the above theorems that $u(k(X)) \leq 12$. With some further work it was proved in [PS1] that $u(k(X)) = 10$. More recently in [PS2] the estimated value $u(k(X)) = 8$ was proved. For an alternate approach to $u(k(X)) = 8$, we refer to ([HH], [HHK], [CTPS]).

5. Sums of squares

An additional reference for sums of squares is given by H. Cohen at http://www.math.u-bordeaux1.fr/∼cohen/Cohensquares.pdf which is a translation of the original paper [C].

Definition 5.1. An element $f \in \mathbb{R}(X_1, \ldots, X_n)$ is called **positive semi-definite** if $f(a) \geq 0$ for all $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ where f is defined.

Hilbert's seventeenth problem:
Let $\mathbb{R}(X_1, \ldots, X_n)$ be the rational function field in n variables over the field $\mathbb{R}$ of real numbers. Hilbert's seventeenth problem asks whether every positive semi-definite $f \in \mathbb{R}(X_1, \ldots, X_n)$ is a sum of squares in $\mathbb{R}(X_1, \ldots, X_n)$. E. Artin settled this question in the affirmative and Pfister gave an effective version of Artin's result (cf. [Pf], chapter 6).

Theorem 5.2 (Artin,Pfister)**.** *Every positive semi-definite function $f \in \mathbb{R}(X_1, \ldots, X_n)$ can be written as a sum of 2^n squares in $\mathbb{R}(X_1, \ldots, X_n)$.*

For $n \leq 2$ the above was due to Hilbert himself. If one asks for expressions of positive definite polynomials in $\mathbb{R}[X_1, \ldots, X_n]$ as sums of 2^n squares in $\mathbb{R}[X_1, \ldots, X_n]$, there are counter examples for $n = 2$; the Motzkin polynomial

$$f(X_1, X_2) = 1 - 3X_1^2X_2^2 + X_1^4X_2^2 + X_1^2X_2^4$$

is positive semi-definite but not a sum of 4 squares in $\mathbb{R}[X_1, X_2]$ ([CEP], [CT]). In fact, Pfister's result has the following precise formulation.

Theorem 5.3 (Pfister)**.** *Let $\mathbb{R}(X)$ be a function field in n variables over $\mathbb{R}$. Then every n fold Pfister form in $\mathbb{R}(X)$ represents every sum of squares in $\mathbb{R}(X)$.*

We sketch a proof of this theorem below.

Proposition 5.4 (Pure Subform Theorem)**.** *Let k be any field and ϕ an anisotropic n-fold Pfister form over k. Let $\phi = 1 \perp \phi'$. If b_1 is any value of ϕ', then $\phi \cong \langle\langle b_1, \ldots, b_n \rangle\rangle$.*

Proof. The proof is by induction on n; for $n = 1$ the statement is clear. Let $n > 1$. We assume the statement holds for all $(n-1)$-fold Pfister forms. let $\phi = \langle\langle a_1, \ldots, a_n \rangle\rangle$, $\psi = \langle\langle a_1, \ldots, a_{n-1} \rangle\rangle$, and let ϕ', ψ' denote the pure subforms of ϕ and ψ repsectively. We have $\phi = \psi \perp a_n\psi$, $\phi' = \psi' \perp a_n\psi$. Let b_1 be a value of ϕ'. We may write $b_1 = b_1' + a_nb$, with b_1' a value of ψ' and b a value of ψ. The only nontrivial case to discuss is when $b \neq 0$ and $b'_1 \neq 0$. By induction we may assume $\psi \cong \langle\langle b'_1, b_2, \ldots, b_{n-1}\rangle\rangle$ and $b\psi \cong \psi$. We thus have

$$\begin{aligned} \phi &\cong \langle\langle b'_1, b_2, \ldots, b_{n-1}, a_n\rangle\rangle \cong \langle\langle b'_1, b_2, \ldots, b_{n-1}, a_n b\rangle\rangle \\ &\cong \langle\langle b'_1, a_n b\rangle\rangle \otimes \langle\langle b_2, \ldots, b_{n-1}\rangle\rangle \end{aligned}$$

Since $b_1 = b'_1 + a_n b$, $\langle b'_1, a_n b\rangle \cong \langle b_1, b_1 b'_1 a_n b\rangle$ and we have

$$\begin{aligned} \langle\langle b'_1, a_n b\rangle\rangle &= \cong \langle 1, b'_1, a_n b, a_n b b'_1\rangle \\ &= \langle 1, b_1, b_1 b'_1 a_n b, a_n b b'_1\rangle \\ &= \langle\langle b_1, c_1\rangle\rangle, \end{aligned}$$

$$\phi \cong \langle\langle b_1, c_1, b_2, \cdots, b_{n-1}\rangle\rangle.$$

□

Proof of Pfister's theorem. Let ϕ be an anisotropic n fold Pfister form over $K = \mathbb{R}(X)$. Let $b = b_1^2 + \cdots + b_m^2$, $b_i \in K$. We show that ϕ represents b by induction on m. For $m = 1$, b is a square and is represented by ϕ. Suppose $m = 2$, $b = b_1^2 + b_2^2$, $b_1 \neq 0$, $b_2 \neq 0$. The field $K(\sqrt{-1})$ is a function field in n variables over $\mathbb{C}$ and is C_n. Then ϕ is universal over $K(\sqrt{-1})$ and hence represents $\beta = b_1 + ib_2$. Let $v, w \in K^{2^n}$ such that $\phi_{K(\sqrt{-1})}(v + \beta w) = \beta$. Hence

$$\phi(v) + \beta^2\phi(w) + \beta(2\phi(v, w) - 1) = 0.$$

The irreducible polynomial of β over K is

$$\phi(w)X^2 + (2\phi(v, w) - 1)X + \phi(v)$$

and hence $N(\beta) = b = \frac{\phi(v)}{\phi(w)}$ is a value of ϕ since ϕ is multiplicative.

Suppose $m > 2$. We argue by induction on m. Suppose ϕ represents all $m-1$ sums of squares. Let b be a sum of m squares. After scaling b by a square, we may assume that $b = 1 + c$, $c = c_1^2 + \cdots + c_{m-1}^2$, $c \neq 0$. Let $\phi \cong 1 \perp \phi'$. By induction hypothesis, ϕ represents c. Let $c = c_0^2 + c'$, c' a value of ϕ'. Let $\psi = \phi \otimes \langle 1, -b\rangle$ and $\psi = 1 \perp \psi'$ with $\psi' = \langle -b\rangle \perp \phi' \perp -b\phi'$. The form ψ' represents $c' - b = (c - c_0^2) - (1 + c) = -1 - c_0^2$. Thus, by the Pure Subform Theorem,

$$\psi \cong \langle\langle -1 - c_0^2, d_1, \ldots, d_n\rangle\rangle = \langle 1, -1 - c_0^2\rangle \otimes \langle\langle d_1, \ldots, d_n\rangle\rangle.$$

By induction, the n fold Pfister form $\langle\langle d_1, \ldots, d_n\rangle\rangle$ represents $1 + c_0^2$ which is a sum of 2 squares; thus ψ is isotropic, hence hyperbolic. Thus $\phi \cong b\phi$ represents b. □

Corollary 5.5. *Let $K = \mathbb{R}(X)$ be a function field in n variables over $\mathbb{R}$. Then every sum of squares in K is a sum of 2^n squares.*

Proof. Set $\phi = \langle 1, 1\rangle^{\otimes n}$ in the above theorem. $\square$

Definition 5.6. The **Pythagoras number** $p(k)$ of a field k is the least positive integer n such that every sum of squares in k^* is a sum of at most n squares.

Example 5.7. If $\mathbb{R}$ is the field of real numbers, $p(\mathbb{R}) = 1$.

Example 5.8. If $\mathbb{R}(X_1, \ldots, X_n)$ is a function field in n variables over $\mathbb{R}$ by Pfister's theorem $p(\mathbb{R}(X_1, \ldots, X_n)) \leq 2^n$.

5.1. **Effectiveness of the bound** $p(\mathbb{R}(X)) \leq 2^n$**.** Let

$$K = \mathbb{R}(X_1, \ldots, X_n)$$

be the rational function field in n variables over $\mathbb{R}$. For $n = 1$ the bound is sharp. For $n = 2$ the Motzkin polynomial

$$f(X_1, X_2) = 1 - 3X_1^2X_2^2 + X_1^4X_2^2 + X_1^2X_2^4$$

is positive semi definite; Cassels-Ellison-Pfister [CEP] show that this polynomial is not a sum of three squares in $\mathbb{R}(X_1, X_2)$. Therefore $p(\mathbb{R}(X_1, X_2)) = 4$.

Lemma 5.9 (Key Lemma)**.** *Let k be a field and $n = 2^m$. Let $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n) \in k^n$ be such that $u.v = \sum_{1\leq i\leq n} u_iv_i = 0$. Then there exist $w_j \in k,\ 1 \leq j \leq n-1$ such that*

$$(\sum_{1\leq i\leq n} u_i^2)(\sum_{1\leq i\leq n} v_i^2) = (\sum_{1\leq j\leq n-1} w_j^2).$$

Proof. Let $\lambda = \sum_{1\leq i\leq n} u_i^2$, $\mu = \sum_{1\leq i\leq n} v_i^2$. We may assume without loss of generality that $u \neq 0$ and $v \neq 0$. The elements λ and μ are values of $\phi_m = \langle 1, 1\rangle^{\otimes m}$ and $\lambda\phi_m \cong \phi_m$ and $\mu\phi_m \cong \phi_m$. We choose isometries $f : \lambda\phi_m \cong \phi_m$, $g : \mu\phi_m \cong \phi_m$ such that $f(1, 0, \ldots, 0) = u$ and $g(1, 0, \ldots, 0) = v$. If U and V are matrices representing f, g respectively, we have

$$UU^t = \lambda^{-1},\ \ VV^t = \mu^{-1},\ \ \lambda^{-1}\mu^{-1} = \lambda^{-1}VV^t = (VU^t)(VU^t)^t.$$

The first row of VU^t is of the form $(0, w_2, \ldots, w_n)$ since $u.v = 0$. Thus $\lambda^{-1}\mu^{-1} = \sum_{2\leq i\leq n} w_i^2$. $\square$

Corollary 5.10. *Let k be an ordered field with $p(k) = n$. Then $p(k(t)) \geq n + 1$.*

Proof. Let $\lambda \in k^*$ be such that λ is a sum of n squares and not a sum of less than n squares. Suppose $\lambda + t^2$ is a sum of n squares in $k(t)$. By Cassels-Pfister theorem,

$$\lambda + t^2 = (\mu_1 + \nu_1 t)^2 + \cdots + (\mu_n + \nu_n t)^2$$

with $\mu_i, \nu_i \in k^*$. If $u = (\mu_1, \ldots, \mu_n)$, $v = (\nu_1, \ldots, \nu_n)$, then $u.v = 0$, $\sum_{1\leq i\leq n} \mu_i^2 = \lambda$, $\sum_{1\leq i\leq n} \nu_i^2 = 1$. Thus $\lambda = (\sum_{1\leq i\leq n} \mu_i^2)(\sum_{1\leq i\leq n} \nu_i^2)$ is a sum of $n-1$ squares by the key lemma 5.9 contradicting the choice of λ. ☐

Corollary 5.11. *$p(\mathbb{R}(X_1, \ldots, X_n)) \geq n+2$. Thus*

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

Proof. By [CEP], we know that $p(\mathbb{R}(X_1, X_2)) = 4$. The fact that $n+2 \leq p(\mathbb{R}(X_1, \ldots, X_n)$ now follows by Corollary 5.10 and induction. ☐

Remark 5.12. It is open whether $p(\mathbb{R}(X_1, X_2, X_3)) = 5, 6, 7$ or 8.

Remark 5.13. The possible values of the Pythagoras number of a field have all been listed ([H], [Pf, pg. 97]).

Proposition 5.14. *If k is a non-formally real field, $p(k) = s(k)$ or $s(k) + 1$.*

Proof. If $s(k) = n$, -1 is not a sum of less than n squares, so that $p(k) \geq s(k)$. For $a \in k^*$,

$$a = (\frac{a+1}{2})^2 + (-1)(\frac{a-1}{2})^2$$

is a sum of $n+1$ squares if -1 is a sum of n squares. Thus $p(k) \leq s(k) + 1$. ☐

Let k be a p-adic field and $K = k(X_1, \ldots, X_n)$ a rational function field in n variables over k. Then $s(k) = 1, 2$ or 4 so that $s(K) = 1, 2$, or 4. Thus $p(K) \leq 5$ (in fact it is easy to see that if $s(k) = s$, $p(K) = s + 1$).

Thus we have bounds for $p(k(X_1, \ldots, X_n))$ if k is the field of real or complex numbers or the field of p adic numbers. The natural questions concern a number field k.

Conjecture. (Pfister) Let k be a number field and $F = k(X)$ a function field in d variables over k. Then

(1) for $d = 1$, $p(F) \leq 5$.
(2) for $d \geq 2$, $p(F) \leq 2^{d+1}$.

Let k be a number field and $F = k(t)$ the rational function field in one variable over k. In this case $p(k(t)) = 5$ is a theorem [La], [Po]. The fact that $p(k(t)) \leq 8$ can be easily deduced from the following injectivity in the Witt groups ([CTCS], Prop. 1.1):

$$W(k(t)) \longrightarrow \prod_{w \in \Omega(k)} W(k_w(t)),$$

with $\Omega(k)$ denoting the set of places of k. In fact, if $f \in k(t)$ is a sum of squares, f is a sum of at most two squares in $k_w(t)$ for a real place w, by Pfister's theorem (which in the case of function fields of curves goes back to Witt). Further, for a finite place w of k or a complex place, $\langle 1, 1\rangle^{\otimes 3} = 0$ in $W(k_w)$. Thus $\langle 1, 1\rangle^{\otimes 3} \otimes \langle 1, -f\rangle$ is hyperbolic over $k_w(t)$ for all $w \in \Omega(k)$.

By the above injectivity, this form is hyperbolic over $k(t)$, leading to the fact that f is a sum of at most eight squares in $k(t)$.

For a general function field $k(X)$ in one variable over k, $(d = 1)$, the best possible known result is due to F. Pop, $p(F) \leq 6$ [P]. We sketch some results and conjectures from the arithmetic side which could lead to a solution of the conjecture for $d \geq 2$ (see Colliot-Thélène, Jannsen [CTJ] for more details).

For any field k, by Voevodsky's theorem, we have an injection

$$e_n : P_n(k) \to H^n(k, \mathbb{Z}/2\mathbb{Z}).$$

In fact, for any field k, if ϕ_1, $\phi_2 \in P_n(k)$ have the same image under e_n then $\phi_1 \perp -\phi_2 \in \ker(e_n) = I^{n+1}(k)$. Then in $W(k)$, the classes $\phi_1 \perp -\phi_2 = \phi_1' \perp -\phi_2'$ where ϕ_1' and ϕ_2' are the pure subforms of ϕ_1 and ϕ_2. Moreover, $\dim(\phi_1' \perp -\phi_2')_{\mathrm{an}} \leq 2^{n+1} - 2 < 2^{n+1}$. By the Arason-Pfister Haupt-Satz, (Theorem 3.1), anisotropic forms in $I^{n+1}(k)$ must have dimension at least 2^{n+1}, therefore $\phi_1 = \phi_2$.

Let k be a number field and $F = k(X)$ be a function field in d variables over k. Let $f \in F$ be a function which is a sum of squares in F. One would like to show that f is a sum of 2^{d+1} squares. Let $\phi_{d+1} = \langle 1, 1\rangle^{\otimes(d+1)}$ and $q = \phi_{d+1} \otimes \langle 1, -f\rangle$. This is a $d+1$ fold Pfister form and ϕ_{d+1} represents f if and only if q is hyperbolic or equivalently, by the injectivity of e_n above, $e_{d+2}(\phi_{d+1} \otimes \langle 1, -f\rangle) = 0$.

We look at this condition locally at all completions k_v at places v of k. Let $k_v(X)$ denote the function field of X over k_v. (We may assume that X is geometrically integral). Let v be a complex place. The field $k_v(X)$ has cohomological dimension d so that $H^m(k_v(X), \mathbb{Z}/2\mathbb{Z}) = 0$ for $m \geq d + 1$. Hence $e_{d+2}(\phi_{d+1} \otimes \langle 1, -f\rangle) = 0$ over $k_v(X)$. Let v be a real place. Then over $k_v(X)$, f is a sum of squares, hence a sum of at most 2^d squares (by Pfister's Theorem 5.3) so that $\phi_d \otimes \langle 1, -f\rangle$ is hyperbolic over $k_v(X)$. Hence $e_{d+2}(\phi_{d+1} \otimes \langle 1, -f\rangle) = 0$.

Let v be a non-dyadic p-adic place of k. Then ϕ_2 is hyperbolic over k_v so that $\phi_{d+1} \otimes \langle 1, -f\rangle = 0$ and $e_{d+2}(\phi_{d+1} \otimes \langle 1, -f\rangle) = 0$.

Let v be a dyadic place of k. Over k_v, ϕ_3 is hyperbolic so that $e_{d+2}(\phi_{d+1} \otimes \langle 1, -f\rangle) = 0$. Thus for all completions v of k, $e_{d+2}(\phi_{d+1} \otimes \langle 1, -f\rangle)$ is zero. The following conjecture of Kato implies Pfister's conjecture for $d \geq 2$.

Conjecture. (Kato) Let k be a number field, X a geometrically integral variety over k of dimension d. Then the map

$$H^{d+2}(k(X), \mathbb{Z}/2\mathbb{Z}) \to \prod_{v \in \Omega_k} H^{d+2}(k_v(X), \mathbb{Z}/2\mathbb{Z})$$

has trivial kernel.

The above conjecture is the classical Hasse-Brauer-Noether theorem if dimension of X is zero, i.e., the injectivity of the Brauer group map:

$$Br(k) \hookrightarrow \bigoplus_{v \in \Omega_k} Br(k_v).$$

For $\dim X = 1$, the conjecture is a theorem of Kato ([K]). For $\dim X = 2$, Kato's conjecture was proved by Jannsen ([Ja]). Using Jannsens's theorem Colliot-Thélène-Jannsen [CTJ] derived Pfister's conjecture: every sum of squares in $k(X)$, X a surface over a number field, is a sum of at most 8 squares.

The case X is a curve is more delicate. We explain how Kato's theorem was used by Colliot-Thélène to derive $p(k(X)) \leq 7$ for a curve X over a number field.

Suppose $K = k(X)$ has no ordering. We claim that $s(K) \leq 4$. To show this it suffices to show that $\langle 1, 1\rangle^{\otimes 3}$ is zero over $k_v(X)$ for every place v of k. At finite places v, $\langle 1, 1\rangle^{\otimes 3}$ is already zero in k_v. If v is a real place of k, $k_v(X)$ is the function field of a real curve over the field of real numbers which has no orderings. By a theorem of Witt, $Br(k_v(X)) = 0$ and every sum of squares is a sum of two squares in $k_v(X)$. Thus -1 is a sum of two squares in $k_v(X)$ and $\langle 1, 1\rangle^{\otimes 3} = 0$ over $k_v(X)$. Since $H^3(k(X), \mathbb{Z}/2\mathbb{Z}) \to \prod_{v \in \Omega_k} H^3(k_v(X), \mathbb{Z}/2\mathbb{Z})$ is injective by Kato's theorem, $e_3(\langle 1, 1\rangle^{\otimes 3}) = 0$ in $H^3(k(X), \mathbb{Z}/2\mathbb{Z})$. Since e_3 is injective on 3 fold Pfister forms, $\langle 1, 1\rangle^{\otimes 3} = 0$ in $k(X)$. Thus $s(k(X)) \leq 4$. In this case, $p(k(X)) \leq 5$.

Suppose K has an ordering. Let $f \in K^*$ be a sum of squares in K. Then $K(\sqrt{-f})$ has no orderings and hence -1 is a sum of 4 squares in

$K(\sqrt{-f})$. Let $a_i, b_i \in K$ be such that

$$-1 = \sum_{1\leq i\leq 4} (a_i + b_i\sqrt{-f})^2, \quad a_i, b_i \in K.$$

Then

$$1 + \sum_{1\leq i\leq 4} a_i^2 = f(\sum_{1\leq i\leq 4} b_i^2), \quad \sum_{1\leq i\leq 4} a_i b_i = 0.$$

By the Key Lemma, 5.9, $(1+\sum_{1\leq i\leq 4} a_i^2). \sum_{1\leq i\leq 4} b_i^2$ is a sum of at most 7 squares.

References

[A] Albert, A.A., *Normal division algebras of degree four over an algebraic field, Trans. Amer. Math. Soc.*, **34** (1932), 363-372.

[Ar] Arason, J. K., *Cohomologische Invarianten quadratischer Formen, J. Algebra*, **36** (1975), 448-491.

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

[C] Cohen, H., *Représentations comme sommes de carrés, Séminaire de Théorie des Nombres, 1971–1972 (Univ. Bordeaux I, Talence), Exp. No. 21*, (1972). English translation available at http://www.math.u-bordeaux1.fr/∼cohen/Cohensquares.pdf

[CT] Colliot-Thélène, J.-L., *The Noether-Lefschetz theorem and sums of 4 squares in the rational function field* $\mathbb{R}(x, y)$*, Compos.Math.* **86** (1993) 235-243.

[CTCS] Colliot-Thélène, J.-L., Coray, D., Sansuc, J.-J, *Descente et principe de Hasse pour certaines variétés rationnelles, J. Reine Angew. Math.*, **320**, (1980), pg. 150–191.

[CTJ] Colliot-Thélène, J.-L., Jannsen, U., *Sommes de carrès dans les corps de fonctions. C.R. Acad.Paris, Sèr. I* **312** (1991), 759-762.

[CTPS] Colliot-Thélène, J.-L., Parimala, R., Suresh, V., *Patching and local-global principles for homogeneous spaces over function fields of p-adic curves*, preprint arXiv:0812.3099.

[CTSaSw] Colliot-Thélène, J.-L., Sansuc, J.-J., Swinnerton-Dyer, P., *Intersections of two quadrics and Châtelet surfaces. II, J. Reine Angew. Math.*, **374**, (1987), pp. 72 - 168.

[EKM] Elman, R., Karpenko, N., Merkurjev, A., *The algebraic and geometric theory of quadratic forms, American Mathematical Society Colloquium Publications*, **56** (2008).

[EL] Elman, R. and Lam,T.-Y., *Quadratic forms and the u-invariant I, Math.Z.* **131** (1973), 238-304.

[H] Hoffmann, D.W., *Pythagoras numbers of fields, J. Amer. Math. Soc.*, **12**, (1999), 3, 839–848.

[HH] Harbater, D., Hartmann, J., *Patching over fields*, preprint arXiv:0710.1392, to appear in Israel Journal of Mathematics.

[HHK] Harbater, D., Hartmann, J., Krashen, D., *Applications of patching to quadratic forms and central simple algebras*, preprint arXiv:0809.4481.

[I] Izhboldin, Oleg T., *Fields of u-invariant* 9. *Ann. of Math. (2)***154** (2001).

[Ja] Jannsen, U., *Principe de Hasse cohomologique, Séminaire de Théorie des Nombres, Paris, 1989–90, Progr. Math., Birkhäuser Boston*, **102**, 121–140 (1992)

[K] Kato, Kazuya, *A Hasse principle for two-dimensional global fields, with an appendix by Jean-Louis Colliot-Thélène, J. Reine Angew. Math.* **366**, (1986).

[L] Lam, T.-Y., *Introduction to Quadratic forms over fields, GSM, AMS* **67** (2004).

[La] Landau, E., *Über die Darstellung definiter Funktionen durch Quadrate, Math. Ann.*, **62**,(1906), 2, pg.272–285.

[Li] Lichtenbaum, S., *Duality theorems for curves over p-adic fields, Invent. Math.* **7** (1969), 120-136.

[OVV] Orlov, D., Vishik, A., and Voevodsky, V., *An exact sequence for $K_*^M/2$ with applications to quadratic forms, Annals of Math.* **165** (2007), 1-13.

[M1] Merkurjev, A. S. *On the norm residue symbol of degree* 2. *Dokl. Akad. Nauk SSSR* **261** (1981), 542–547.

[M2] Merkurjev, A. S., *S*imple algebras and quadratic forms. Izv. Akad. Nauk SSSR Ser. Mat. **55** (1991), 218–224; *translation in Math. USSR-Izv. 38 (1992), no. 1, 215–221.*

[Pf] Pfister, A., *Quadratic forms with applications to algebraic geometry and topology, London Mathematical Society Lecture Note Series, Cambridge University Press*, **217** (1995)

[P] Pop, F., *Summen von Quadraten in arithmetischen Funktionenkorpern*, preprint (http://www.math.upenn.edu/ pop/Research/Papers.html)

[PS1] Parimala, R. and Suresh, V., *Isotropy of quadratic forms over function fields in one variable over p-adic fields, Publ. de I.H.É.S.* **88** (1998), 129-150.

[PS2] Parimala, R. and Suresh, V., *The u-invariant of the function fields of p-adic curves*, arxiv:0708.3128v1

[Po] Pourchet, Y.,*Sur la représentation en somme de carrés des polynômes à une indéterminée sur un corps de nombres algébriques, Acta Arith., Polska Akademia Nauk. Instytut Matematyczny. Acta Arithmetica*, **19**, (1971), 89–104.

[RST] Rowen, L.H., Sivatski, A.S., Tignol, J.-P., *Division algebras over rational function fields in one variable, in Algebra and Number Theory, 158 - 180, Hindustan Book Agency, Dheli* (2005).

[Sa] Saltman, D., *Division algebras over p-adic curves with an appendix by William Jacob and J.-P. Tignol, Journal of Ramanujan Math. Soc.* **12**, pp 25 - 47, (1997).

[S] Scharlau, W., *Quadratic and Hermitian forms, Springer-Verlag* **270** (1985).

[Se] Serre, J-P., *Galois cohomology, Springer-Verlag* (1997).

[V] Vishik, A., *Fields of u-invariant 2^r+1. Algebra, Arithmetic and Geometry - Manin Festschrift, Birkhauser*, (2007), in press.