Polynomial equivalence
An application of Witt's theorem to decide rational quadratic form equivalence

Lars A. Wallenborn
lars@wallenborn.net

09./12. July 2012

Abstract

This is the handout for a seminar talk on 09. and 12. July 2012 given in the seminar on "Algorithms in Real Algebraic Geometry" at the Mathematical Institute of the University Bonn that was organized by Prof. Nitin Saxena. In my talk I describe a well-known algorithm for deciding quadratic form equivalence over different interesting fields by using Witt's theorem. Here "form equivalence" means that two forms are equivalent if and only if there exists an invertible linear transformation on the variables such that one of the forms becomes equal to the other. The desciription of the algorithm to decide this problem was also given in [AS06b] and the proof of Witt's theorem is based on the proof from [Ser73].

Contents

Algorithms

1 Introduction

1.1 Basic Definitions

Definition 1.1 (Polynomial Equivalence)**.** Let $\mathbb{K}/\mathbb{F}$ be a field extension. Two polynomials $f, g \in \mathbb{F}[x_1, \ldots, x_n]$ are said to be **equivalent over** $\mathbb{K}$ if there exists an invertible linear transformation τ sending each x_i to a lineare combination of the $x_1, \ldots, x_n$ with coefficients in $\mathbb{K}$:

$$f(\tau(x_1), \ldots, \tau(x_n)) = g(x_1, \ldots, x_n).$$

We then write $f \sim_{\mathbb{K}} g$. If $\mathbb{K} = \mathbb{F}$, we simply write $f \sim g$ and say that f and g are **equivalent**.

Remark 1.2. An invertible linear transformation τ on the variables of a polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ can also be expressed by a matrix $A \in \mathrm{Gl}_n(\mathbb{K})$ acting on $\mathbb{F}[x_1, \ldots, x_n]$. So abbusing notation a little bit, we can say that

$$f \sim_{\mathbb{K}} g \Leftrightarrow \exists A \in \mathrm{Gl}_n(\mathbb{K}) : f \circ A = g.$$

In this case, we say that f **is equivalent to** g **via** A.

Fact 1.3. *The equivalence of polynomials is indeed an equivalence relation.*

Proof. For all $f, g, h \in \mathbb{F}[x_1, \ldots, x_n]$:

- **Reflexivity:** $f \sim_{\mathbb{K}} f$ via the identity matrix in $\mathrm{Gl}_n(\mathbb{K})$.
- **Symmetry:** If $f \sim_{\mathbb{K}} g$ via $A \in \mathrm{Gl}_n(\mathbb{K})$ then $g \sim_{\mathbb{K}} f$ via A^{-1}.
- **Transitivity:** If $f \sim_{\mathbb{K}} g$ via $A \in \mathrm{Gl}_n(\mathbb{K})$ and $g \sim_{\mathbb{K}} h$ via $B \in \mathrm{Gl}_n(\mathbb{K})$, then $f \sim_{\mathbb{K}} h$ via $B \cdot A \in \mathrm{Gl}_n(\mathbb{K})$.

□

Example 1.4. Let $f(x, y) = x^2 + y^2$ and $g(x, y) = 2x^2 + 2y^2$ be polynomials over $\mathbb{Q}$. The map

$$\tau : \begin{cases} x \mapsto x + y \\ y \mapsto x - y \end{cases}$$

is an invertible linear transformation as in the above definition 1.1 and $\tau \circ f = g$, so $f \sim g$ over rationals. We could also say, that

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \in \mathrm{Gl}_n(\mathbb{Q})$$

is invertible and calculate

$$f\left(\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\right) = f\begin{pmatrix} x+y \\ x-y \end{pmatrix} = (x+y)^2 + (x-y)^2 = 2x^2 + 2y^2 = g(x, y)$$

Example 1.5. Consider $f, g \in \mathbb{Q}[x]$ with $f(x) = x^2$ and $g(x) = 2x^2$. Then f and g are not equivalent over $\mathbb{Q}$ but they are equivalent over $\mathbb{R}$ via $\tau : x \mapsto \sqrt{2}x$.

Notation 1.6. Denote by $[\mathbb{N}]$ the set of multi-indices and for $d \in \mathbb{N}$ by

$$[\mathbb{N}]^{=d} := \{\, \alpha \in \mathbb{N}^n \mid |\alpha|_1 = d \,\}$$

the set of multi-indices of norm d and

$$[\mathbb{N}]^{\leq d} := \{\, \alpha \in \mathbb{N}^n \mid |\alpha|_1 \leq d \,\}.$$

We denote the set of natural numbers from 0 respectively 1 to $n \in \mathbb{N}$ by

$$[n]_0 := [0, n] \cap \mathbb{Z} \quad \text{respectively} \quad [n] := [1, n] \cap \mathbb{Z}.$$

Definition 1.7 (Total Degree of a Polynomial)**.** Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a polynomial and write $f(x_1, \ldots, x_n) = \sum_{\alpha \in [\mathbb{N}]} a_\alpha x^\alpha$ where only finitely many $a_\alpha \neq 0$, then the **(total) degree of** f is given by

$$\deg(f) = \sup \{\, |\alpha| \mid a_\alpha \neq 0 \,\}.$$

Remark 1.8. Note that non-zero constants (elements of $\mathbb{F}^*$) have degree 0 and, since $\sup\{\emptyset\} := -\infty$, we have $\deg(0) = -\infty$.

Fact 1.9. *Equivalent polynomials have the same degree.*

Proof. Let f and g be equivalent polynomials via $A \in \mathrm{Gl}_n(\mathbb{K})$. So A replaces every variable by a linear combination of of x_i which does not change the degree. □

Definition 1.10. For $d \in \mathbb{N}_{>0}$ a polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ of the form

$$f(x_1, \ldots, x_n) = \sum_{\alpha \in [\mathbb{N}]^{=d}} a_\alpha x^\alpha$$

is called **homogeneous polynomial of degree** d or **form of degree** d. Furthermore define:

- $\mathbb{F}[x_1, \ldots, x_n]^{=d}$ the forms of degree d.
- $\mathbb{F}[x_1, \ldots, x_n]^{\leq d}$ the forms of degree at most d.

Definition 1.11 (Polynomials as input)**.** We assume every polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ with total degree d to be given in **expanded** form:

$$f(x_1, \ldots, x_n) = \sum_{0 \leq i_1 + \ldots + i_n \leq d} a_{i_1, \ldots, i_n} x_1^{i_1} \cdot \cdots \cdot x_n^{i_n}$$

with $a_{i_1, \ldots, i_n} \in \mathbb{F}$. This can also be written in a little bit more elegant way:

$$f(x_1, \ldots, x_n) = \sum_{\alpha \in [\mathbb{N}]^{\leq d}} a_\alpha x^\alpha$$

where $x = (x_1, \ldots, x_n)$ and $a_\alpha \in \mathbb{F}$.

Definition 1.12. We define the following decision problems

$$\textsc{PolyEquiv}_{d,\mathbb{F}} := \left\{ (f,g) \in \mathbb{F}[x_1,\ldots,x_n]^2 \;\middle|\; n \in \mathbb{N}, \deg(f) = d = \deg(g), f \sim g \right\}$$
$$\textsc{FormEquiv}_{d,\mathbb{F}} := \left\{ (f,g) \in \left(\mathbb{F}[x_1,\ldots,x_n]^{=d}\right)^2 \;\middle|\; n \in \mathbb{N}, f \sim g \right\}$$

and for shortage of notation also

$$\begin{aligned}
\textsc{QuadraticPolyEquiv}_{\mathbb{F}} &:= \textsc{PolyEquiv}_{2,\mathbb{F}} \\
\textsc{QuadraticFormEquiv}_{\mathbb{F}} &:= \textsc{FormEquiv}_{2,\mathbb{F}} \\
\textsc{CubicPolyEquiv}_{\mathbb{F}} &:= \textsc{PolyEquiv}_{3,\mathbb{F}} \\
\textsc{CubicFormEquiv}_{\mathbb{F}} &:= \textsc{FormEquiv}_{3,\mathbb{F}}
\end{aligned}$$

1.2 Connection to other problems

In this chapter I want to present some results about the complexity of the previously defined form equivalence problem and it's special cases – all without proofs. Since the algebra isomorphism problem plays a major role, we first define how we want to give an $\mathbb{F}$-algebra as input to an algorithm.

Definition 1.13 (Commutative $\mathbb{F}$-algebra)**.** A commutative ring containing a field $\mathbb{F}$ is called **commutative $\mathbb{F}$-algebras**.

Definition 1.14 ($\mathbb{F}$-algebras as input)**.** Let $\mathbb{F}$ be a field and A be a finitely generated commutative $\mathbb{F}$-algebra with additive basis $b_1,\ldots,b_n \in A$ (such an algebra is also called a commutative affine algebra). We now want to capture the multiplicative structure of the algebra and therefore write every product of base elements as a linear combination of all base elements:

$$\forall i,j,k \in [n]\ \exists a_{ijk} \in \mathbb{F} : b_i b_j = \sum_{k=1}^{n} a_{ijk} b_k.$$

The a_{ijk} are called **structure coefficients**.

Fact 1.15. *Let A be an $\mathbb{F}$-algebra with additive basis $\{ b_i \}_{i\in[n]}$ and structure coefficients $\{ a_{ijk} \}_{i,j,k\in[n]}$ then:*

$$A \cong \mathbb{F}[x_1,\ldots,x_n] \Big/ \left(x_i x_j - \sum_{k=1}^{n} a_{ijk} x_k \right)_{i,j\in[n]}.$$

To specify an isomorphism $\psi : A \to B$ it is sufficient to write for every i the element $\psi(b_i)$ as linear combination of $b_1,\ldots,b_n$ in B.

Definition 1.16.

$$\begin{aligned}
\textsc{CommAlgIso}_{\mathbb{F}} := \{(A,B) \mid\ & A, B \text{ commutative } \mathbb{F}\text{-algebras with} \\
& \text{basis } b_1,\ldots,b_n \text{ and } A \cong B\}
\end{aligned}$$

Theorem 1.17.

(i). $\textsc{CommAlgIso}_{\mathbb{F}_q} \in \mathbf{NP} \cap \mathbf{coAM}$ for a prime power q

(ii). $\textsc{CommAlgIso}_{\mathbb{R}} \in \mathbf{EEXP}$

(iii). $\textsc{CommAlgIso}_{\mathbb{F}} \in \mathbf{PSPACE}$ if $\mathbb{F} = \overline{\mathbb{F}}$

Proof. A proof can be found in

(i). [KS05, Theorem 3.1.]

(ii). [DH88]

(iii). [Bro06]

□

Theorem 1.18. *For every field $\mathbb{F}$ one has:*

(i). $\textsc{GraphIso} \leq_T^p \textsc{CommAlgIso}_{\mathbb{F}}$

(ii). $\textsc{GraphIso} \leq_T^p \textsc{CubicFormEquiv}_{\mathbb{F}}$

Proof. A proof can be found in

(i). [KS05, Theorem 3.2.] or [AS05, Theorem 2] or [AS06b, Lemma 6.13].

(ii). [AS05, Theorem 4]

□

Theorem 1.19.

(i). $\textsc{PolyEquiv}_{d,\mathbb{F}_q} \in \mathbf{NP} \cap \mathbf{coAM}$ for a prime power q

(ii). $\textsc{PolyEquiv}_{d,\mathbb{R}} \in \mathbf{EEXP}$

(iii). $\textsc{PolyEquiv}_{d,\mathbb{F}} \in \mathbf{PSPACE}$ if $\mathbb{F} = \overline{\mathbb{F}}$

Proof. The proof is given in [AS06b, Theorem 2.1]. □

Theorem 1.20.

(i). $\textsc{CommAlgIso}_{\mathbb{F}} \leq_T^p \textsc{CubicFormEquiv}_{\mathbb{F}}$

(ii). $\textsc{CommAlgIso}_{\mathbb{F}} \leq_T^p \textsc{CubicPolyEquiv}_{\mathbb{F}}$

(iii). $\textsc{FormEquiv}_{d,\mathbb{F}} \leq_T^p \textsc{CommAlgIso}_{\mathbb{F}}$ (if $\mathbb{F}$ contains d-th roots)

Proof. A proof can be found in

(i). [AS06a, Theorem 4.1] or [AS06b, Theorem 3.10]

(ii). [AS06b, Theorem 2.7]

(iii). [AS06b, Theorem 2.3]

□

2 Witt's theorem

The goal of this chapter is to understand an algorithm to decide quadratic form equivalence. This case is significantly easier than equivalences of higher degrees because quadratic modules are well studied and have a lot of structure. We will define a quadratic module associated to a quadratic form and the associated modules to two quadratic forms will turn out to be isomorphism if and only if the two forms are equivalence. In the end we will prove Witt's theorem that gives us two usefull corollaries:

(i). Every quadratic form is equivalent to a form $\sum_{k=1}^{n} a_k x_k^2$.

(ii). We have a cancelation rule for quadratic forms (the $\hat{\oplus}$ will be defined on the way):

$$f \hat{\oplus} h \sim g \hat{\oplus} h \;\Rightarrow\; f \sim g$$

2.1 Definitions

Definition 2.1 (The category of quadratic modules)**.** Let V be a module over a commutative ring R. A function $Q : V \to R$ is called a quadratic form on V if:

(i). $\forall r \in R, x \in V : Q(ax) = a^2 Q(x)$.

(ii). $\Theta_Q : V \times V \to \mathbb{F}, (x, y) \mapsto Q(x + y) - Q(x) - Q(y)$ is bilinear.

The pair (V, Q) is called **quadratic module**. The set of all quadratic forms on V is denoted by $\mathrm{Quad}(V)$. Let (V', Q') be another quadratic module. A linear map $f : V \to V'$ is called **morphism of quadratic modules** or **metric morphism** if $Q' \circ f = Q$, which means the following diagram commutes:

$$\begin{array}{ccc} V & \xrightarrow{f} & V' \\ & \searrow^{Q} & \downarrow{\scriptstyle Q'} \\ & & R \end{array}$$

We write $f : (V, Q) \to (V', Q')$.

Remark 2.2. A form $\phi : V^2 \to R$ is called bilinear, if

(i). $\forall a, b, c, d \in V : \phi(a + b, c + d) = \phi(a, c) + \phi(a, d) + \phi(b, c) + \phi(b, d)$

(ii). $\forall \lambda \in R, a, b \in V : \phi(\lambda a, b) = \lambda \phi(a, b) = \phi(a, \lambda b)$

In our situation, the ring R will always be a field $\mathbb{F}$ with $\mathrm{char}(\mathbb{F}) \neq 2$ and the module V will therefore be a vectorspace. We will furthermore assume that V is finite-dimensional.

Definition 2.3. Let (V, Q) be a quadratic vectorspace, define $\forall x, y \in V$:

$$x.y = \frac{\Theta_Q(x, y)}{2}.$$

Definition/Proposition 2.4. *Let (V, Q) be a quadratic vectorspace. Then:*

(i). $\forall x \in V : Q(x) = x.x$ *and therefore there is a bijective correspondence between quadratic forms and symmetric bilinear forms.*

(ii). $(x, y) \mapsto x.y$ *is a symmetric bilinear form on* V *called the* ***scalar product associated to*** Q*.*

(iii). *For a metric morphism* $f : (V, Q) \to (V', Q')$ *it holds that* $\forall x, y \in V : f(x).f(y) = x.y$*.*

Proof.

(i). $\forall x \in V : x.x = \frac{Q(2x)-Q(x)-Q(x)}{2} = \frac{4Q(x)-Q(x)-Q(x)}{2} = Q(x)$.

(ii). Symmetry is obvious and linearity follows from the properties of Q.

(iii). This directly follows from $Q' \circ f = Q$.

□

Remark 2.5. Even though $(x, y) \mapsto x.y$ is called scalar product, there is no such thing as positive-definiteness since $\mathbb{F}$ does not need to be ordered.

Notation 2.6. For a basis $B = \{ b_1, \ldots, b_n \}$ of V and $x \in V$ one can of course write $x = \sum_{i=1}^{n} x_i b_i$ where $\forall i \in [n] : x_i \in \mathbb{F}$. We denote by $\overline{x}$ the vector of coefficients $(x_1, \ldots, x_n)^T$ with respect to a given basis B.

Definition 2.7 (Matrix associated to a quadratic form)**.** Let (V, Q) be a quadratic vectorspace and $B = \{ b_1, \ldots, b_n \}$ be a basis of V. The **matrix of Q with respect to** B is defined by $(a_{ij})_{ij}$ where $a_{ij} := b_i.b_j$.

Remark 2.8. The matrix associated to Q is symmetric and we have:

$$Q(x) \overset{2.4(i)}{=} x.x = \left(\sum_{i=1}^{n} \overline{x}_i b_i \right) . \left(\sum_{j=1}^{n} \overline{x}_j b_j \right) = \sum_{i,j=1}^{n} \overline{x}_i \overline{x}_j (b_i.b_j) = \sum_{i,j=1}^{n} a_{ij} \overline{x}_i \overline{x}_j .$$

Hence Q is a quadratic form in the variables $\overline{x}_1, \ldots, \overline{x}_n$ in the usual sense. Furthermore we can calculate for the coefficient vectors:

$$\begin{aligned}
\overline{x.y} &= \frac{1}{2} \overline{\left(Q\left(x + y\right) - Q\left(x\right) - Q\left(y\right) \right)} \\
&= \frac{1}{2} \left(\left(\overline{x} + \overline{y} \right)^T A \left(\overline{x} + \overline{y} \right) - \overline{x}^T A \overline{x} - \overline{y}^T A \overline{y} \right) \\
&= \frac{1}{2} \left(\overline{x}^T A \overline{x} + \overline{x}^T A \overline{y} + \overline{y}^T A \overline{x} + \overline{y}^T A \overline{y} - \overline{x}^T A \overline{x} - \overline{y}^T A \overline{y} \right) \\
&= \frac{1}{2} \left(\overline{x}^T A \overline{y} + \overline{y}^T A \overline{x} \right) \\
&= \overline{x}^T A \overline{y}
\end{aligned}$$

Definition 2.9. We define a subgroup of the multiplicative group of $\mathbb{F}$ by

$$\mathbb{F}^{*n} := \left\{ x^n \mid x \in \mathbb{F}^* \right\}.$$

Definition 2.10 (Discriminant of a quadratic form)**.** Let (V, Q) be a quadratic vectorspace and let A be a matrix associated to Q. Denote the projection

$$\mathbb{F} \to \mathbb{F}/\mathbb{F}^{*2}$$

by π, then define the **discrimanent of** Q by $\operatorname{disc}(Q) := \pi(\det(A))$.

Remark 2.11. If one changes the basis that defined A by $X \in \mathrm{Gl}_n(\mathbb{F})$, the matrix A' with respect to this new basis is $X \cdot A \cdot X^t$, which means

$$\det(A') = \det(A) \det(X)^2$$

and therefore $\det(A)$ is determined up to multiplication by a square in $\mathbb{F}^*$, hence $\operatorname{disc}(Q)$ is independent of the choice of a basis.

2.2 Orthogonality

Definition/Proposition 2.12 (Orthogonality)**.** *Two elements x and y of V are called **orthogonal** if $x.y = 0$. For a subset $H \subseteq V$, we define the **orthogonal complement of** H by*

$$H^\perp := \{\, x \in V \mid \forall y \in H : x.y = 0 \,\}.$$

*Two subspaces $U, W \subseteq V$ are called **orthogonal** if $U \subseteq W^\perp$ i.e. if $x \in U, y \in W$ implies $x.y = 0$. The orthogonal complement $V^\perp$ of the whole space V is called **radical** or **kernel of** V and is denoted by $\operatorname{rad}(V)$. Its codimension i.e. $\dim(V) - \dim(\operatorname{rad}(V))$ is called **rank of** Q. If $\operatorname{rad}(V) = \{0\}$, we say that (V, Q) is **nondegenerate** (we may leave out V or Q if it is clear from the context and just say that V is nondegenerate or Q is nondegenerate).*

Fact 2.13.

(i). The orthogonal complement $H^\perp$ of any set $H \subseteq V$ is a subspace of V.

(ii). $H \subseteq H^{\perp\perp}$.

(iii). Q is nondegenerate if and only if $\operatorname{disc}(Q) \neq 0$.

Proof.

(i). This is clear by definition/proposition 2.4(ii).

(ii). Let $x \in H$, to show that $x \in H^{\perp\perp}$, we have to show that $\forall y \in H^\perp$ we have that $x.y = 0$. So let $y \in H^\perp$ be arbitrary, by definition of $H^\perp$, we have that $\forall z \in H : y.z = 0$, especially for $z = x$.

(iii). Choose a basis B and check

$$\begin{aligned}
\operatorname{disc}(Q) = 0 &\Leftrightarrow \det(A) = 0 \\
&\Leftrightarrow \exists x \in V \setminus \{0\} : A\overline{x} = 0 \\
&\overset{*}{\Leftrightarrow} \exists x \in V \setminus \{0\} : \forall y \in V : \overline{y}^T A \overline{x} = 0 \\
&\Leftrightarrow Q \text{ is not nondegenerate}
\end{aligned}$$

The implication "'$\Leftarrow$"' at $*$ can be seen like this:

$$\begin{aligned} &\exists x \in V \setminus \{0\} : \forall y \in V : \overline{y}^T A\overline{x} = 0 \\ \Rightarrow &\exists x \in V \setminus \{0\} : \forall j \in [n] : \left(\overline{b}^T\right)_j A\overline{x} = 0 \end{aligned}$$

And hence $\forall j \in [n]$ the j-th coordinate of $A\overline{x}$, namely $\left(\overline{b}^T\right)_j A\overline{x}$ is zero, hence $A\overline{x}$ is zero.

□

Example 2.14. Being nondegenerate is not passed on to subspaces: Let for example $Q : \mathbb{R}^3 \to \mathbb{R}$ be given by

$$\begin{aligned} Q(x_1, x_2, x_3) &= x_1^2 + x_3^2 + 2x_2x_3 + 2x_1x_3 \\ &= (x_1, x_2, x_3) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \end{aligned}$$

Since $\operatorname{disc}(Q) = -1$ we get that the quadratic space $(\mathbb{R}^3, Q)$ is nondegenerate. But the subspace

$$U := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \in \mathbb{R}^3 \right\}$$

with the restriction $Q\Big|_U$ is not nondegenerate since $Q(x_1, x_2, 0) = x_1^2$ and therefore

$$\begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} . \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0 \quad \forall \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \in U$$

which means that $(0, 1, 0)^T$ is orthogonal to every other element of U.

Definition 2.15. Let $U \subseteq V$ be subspace and denote the **dual space** by

$$U^* := \{\, \phi : U \to \mathbb{F} \mid \phi \text{ is linear} \,\}.$$

Furthermore define

$$\begin{array}{rccl} q_U : V & \longrightarrow & U^* \\ x & \longmapsto & (y \in U \mapsto x.y) \end{array}$$

Fact 2.16.

(i). $\ker(q_U) = U^\perp$

(ii). Q is nondegenerate if and only if $q_V : V \to V^*$ is an isomorphism.

Proof.

(i). For $x \in U$ with $q_U(x) = 0 \in V^*$ we have

$$\forall y \in V : 0 = (q_U(x))(y) = x.y$$

which exactly defines $U^\perp$.

(ii). $\ker(q_V) \overset{(i)}{=} V^\perp$ which is by definition $\{0\}$ if and only if Q is nondegenerate, therefore q_V is injective, but since $V \cong V^*$ it is also surjective.

□

Definition 2.17. Let $U_1, \ldots, U_m \subseteq V$ be subspaces. We say that V is the **orthogonal direct sum of the** U_i if they are pairwise orthogonal and if V is the direct sum of them, we then write:

$$V = U_1 \hat{\oplus} \ldots \hat{\oplus} U_m.$$

Remark 2.18. Let $V = U_1 \hat{\oplus} \ldots \hat{\oplus} U_m$ and decompose $x \in V$ into it's components $x_i \in U_i$, then

$$Q(x) = Q_1(x_1) + \ldots + Q_m(x_m) \tag{2.1}$$

where $Q_i := Q\Big|_{U_i}$ are the restrictions of Q to U_i. Conversely if (U_i, Q_i) for $i =\in [m]$ are quadratic modules, we can define a quadratic module (V, Q) where $V = \bigoplus_{i=1}^m U_i$ by eq. (2.1) above and have:

$$V = U_1 \hat{\oplus} \ldots \hat{\oplus} U_m.$$

Example 2.19. If $U \subseteq V$ is a supplementary subspace of $\operatorname{rad}(V)$ (i.e. $V = U \oplus \operatorname{rad}(V)$) then

$$V = U \hat{\oplus} \operatorname{rad}(V).$$

Proposition 2.20. *Let (V, Q) be nondegenerate. Then the following statements hold:*

(i). All metric morphisms of V into a quadratic module (V', Q') are injective.

(ii). For all subspaces $U \subseteq V$, we have:

- *(a)* $\dim(U) + \dim\left(U^\perp\right) = \dim(V)$
- *(b)* $U^{\perp\perp} = U$
- *(c)* $\operatorname{rad}(U) = \operatorname{rad}\left(U^\perp\right) = U \cap U^\perp$

The quadratic module $\left(U, Q\Big|_U\right)$ is nondegenerate if and only if $\left(U^\perp, Q\Big|_{U^\perp}\right)$ is nondegenerate in which case $V = U \hat{\oplus} U^\perp$.

(iii). If V is the orthogonal direct sum of two subspaces, they are nondegenerate and each of them is orthogonal to the other.

Proof.

(i). If $f : V \to V'$ is a metric morphism and if $f(x) = 0$, we have

$$x.y = f(x).f(y) = 0 \quad \forall y \in V$$

this implies $x = 0$ because (V, Q) is nondegenerate.

(ii). Let $U \subseteq V$ be a subspace. Note that $q_U = q_V \circ \pi_{U^*}$ where $\pi_{U^*} : V^* \to U^*$ is the canonical projection. Since q_V is bijective (by fact 2.16(i)), q_U is surjective, thus with the canonical injection $\iota : U^\perp \to V$ the following sequence is exact:

$$\{0\} \longrightarrow U^\perp \overset{\iota}{\hookrightarrow} V \overset{q_U}{\twoheadrightarrow} U^* \longrightarrow \{0\}$$

hence

$$\dim(V) = \dim(U^*) + \dim\left(U^\perp\right) = \dim(U) + \dim\left(U^\perp\right).$$

Taking $U^\perp$ as the subspace in this argument we also get

$$\dim(V) = \dim\left(U^\perp\right) + \dim\left(U^{\perp\perp}\right)$$

which implies that

$$\dim(U) + \dim\left(U^\perp\right) = \dim(V) = \dim\left(U^\perp\right) + \dim\left(U^{\perp\perp}\right)$$

giving that $\dim(U) = \dim\left(U^{\perp\perp}\right)$. Fact 2.13(ii) now implies $U = U^{\perp\perp}$. By the definitions:

$$\begin{aligned} \operatorname{rad}(U) &:= \{\, x \in U \mid \forall\, y \in U : x.y = 0 \,\} \\ U^\perp &:= \{\, x \in V \mid \forall\, y \in U : x.y = 0 \,\} \end{aligned}$$

we clearly get $U \cap U^\perp = \operatorname{rad}(U)$. Applying this formular to $U^\perp$, we get $U^\perp \cap U^{\perp\perp} = \operatorname{rad}\left(U^\perp\right)$ and calculate

$$\operatorname{rad}\left(U^\perp\right) = U^\perp \cap U^{\perp\perp} \overset{(ii)b}{=} U^\perp \cap U = \operatorname{rad}(U).$$

(iii). This statement is finally trivial, because if $V = U \hat{\oplus} W$ is nondegenerate, none of U and W can be not nondegenerate and the orthogonality directly follows from the defintion of the orthogonal direct sum.

□

Example 2.21. Example 2.14 does not yield a counter example to proposition 2.20(iii) since although with

$$U := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \in \mathbb{R}^3 \right\}, \quad W := \left\{ \begin{pmatrix} 0 \\ 0 \\ y_3 \end{pmatrix} \in \mathbb{R}^3 \right\}$$

we have that $\mathbb{R}^3 = U \oplus W$ we also calculate

$$\begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} . \begin{pmatrix} 0 \\ 0 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ y_3 \end{pmatrix} = (x_1 + x_2) y_3$$

which gives us that for example $(1,0,0)^T \in U$ and $(0,0,1)^T \in W$ are not orthogonal, which means that $\mathbb{R}^3$ is not the orthogonal sum of U and W.

2.3 Isotropic vectors

Definition 2.22. An element $x \in V$ is called isotropic if $Q(x) = 0$. A subspace $U \subseteq V$ is called isotropic if all its elements are **isotropic**.

Example 2.23. A nondegenerate space can contain isotropic vectors: Consider the quadratic form $Q(x, y) = 2xy$ over $\mathbb{R}^2$. The associated matrix is

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

One has $Q(1, 0) = 0$ which means that $(1, 0)^T$ is isotropic but $\det(A) = -1 \neq 0$ which means that $(\mathbb{R}^2, Q)$ is nondegenerate.

Fact 2.24.

$$U \text{ isotropic} \quad \Leftrightarrow \quad U \subseteq U^\perp \quad \Leftrightarrow \quad Q\Big|_U = 0$$

Definition 2.25. A quadratic module having a basis formed of two isotropic elements $x, y \in V$ such that $x.y \neq 0$ is called **hyperbolic plane**.

Remark 2.26. Without loss of generality, we can assume that $x.y = 1$: Just multiply y by $\frac{1}{x.y}$. Then the matrix of the quadratic form with respect to the basis $\{ x, y \}$ is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. The discriminant then is $\operatorname{disc}(Q) = -1$, in particular Q is nondegenerate.

Proposition 2.27. *Let $x \in V \setminus \{0\}$ be isotropic and Q be nondegenerate. Then there exists a subspace $U \subseteq V$ which contains x and is a hyperbolic plane.*

Proof. Since V is nondegenerate, there exists $z \in V$ such that $x.z = 1$. The element $y = 2z - (z.z)x$ is isotropic and $x.y = 2$. The subpsace $U = \langle x, y \rangle$ has the desired property. □

Corollary 2.28. *If (V, Q) is nondegenerate and contains a nonzero isotropic element, we have $Q(V) = \mathbb{F}$.*

Proof. We have to show that $\forall a \in \mathbb{F}\ \exists v \in V$ such that $Q(v) = a$. Without loss of generalty, we may assume that V is a hyperbolic plane: Let $x \in V$ be the nonzero isotropic element, then proposition 2.27 we get $y \in V$ such that $\langle x, y \rangle$ is a hyperbolic plane. Furthermore we can assume that x and y are isotropic and $x.y = 1$ (see remark 2.26). Now for $a \in \mathbb{F}$ one calculates

$$\begin{pmatrix} 1 & \frac{a}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \frac{a}{2} \end{pmatrix} = \begin{pmatrix} 1 & \frac{a}{2} \end{pmatrix} \begin{pmatrix} \frac{a}{2} \\ 1 \end{pmatrix} = a$$

and therefore get we get $a = Q\left(x + \frac{a}{2}y\right)$. □

2.4 Orthogonal basis

Definition 2.29. A Basis $\{ b_1, \ldots, b_n \}$ is called **orthogonal** if its elements are pairwise orthogonal i.e.

$$V = \langle b_1 \rangle \mathbin{\hat{\oplus}} \ldots \mathbin{\hat{\oplus}} \langle b_n \rangle .$$

Remark 2.30. This is equivalent to saying that the matrix associated to Q with respect to the basis $B = \{ b_1, \ldots, b_n \}$ is a diagonal matrix with diagonal entries $a_1, \ldots, a_n \in \mathbb{F}^*$:

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}.$$

If $\overline{x} = (x_1, \ldots, x_n)^T$ is the coordinate vector of $x \in V$ with respect to the basis B, we have that $Q(x) = a_1 x_1^2 + \ldots + a_n x_n^2$.

Fact 2.31. *Let (V, Q) be a nondegenreate quadratic space with an orthogonal basis $\{ b_1, \ldots, b_n \}$, then $\forall i \in [n] : (b_i.b_i) \neq 0$.*

Proof. Write

$$V = \langle b_1 \rangle \hat{\oplus} \cdots \hat{\oplus} \langle b_n \rangle .$$

Now proposition 2.20(iii) gives that each $\langle b_i \rangle$ is nondegenerate which makes it impossible for b_i to be isotropic. □

Theorem 2.32. *Every quadratic module has an orthogonal basis.*

Proof. We prove this by induction on the dimension $n := \dim(V)$. The case $n = 0$ is trivial. Now let n be arbitrary. If V is isotropic, all bases of V are orthogonal. Otherwise, choose an element $b \in V$ such that $b.b \neq 0$. Now the orthogonal complement $U := \{ b \}^\perp$ is a hyperplane (i.e. has dimension $n - 1$) and since $b \notin U$, one has $V = \langle b \rangle \hat{\oplus} U$. By induction hypothesis U has an orthogonal basis B and $\{ b \} \cup B$ is an orthogonal basis. □

Definition 2.33. Two orthogonal bases

$$B = \{ b_1, \ldots, b_n \} \quad \text{and} \quad C = \{ c_1, \ldots, c_n \}$$

of V are called **contiguous** if they have an element in common (i.e. if there exist i and j with $b_i = c_j$). A sequence of basis $B_0, B_1, \ldots, B_m$ is called a **chain contigouously relating B and C** if

- $B_i \subseteq V$ is an orthogonal basis for $1 \leq i \leq m$,
- $B_0 = B$ and $B_m = C$,
- B_i and B_{i+1} are contiguous for $0 \leq i < m$.

Lemma 2.34. *Let (V, Q) be a nondegenerate quadratic module, $a, b \in V \setminus \{0\}$ and define $P := \langle a, b \rangle$. Then*

$$(a.a)(b.b) \neq (a.b)^2 \Leftrightarrow \dim(P) = 2 \text{ and } P \text{ is nondegenerate.}$$

Proof. We will prove the equivalent statement

$$(a.a)(b.b) = (a.b)^2 \Leftrightarrow \dim(P) < 2 \text{ or } P \text{ is degenerate.}$$

"$\Leftarrow$": Asume that $\dim(P) < 2$, then there exists $\lambda \in \mathbb{F}$ with $a = \lambda b$ implying:

$$(a.a)(b.b) - (a.b)^2 = (\lambda b.\lambda b)(b.b) - (\lambda b.b)^2 = \lambda^2(b.b)^2 - \lambda^2(b.b)^2 = 0.$$

Now if P is degenerate, then there exists $v = \lambda a + \mu b \in P \setminus \{0\}$ with the property that $\forall w \in P : (v.w) = 0$. Now calculate

$$0 = (v.a) = \lambda(a.a) + \mu(b.a) \quad \Leftrightarrow \quad -\mu(b.a) = \lambda(a.a) \tag{2.2}$$
$$0 = (v.b) = \lambda(a.b) + \mu(b.b) \quad \Leftrightarrow \quad -\lambda(a.b) = \mu(b.b). \tag{2.3}$$

Now since $v \neq 0$ at least one of μ and λ is not zero. If $\lambda \neq 0$ we get by (2.3) that $(a.b) = -\frac{\mu}{\lambda}(b.b)$ which, plugged into (2.2) yields

$$\mu\frac{\mu}{\lambda}(b.b) = \lambda(a.a) \quad \Leftrightarrow \quad \frac{\mu^2}{\lambda^2}(b.b) = (a.a)$$

putting this all together we get

$$(a.a)(b.b) - (a.b)^2 = \frac{\mu^2}{\lambda^2}(b.b)(b.b) - \left(-\frac{\mu}{\lambda}(b.b)\right)^2 = 0.$$

The same works for $\mu \neq 0$ because (2.2) and (2.3) are symmetric in a and b.

"$\Rightarrow$": Define the element

$$c := (b.b)a - (a.b)b \in P$$

and observe that

$$c.a = (b.b)(a.a) - (a.b)(b.a) = 0 \text{ by assumption}$$
$$c.b = (b.b)(a.b) - (a.b)(b.b) = 0$$

So $c \in \operatorname{rad}(P)$. So c is zero or P is degenerate (in the latter case, we are done). So let $c = 0$ we get that $(b.b)a - (a.b) = 0$ which is a linear combination of 0 in a and b (which generate P). So either is $\dim(P) < 2$ (in which case we are done again) or $(b.b) = 0$ and $(a.b) = 0$, which implies that $b \in \operatorname{rad}(P)$. Now if $b = 0$, we again get that $\dim(P) < 2$ and if $b \neq 0$, P is degenerate.

□

Lemma 2.35 (Gram-Schmidt)**.** *Let (V, Q) be nondegenerate, $a, b \in V$ be linearly independent and a be nonisotropic. Then there exists $c \in V$ such that*

$$\langle a, b\rangle = \langle a\rangle \hat{\oplus} \langle c\rangle$$

Proof. Set $p := \frac{a.b}{a.a}a$, $c := b - p$ and calculate:

$$a.c = a.(b - p) = a.b - a.p = a.b - a.a\frac{a.b}{a.a} = 0.$$

Let $\lambda, \mu \in \mathbb{F}$ and calculate

$$0 = \lambda a + \mu c = \lambda a + \mu b - \mu p = \lambda a + \mu b - \mu\frac{a.b}{a.a}a = \left(\lambda - \mu\frac{a.b}{a.a}\right)a + \mu b$$

now since a and b are linearly independent, we get $\mu = 0$ and $\lambda - \mu\frac{a.b}{a.a} = 0$ i.e. $\lambda = 0$, meaning that a and c are linearly independent too. □

Theorem 2.36. *Let (V, Q) be a nondegenerate quadratic module of dimension $\dim(V) \geq 3$ with two orthogonal basis B and C then there exists a chain contigously relating B and C.*

Proof. Define $\mu_i := (b_1.b_1)(c_i.c_i) - (b_1.c_i)^2$ and distinguish the cases where $\mu_i = 0$ for $i \in \{1, 2\}$ and where $\mu_i \neq 0$.

Case 1. ($\mu_i \neq 0$ for $i = 1, 2$) By assumption and lemma 2.34 (applied to b_1 and c_i) $P := \langle b_1, c_i \rangle$ has dimension 2 and is nondegenerate. Since B and C are orthogonal basis by fact 2.31 we know that b_1 and c_i are both nonisotropic and lemma 2.35 therefore yields $x, y \in V$ with

$$P = \langle b_1 \rangle \hat{\oplus} \langle x \rangle \quad \text{and} \quad P = \langle c_i \rangle \hat{\oplus} \langle y \rangle$$

Additionally by proposition 2.20 we get, that $P^\perp$ is nondegenerate too. And ultimately $V = P \hat{\oplus} P^\perp$. Now let $\{ d_3, \dots d_n \}$ be an orthogonal basis of $P^\perp$ (which exists because of theorem 2.32). Then the sequence

$$B \quad , \quad \{ b_1, x, d_3, \dots, d_n \} \quad , \quad \{ c_i, y, d_3, \dots, d_n \} \quad , \quad C$$

contigouously relates B and C.

Case 2. ($\mu_i = 0$ for $i = 1, 2$) We first prove the following claim:

Claim 2.36.1. *$\exists \lambda \in \mathbb{F} : e_\lambda := c_1 + \lambda c_2$ is nonisotropic and $\langle e_\lambda, b_1 \rangle$ is a nondegenerate plane.*

Proof. For e_λ being nonisotropic, we need to ensure that $0 \neq e_\lambda.e_\lambda$. So calculate

$$(e_\lambda.e_\lambda) = (c_1.c_1) + 2\lambda(c_1.c_2) + \lambda^2(c_2.c_2) = (c_1.c_1) + \lambda^2(c_2.c_2)$$

since C is orthogonal. We know that $(c_i, c_i) \neq 0$ (fact 2.31) and therefore have that e_λ is nonisotropic if and only if $\lambda^2 \neq -\frac{c_1.c_1}{c_2.c_2}$.
Applying lemma 2.34 to e_λ and b_1 yields that it is necessary and sufficient for them to generate a nondegenerate plane that

$$(b_1.b_1)(e_\lambda.e_\lambda) - (b_1.e_\lambda)^2 \neq 0$$

So calculate

$$\begin{aligned}(b_1.b_1)(e_\lambda.e_\lambda) &= (b_1.b_1)((c_1.c_1) + \lambda^2(c_2.c_2)) \\ &= (b_1.b_1)(c_1.c_1) + \lambda^2(b_1.b_1)(c_2.c_2) \\ &= (b_1.c_1)^2 + \lambda^2(b_1.c_2)^2 \qquad \text{since } \mu_i = 0\end{aligned}$$

and

$$\begin{aligned}(b_1.e_\lambda)^2 &= (b_1.c_1 + \lambda(b_1.c_2))^2 \\ &= (b_1.c_1)^2 + 2\lambda(b_1.c_1)(b_1.c_2) + \lambda^2(b_1.c_2)^2\end{aligned}$$

leading to

$$0 \neq (b_1.b_1)(e_\lambda.e_\lambda) - (b_1.e_\lambda)^2 = -2\lambda(b_1.c_1)(b_1.c_2)$$

By fact 2.31 and $\mu_i = 0$, we get that $(b_1.c_i)^2 \neq 0$ and tehrefore that $\lambda \neq 0$. Summarized, e_λ verifies the conditions of claim 2.36.1 if and only if $\lambda^2 \neq -\frac{c_1.c_1}{c_2.c_2}$ and $\lambda \neq 0$. This rules out only 3 values for $\lambda \in \mathbb{F}$, so if $\mathbb{F}$ has at least 4 elements, we are done. So we are left to show the statement for the case $\mathbb{F} = \mathbb{F}_3$ ($\mathbb{F} = \mathbb{F}_2$ is excluded, since $\operatorname{char}(\mathbb{F}) \neq 2$): In $\mathbb{F}_3$, all nonzero squares are 1, so $\mu_i = 0$ is equivalent to $(b_1.b_1)(c_i.c_i) = 1$. Now calculate

$$\lambda^2 \neq -\frac{c_1.c_1}{c_2.c_2} = -\frac{(b_1.b_1)(c_1.c_1)}{(b_1.b_1)(c_2.c_2)} = -1$$

and see that $\lambda = 1$ realizes the conditions $\lambda^2 \neq -1$ and $\lambda \neq 0$ which finishes the proof of claim 2.36.1. □

Let $e_\lambda \in \mathbb{F}$ be such as in claim 2.36.1 and since e_λ is not isotropic there is $y \in \mathbb{F}$ such that $\{ e_\lambda, y \}$ is an orthogonal basis of $\langle e_\lambda, y \rangle$.

Finish Proof: theorem for contigous bases

□

2.5 Proof of Witt's Theorem

Let (V, Q) and (V', Q') be two nondegenerate quadratic spaces, $U \subseteq V$ a subspace of V and $s : U \to V'$ be an injective metric morphism in this section. The goal is to extend s to a subspace larger than U and if possible to all of V.

Proposition 2.37. *If U is degenerate, there exists $U_1 \subseteq V$ containing U with*

$$\dim(U_1) = \dim(U) + 1$$

extending s to an injective metric morphism $s_1 : U_1 \hookrightarrow V'$ with $s_1\big|_U = s$.

Proof. Let $x \in \operatorname{rad}(U) \setminus \{0\}$ and $g : U \to \mathbb{F}$ be linear such that $g(x) = 1$. Since U is nondegenerate, fact 2.16(ii) implies that q_V is an isomorphism and therefore surjective i.e. exists $y \in V$ such that $q_V(y)\big|_U = g$ or in other words for all $u \in U : g(u) = u.y$. Since $x \in \operatorname{rad}(U)$ and $y.x = 1 \neq 0$ we get that $y \notin U$ and therefore that $U_1 := U \oplus \langle y \rangle$ contains U as a hyperplane.

Replacing y by $y - \lambda x$ with $\lambda = (y.y)/2$ does not change g since for any $u \in U$:

$$u.(y - \lambda x) = u.y - \lambda \underbrace{u.x}_{\substack{=0 \text{ since} \\ x \in \operatorname{rad}(U)}} = u.y.$$

But the replacement yields that $y.y = 0$ since

$$(y - \lambda x).(y - \lambda x) = (y.y) - 2\lambda \underbrace{(y.x)}_{\substack{=1 \text{ since} \\ l(x)=1}} + \lambda^2 \underbrace{(x.x)}_{\substack{=0 \text{ since} \\ x \in \operatorname{rad}(U)}} = (y.y) - 2\frac{(y.y)}{2} = 0.$$

The same constructions works for $U' := s(U)$, $x' = s(x)$ and $g' = g \circ s^{-1}$ yielding $y' \in V'$ and $U_1' = U' \oplus \langle y' \rangle$. Now define $s_1 : U_1 \to U_1'$ by

$$\begin{aligned} s_1 : U \oplus \langle y \rangle &\to U' \oplus \langle y' \rangle \\ (u, \alpha y) &\mapsto (s(u), \alpha y'). \end{aligned}$$

Now we claim that s_1 is a metric isomorphism. s_1 is indeed well-defined, linear, injective and surjective by definition and we now check that it is metric. Let $u \in U$ and $\alpha y \in \langle y \rangle$, then

$$Q'\left(s_1(u, \alpha y)\right) = Q'\left(s(u), \alpha y'\right) \overset{2.18}{=} Q'\Big|_U (s(u)) + \underbrace{Q'\Big|_{\langle y \rangle} (\alpha y')}_{=0 \text{ since } y'.y'=0} = Q'(s(u)).$$

And since s is metric, $Q' \circ s = Q$ implying that $Q' \circ s_1 = Q$ and finally that s_1 is metric. □

Theorem 2.38 (Witt)**.** *If (V, Q) and (V', Q') are isomorphic and nondegenerate, every injective metric morphism*

$$s : U \hookrightarrow V'$$

from a subspace $U \subseteq V$ can be extended to a metric isomorphism of V onto V'.

Proof. Since V and V' are isomorphic, we can without loss of generality assume that $V = V'$. If V is degenerate, we can apply proposition 2.37 to be finished or to be left with a non-degenerate subspace $U \subseteq V$. We now argue by induction on $\dim(U)$.

If $\dim(U) = 1$, U is generated by a nonisotropic element $x \in U$. If $y = s(x)$, we have $y.y = s(x).s(x) = x.x$. Now one can choose $\epsilon = \pm 1$ such that $x + \epsilon y$ is nonisotropic too since otherwise we whould have:

$$\begin{aligned} 0 &= (x+y).(x+y) = x.x + 2x.y + y.y = 2x.x + 2x.y \\ 0 &= (x-y).(x-y) = x.x - 2x.y + y.y = 2x.x - 2x.y \\ \Rightarrow 0 &= 4x.x \\ \Rightarrow 0 &= x.x \end{aligned}$$

Now define $z = x + \epsilon y$ and let $H = \langle z \rangle^{\perp}$. Now we have $V = \langle z \rangle \hat{\oplus} H$ by proposition 2.20(iii) since (U, Q) is nondegenerate. Now let $\sigma : V \to V$ be the unique automorphism defined by $\sigma\Big|_H = \mathrm{id}_H$ and $\sigma(z) := -z$. We have

$$\begin{aligned} \sigma(x - \epsilon y) &= x - \epsilon y && \text{since } x - \epsilon y \in H \\ \sigma(x + \epsilon y) &= -x - \epsilon y && \text{by definition} \end{aligned}$$

yielding

$$\sigma(2x) = \sigma(x - \epsilon y) + \sigma(x + \epsilon y) = x - \epsilon y - x - \epsilon y = -2\epsilon y$$

and ultimately $\sigma(x) = -\epsilon y$, hence the automorphism $-\epsilon\sigma$ extends s.

If $\dim(U) > 1$, we decompose U in the form $U_1 \hat{\oplus} U_2$ with $U_1, U_2 \neq \{0\}$. By induction hypothesis, the restriction s_1 of s to U_1 extends to an automorphism σ_1 of V. After replaceing s by $\sigma_1^{-1} \circ s$ one can thus suppose that s is the identity on U_1. Then the morphism s carries U_2 into the orthogonal complement V_1 of U_1. Again by induction hypothesis, the restriction of s to U_2 extends to an automorphism σ_2 of V_1. Now define σ by $\sigma\Big|_{U_1} = \mathrm{id}_{U_1}$ and $\sigma\Big|_{V_1} = \sigma_2$ has the desired property. □

Corollary 2.39. *Two isomorphic subspaces of a nondegenerate quadratic module have isomorphic orthogonal complements.*

Proof. Let $U, W \subseteq V$ be two isomorphic subspaces, by theorem 2.38 we can extend the isomorphism between them to an automorphism of the whole space and restrict it to the orthogonal complement $U^\perp$ yielding an isomorphism between $U^\perp$ and $W^\perp$. □

2.6 Application to quadratic form equivalence

Definition 2.40. Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a quadratic form with

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{i<j} a_{ij} x_i x_j \quad \forall i \leq j \in [n] : a_{ij} \in \mathbb{F}$$

then $(\mathbb{F}^n, f)$ is the **quadratic module associated to** f.

Proposition 2.41. *Quadratic forms in the same number of variables are equivalent if and only if the associated quadratic modules are isomorphic.*

Proof. Let $f, g \in \mathbb{F}[x_1, \ldots, x_n]$ and let A be the matrix associated to f and B be the matrix associated to g with respect to some basis (for example the canonical one) of $\mathbb{F}^n$. An isomorphism of quadratic modules $\phi : (\mathbb{F}^n, f) \xrightarrow{\cong} (\mathbb{F}^n, g)$ can now be represented by a matrix $P \in \mathrm{Gl}_n(\mathbb{F})$ such that $f \circ P = g$ which is exactly the definition of form equivalence. □

Remark 2.42. Let $f, g \in \mathbb{F}[x_1, \ldots, x_n]$ be two quadratic forms with corresponding matrices A and B. Saying $f \sim g$ amounts to saying that there exists $X \in \mathrm{Gl}_n(\mathbb{F})$ with $B = XAX^T$ (by remark 2.11).

Definition 2.43. Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ and $g \in \mathbb{F}[x_1, \ldots, x_m]$ be two quadratic forms, then we define the **orthogonal sum** $f \hat{\oplus} g \in \mathbb{F}[x_1, \ldots, x_{n+m}]$ by

$$(f \hat{\oplus} g)(x_1, \ldots, x_{n+m}) := f(x_1, \ldots, x_n) + g(x_{n+1}, \ldots, x_{n+m}).$$

Correspondingly we write $f \hat{\ominus} g := f \hat{\oplus} (-g)$.

Fact 2.44. *The orthogonal sum of forms corresponds to the orthogonal sum of quadratic spaces, i.e.* $\forall\ f \in \mathbb{F}[x_1, \ldots, x_n], g \in \mathbb{F}[x_1, \ldots, x_m]$*:*

$$\left(\mathbb{F}^{n+m}, f \hat{\oplus} g\right) \cong (\mathbb{F}^n, f) \ \hat{\oplus}\ (\mathbb{F}^m, g).$$

Proof. Define the map $\varphi : \mathbb{F}^n \oplus \mathbb{F}^m \to \mathbb{F}^{n+m}$ component wise by

$$\mathbb{F}^n \ni (x_1, \ldots, x_n) \ \mapsto\ (x_1, \ldots, x_n, 0, \ldots, 0) \in \mathbb{F}^{n+m}$$

and

$$\mathbb{F}^m \ni (x_1, \ldots, x_m) \ \mapsto\ (0, \ldots, 0, x_1, \ldots, x_m) \in \mathbb{F}^{n+m}.$$

This map induces an isomorphism of vectorspaces which is also a metric morphism by definition of $f \hat{\oplus} g$. □

Fact/Definition 2.45. A form $f \in \mathbb{F}[x,y]$ is called **hyperbolic** if and only if:

$$f \sim xy \sim x^2 - y^2.$$

This means that the quadratic space $(\mathbb{F}^2, f)$ is a hyperbolic plane.

Proof. First note that xy and $x^2 - y^2$ are equivalent via

$$\tau : \begin{cases} x & \mapsto \frac{x+y}{2} \\ y & \mapsto \frac{x-y}{2} \end{cases},$$

since

$$\tau(x)^2 - \tau(y)^2 = \left(\frac{x+y}{2}\right)^2 - \left(\frac{x+y}{2}\right)^2 = \frac{x^2+2xy+y^2}{4} - \frac{x^2-2xy+y^2}{4} = xy.$$

Let now $f \in \mathbb{F}[x,y]$ be hyperbolic. We now need to find a basis $\{v, w\}$ of isotropic vectors such that $v.w \neq 0$. Since $f \sim xy$, there exists $A \in \mathrm{Gl}_n(\mathbb{F})$ such that

$$f\left(A\begin{pmatrix} x \\ y \end{pmatrix}\right) = xy.$$

Now choose

$$v = A\begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } w = A\begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

which is, since A is bijective, a basis of $\mathbb{F}^2$ and calculate

$$\begin{aligned} f(v) &= f\left(A\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = 0 \\ f(w) &= f\left(A\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = 0 \\ f(v+w) &= f\left(A\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = 1 \end{aligned}$$

yielding that

$$\begin{aligned} v.v &= \frac{1}{2}\left(f(v+v) - f(v) - f(v)\right) = \frac{1}{2}\left(4f(v) - f(v) - f(v)\right) = 0 \\ w.w &= \frac{1}{2}\left(f(w+w) - f(w) - f(w)\right) = \frac{1}{2}\left(4f(w) - f(w) - f(w)\right) = 0 \\ v.w &= \frac{1}{2}\left(f(v+w) - f(v) - f(w)\right) = \frac{1}{2} \neq 0. \end{aligned}$$

This means that $\left(\mathbb{F}^2, f\right)$ is a hyperbolic plane. □

Definition 2.46. A form $f \in \mathbb{F}[x_1, \ldots, x_n]$ **represents** an element $a \in \mathbb{F}$ if there exists $x \in \mathbb{F}^n \setminus \{0\}$ with $f(x) = a$.

Remark 2.47. A form represents zero if and only if the corresponding quadratic space contains a non-zero isotropic element.

Proposition 2.48. *If f represents 0 and is nondegenerate, one has $f \sim h + g$ where h is hyperbolic. Moreover, f represents all elements of $\mathbb{F}$.*

Proof. Basically this is the translation of proposition 2.27 and corollary 2.28 in the language of quadratic form equivalence: Since f represents zero, there exists a non-zero isotropic element x. Then by proposition 2.27 there exists a hyperbolic plane $U \subseteq V$ containing x, which by remark 2.26 is nondegenreate. This implies with proposition 2.20(iii) that $\mathbb{F} = U \hat{\oplus} U^{\perp}$ where U is a hypberbolic plane. By fact/definition 2.45, we now get that there is h that is hyperbolic. Finally by corollary 2.28 we have that $f(\mathbb{F}^n) = \mathbb{F}$ which means that every element of $\mathbb{F}$ is represented. □

Corollary 2.49. *Let $g \in \mathbb{F}[x_1, \ldots, x_{n-1}]$ be a nondegenerate quadratic form and let $a \in \mathbb{F}^*$, then the following properties are equivalent:*

(i). g represents a.

(ii). $\exists\, h \in \mathbb{F}[x_1, \ldots, x_{n-2}] : g \sim h \hat{\oplus} ax_n^2$.

(iii). $g \hat{\ominus} ax_n^2$ represents zero.

Proof.

(ii)⇒(i): and (ii)⇒(iii): Let h be as in the statement. Then $t := (t_1, \ldots, t_{n-2}, t_n)$ with $t_i = 0$ for $i \in [n-2]$ and $t_n = 1$ gives

$$\left(h \hat{\oplus} ax_n^2\right)(t) = a.$$

Let τ be an invertible linear transformation on the variables that takes g to $h \hat{\oplus} ax_n^2$ then we have

$$g(\tau(t)) = \left(h \hat{\oplus} ax_n^2\right)(t) = a.$$

And since τ is invertible and linear from $t \neq 0$ it follows that $\tau(t) \neq 0$. Hence g represents a. This immediately gives

$$\left(g \hat{\ominus} ax_n^2\right)(\tau(t), 1) = a - a = 0.$$

(i)⇒(ii): Since g represents a, the quadratic space corresponding to g contains an element x such that $x.x = a$. By proposition 2.20(ii) we then can write $\mathbb{F}^{n-1} = \{x\}^{\perp} \hat{\oplus} \langle x \rangle$. Now let $\{b_1, \ldots, b_{n-2}\}$ be a basis of $\{x\}^{\perp}$ and define the quadratic form h by

$$h(x_1, \ldots, x_{n-2}) := \sum_{i=1}^{n-2} b_i.b_i x_i + 2 \sum_{i<j} b_i.b_j x_i x_j.$$

Fact 2.44 then implies the statement.

(iii)⇒(i): If the form $f := g \hat{\ominus} ax_n^2$ represents 0, there exists a non-trivial zero $(x_1, \ldots x_n)$ of f. then either $x_n = 0$, which implies that $(x_1, \ldots, x_{n-1})$ is a non-trivial zero of g and by proposition 2.48 we then get that g represents a or $x_n \neq 0$ in which case

$$0 = \frac{f(x_1, \ldots, x_n)}{x_n^2} = f\left(\frac{x_1}{x_n}, \ldots, \frac{x_{n-1}}{x_n}, 1\right) = g\left(\frac{x_1}{x_n}, \ldots, \frac{x_{n-1}}{x_n}\right) - a.$$

□

Theorem 2.50 (Decomposition into Sums of squares)**.** *It holds that*

$$\forall\ f \in \mathbb{F}\left[x_1, \ldots, x_n\right]^{=2} : \exists\ a_1, \ldots, a_n \in \mathbb{F} : f \sim \sum_{i=1}^{n} a_i x_i^2.$$

Proof. Theorem 2.32 together with fact 2.44 gives the statement. □

Fact 2.51. *Let $f \in \mathbb{F}\left[x_1, \ldots, x_n\right]^{=2}$, by theorem 2.50 there exist $a_1, \ldots, a_n \in \mathbb{F}$ such that $f \sim \sum_{i=1}^{n} a_i x_i^2$. The* $\operatorname{rank}(f)$ *defined in definition/proposition 2.12 coincides with the number*

$$\left|\left\{\, i \in [n] \mid a_i \neq 0 \,\right\}\right|.$$

Two isomorphic quadratic modules $(\mathbb{F}^n, f) \cong (\mathbb{F}^n, g)$ are of the same rank.

Corollary 2.52. *Let g and h be two nondegenerate forms of rank ≥ 1 and $f := g \hat{\ominus} h$. The following properties are equivalent:*

(i). f represents zero.

(ii). $\exists\ a \in \mathbb{F}^*$ which is represented by g and by h.

(iii). $\exists\ a \in \mathbb{F}^*$ such that $g \hat{\ominus} aZ^2$ and $h \hat{\ominus} aZ^2$ represent zero.

Proof. **(ii)⇔(iii):** This follows from corollary 2.49.

(ii)⇒(i): Since f is defined as the difference of g and h it represents 0.

(i)⇒(ii): A nontrivial zero of f can be writen as (x, y) with $f(x) = g(x)$ (by definition of $\hat{\ominus}$). If $a = g(x) = h(y)$ is $\neq 0$, we are done. So let $a = 0$ which means that at least one of the forms g and h represents 0, thus by proposition 2.48 all elements of $\mathbb{F}$ – in particular all non-zero values taken by h.

□

Theorem 2.53 (Witt's cancelation theorem)**.** *Let $f = g \hat{\oplus} h$ and $f' = g' \hat{\oplus} h'$ be two nondegenerate quadratic forms. If $f \sim f'$ and $h \sim h'$, one has $g \sim g'$.*

Proof. Corollary 2.39 gives the statement. □

Corollary 2.54. *If f is nondegenerate, then there exist hyperbolic $g_1, \ldots, g_m$ and h that does not represent zero with:*

$$f \sim g_1 \hat{\oplus} \ldots \hat{\oplus} g_m \hat{\oplus} h$$

and this decomposition is unique up to equivalence.

Proof. The existence follows from proposition 2.48 and uniqueness from theorem 2.53. □

3 The algorithm

In the main algorithm QUADRATIC-FORM-EQUIVALENCE presented in section 3.2 we will need to find rational solutions for so called "diagonal quadratic equations" which have the form:

$$\sum_{i=0}^{n} a_i x_i^2 = b \quad \text{where } a_i, b \in \mathbb{Q}.$$

For this we'll need some elementary number theory that will be presented in section 3.1. Additionally many results from section 2 will then prove the correctness of the algorithm.

3.1 Quadratic diagonal equations

Notation 3.1 (Square decomposition)**.** For $a \in \mathbb{Z}$, denote by $\widetilde{a} \in \mathbb{N}_{>0}$ the maximal number such that for some $\overline{a} \in \mathbb{Z}$ we can write

$$a = \widetilde{a}^2 \overline{a}.$$

Fact 3.2. *For $a \in \mathbb{Z}$, $\overline{a}$ is square-free.*

Lemma 3.3. *For all $\alpha = \frac{\alpha_1}{\alpha_2}, \beta = \frac{\beta_1}{\beta_2}, \gamma = \frac{\gamma_1}{\gamma_2} \in \mathbb{Q}$ there exist square free $a, b \in \mathbb{Z}$ such that*

$$\alpha x^2 + \beta y^2 = \gamma \neq 0 \tag{3.1}$$

is solvable over rationals if and only if

$$ax^2 + by^2 = z^2 \tag{3.2}$$

is solvable over integers with pairwise coprime x, y, z.
In great detail we have with

$$A := \alpha_1\beta_2\gamma_2 \qquad B := \alpha_2\beta_1\gamma_2 \qquad C := \alpha_2\beta_2\gamma_1$$

that

$$a = \overline{A} \cdot \overline{C} \qquad b = \overline{B} \cdot \overline{C}$$

such that one can obtain a solution $(x, y) = \left(\frac{u}{w}, \frac{v}{w}\right) \in \mathbb{Q}^2$ of eq. (3.1) from a solution $(\hat{x}, \hat{y}, \hat{z}) \in \mathbb{Z}^3$ of eq. (3.2) and vice versa by the following relations:

$$x = \frac{\hat{x}\overline{C}\widetilde{C}}{\hat{z}\widetilde{A}} \qquad y = \frac{\hat{y}\overline{C}\widetilde{C}}{\hat{z}\widetilde{B}}$$

$$\hat{x} = \frac{\widetilde{A}}{\widetilde{C}}u \qquad \hat{y} = \frac{\widetilde{B}}{\widetilde{C}}v \qquad \hat{z} = \overline{C}w.$$

Proof. Equation (3.1) is solvable over rationals if and only if

$$Ax^2 + By^2 = C$$

is solvable over rationals. The solvability of the last equation is equivalent to the solvability of the following equation:

$$\overline{A}\left(\frac{\widetilde{A}}{\widetilde{\overline{C}}}x\right)^2 + \overline{B}\left(\frac{\widetilde{B}}{\widetilde{\overline{C}}}y\right)^2 = \overline{C}. \tag{3.3}$$

Hence with the substitution $\hat{x} := \frac{\widetilde{A}}{\widetilde{\overline{C}}}x$ and $\hat{y} := \frac{\widetilde{B}}{\widetilde{\overline{C}}}y$ we want to solve

$$\overline{A}\hat{x}^2 + \overline{B}\hat{y}^2 = \overline{C} \tag{3.4}$$

over rationals. Now we use homogenization to switch to integers: Having a solution for eq. (3.4) means that there exists $u, v, w \in \mathbb{Z}$ with $w \neq 0$ such that

$$\overline{A}\left(\frac{u}{w}\right)^2 + \overline{B}\left(\frac{v}{w}\right)^2 = \overline{C}$$

or, equivalently

$$\overline{A}u^2 + \overline{B}v^2 = \overline{C}w^2$$

multiplying by $\overline{C} \neq 0$ yields

$$(\overline{C} \cdot \overline{A})u^2 + (\overline{C} \cdot \overline{B})v^2 = \overline{C}^2 w^2. \tag{3.5}$$

Now eq. (3.5) has an integer solution with $w \neq 0$ if and only if

$$(\overline{C} \cdot \overline{A})u^2 + (\overline{C} \cdot \overline{B})v^2 = z^2$$

has an integer solution with $z \neq 0$. Since $\overline{C}(\overline{a}u^2 + \beta v^2) = z^2$ we have that $\overline{C} \mid z^2$ and since $\overline{C}$ is square free $\overline{C} \mid z$ and therefore $w^2 = \frac{z^2}{\overline{C}^2} \in \mathbb{Z}$. □

Proposition 3.4. *For $x, y, z, d \in \mathbb{Z}$ with $x + y = z$ we have: If d divides two elements of the set $\{x, y, z\}$, then d divides all three elements of $\{x, y, z\}$.*

Proof. Without loss of generality assume that $d \mid x$ and $d \mid y$ Then there exist elements $u, v \in \mathbb{Z}$ such that $x = du$ and $y = dv$, hence we have:

$$z = du + dv = d(u + v) \quad \Rightarrow \quad d \mid z.$$

□

Proposition 3.5. *If a prime divides a product of integrals, it divides at least one of the factors.*

Proof. Let p be a prime and $a, b \in \mathbb{Z}$ with $p \mid ab$. We want to prove:

$$p \nmid a \Rightarrow p \mid b.$$

Set $g := \gcd(a, p)$. Then of course $g \mid p$. Since p is prime, we have that $g = 1$ or $g = p$. If $g = p$, since $g \mid a$ too, $p \mid a$ which is a contradiction. So $g = 1$. By the euclidean algorithm, there exist

$$x, y \in \mathbb{Z} : px + ay = 1.$$

Multiplying by b gives:

$$bpx + bay = b.$$

Observe that $p \mid pxb$ and $p \mid aby$ (since it is assumed that $p \mid ab$). Therefore, by proposition 3.4 it follows that $p \mid b$. □

Definition 3.6 (Norm of an element in a number field)**.** Note that for $a \in \mathbb{Q}$ an element of the number field $\mathbb{Q}(\sqrt{a})$ can be written as $\alpha + \beta\sqrt{a}$ with $\alpha, \beta \in \mathbb{Q}$. Now define the **norm** by

$$\begin{array}{rcl} N : \mathbb{Q}(\sqrt{a}) & \longrightarrow & \mathbb{Q} \\ \alpha + \beta\sqrt{a} & \longmapsto & \alpha^2 - a\beta^2. \end{array}$$

The norm is a multiplicative function.

Fact 3.7. *The norm is a multiplicative function.*

Proof. Let $\alpha + \beta\sqrt{a}, \alpha' + \beta'\sqrt{a} \in \mathbb{Q}(\sqrt{a})$ and calculate

$$\begin{aligned} N\left(\alpha + \beta\sqrt{a}\right) N\left(\alpha' + \beta'\sqrt{a}\right) &= \left(\alpha^2 - a\beta^2\right)\left(\alpha'^2 - a\beta'^2\right) \\ &= \alpha^2\alpha'^2 - a\alpha^2\beta'^2 - a\beta^2\alpha'^2 + a^2\beta^2\beta'^2 \\ &= \alpha^2\alpha'^2 + a^2\beta^2\beta'^2 - a\left(\alpha^2\beta'^2 + \beta^2\alpha'^2\right) \\ &= \alpha^2\alpha'^2 - 2\alpha\alpha'\beta\beta'a + a^2\beta^2\beta'^2 \\ &\quad - a\left(\alpha^2\beta'^2 - 2\alpha\beta'\alpha'\beta + \alpha'^2\beta^2\right) \\ &= \left(\alpha\alpha' + a\beta\beta'\right)^2 - a\left(\alpha\beta' + \alpha'\beta\right)^2 \\ &= N\left(\alpha\alpha' + a\beta\beta' + (\alpha\beta' + \alpha'\beta)\sqrt{a}\right) \\ &= N\left(\left(\alpha + \beta\sqrt{a}\right)\left(\alpha' + \beta'\sqrt{a}\right)\right) \end{aligned}$$

□

Lemma 3.8. *Let $a, b \in \mathbb{Z}$ be square-free with $|a| < |b|$ and $1 < |b|$. Then there exists $b' \in \mathbb{Z}$ with $|b'| < |b|$ such that*

$$ax^2 + by^2 = z^2 \tag{3.6}$$

has a solution if and only if

$$ax^2 + b'y^2 = z^2 \tag{3.7}$$

has a solution and this solution can be converted into each other effectively.

Proof. If eq. (3.6) has a solution, then for any $p \mid b$ we have that p cannot divide x, since otherwise:

$$\begin{array}{rcl} p \mid b \text{ and } p \mid x & \Rightarrow & 0 \equiv_p ax^2 + by^2 = z^2 \\ & \Rightarrow & p \mid z^2 \\ & \overset{3.5}{\Rightarrow} & p \mid z \\ & \Rightarrow & p^2 \mid z^2, z^2 = ax^2 + by^2 \\ & \overset{p \mid x}{\Rightarrow} & p^2 \mid by^2 \end{array}$$

This means that $\exists\, n \in \mathbb{Z} : np^2 = by^2$. We furthermore know that $p \mid b$, meaning that $\exists\, m \in \mathbb{Z} : mp = b$ with the additional property that $p \nmid m$ since b is square

free. Putting this together, we get

$$\begin{aligned} np^2 = by^2 = mpy^2 \quad &\Leftrightarrow \quad np = my^2 \\ &\Rightarrow \quad p \mid my^2 \\ &\overset{3.5}{\Rightarrow} \quad p \mid m \text{ or } p \mid y^2 \\ &\overset{p \nmid m}{\Rightarrow} \quad p \mid y^2 \\ &\overset{3.5}{\Rightarrow} \quad p \mid y \end{aligned}$$

which is a contradiction since x, y, z were assumed to be coprime. So $p \nmid x$ which means that $x \in \mathbb{F}_p^*$ is invertible and therefore we get

$$z^2 = ax^2 + by^2 \equiv_p ax^2 \quad \Leftrightarrow \quad z^2 \left(x^{-1}\right)^2 \equiv_p a$$

i.e. a is a square modulo p. By the chinese remainder theorem $a \in \mathbb{Z}$ is a square. Thus there is a $t \in \mathbb{Z}$ such that $|t| \leq \frac{|b|}{2}$ and $a \equiv_b t^2$. Let $b' \in \mathbb{Z}$ be such that

$$t^2 = a + bb'. \tag{3.8}$$

We now claim that $ax^2 + by^2 = z^2$ has a solution if and only if $ax^2 + b'y^2 = z^2$ has a solution: If $ax^2 + by^2 = z^2$ has a solution then

$$\begin{aligned} & N\left(\frac{z + x\sqrt{a}}{y}\right) = \frac{z^2}{y^2} - a\frac{x^2}{y^2} \\ \iff & y^2 N\left(\frac{z + x\sqrt{a}}{y}\right) = z^2 - ax^2 \\ \iff & ax^2 + y^2 N\left(\frac{z + x\sqrt{a}}{y}\right) = z^2 \\ \implies & N\left(\frac{z + x\sqrt{a}}{y}\right) = b. \end{aligned}$$

Also from eq. (3.8) we get:

$$\begin{aligned} & bb' = t^2 - a = N\left(t + \sqrt{a}\right) \\ \implies & b' = \frac{N\left(t + \sqrt{a}\right)}{b} = \frac{N\left(t + \sqrt{a}\right)}{N\left(\frac{z + x\sqrt{a}}{y}\right)} \overset{\text{fact } 3.7}{=} N\left(\frac{yt + y\sqrt{a}}{z + x\sqrt{a}}\right). \end{aligned}$$

Which, by expanding the fraction by $z - x\sqrt{a}$ i.e. rationalizing the denominator, effectively gives an integral solution of eq. (3.7). Since the argument is symmetric in b and b' we get a solution of eq. (3.6) from a solution of eq. (3.7).

We are left to show that $|b'| < |b|$:

$$\begin{aligned} |a| + |b'| &= |a| + \left|\frac{t^2 - a}{b}\right| \\ &\leq |a| + \left|\frac{t^2}{b}\right| + \left|\frac{a}{b}\right| \\ &\overset{|a|<|b|}{\leq} |a| + \left|\frac{t^2}{b}\right| + 1 \\ &\overset{|t|\leq\frac{|b|}{2}}{\leq} |a| + \frac{|b|}{4} + 1 \\ &\overset{1<|b|}{<} |a| + |b| \end{aligned}$$

□

Corollary 3.9. *Let $a, b \in \mathbb{Z}$ be square-free with $|a| < |b|$ and $1 < |b|$. Then to determine if a solution of*

$$ax^2 + by^2 = z^2$$

exists and to calculate it, can be done effectivly.

Proof. Repeatedly apply lemma 3.8 to end up with one of the following equations:

$$\pm x^2 \pm y^2 = z^2 \text{ or } \pm x^2 = z^2.$$

Their solvability over $\mathbb{Z}$ is easy to check and they are solvable just as easy. Since lemma 3.8 gives an effective way to convert solutions, the whole process is effective too. □

3.2 Rational quadratic forms

Theorem 3.10. *A generalization of lemma 3.3 and corollary 3.9 for arbitrary $n \in \mathbb{N}$ give an algorithm to effectively compute a solution for a diagonal quadratic equation.*

Fact 3.11. *For any $f \in \mathbb{F}[x_1, \ldots, x_n]^{=2}$, we can write*

$$f(x_1, \ldots, x_n) = (x_1, \ldots, x_n)A(x_1, \ldots, x_n)^T$$

for a symmetric $A \in \mathbb{F}^{n\times n}$ with entries $a_{ij} \in \mathbb{F}$. Since A is a symmetric matrix over a field of characteristic not equal to 2, we can apply Gaussian elimination to optain $C \in \mathrm{Gl}_n(\mathbb{F})$ such that CAC^T is diagonal. Call the diagonal elements $b_i \in \mathbb{F}$. Then we have

$$f((x_1, \ldots, x_n)C) = (x_1, \ldots, x_n)CAC^T(x_1, \ldots, x_n)^T = \sum_{i=1}^{n} b_i x_i^2.$$

This process makes theorem 2.50 explicit and effective.

Fact 3.12. *For $f, g \in \mathbb{F}[x]^{=2}$ with $f(x) = ax^2$ and $g(x) = bx^2$ we have that*

$$f \sim g \Leftrightarrow \frac{a}{b} \in \mathbb{F}^{*2}.$$

This criterium can be checked in polynomial time for $\mathbb{F} = \mathbb{Q}$.

Proof. Being equivalent for this two forms means that there exists a number $\lambda \in \mathbb{F}^*$ such that $g(\lambda x) = f(x)$:

$$g(\lambda x) = f(x) \Leftrightarrow b(\lambda x)^2 = ax^2 \Leftrightarrow b\lambda^2 x^2 = ax^2 \Leftrightarrow b\lambda^2 = a \Leftrightarrow \lambda^2 = \frac{a}{b}.$$

The last equation is solvable if and only if $\frac{a}{b}$ is a square in $\mathbb{F}$ i.e. $\frac{a}{b} \in \mathbb{F}^{*2}$. For $\mathbb{F} = \mathbb{Q}$ this is the case if and only if the nominator and denominator are squares in $\mathbb{Z}$. To check this one can simply perform a binary search which can be done in linear time in the number of digits see the following Check-Perfect-Square-Algorithm:

Algorithm 1 Check-Perfect-Square-Algorithm

Input: $z \in \mathbb{Z}$.
Output: **true** if $\sqrt{z} \in \mathbb{Z}$, **false** else.

```
1: if z < 0 then
2:     return false
3: end if
4: set x := z DIV 2, S := { x }
5: while x^2 ≠ z do
6:     set x := (x + (z DIV x)) DIV 2
7:     if x ∈ S then
8:         return false
9:     end if
10:    set S := S ∪ { x }
11: end while
12: return true
```

□

Now we come to the main algorithm that I want to present in this script. It can decide the QuadraticFormEquiv$_{\mathbb{Q}}$-Problem in exponential time and can also be used to find such an equivalence. It is noteworthy that it can be generalized for other fields too, for example $\mathbb{F} = \mathbb{F}_q$ for a prime power q or $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Except for step 10 the algorithm boils down to linear algebra and the results from the previous section. The algorithm for 10 in the case $\mathbb{F} = \mathbb{Q}$ is given above. For other fields consult the following references:

- $\mathbb{F} = \mathbb{F}_q$ for a prime power q: by a classical theorem of Weil (see [Bac96]) for a random choice of $x_1, \ldots, x_n \in \mathbb{F}_q$ there exists $x_n \in \mathbb{F}_q$ solving the equation.
- $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$: One can just choose i such that $a_i \neq 0$, set $x_i = \sqrt{b/a_i}$ and $x_j = 0$ for $j \neq i$.

Algorithm 2 QUADRATIC-FORM-EQUIVALENCE

Input: $f, g \in \mathbb{Q}[x_1, \ldots, x_n]^{=2}$.
Output: **true** if $f \sim g$, **false** else.

1: By fact 3.11 assume $f = \sum_{i=1}^{n} a_i x_i^2$ and $g = \sum_{i=1}^{n} b_i x_i^2$ with $a_i, b_i \in \mathbb{Q}$.
2: Without loss of generality **set** $n = \operatorname{rank}(f)$ and permute the variables such that $a_i, b_i \in \mathbb{Q}^* \ \forall i \in [n]$.
3: **if** $\operatorname{rank}(f) \neq \operatorname{rank}(g)$ **then**
4: **return false**
/* fact 2.51 */
5: **end if**
6: **if** $\operatorname{rank}(f)$=1 **then**
7: Write $f(x) = ax^2$ and $g(x) = bx^2$
8: **return** truth value of $\frac{a}{b} \in \mathbb{Q}^{*2}$
/* fact 3.12 */
9: **end if**
10: Theorem 3.10 gives a solution $\alpha \in \mathbb{Q}^n$ of the diagonal quadratic equation $f(x_1, \ldots, x_n) = b_n$.
11: The subspace $U := \langle \alpha \rangle^{\perp}$ is nondegenerate since $b_n \neq 0$, which means by proposition 2.20(ii) that we have

$$V = \langle \alpha \rangle \hat{\oplus} U$$

So every $v \in V$ can be written as $v = \lambda\alpha + u$ with $\lambda \in \mathbb{Q}$ and $u \in U$. Thus

$$\begin{aligned} f(v) = v.v &= (\lambda\alpha + u).(\lambda\alpha + u) = \lambda^2 \alpha.\alpha + u.u \\ &= \lambda^2 f(\alpha) + f(u) = \lambda^2 b_n + f(u). \end{aligned}$$

This simply means that $f \sim b_n x_n^2 \hat{\oplus} f_1(x_1, \ldots, x_{n-1})$ for some quadratic form $f_1 \in \mathbb{Q}[x_1, \ldots, x_{n-1}]$.
12: Now we have

$$\begin{aligned} f &\sim b_n x_n^2 \hat{\oplus} f_1(x_1, \ldots, x_{n-1}) \\ g(x_1, \ldots, x_n) &= b_n x_n^2 \hat{\oplus} \sum_{i=1}^{n-1} b_i x_i^2 \end{aligned}$$

Theorem 2.53 (Witt's cancelation theorem) then says that:

$$\begin{aligned} b_n x_n^2 \hat{\oplus} f_1(x_1, \ldots, x_{n-1}) &\sim b_n x_n^2 \hat{\oplus} \sum_{i=1}^{n-1} b_i x_i^2 \\ \iff \quad f_1(x_1, \ldots, x_{n-1}) &\sim \sum_{i=1}^{n-1} b_i x_i^2 \end{aligned}$$

13: **set** $f := f_1, g := \sum_{i=1}^{n-1} b_i x_i^2$ and **goto** 1.

Theorem 3.13. *We have that*

$$\textsc{QuadraticFormEquiv}_{\mathbb{Q}} \in \mathbf{EXP}$$

and the equivalence can also be found in **EXP**.

References

[AS05] Manindra Agrawal and Nitin Saxena. Automorphisms of finite rings and applications to complexity of problems. pages 1–17, 2005.

[AS06a] Manindra Agrawal and Nitin Saxena. Equivalence of $\mathbb{F}$-algebras and cubic forms. pages 115–126, 2006.

[AS06b] Manindra Agrawal and Nitin Saxena. On the complexity of cubic forms. to be submitted, 2006.

[Bac96] Eric Bach. Weil bounds for singular curves. *Applicable Algebra in Engineering, Communication and Computing*, 7:289–298, 1996.

[Bro06] W. D. Brownawell. Bounds for the degrees in the nullstellensatz. pages 577–591, 2006.

[DH88] J. Davenport and J. Heintz. Real quantifier elimination is doubly exponential. pages 29–35, 1988.

[KS05] Neeraj Kayal and Nitin Saxena. On the ring isomorphism and automorphism problems. pages 2–12, 2005.

[Ser73] Jean-Pierre Serre. *A Course in Arithmetic*. Springer, 1973.