RATIONAL POINTS AND GENERALIZED TRACE FORMS ON A FINITE ALGEBRA OVER A REAL CLOSED FIELD

DILIP P. PATIL\textsuperscript{1*} AND JUGAL VERMA\textsuperscript{2}

ABSTRACT. The core of this article is the equality of the number of $K$-rational points with the signature of the trace form of a finite $K$-algebra over a real closed field $K$. The main tools are symmetric bilinear forms, hermitian forms, trace forms, generalized trace forms and their types and signatures. Further, we prove a criterion for the existence of $K$-rational points by using generalized trace forms. As an application we prove the Pederson-Roy-Szpirglas theorem about counting common real zeros of real polynomial equations. We also deduce this theorem as a special case of a similar classical theorem for real closed fields.

\section{Introduction}

The objective of this paper is to present an exposition of classical and modern results concerning the number of real or complex points in the solution space of a finite system of polynomial equations with real coefficients in arbitrary number of variables. More precisely, for $F_1, \ldots, F_m \in \mathbb{R}[X_1, \ldots, X_n]$, assume that the set of common zeros $V_{\mathbb{R}}(F_1, \ldots, F_m)$ (resp. $V_{\mathbb{C}}(F_1, \ldots, F_m)$) of $F_1, \ldots, F_m$ in $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) is finite, or equivalently, the $\mathbb{R}$-algebra $\mathbb{R}[X_1, \ldots, X_n]/(F_1, \ldots, F_m)$ (resp. the $\mathbb{C}$-algebra $\mathbb{C}[X_1, \ldots, X_n]/(F_1, \ldots, F_m)$) is finite dimensional over $\mathbb{R}$ (resp. over $\mathbb{C}$). By the classical Hilbert’s nullstellensatz $V_{\mathbb{C}}(F_1, \ldots, F_m) \neq \emptyset$ if and only if the ideal $(F_1, \ldots, F_m)$ generated by $F_1, \ldots, F_m$ in $\mathbb{C}[X_1, \ldots, X_n]$ is a non-unit ideal. However this is not true over the field $\mathbb{R}$ or more generally over real closed fields. Therefore the natural questions one deals with are when exactly $V_K(F_1, \ldots, F_m) \neq \emptyset$ and how to find its cardinality, where $K$ is an arbitrary real closed field.

Many researchers have been studying the above problems by using various methods including effectiveness and complexity analysis. For example, already in the 19th century Sturm, Jacobi, Borchardt, Sylvester, Hermite, Hurwitz proved fundamental results for counting real points (in small number of variables $n = 1$ or $n = 2$) by using the signature of appropriate quadratic forms. Section 2 and Section 3 are essentially expository in nature. For the sake of completeness, in these sections we collect standard results on symmetric and Hermitian bilinear forms over an arbitrary real closed field $K$ and its algebraic closure $\mathbb{C}_K = K[i]$ with $i^2 = -1$. In particular, we provide complete proofs of Sylvester’s law of inertia and Hurwitz’s criteria. Moreover, using strong topology on $K^n$ we present very useful and important Rigidity theorem for quadratic forms which will be used crucially in the proof of Theorem 4.6 in section 4.

In Section 4, we use elementary commutative algebra to study trace form and generalized trace forms on finite dimensional algebra $A$ over a real closed field $K$. We give a (new) proof of Theorem 4.6 which relate the $K$-rational points of $A$ with the type of the trace form $\text{tr}_A$ on $A$ and derive some consequences. Further, in Theorem 4.11, we prove a criterion for the existence of $K$-rational points by using generalized trace forms on $A$.

\textsuperscript{1}This expository article is an expanded version of the lecture on the Pederson-Roy-Szpirglas theorem about counting real roots of real polynomial equations, given by the second author at the International Conference on Computational Mathematics held at Banasthali Vidyapeeth at Niwai (Rajasthan) in 2017. Most of the exposition is influenced by the discussions of the first author with late Prof. Dr. Uwe Storch (1940-2017) and the lecture courses delivered by him. Prof. Uwe Storch was known for his work in commutative algebra, analytic and algebraic geometry, in particular derivations, divisor class group and resultants.

\textsuperscript{*}During the preparation of this work the second author was visiting Department of Mathematics, Indian Institute of Technology Bombay. He would like to express his gratitude for the generous financial support from IIT Bombay and encouraging cooperation.

\textit{Key words and phrases.} Real closed fields, Finite $K$-algebra, Hermitian forms, Quadratic forms, Sylvester’ law of inertia, Type, Signature, Trace forms.
In Section 5, we compute the cardinality of the $K$-rational points of finite algebra over real closed field $K$. The main ingredient in this section is the Shape Lemma 5.3 which gives a very useful presentation of the radical ideal $\mathfrak{A} \subseteq K[X_1, \ldots, X_n]$ from which one can read better information about the set of $K$-rational points $V_K(\mathfrak{A})$ which reduces the problem one variable case. In Theorem 5.4 we relate type, signature and rank of a generalized trace forms on $A = K[X_1, \ldots, X_n]/\mathfrak{A}$ with the number of points in $V_K(\mathfrak{A})$ and in $V_{\mathfrak{F}}(\mathfrak{A})$. Finally, we give a proof of theorem of Pederson-Roy-Szpirglas [11] Theorem 2.1 which was the main goal of the lecture delivered by the second author at the International Conference on Computational Mathematics held at Banasthali Vidyapeeth at Niwai (Rajasthan) in 2017.

§2 Symmetric bilinear forms

In this section we collect basic results which are used in this article. Most of these results can be found in standard graduate text books, for instance see [13] Ch. V, §12], [12] Ch. IX or [7] Ch. 11. However, for setting the notation, terminology and for the sake of completeness, we recall them in the format that they are used with sketchy proofs and leave the routine details for the reader to verify.

2.1 Definitions Let $V$ and $W$ be vector spaces over an arbitrary field $K$. A function $\Phi : V \times W \to K$ is called bilinear if $\Phi$ is $K$-linear in both the components, i.e. if for all $a, a' \in K$ and all $x, x' \in V, y, y' \in W$, we have:

1. $\Phi(ax + a'x', y) = a\Phi(x, y) + a'\Phi(x', y)$.
2. $\Phi(x, ay + a'y') = a\Phi(x, y) + a'\Phi(x, y')$.

The set of bilinear functions $V \times W \to K$ is denoted by $\text{Mult}_K(V, W)$ which is clearly a subspace of the $K$-vector space $K^{V \times W}$. A bilinear function $\Phi : V \times V \to K$ is called symmetric bilinear form on $V$ if for all $x, y \in V$, we have $\Phi(x, y) = \Phi(y, x)$.

2.2 Gram’s Matrix Let $V$ and $W$ be finite dimensional $K$-vector spaces with bases $\mathfrak{r} := \{x_i \mid i \in I\}$, and $\mathfrak{r}' := \{y_j \mid j \in J\}$, respectively. For the bilinear function $\Phi : V \times W \to K$, from definition it follows immediately that $\Phi$ is uniquely determined by the values $\Phi(x_i, y_j), (i, j) \in I \times J$. Conversely, for arbitrary family $c_{ij} \in K, (i, j) \in I \times J$, there is a (unique) bilinear function $\Phi : V \times W \to K$ defined by $\Phi(\sum_{i \in I} a_i x_i, \sum_{j \in J} b_j y_j) := \sum_{(i, j) \in I \times J} a_i b_j c_{ij}$. Therefore, the map

$$\text{Mult}_K(V, W) \longrightarrow K^{I \times J} = M_{I,J}(K), \Phi \longmapsto (\Phi(x_i, y_j))_{(i, j) \in I \times J},$$

an isomorphism of $K$-vector spaces.

For a bilinear function $\Phi : V \times W \to K$ the $I \times J$-matrix $\mathcal{G}^{\mathfrak{r}, \mathfrak{r}'}(\Phi) := (\Phi(x_i, y_j))_{(i, j) \in I \times J} \in M_{I,J}(K)$ is called the Gram’s matrix\footnote{Gram, J. P. (1850-1916) was a Danish actuary and mathematician who is best remembered for the Gram-Schmidt orthogonalization process published in 1883. He was not however the first to use this method. The process seems to be a result of Laplace and it was essentially used by Cauchy in 1836.} or the fundamental matrix of $\Phi$ with respect to the bases $\mathfrak{r}$ and $\mathfrak{r}'$. In the case $V = W$ and $\mathfrak{r} = \mathfrak{r}'$, we simply write $\mathcal{G}^{\mathfrak{r}}(\Phi)$. Further, in the case $I = J$, the determinant $G_{\Phi}(\mathfrak{r}, \mathfrak{r}') := \det \mathcal{G}^{\mathfrak{r}, \mathfrak{r}'}(\Phi)$ is called the Gram’s determinant with respect to the bases $\mathfrak{r}$ and $\mathfrak{r}'$. In the case $V = W$ and $\mathfrak{r} = \mathfrak{r}'$, we simply write $G_{\Phi}(\mathfrak{r})$.

Let $\mathfrak{r} = (x_i)_{i \in I}, \mathfrak{r}' = (x'_i)_{i \in I}$ be $K$-bases of $V$ and $\mathfrak{r} = (y_i)_{i \in I}, \mathfrak{r}' = (y'_i)_{i \in I}$ be $K$-bases of $W$. Let $\mathfrak{A} = (a_{ij}) \in \text{GL}_I(K)$ and $\mathfrak{B} = (b_{ij}) \in \text{GL}_J(K)$ be the transition matrices of the bases from $\mathfrak{r}$ to $\mathfrak{r}'$ and from $\mathfrak{r}$ to $\mathfrak{r}'$, respectively. Then for a bilinear function $\Phi : V \times W \to K$, we have the transformation formula: $\mathcal{G}_{\Phi}(\mathfrak{r}') = \mathfrak{A} \mathcal{G}_{\Phi}(\mathfrak{r}) \mathfrak{B}$. In particular, if $V = W$, then the above transformation formula has the form: $\mathcal{G}_{\Phi}(\mathfrak{r}') = \mathfrak{A} \mathcal{G}_{\Phi}(\mathfrak{r}) \mathfrak{A}$, where $\mathfrak{A} \in \text{GL}_I(K)$ is the transition matrix of the change of the basis (from the basis $\mathfrak{r}$ to the basis $\mathfrak{r}'$).
2.3 Examples (a) (Standard forms) Let $K$ be a field, and $I, J$ be sets. The bilinear form on $K^{(I)}$ whose Fundamental matrix with respect to the standard basis is the unit matrix called the standard form on $K^{(I)}$ and is denoted by $\langle -,- \rangle$. Therefore $\langle (a_i), (b_i) \rangle = \sum_{i \in I} a_i b_i = a \cdot b$, where $a := (a_i), b := (b_i) \in K^{(I)}$. In particular, for $I = \{1, \ldots, n\}$, it is

$$
\left( \begin{array}{c} a_1 \\ \vdots \\ a_n \\
\end{array} \right), \left( \begin{array}{c} b_1 \\ \vdots \\ b_n \\
\end{array} \right) = (a_1, \ldots, a_n) \left( \begin{array}{c} b_1 \\ \vdots \\ b_n \\
\end{array} \right) = a_1 b_1 + \cdots + a_n b_n.
$$

(b) (Natural Duality) Let $V$ be an arbitrary vector space over a field $K$ and let $V^* = \text{Hom}_K(V,K)$ be the dual space of $V$. Then the evaluation map $V \times V^* \rightarrow K, (x,f) \mapsto f(x)$, is a bilinear map $V \times V^* \rightarrow K$. This map is called the natural duality between $V$ and $V^*$. If $V$ is finite dimensional with basis $x_i, i \in I$, and if $x_i^*, i \in I$, the corresponding dual basis, then the fundamental matrix of the natural duality with respect to these bases is clearly the unit matrix $\mathbf{E}$.

2.4 Non-degeneracy and Complete Duality An important motivation for the study of bilinear functions is the description of linear form through vectors. Let $\Phi: V \times W \rightarrow K$ be a bilinear function. Then for every $y \in W$, the map $\Phi(-,y): V \rightarrow K, x \mapsto \Phi(x,y)$, is a linear form on $V$ and for every $x \in V$, then map $\Phi(x,-): W \rightarrow K, y \mapsto \Phi(x,y)$ is a linear form $W$. Therefore we have the canonical maps: $\Phi_1: W \rightarrow V^*$ and $\Phi_2: V \rightarrow W^*$. Both these maps $\Phi_1$ and $\Phi_2$ are linear and from $\Phi_1$ or from $\Phi_2$, we can recover $\Phi$: $\Phi(x,y) = (\Phi_1(y))(x) = (\Phi_2(x))(y)$.

Suppose that $V$ and $W$ are finite dimensional with bases $\mathbf{x} = (x_i)_{i \in I}$ and $\mathbf{y} = (y_j)_{j \in J}$ of $V$ and $W$, respectively. The matrices of $\Phi_1$ and $\Phi_2$ with respect to bases $\mathbf{x}$ and $\mathbf{y}$ and their dual bases $\mathbf{x}^*$ and $\mathbf{y}^*$ are given by: $
abla_{\mathbf{y}}^{\mathbf{x}}(\Phi_1) = \Theta^{\mathbf{x}^* \mathbf{y}}(\Phi)$, and $\nabla_{\mathbf{y}}^{\mathbf{x}}(\Phi_2) = \Theta^{\mathbf{x}^* \mathbf{y}}(\Phi)$.

Note that $\Phi_1$ and $\Phi_2$ have the same rank, since by taking transpose its rank is unaltered. In this situation, the common rank of the maps $\Phi_1, \Phi_2$ is called the rank of the bilinear function $\Phi$ and is denoted by $\text{rank} \Phi$. Therefore, $\text{rank} \Phi$ is the rank of the fundamental matrix of $\Phi$ with respect to arbitrary bases of $V$ and $W$. The case when $\Phi_1$ and $\Phi_2$ are injective or bijective are important:

2.5 Definition Let $\Phi: V \times W \rightarrow K$ be a bilinear function. We say that

(i) $\Phi$ is non-degenerate if $\Phi_1$ and $\Phi_2$ are injective.

(ii) $\Phi$ defines a complete duality (between $V$ and $W$) if $\Phi_1$ and $\Phi_2$ are bijective.

Complete duality is possible only in the case when both vector spaces $V$ and $W$ are finite dimensional. Moreover, in the finite dimensional case we prove the following:

2.6 Theorem For a bilinear function $\Phi: V \times W \rightarrow K$, with $V$ and $W$ finite dimensional $K$-vector spaces, the following statements are equivalent:

(i) $\Phi$ defines a complete duality.

(ii) $\Phi$ is non-degenerate.

(iii) $\text{Dim}_K V = \text{Dim}_K W$, and one of the canonical maps $\Phi_1$ and $\Phi_2$ is injective.

(iv) $\text{Dim}_K V = \text{Dim}_K W$, and for one (and hence every) basis $\mathbf{x} = (x_i)_{i \in I}$ of $V$ and $\mathbf{y} = (y_i)_{i \in J}$ of $W$ the Fundamental matrix $\Theta^{\mathbf{x}^* \mathbf{y}}(\Phi)$ is invertible.

(v) $\text{Dim}_K V = \text{Dim}_K W = \text{Rank} \Phi$.

Proof The implications (i)$\Rightarrow$(ii), (iii)$\Rightarrow$(iv), (iv)$\Rightarrow$(v) and (v)$\Rightarrow$(i) are clear by the above discussion. For the implication (ii)$\Rightarrow$(iii) it remains to prove that $\text{Dim}_K V = \text{Dim}_K W$. From the injectivity of $\Phi_1$ it follows $\text{Dim}_K V \leq \text{Dim}_K V^* = \text{Dim}_K V$. Analogously, from the injectivity of $\Phi_2$, we have the inequality $\text{Dim}_K V \leq \text{Dim}_K W$.

2.7 Example For a finite dimensional vector space $V$ the natural duality between $V$ and $V^*$ is a complete duality, since its Fundamental matrix with respect to dual bases is the unit matrix. In particular, in the case of finite dimensional $V$ every linear form $\phi: V^* \rightarrow K, f \mapsto \phi(f) = f(x)$, where the vector $x \in V$ is uniquely determined by $\phi$; therefore $\phi$ is the evaluation of the linear forms in $V^*$ at a vector $x \in V$ (which is dependent on $\phi$). If $V$ is not finite dimensional, then the natural duality is non-degenerate but never complete. This means: for every $x \in V, x \neq 0$, there exists a linear form $f \in V^*$ with $\langle x, f \rangle = f(x) \neq 0$. This follows from the
fact that \( x \) (also in the infinite dimensional case) can be extended to a basis \( V \) (one can then assign arbitrary given values to \( f \) on this basis).

### 2.8 Definition

Let \( \Phi: V \times W \rightarrow K \) be a bilinear function.

1. The vectors \( x \in V \) and \( y \in W \) are called **orthogonal** or **perpendicular** to each other with respect to \( \Phi \) if \( \Phi(x,y) = 0 \). In this case we write \( x \perp y \).

2. Two subsets \( M \subseteq V \) and \( N \subseteq W \) are called **orthogonal** if \( x \perp y \) for all \( x \in M \) and for all \( y \in N \). In this case we write \( M \perp N \). Further, let

\[
M^\perp := \{ y \in W \mid M \perp \{ y \} \} \quad \text{and} \quad N^\perp := \{ x \in V \mid \{ x \} \perp N \}
\]

Obviously, \( M^\perp \) and \( N^\perp \) are \( K \)-subspaces of \( W \) and \( V \), respectively.

### 2.9 Example (Gradient of a linear form)

Let \( V \) and \( W \) be \( K \)-vector spaces. Suppose that a bilinear function \( \Phi: V \times W \rightarrow K \) defines a complete duality. Then since the canonical maps \( \Phi_1: V \rightarrow V^* \) and \( \Phi_2: W \rightarrow W^* \) are bijective. For every linear form \( f \in W^* \) (resp. \( e \in V^* \)) there exists a unique vector \( x_f \in V \) (resp. \( y_e \in W \), such that \( f = \Phi(x_f, -) \) (resp. \( e = \Phi(-, y_e) \)). The vectors \( x_f \in V \) (resp. \( y_e \in W \)) is called the **gradient** of \( f \) (resp. \( e \)) with respect to \( \Phi \) and is denoted by \( \text{grad} f \) (resp. \( \text{grad} e \)). The linear forms on \( W \) (resp. \( V \)) correspond to its respective gradients: It is \( f = \Phi_2(\text{grad} f) \) and \( e = \Phi_1(\text{grad} e) \). Further, \( \text{Ker} f \) (resp. \( \text{grad} f \)) and \( \text{Ker} e \) is \( \text{grad} e \).

### 2.10 Example (Orthogonal direct sums)

Let \( V_i, i \in I; W_i, i \in I \) be two families of \( K \)-vector spaces and let \( \Phi_i: V_i \times W_i \rightarrow K \) be a family of bilinear functions. Then the map

\[
\Phi: \left( \bigoplus_{i \in I} V_i \right) \times \left( \bigoplus_{i \in I} W_i \right) \rightarrow K,
\]

is a bilinear map and its restrictions \( \Phi|_{V_i \times W_i} = \Phi_i \) for all \( i \in I \), where \( V_i \) (resp. \( W_i \)) is considered canonically as subspace of \( \bigoplus_{i \in I} V_i \) (resp. \( \bigoplus_{i \in I} W_i \)). Further, \( V_i \perp W_i \) with respect to \( \Phi \) for all \( i, j \in I, i \neq j \). This bilinear function \( \Phi \) is called the **orthogonal direct sum** of the family \( \Phi_i, i \in I \) and is denoted by \( \bigoplus_{i \in I} \Phi_i \). Conversely, if \( \Phi: V \times W \rightarrow K \) is a bilinear function and \( V \) (resp. \( W \)) is a direct sum of the \( K \)-subspaces \( V_i, i \in I \) (resp. \( W_i, i \in I \)) with \( V_i \perp W_i \) for all \( i, j \in I, i \neq j \) with respect to \( \Phi \), such that \( \Phi(\sum_{i \in I} v_i, \sum_{j \in J} w_j) = \sum_{i \in I} \Phi_i(v_i, w_j) \) for \( v_i \in V_i, w_j \in W_j \). Then we say that \( \Phi \) is the orthogonal direct sum of the \( \Phi_i, i \in I \). In particular, if \( V = W \) and \( V_i = W_i, i \in I \), then \( V \) is the orthogonal direct sum of the subspaces \( V_i, i \in I \), with respect to \( \Phi \) and is denoted by \( V = \bigoplus_{i \in I} V_i \).

### 2.11 Theorem

Let \( V \) and \( W \) be finite dimensional \( K \)-vector spaces and \( \Phi: V \times W \rightarrow K \) be a non-degenerate bilinear function. Then the maps \( V \rightarrow V^\perp \) and \( W \rightarrow W^\perp \) are anti-isomorphisms between the lattices of \( K \)-subspaces of \( V \) resp. of \( W \) which are inverses of each other.

**Proof** Obviously, both these maps are inclusion reversing. Further, for subspaces \( V' \subseteq V \) and \( W' \subseteq W \), using Theorem 2.6, it is easy to check that:

\[
\text{Dim}_K V' + \text{Dim}_K V'^\perp = \text{Dim}_K V = \text{Dim}_K W' + \text{Dim}_K W'^\perp = (V^\perp)^\perp = V' \quad \text{and} \quad (W^\perp)^\perp = W'.
\]

Now, combining all these equalities, the assertion follows.

### 2.12 Perpendicularity relation

Let \( V \) be a vector space over a field \( K \) and let \( \Phi: V \times V \rightarrow K \) be a bilinear form on \( V \). Then for a subset \( M \subseteq V \), the subsets \( M^\perp = \{ y \in V \mid M \perp \{ y \} \} \) and \( \perp M = \{ \{ x \} \subseteq V \mid \{ x \} \perp M \} \) are not equal in general, since the relation of perpendicularity is not symmetric. To remove this difference, we consider symmetric bilinear forms. For this note that \( \Phi \) on \( V \) is symmetric if and only if \( \Phi_1 = \Phi_2 \), where \( \Phi_1: V \rightarrow V^* \) and \( \Phi_2: V \rightarrow V^* \) are the canonical maps associated to \( \Phi \), see 2.3. Further, for a symmetric form \( \Phi \) on \( V \), the relation \( \perp \) on \( V \) is symmetric. In this case the subspace \( \perp V = V^\perp = \text{Ker} \Phi_1 = \text{Ker} \Phi_2 \) is called the **radical** of \( \Phi \) and is also denoted by \( \text{Rad}(V, \Phi) \).

### 2.13 Proposition

Let \( \Phi: V \times V \rightarrow K \) be a bilinear form on \( V \) and \( \{ x_i \}_{i \in I} \) is a basis of \( V \). Then \( \Phi \) is symmetric if and only if the Gram’s matrix \( \Phi_\Phi(\{ x \}) = (\Phi(\{ x_i \}, \{ x_j \}) \in M_I(K) \) is symmetric.

### 2.14 Proposition

Let \( \Phi: V \times V \rightarrow K \) be a bilinear function on the finite dimensional \( K \)-vector space \( V \) and let \( \{ x_i \}_{i \in I} \) be a basis of \( V \). Then the map \( \Phi \mapsto \Phi_\Phi(\{ x \}) \) is a \( K \)-linear isomorphism of the \( K \)-vector space of symmetric bilinear forms on \( V \) onto the \( K \)-vector space of symmetric matrices in \( M_I(K) \). Moreover, if \( \{ x_i \}_{i \in I} \) is another basis of \( V \) with transition matrix \( \mathcal{A} = (a_{ij}) \in \text{GL}_I(K) \),
i.e. \( x_i = \sum_{i \in I} a_{ij} x'_j \), then the Gram’s matrices \( \Phi(x) \) and \( \Phi(x') \) are related by the rule: \( \Phi(x) = {}^t \mathbf{A} \Phi(x') \mathbf{A} \) resp. \( \Phi(x') = {}^t \mathbf{A}^{-1} \Phi(x) \mathbf{A}^{-1} \).

In important cases a bilinear form \( \Phi : V \times V \to K \) is completely determined by its values on the diagonal \( \Delta_V = \{ (x, x) \mid x \in V \} \). More precisely:

2.15 **Polarisation Theorem** Let \( V \) be a vector space over the field \( K \) of characteristic \( \neq 2 \). Then for every symmetric bilinear form \( \Phi : V \times V \to K \) on \( V \), for all \( x, y \in V \), we have: \( \Phi(x, y) = \frac{1}{2} \left( \Phi(x + y, x + y) - \Phi(x, x) - \Phi(y, y) \right) \). In particular, every symmetric bilinear form on \( V \) is determined uniquely by its values on the diagonal.

**Proof** Immediate from \( \Phi(x \pm y, x \pm y) = \Phi(x, x) \pm 2 \Phi(x, y) + \Phi(y, y) \).

2.16 **Corollary** Let \( V \) be a vector space over the field \( K \) of characteristic \( \neq 2 \). Then the symmetric bilinear form \( \Phi : V \times V \to K \) on \( V \) is the zero form if and only if \( \Phi(x, x) = 0 \) for all \( x \in V \).

**Proof** Immediate from the above the Polarisation Theorem.

2.17 **Quadratic forms** Let \( V \) be a vector space over the field \( K \). The restriction of the bilinear form \( \Phi : V \times V \to K \) to the diagonal \( \Delta_V \), defines a map \( Q_\Phi : V \to K \), \( x \mapsto \Phi(x, x) \), which is called the **quadratic form associated to \( \Phi \).**

(a) If \( \mathfrak{r} = \{ x_i \mid i \in I \} \) is a (finite) basis of \( V \) over \( K \) and if \( \mathbf{C} = \Phi(\mathfrak{r}) = (c_{ij}) \in M_I(K) \) is the Gram’s matrix of \( \Phi \) with respect to \( \mathfrak{r} \), then the quadratic form associated to \( \Phi \) is the (polynomial-) function \( V \to K \), \( \sum_{i,j \in I} a_i x_i \mapsto \sum_{i,j \in I} c_{ij} a_i a_j = {}^t \mathbf{a} \mathbf{C} \mathbf{a} \), where \( \mathbf{a} \) is the column vector \( {}^t \{ a_i \}_{i \in I} \in K^I \).

The set \( Q(V, K) \) of quadratic forms on \( V \) associated to bilinear forms on \( V \) forms a \( K \)-vector space which is isomorphic to the \( K \)-vector space \( K[X_1, \ldots, X_n]_2 \) of homogeneous quadratic polynomials in \( K[X_1, \ldots, X_n] \). This follows from the fact that two homogeneous quadratic polynomials in \( K[X_1, \ldots, X_n] \) are equal if and only if the corresponding polynomial functions are equal.

(b) One can also define (generalized) **quadratic form** on \( V \) independently from the bilinear forms as a function \( Q : V \to K \) which satisfies the following conditions:

1. \( Q(ax) = a^2 Q(x) \) for all \( a \in K \) and all \( x \in V \).
2. The map \( V \times V \to K \), \( (x, y) \mapsto Q(x+y) - Q(x) - Q(y) \), is a (symmetric) bilinear form on \( V \).

The quadratic form \( Q \) associated to a bilinear form \( \Phi : V \times V \to K \) is a generalized quadratic form in the above sense: in this case the symmetric bilinear form satisfying the above conditions (2) is nothing but the **symmetrization** \( (x, y) \mapsto \Phi(x, y) + \Phi(y, x) \) of \( \Phi \). **All generalized quadratic forms arise from bilinear forms.** (Proof).

(c) If \( 2 \in K^\times \) and \( Q : V \to K \) is a quadratic form on \( V \), then the symmetric bilinear form \( \Phi : V \times V \to K \), \( (x, y) \mapsto Q(x+y) - Q(x) - Q(y) \) satisfies \( \Phi(x, x) = 2 Q(x) \), or \( Q(x) = \frac{1}{2} \Phi(x, x) \), i.e. \( Q \) is associated to the symmetric bilinear form \( \frac{1}{2} \Phi \). Conversely, every symmetric bilinear form \( \Phi \) on \( V \) is uniquely determined by the quadratic form associated to \( \Phi \), see the Polarisation Theorem.

Therefore: **The map** \( \Sym_K(V, K) \to Q(V, K) \), \( \Phi \mapsto Q_\Phi \), **is an isomorphism of** \( K \)-**vector spaces** from the \( K \)-**vector space** \( \Sym_K(V, K) \) of symmetric bilinear forms on \( V \) onto the \( K \)-**vector space** \( Q(V, K) \) of quadratic forms on \( V \). With this isomorphism, if \( 2 \in K^\times \), then one can identify the symmetric bilinear forms with the quadratic forms. Moreover, if \( \mathfrak{r} = \{ x_1, \ldots, x_n \} \) is a \( K \)-basis of \( V \) and if the quadratic form \( Q \) on \( V \) with respect to the basis \( \mathfrak{r} \) is identified with the homogeneous polynomial \( \sum_{1 \leq i, j \leq n} b_{ij} X_i X_j \), then the corresponding symmetric bilinear form with respect to \( \mathfrak{r} \) is defined by the Gram’s matrix \( (c_{ij}) \in M_n(K) \), where \( c_{ij} = c_{ji} = \frac{1}{2} b_{ij} \) for \( 1 \leq i < j \leq n \) and \( c_{ii} = b_{ii} \) for \( 1 \leq i \leq n \).

If \( 2 \notin K^\times \), then, in general, not every quadratic form arise from a symmetric bilinear form.

2.18 **Definition** Let \( \Phi \) be a symmetric bilinear form the \( K \)-vector space \( V \).

1. A vector \( x \in V \) is called **isotropic** with respect to \( \Phi \) if \( \Phi(x, x) = 0 \), i.e. \( x \perp x \). The set of all isotropic vectors in \( V \) is called the **isotropic cone** or **lightcone** (with respect to \( \Phi \).)

If the characteristic of \( K \) is \( \neq 2 \), then it is the zero-set of the quadratic form associated to \( \Phi \) and contains the whole line \( Kx \) for every vector \( x \) in it.
(2) A subspace $W \subseteq V$ is called isotropic (resp. totally isotropic) if the restriction $\Phi|W$ of $\Phi$ on $W$ is degenerate (res. $\Phi|W = 0$ (the zero form on $W$). A vector $x \in V$, $x \neq 0$ is isotropic if and only if the line $Kx \subseteq V$ is an isotropic subspace of $V$.

(3) The form $\Phi$ is called anisotropic if there is no non-zero vector $x \in V$ which is isotropic. The form $\Phi$ is anisotropic if and only if its isotropic-cone is trivial. Further, the form $\Phi$ is anisotropic on $V$ if and only if every subspace $W \subseteq V$ is non-isotropic. For example, the standard symmetric bilinear forms on $\mathbb{R}^l$ are anisotropic. The standard complex-symmetric form on $\mathbb{C}^l$ are not anisotropic if $|l| \geq 2$.

2.19 Example (Lorentz Forms) An important example of a form with non-trivial isotropic vectors is the so-called standard Lorentz form on the $\mathbb{R}^{n+1}$. Its Gram’s matrix with respect to the standard basis is the $(n+1) \times (n+1)$-matrix

$$
\begin{pmatrix}
-c^2 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{pmatrix},
$$

where $c > 0$ denote the velocity of light which we normalise to $c = 1$. Then the light-cone is the set of points $(t,a_1, \ldots, a_n) \in \mathbb{R}^{n+1} = \mathbb{R} \times \mathbb{R}^n$ with $a_1^2 + \cdots + a_n^2 = t^2$.

In the case $n = 3$, by using a Cartesian coordinate system for the space-time $(t,x) \in \mathbb{R} \times \mathbb{R}^3$. Portions of the light cone corresponding to $t > 0$ are typically referred to as the “forward light cone”, while the $t < 0$ portion is referred to as the “backward light cone”.

In the case $n = 1$, the light-cone is the double-line $a_1^2 = t^2$. For a line $W = \mathbb{R}(t,a)$ in $\mathbb{R} \times \mathbb{R}$, $(t,a) \neq 0$, $W^\perp$ is the line $W^\perp = \mathbb{R}(a,t)$. Therefore the diagonals $\Delta$ (along the anti-diagonal $\mathbb{R}(1,-1)$) is mirrored line. The both lines of the light-cone, namely, $\Delta = \mathbb{R}(1,1)$ and $\mathbb{R}(1,-1)$, are self-perpendicular. In the case $n = 2$, the light-cone corresponding the Lorentz–form is the usual double-cone:

The finite dimensional non-isotropic subspaces $W \subseteq V$ are characterized by the Lemma 2.20 below, which is important for the proof of the Decomposition Theorem 2.24 of a symmetric bilinear form into the orthogonal direct sums, see Example 2.10:

2.20 Lemma Let $\Phi : V \times V \rightarrow K$ be a symmetric bilinear form on the $K$-vector space $V$ and let $W \subseteq V$ be a finite dimensional subspace of $V$. Then the following statements are equivalent:

(i) $W$ is not isotropic. (ii) $W \cap W^\perp = 0$. (iii) $V = W \oplus W^\perp$.

Proof We shall prove (i) $\iff$ (ii) and (i)$\Rightarrow$(iii)$\Rightarrow$(ii) (i)$\Rightarrow$(ii): Let $x \in W \cap W^\perp$, i.e. $\Phi(y,x) = 0$ for every $y \in W$. Since $\Phi|W$ is non-degenerate on $W$, it follows $x = 0$. (ii)$\Rightarrow$(i): Let $(\Phi|W)_1 : W \rightarrow W^*$ be the linear map corresponding to $\Phi|W$. Then $(\Phi|W)_1$ is injective, since its kernel is $W \cap W^\perp = 0$ by (ii) and hence $\Phi|W$ is non-degenerate on $U$ by Theorem 2.6 (iii)$\Rightarrow$(ii) is trivial. Finally, (i)$\Rightarrow$(iii): Since the implication (i)$\Rightarrow$(ii) is already proved, to prove (iii), it is enough to prove that $V = W + W^\perp$. Let $z \in V$. Now, since (already proved) $(\Phi|W)_1 : W \rightarrow W^*$ is bijective, for the linear form $W \rightarrow K$, $x \mapsto \Phi(x,z)$, there exists $y \in W$ with $\Phi(x,z) = \Phi(x,y)$ for every $x \in W$. Then $z - y \in W^\perp$ and $z = y + (z - y) \in W + W^\perp$. 

\footnote{Lorentz, Hendrik Antoon (1853-1928) was a Dutch physicist who shared the 1902 Nobel Prize in Physics with Pieter Zeeman for the discovery and theoretical explanation of the Zeeman effect. It is said that Lorentz was regarded by all theoretical physicists as the world’s leading spirit, who completed what was left unfinished by his predecessors and prepared the ground for the fruitful reception of the new ideas based on the quantum theory.

The set of all Lorentz transformations of Minkowski space-time form a group — called the Lorentz group. It is a subgroup of the Poincard group (the group of all isometries of Minkowski space-time). The Lorentz group is an isoptry subgroup at the origin of the isometry group of Minkowski space-time. The Lorentz group is also closely related to the projective special linear group $PSL(2, \mathbb{C})$ which is isomorphic to the Mbius group, the symmetry group of conformal geometry on the Riemann sphere.}
2.21 Definition Let $\Phi$ be a symmetric bilinear form on the $V$ and $W \subset V$ be a subspace of $V$. If $V = W \oplus W^\perp$, then $W^\perp$ is called the **orthogonal complement** of $W$ in $V$. The projection of $V$ onto $W$ along $W^\perp$ is called the **orthogonal projection** onto $W$ and is denoted by $p_W$.

If $W^\perp$ is an orthogonal complement of $W$ in $V$, i.e. if $V = W \oplus W^\perp$, then $W$ is non-isotropic. By Lemma 2.20 the converse holds for every finite dimensional, non-isotropic subspace $W \subset V$ with the orthogonal complement $W^\perp$. If $\Phi$ is non-degenerate on $V$ and if $V = W \oplus W^\perp$ for a subspace $W \subset V$, then $W = (W^\perp)^\perp$ is the orthogonal complement of $W^\perp$ in $V$. In this case $p_W + p_W^\perp = \text{id}_V$.

2.22 Example (Conjugate Lines) Let $\Phi$ be a non-degenerate symmetric bilinear form on the two dimensional $K$-vector space $V$ and let $W = Kw$, $w \neq 0$, be a line in $V$. Then $W^\perp$ is also a line in $V$ and is called the conjugate line corresponding to $W$. We distinguish the following two cases:

1. $w$ is not isotropic i.e. $\Phi(w, w) \neq 0$. Then $V = W \oplus W^\perp$ and every generating vector of $W^\perp$ is also not isotropic.
2. $w$ is isotropic, i.e. $\Phi(w, w) = 0$. Then $W = W^\perp$, and hence $W$ is conjugate to itself.

If the form $\Phi$ is anisotropic, then two cases do not occur. If $V$ is a real vector space and if $V$ is oriented then corresponding to every generating vector $w$ of $W$, there is a *canonical* vector $w^\perp$ which generates $W^\perp$.

In the case of the standard form $(-, -)$ on $\mathbb{R}^2$, the conjugate line corresponding to $Rw$, $w = (w_1, w_2) \neq 0$, is the perpendicular line $R(w_2, -w_1)$. More generally, if $\Phi$ is defined by the Gram’s matrix $\text{Diag}(1/a^2, 1/b^2)$, $a, b > 0$, with respect to the standard basis of $\mathbb{R}^2$. Then $R(a^2w_2, -b^2w_1)$ is the conjugate line corresponding to $Rw$. Note that the linear automorphism $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f(e_1) = ae_1, f(e_2) = be_2$ transforms the standard form to $\Phi$, i.e. for all $x, y \in \mathbb{R}^2$, we have $\Phi(f(x), f(y)) = (x, y)$. Therefore $f$ maps conjugate lines with respect to the standard form to conjugate lines with respect to $\Phi$ and hence the conjugate pair of lines $R(w_1, w_2), R(w_2, -w_1)$ with respect to the standard form correspond to the conjugate pair of lines $R(aw_1, bw_2), R(aw_2, -bw_1)$ with respect to $\Phi$.

![Diagram of conjugate lines](image)

For the formulation of the Decomposition Theorem we make the following definition:

2.23 Definition We say that a family of vectors $x_i$, $i \in I$, in a $K$-vector space $V$ with a symmetric bilinear $\Phi$, is **orthogonal** with respect to $\Phi$ if $x_i \perp x_j$ for all $i, j \in I$ with $i \neq j$. Moreover, if $\Phi(x_i, x_i) = 1$ for all $i \in I$, then the family $x_i, i \in I$ is called **orthonormal** with respect to $\Phi$.

An **orthonormal family** of vectors $x_i$, $i \in I$ in $V$, such that $\Phi(x_i, x_i) \neq 0$, is linearly independent over $K$. For, if $\sum_i a_ix_i = 0$, then $0 = \Phi(\sum_i a_ix_i, x_j) = \sum_i a_i \Phi(x_i, x_j) = a_j \Phi(x_j, x_j)$ and so $a_j = 0$ for all $j \in I$. In particular, an orthonormal family is linearly independent. Therefore an orthonormal generating system is an orthonormal basis.

If $x_i$, $i \in I$, is an orthogonal basis of $V$ with respect to $\Phi$, then $V$ is the orthogonal direct sum of the 1-dimensional subspaces $Kx_i$, $i \in I$. Moreover, if $I$ is finite, then the Gram’s matrix of $\Phi$ with respect to the basis $x_i$, $i \in I$, is the diagonal matrix $\text{Diag}(\Phi(x_i, x_i))_{i \in I}$.

---

3 Recall that an **oriented vector space** over $\mathbb{R}$ is a finite dimensional vector space over $\mathbb{R}$ with a fixed ordered $\mathbb{R}$-basis $\tau = (x_1, \ldots, x_n)$. Two bases $\tau = (x_1, \ldots, x_n)$ and $\eta = (y_1, \ldots, y_n)$ defines the same orientation on $V$ if the determinant of the transition matrix from the basis $\tau$ to the basis $\eta$ is a positive real number.
2.24 Decomposition Theorem  Let \( \Phi \) be a symmetric bilinear form on the finite dimensional vector space \( V \) over the field \( K \) of characteristic \( \neq 2 \). Then there exists an orthogonal basis for \( V \) with respect to \( \Phi \).

Proof  By induction on \( n := \dim K V \). The assertion is trivial if \( \dim K V \leq 1 \). Now, assume that \( n := \dim K V \geq 2 \) and \( \Phi \) is not the zero form. Then by Corollary 2.26 there exists a non-isotropic vector \( x_1 \in V \). Put \( W := K x_1 \). Therefore by Lemma 2.20 \( V = W \perp \) is the orthogonal direct sum of \( W \) and by induction hypothesis there exists an orthogonal basis \( x_2, \ldots, x_n \) of \( W \perp \). Then \( x_1, x_2, \ldots, x_n \) is an orthogonal basis of \( V \).

For matrices one can formulate the above Decomposition Theorem 2.24 as follows:

2.25 Corollary  Let \( K \) be a field of characteristic \( \neq 2 \) and \( \mathcal{C} \in M_I(K) \), \( I \) finite set, be a symmetric matrix. Then there exists an invertible matrix \( \mathcal{A} \in \text{GL}_I(K) \) such that \( t^T \mathcal{A} \mathcal{C} \mathcal{A} \) is a diagonal matrix.

2.26 Remark (Gram-Schmidt orthogonalisation process) In the following case the construction of an orthogonal basis is simple: Let \( x_1, \ldots, x_n \in V \) be a basis of the \( K \)-vector space \( V \) such that the subspaces \( V_m := K x_1 + \cdots + K x_m, m = 1, \ldots, n \), are not isotropic with respect to the symmetric bilinear form \( (\cdot, \cdot) : V \times V \rightarrow K \) on \( V \). We construct recursively an orthogonal basis \( x_1', \ldots, x_n' \) such that \( V_m = K x_1' + \cdots + K x_m', m = 1, \ldots, n \). For this, put \( x_1' := x_1 \) and assume that \( x_1', \ldots, x_{m-1}', \) \( m < n \), are already constructed. Then \( (x_i', x_j') = \delta_{ij} c_j \) with \( c_j \neq 0 \), \( i, j = 1, \ldots, m \) and for \( x_{m+1}' := a_1 x_1' + \cdots + a_m x_m' + x_{m+1} \), the conditions \( (x_{m+1}', x_j') = 0 \) for \( j = 1, \ldots, m \) are equivalent with the equations \( a_j (x_j, x_j') + (x_{m+1}, x_j') = 0 \) or with

\[
x_{m+1}' := x_{m+1} - \sum_{j=1}^m \frac{(x_{m+1}, x_j')}{(x_j', x_j')} x_j' .
\]

This construction is called the Gram-Schmidt orthogonalisation process.

The Decomposition Theorem 2.24 on the existence of orthogonal bases is the starting point for the classification of symmetric bilinear forms. For this we make the following definition:

2.27 Definition  Let \( \Phi \) and \( \Psi \) be two bilinear forms on the \( K \)-vector spaces \( V \), resp. \( W \). A map \( f : V \rightarrow W \) is called a homomorphism of \( (V, \Phi) \) in \( (W, \Psi) \) if it is \( K \)-linear and is compatible with the forms, i.e. \( \Phi(x, y) = \Psi(f(x), f(y)) \) for all \( x, y \in V \). A bijective homomorphism \( f : (V, \Phi) \rightarrow (W, \Psi) \) is called an isomorphism.

The homomorphisms \( (V, \Phi) \rightarrow (V, \Phi) \) are called the endomorphisms of \( (V, \Phi) \) or the endo-

morphisms of \( \Phi \). They form a monoid (with the binary operation composition). The isomorphisms \( (V, \Phi) \rightarrow (V, \Phi) \) are called the automorphisms of \( (V, \Phi) \) or \( \Phi \). They form the automorphism group of \( \Phi \), and is denoted by \( \text{Aut}_I(V, \Phi) = \text{Aut}_I \Phi \) or \( \text{GL}_K(V, \Phi) = \text{GL}_K(\Phi) \).

If there exists an isomorphism from \((V, \Phi)\) onto \((W, \Psi)\), then \((V, \Phi)\) and \((W, \Psi)\) or also the forms \( \Phi \) and \( \Psi \) are called congruent. If \( f : (V, \Phi) \rightarrow (W, \Psi) \) is an isomorphism, then the map \( \text{Aut}_V \Phi \rightarrow \text{Aut}_W \Psi \), \( g \mapsto f g f^{-1} \) is an isomorphism of groups.

We explicitly note the following:

2.28 Theorem  Let \( V \) and \( W \) be finite dimensional \( K \)-vector spaces with \( \dim_K V = \dim_K W \) and with bases \( \mathcal{r} = (x_i)_{i \in I} \), resp. \( \mathcal{r} = (y_i)_{i \in I} \). Let \( \Phi \) and \( \Psi \) be bilinear forms on \( V \) resp. \( W \) with the Gram’s matrices \( \mathcal{G}_\Phi(\mathcal{r}) \) and \( \mathcal{G}_\Psi(\mathcal{r}) \). Then the bijective \( K \)-linear map \( f : V \rightarrow W, x_j \mapsto \sum_{i \in I} a_{ij} y_i \) with the matrix \( \mathcal{A} = (a_{ij}) \in \text{GL}_I(K) \) is an isomorphism \( (V, \Phi) \rightarrow (W, \Psi) \) if and only if

\[
\mathcal{G}_\Phi(\mathcal{r}) = \mathcal{A}^T \mathcal{G}_\Psi(\mathcal{r}) \mathcal{A} \quad \text{resp.} \quad \mathcal{G}_\Phi(\mathcal{r}) = \mathcal{A}^{-1} \mathcal{G}_\Psi(\mathcal{r}) \mathcal{A}^{-1} .
\]

In particular, the bilinear forms \( \Phi \) and \( \Psi \) are congruent if and only if there exists \( \mathcal{A} \in \text{GL}_I(K) \) with \( \mathcal{G}_\Phi(\mathcal{r}) = \mathcal{A}^T \mathcal{G}_\Psi(\mathcal{r}) \mathcal{A} \).

Proof  The proof follows from the transformation formula in 2.22 and the definition of the compatibility of \( f \) with \( \Phi \) and \( \Psi \).
2.29 Classification Problem Recall that two square matrices \( \mathcal{C}, \mathcal{C}' \in M_l(K) \) are said to be congruent if there exists an invertible matrix \( \mathfrak{A} \in \text{GL}_l(K) \) with \( \mathcal{C} = \mathfrak{A} \mathcal{C}' \mathfrak{A}^{-1} \). Therefore, two \( l \times l \)-matrices describe the same bilinear form \( V \times V \rightarrow K \) if they are congruent. The classification problem for the symmetric bilinear forms on finite dimensional \( K \)-vector spaces is by Theorem 2.28 is equivalent to the problem of finding a well arranged representative system for the classes of congruent matrices over \( K \).

For example, by Decomposition Theorem 2.24 or by Corollary 2.25: If \( K \) is a field of characteristic \( \neq 2 \), then every symmetric matrix \( \mathcal{C} \in M_l(K) \) is congruent to a diagonal matrix \( \text{Diag}(c_i)_{i \in I} \). Further, the form defined by the matrix \( \mathcal{C} \) is congruent to the form \( K^l \times K^l \rightarrow K \), \( (e_i, e_j) \mapsto \delta_{ij}c_i \), on \( K^l \), where \( e_i, i \in I \), is the standard basis of \( K^l \) and hence also with the form \( \sum_{i \in I} a_i b_i c_i \). This form (and also every other form which is congruent this form) is denoted by \( [c_i]_{i \in I} \) and for \( I = \{1, \ldots, n\} \) also by \( [c_1, \ldots, c_n] \). The form \([c_i]_{i \in I}\) is the orthogonal direct sum of the forms \([c_i] : (a,b) \mapsto abc_i, \) on \( K \), therefore: \([c_i]_{i \in I} = \mathfrak{S}_{i \in I} [c_i] \).

In general, it is difficult to classify the forms \([c_i]_{i \in I}\) up to congruence. Obviously, the form \([c_i]_{i \in I}\) is congruent to the form \([a_i^2 c_i]_{i \in I}\), where \( a_i \in K^\times, i \in I \), since this is the transition of the basis \( e_i, i \in I \), to the basis \( a_i e_i, i \in I \). Therefore, one can replace the elements \( c_i \) by their images in the residue class group \( K^\times/2K^\times \) of \( K^\times \) modulo of the subgroup of quadratic-units in \( K \).

Therefore from the Decomposition Theorem 2.24 it follows that:

2.30 Theorem Let \( K \) be a field of characteristic \( \neq 2 \) with \( 2K^\times = K^\times \) (for example, if \( K \) is algebraically closed, \( K = \mathbb{C} \)). Then every symmetric bilinear form of rank \( r \) on an \( n \)-dimensional \( K \)-vector space \( V \) is congruent to the form \( [1, \ldots, 1, 0, \ldots, 0] \) where \( 1 \) occurs \( r \) times and \( 0 \) occurs \( n-r \) times. In particular, every non-degenerate symmetric bilinear form on an \( n \)-dimensional \( K \)-vector space \( V \) is congruent to the standard form \( [1, \ldots, 1] \) on \( K^n \).

Over a field \( K \) of characteristic \( \neq 2 \), in which every element is a square (for example, if \( K \) is algebraically closed, \( K = \mathbb{C} \)), all symmetric matrices in \( M_n(K) \) of equal rank are congruent: The diagonal matrices \( \text{Diag}(0,\ldots,0), \text{Diag}(1,0,\ldots,0), \ldots, \text{Diag}(1,\ldots,1) = \mathfrak{E}_n \) form a complete representative system for the congruence classes of symmetric matrices over \( K \).

From Theorem 2.28 it follows directly the following description of the automorphisms of a symmetric bilinear form on a finite dimensional \( K \)-vector space.

2.31 Theorem Let \( \Phi \) be a symmetric bilinear form on a finite dimensional \( K \)-vector space \( V \) with basis \( \mathfrak{x} = (x_i)_{i \in I} \). Then an invertible linear operator \( f : V \rightarrow V \) is an automorphism of \( (V, \Phi) \) if and only if \( \Phi_\mathfrak{f}(\mathfrak{x}) = \mathfrak{f} \Phi_\mathfrak{x} \mathfrak{f}^{-1} \mathfrak{x} \), where \( \mathfrak{f} = \text{gl}_V(f) \in \text{GL}_l(K) \) is the matrix of \( f \) with respect to the basis \( \mathfrak{x} \).

2.32 Discriminant of a symmetric bilinear form An important invariant of symmetric bilinear forms on a finite dimensional vector spaces is the discriminant. Let \( V \) be a finite dimensional \( K \)-vector space with the basis \( \mathfrak{x} = (x_i)_{i \in I} \), and \( \Phi \) a symmetric bilinear form on \( V \). The Gram’s determinant \( \Phi_\mathfrak{x}(\mathfrak{x}) \) of the basis \( \mathfrak{x} \) is called the discriminant of the basis \( \mathfrak{x} \) with respect to the basis \( \mathfrak{x} \) and denoted by \( \text{Discr}_\mathfrak{x}(\mathfrak{x}) \).

If \( \mathfrak{x}' = (x'_i)_{i \in I} \) is another basis of \( V \) with the transition matrix \( \mathfrak{A} = (a_{ij}) \in \text{GL}_l(K) \), i.e. with \( x'_j = \sum_{i \in I} a_{ij} x_i, \) \( j \in I \), then \( \Phi_\mathfrak{x}'(\mathfrak{x}') = a^2 \Phi_\mathfrak{x}(\mathfrak{x}) \), \( a := \text{Det}\mathfrak{A} = \text{Det}\mathfrak{A} \in K^\times, \) since \( \Phi_\mathfrak{x}'(\mathfrak{x}') = \mathfrak{A} \Phi_\mathfrak{x}(\mathfrak{x}') \mathfrak{A}^{-1} \). The discriminant corresponding to different bases differ only by an element of the group \( K^\times/2K^\times \). Therefore the discriminant of the bases of \( V \) defines a well-defined residue class in the multiplicative residue class monoid \( K^\times/2K^\times \). This is called the discriminant of \( \Phi_\mathfrak{x}(\mathfrak{x}) \) (or also the determinant of the form \( \Phi \) and denoted by \( \text{d} \)).

2.33 Definition Let \( V \) be a finite dimensional \( K \)-vector space. The map \( \text{End}_K V \times \text{End}_K V \rightarrow \text{tr}(fg) \) is a symmetric bilinear form on the \( K \)-vector space \( \text{End}_K V \) of \( K \)-endomorphisms of \( V \) and is called the trace form on \( \text{End}_K V \).

2.34 Proposition Let \( V \) be a finite dimensional \( K \)-vector space. Then the trace form defines a complete duality on \( \text{End}_K V \).
\section*{§ 3 Symmetric bilinear and hermitian forms over real closed fields}

In this section, we classify the symmetric and Hermitian forms on finite dimensional vector spaces over a real closed field and its algebraic closure, up to congruence, see Definition \ref{2.27}.

\subsection*{3.1 Preliminaries}
In this subsection, we recall the concepts of (formally) real fields, ordered fields, strong topology and real closed fields and basic results concerning them. For details the reader is recommended to see \cite[Ch. 11]{7}.

\textbf{(a) Ordered fields} A field $K$ is called \textit{formally real} or just \textit{real} if for all $a_1, \ldots, a_n \in K$, $a_1^2 + \cdots + a_n^2 = 0$ implies $a_1 = \cdots = a_n = 0$. In 1927 Artin-Schreier proved that: A field $K$ is real if and only if there is an order $\leq$ on $K$ such that $(K, \leq)$ is an ordered field.

Now, let $K$ be an arbitrary ordered field, i.e. $K$ is a field with a total order $\leq$, which satisfies the usual rules of monotony for addition and multiplication.

\textbf{(b) Ordered Topology} The order topology is equipped with the order topology, for which the open intervals $(a, b]$, $a, b \in K$, $a < b$, form a base for open sets in the order topology.

\textbf{(c) Product Topology} The vector spaces $K^n$, $n \in \mathbb{N}$, are endowed with the product topology (with a base given by the open cuboids $]a_1, b_1[ \times \cdots \times ]a_n, b_n[$, $a_i < b_i$, $i = 1, \ldots, n$). Addition, multiplication and inverse are continuous functions on $K \times K$ and $K^x = K \setminus \{0\}$, respectively. It follows that polynomial functions and more general rational functions $F/G$ in $n$ variables are continuous $K$-valued functions on $K^n$ outside the (closed) zero set $V_k(G) := \{a \in K : G(a) = 0\}$ of the denominator $G$.

\textbf{(d) Strong Topology} The product topology on $K^n$ transfers uniquely to every $n$-dimensional $K$-vector space by a $K$-linear isomorphism $f : V \to K^n$. Any other isomorphism $g : V \to K^n$ defines the same topology, since $gf^{-1} : K^n \to K^n$ and $(gf^{-1})^{-1} = fg^{-1} : K^n \to K^n$ are continuous (polynomial) maps. Therefore, polynomial and rational functions are also defined on any finite dimensional vector space $V$ by an isomorphism $f : V \to K^n$.

This topology on $V$ may be characterized as the smallest topology for which the $K$-linear functions $V \to K$ are continuous (with respect to the topology on $K$ from above). This topology on $V$ is called the \textit{strong topology} on $V$ (in contrast to the Zariski topology, which is weaker if $V \neq \emptyset$).

\textbf{(e) Line Connected Subsets} For two points $x, y \in V$ the (closed) \textit{line segment} $[x, y] = [y, x]$ connecting $x$ and $y$ is the set $\{(1 - t)x + ty \mid t \in K, 0 \leq t \leq 1\}$. For $x_0, \ldots, x_r \in V$, $r \geq 1$, we denote by $[x_0, \ldots, x_r] = \bigcup_{i=1}^{r+1} [x_{i-1}, x_i]$ the \textit{broken line from} $x_0$ to $x_r$. A subset $V' \subseteq V$ is called \textit{line connected} if for any two points $x, y \in V'$ there is a broken line from $x$ to $y$ which lies entirely in $V'$. Note that, if $K = \mathbb{R}$ and $U \subseteq V$ is open, then the notion “line connected” is equivalent to the topological notion of “connected” whereas the only topologically connected subspaces of $K = \mathbb{Q}$ are the singletons. If $V$ is a line, i.e. 1-dimensional, and if $x \in V$, then $V \setminus \{x\}$ is not line connected. However, if $\dim_K V \geq 2$, then $V \setminus \{x\}$ is always line connected: If $u, v \in V \setminus \{x\}$ are arbitrary points, there is always a point $v \in V \setminus \{x\}$ such that $[u, v, w] \subseteq V \setminus \{x\}$. Moreover, the following more general statement holds:

\textbf{(f) If $U_j = x_j + W_j$ with $x_j$ and $W_j \subseteq V$ subspaces of codimension $\geq 2$ for all $j = 1, \ldots, k$, of a finite-dimensional $K$-vector space $V$ and if $U \subseteq V$ is open and line connected in $V$, then $U \setminus \bigcup_{j=1}^{k} U_j$ is (open and) line connected.}

\textbf{Proof} Let $x = (x_i)_{i \in I}$, be a basis of $V$ and $e_{rs}$, $r, s \in I$, be the basis of $\text{End}_K V$ with $e_{rs}(x_j) = \delta_{sj}x_r$. For the canonical homomorphism $\rho := (\text{tr}: \text{End}_K V \to (\text{End}_K V)^*)$ corresponding to the trace form (see \ref{2.34}), we have $\rho(e_{ij})(e_{rs}) = \text{tr}(e_{ij}e_{rs}) = \text{tr}(\delta_{jr}\delta_{is}) = \delta_{js}\delta_{ir}$ and hence $\rho(e_{ij}) = e_{ij}$. Therefore $\rho$ is an isomorphism.

Obviously, the discriminant of the trace form on $\text{End}_K V$ is the signature of the permutation $(i, j) \mapsto (j, i)$ of $I \times I$, and hence is equal to $(-1)^{I(I)}$, $n := \text{card} I = \dim_K V$.

Analogous (to Proposition \ref{2.34}) assertion also hold for the matrix algebras $M_{\Gamma}(K)$.

Now, let $A$ be a finite (dimensional) $K$-algebra. Then the map $A \times A \to K$, $(x, y) \mapsto \text{tr}_K(xy) = \text{Tr}(\lambda_x \lambda_y)$, defines a symmetric bilinear form on $A$ and is called the \textit{trace form} of the $K$-algebra $A$. It reflects many important properties of the $K$-algebra $A$. If $A = \text{End}_K V$ is the algebra of the endomorphisms of a finite dimensional $K$-vector space $V$, then the trace form on the $K$-algebra $\text{End}_K V$ is different from the above introduced trace form on the $K$-vector space $\text{End}_K V$. Obviously, for every endomorphism $f \in \text{End}_K V$ : $\text{tr}^{\text{End}_K V} f = n \cdot \text{tr} f$, $n := \dim_K V$.
§ 3 Symmetric bilinear and hermitian forms over real closed fields

Proof Since $\dim_K V - \dim_K W_j = \text{Codim}_K W_j \geq 2$ for every $j = 1, \ldots, k$, it follows that each line-segment $[x, y]$ with $x, y \not\in \bigcup_{j=1}^k U_j$ intersects with each $U_j$ only in a finite set of points. With this use induction on $k$ to complete the proof.

(g) Real closed fields A field $K$ is called real closed if it is real and if it has no nontrivial real algebraic extension $L/ K$, $L \neq K$. For example, the field $\mathbb{R}$ of real numbers is real closed. The algebraic closure of $\mathbb{Q}$ in $\mathbb{R}$ is real closed. The field $\mathbb{Q}$ is real, but not real closed.

Theorem (Euler-Lagrange) Let $(K, \leq)$ be an ordered field satisfying the properties: (i) Every polynomial $f \in K[X]$ of odd degree has a zero in $K$. (ii) Every positive element in $K$ is a square in $K$. Then the field $K = K(i)$ obtained from $K$ by adjoining a square root $i$ of $-1$ is algebraically closed. In particular, $K$ itself is real-closed. For a proof see [7, Ch. 11, §11.1].

Remark: Since the field $\mathbb{R}$ of real numbers is ordered and satisfies the properties (i) and (ii), the above theorem proves the Fundamental Theorem of Algebra: The field $C = \mathbb{R}(i)$ of complex numbers is algebraically closed.

The above theorem has a remarkable complement:

Theorem (Artin-Schreier) Let $L$ be an algebraically closed field. If $K \subseteq L$ be a subfield of $L$ such that $L/K$ is finite and $K \neq L$, then $L = K(i)$ with $i^2 = 1$. If $0 = K$ is a real-closed field. For a proof see [7, Ch. 11, §11.7].

Intermediate Value Theorem for polynomial functions If $F \in K[T]$ is a polynomial with coefficients in $K$ such that $F(a)F(b) < 0$ for some $a, b \in K$, then $F$ has a zero in $[a, b]$. In other words, the values $F(t)$, $t \in [a, b]$, have the same sign if $F$ has no zero on $[a, b]$. In particular, every polynomial of odd degree has a zero in $K$. Generally, a field with this property is called a 2-field. Therefore, a real closed field is a 2-field. Furthermore, every monic polynomial $F$ over a real closed field $K$ has a positive zero in $K$ if $F(0) < 0$ (since $F(x)$ for large $x$). For the general theory of real closed fields, we refer to [7, Ch. 11].

3.2 Hermitian forms In this subsection, we consider more general situation than Section 2: Let $K$ be a field with a (given) involution $\sigma : K \rightarrow K$, i.e. $\sigma$ is an automorphism of the field $K$ with $\sigma^2 = id_K$ (often $\sigma$ is denoted by the conjugation notation $\overline{\cdot} : K \rightarrow K$, $a \mapsto \overline{a}$). In the case $\sigma = id_K$, we assume that $char K \neq 2$. The results on symmetric bilinear forms proved in Section 2 are then particular cases of the general results which are stated below. Their proofs are similar to the corresponding results proved in Section 2 and hence omit their proofs.

A function $\Phi : V \times W \rightarrow K$ is called sesquilinear if $\Phi$ is $K$-linear in the first component and $K$-semilinear in the second component, i.e. if for all $a, a' \in K$ and all $x, x', y, y' \in W$, we have: (1) $\Phi(ax + a'x', y) = a\Phi(x, y) + a'\Phi(x', y)$. (2) $\Phi(x, ay + a'y') = \overline{a}\Phi(x, y) + \overline{a'}\Phi(x, y')$. In particular, $\Phi(ax, y) = \Phi(x, ay)$.

(a) Hermitian forms and Hermitian matrices Let $\Phi : V \times V \rightarrow K$ be a sesquilinear form on the $K$-vector space $V$. Then $\Phi$ is called hermitian form if for all $x, y \in V$, we have $\Phi(x, y) = \Phi(y, x)$. If the involution $\sigma$ of the field $K$ is the identity $id_K$, then the Hermitian forms of $V$ are the symmetric bilinear forms.

4 Over the field $\mathbb{C}$ of complex numbers, besides bilinear forms so-called sesquilinear forms play an important role. Already in the case, $V = W = \mathbb{C}$, the distance of a point $z$ from the origin is not given by the bilinear form $(z, w) \mapsto z \cdot \overline{w}$, but by using the map $(z, w) \mapsto z \cdot \overline{w}$, namely, $|z| = \sqrt{z \overline{z}}$. Therefore it is beneficial to consider the concept of bilinear functions relative to given (fixed) involution $\sigma : K \rightarrow K$ (an automorphism of the field $K$ with $\sigma^2 = id_K$) of the field $K$. For example, for arbitrary field $id_K : K \rightarrow K$ is an involution. For $K = \mathbb{R}$, besides identity $id_{\mathbb{R}}$, no other involution. For $K = \mathbb{C}$, the complex conjugation $\sigma = \overline{\cdot} : \mathbb{C} \rightarrow \mathbb{C}$, $z \mapsto \overline{z}$, is an involution of $\mathbb{C}$. The complex conjugation is a special case of the conjugation of a quadratic algebra $A$ over an arbitrary field $K$: If $1, \omega \in A$ is a $K$-basis of $A$ with $\omega^2 = \alpha + \beta \omega$, $\alpha, \beta \in K$, then the conjugation $\sigma : A \rightarrow A$ of $A$ is defined by $\sigma(a + b\omega) = (a + b \overline{\omega}) - b\omega, a, b \in K$. It is easy to see that $\sigma$ is an involution of the $K$-algebra $A$ and $\sigma \neq id_A$. There are many examples of this type, for example, if $L$ is a field and $\sigma \in \text{Aut}_L$ is an involution of $L$ with $\sigma \neq id_L$, then $L$ is a quadratic algebra over the field $K := L^2 := \{a \in L | \sigma(a) = a\}$ and the given involution $\sigma$ of $L$ coincides with the conjugation of $L$ over $K$, which is the only $K$-algebra automorphism of $L$ different from $id_L$, i.e. the Galois group $\text{Gal}(L^L) = \text{Aut}_{L, \text{alg}} L = \{id_L, \sigma\}$. A typical example of this type is the algebraic closure $C_K = K[i]$, where $i^2 = -1$, of a real closed field $K$. See [3, §1].

5 Hermite, Charles (1822-1901) was a French mathematician who worked in number theory, quadratic forms, invariant theory, orthogonal polynomials, elliptic functions, and algebra. Independently from Jacobi, C. G. J. (1804-1851) Hermite arrived at the rules for separation of both real and complex zeros of algebraic equations purely algebraically and generalized Sturm’s theorem (Sturm, J. F. K. (1803-1855) for systems of equations. However, instead of the theory of usual quadratic forms he applied the theory to a more general class of forms which he introduced in science now called Hermitian forms. Hermite polynomials, Hermite interpolation, Hermite normal form, Hermitean operators are named in his honor. One of his students was Poincaré, H. J. (1854-1912). He was the first to prove that the base $e$ of natural logarithms, is a transcendental number. His methods were used later by Lindemann, C. L. F. (1852-1939) who proved that $\pi$ is transcendental.
V × V → K. In the case K = C and σ : C → C is the usual complex conjugation, then the Hermitian forms are called the complex-Hermitian forms.

If \( \mathbf{x} = (x_i)_{i \in I} \) is a basis of V, then \( \Phi \) is Hermitian if and only if for the Gram’s matrix \( \mathcal{G}_\Phi(\mathbf{x}) = (\Phi(x_i, x_j)) \in M_I(K) \) satisfies: \( \mathcal{G}_\Phi(\mathbf{x}) = \overline{\mathcal{G}_\Phi(\mathbf{x})} \). Further, if \( \mathbf{x}' = (x_i')_{i \in I} \) is another basis of V with transition matrix \( A = (a_{ij}) \in GL_I(K) \), i.e. \( x_j = \sum_{i \in I} a_{ij} x'_i \), then the Gram’s matrices \( \mathcal{G}_\Phi(\mathbf{x}) \) and \( \mathcal{G}_\Phi(\mathbf{x}') \) are related by the rule: 

\[
\mathcal{G}_\Phi(\mathbf{x}') = A \mathcal{G}_\Phi(\mathbf{x}) A^\top.
\]

For a matrix \( A = (a_{ij}) \in M_I(K) \) with \( I, J \) finite sets, put \( \overline{A} := (\overline{a_{ij}}) \in M_J(K) \). The map \( M_I(K) \to M_J(K), \quad A \mapsto \overline{A} \), is a \( \sigma \)-semilinear. Further, \( \overline{\overline{A}} = A \) and \( \overline{AB} = \overline{B} \overline{A} \) if \( A \in M_J(K) \), \( B \in M_I(K) \), \( K \) finite set. For a square matrix \( A \in M_n(K) \), \( \det \overline{A} = \overline{\det A} \) and if \( A \in GL_n(K) \), then \( \overline{A^{-1}} = (\overline{A})^{-1} \). A matrix \( A \in M_n(K) \) is called symmetric (resp. Hermitian) if \( A = \overline{A} \) (resp. \( A = \overline{A} \)). If \( A \in M_n(K) \) is Hermitian, then so are \( A, \overline{A} \) and \( A^{-1} \) if \( A \) is invertible. The symmetric matrices are the Hermitian matrices for the case \( \sigma = id_K \). In the case \( K = \mathbb{C} \) with \( \sigma = \) the usual complex-conjugation, the Hermitian matrices are called the complex-Hermitian.

Two square matrices \( A, C \in M_n(K) \) are said to be congruent if there exists an invertible matrix \( B \in GL_n(K) \) with \( C = B^\top A B \).

(b) Let \( \Phi : V \times V \to K \) be a sesquilinear form on the \( K \)-vector space V. Then \( \Phi \) is uniquely determined by its values on the diagonal \( \Delta_V := \{(x, x) \mid x \in V\} \). In particular, \( \Phi \) is Hermitian if and only if \( \Phi(x, x) = \overline{\Phi(x, x)} \) for all \( x \in V \), i.e. \( \Phi(x, y) \in K \) for all \( x, y \in V \). In particular, a complex-sesquilinear form is Hermitian if and only if the values \( \Phi(x, y), x, y \in V \), are all real.

Proof. We may assume \( \sigma \neq id_K \). Then for \( a \in K \) with \( \overline{a} \neq a \) and for all \( x, y \in V \), note that:

\[
\Phi(x, y) + \Phi(y, x) = \Phi(x + y, x + y) - \Phi(x, x) - \Phi(y, y)
\]
and

\[
a \Phi(x, y) + \overline{\sigma} \Phi(y, x) = \Phi(ax + y, ax + y) - a \overline{\sigma} \Phi(x, x) - a \Phi(y, y).
\]

is a linear system of equations in \( \Phi(x, y) \) and \( \Phi(x, y) \) with the determinant of the coefficient matrix equal to \( \overline{a} - a \neq 0 \). Now apply Cramer’s rule to get:

\[
\Phi(x, y) = \frac{1}{\overline{a} - a} \left( \Phi(ax + y, ax + y) - \overline{\sigma}(ax + y, ax + y) - \overline{a}(a - 1) \Phi(x, x) - (1 - \overline{a}) \Phi(y, y) \right)
\]
and

\[
\Phi(y, x) = \frac{1}{\overline{a} - a} \left( \Phi(ax + y, ax + y) - a \overline{\sigma}(x, x) - a(1 - \overline{a}) \Phi(y, y) - (1 - a) \Phi(x, x) \right)
\]

(c) Decomposition Theorem (see [2.24]). Let \( \Phi : V \times V \to K \) be a Hermitian form on the finite dimensional \( K \)-vector space V. Then there exists an orthogonal basis for V with respect to \( \Phi \).

(d) Notation. Let \( K \) be a real closed field. Then Aut \( K = \{id_K\} \) and the field \( C_K := K[i] \), where \( i^2 = -1 \), of complex numbers over \( K \), is the algebraic closure of \( K \) with the Galois group Gal(\( C_K/K \)) = \{id_{C_K}, \sigma = \overline{\cdot} \}, \( \overline{\cdot} : C_K \to C_K \), is the conjugation defined by \( \overline{i} = -i \), see [3.1](g) and Footnote No. 4. Further, we denote by \( K \) either the field \( K \) with the involution \( id_K \), or the field \( C_K \) with the involution \( \overline{\cdot} : C_K \to C_K \). Note that the term “Hermitian” means symmetric if either \( K = K \) or \( K = C_K \) with \( \sigma = id_{C_K} \), and \( C_K \)-Hermitian if \( K = C_K \) with \( \sigma \neq id_{C_K} \).

(e) With the notation as in (d), let \( \Phi : V \times V \to K \) be a Hermitian form on the finite dimensional \( K \)-vector space V and let \( x_1, \ldots, x_n \), \( n := \dim_K V \) be an orthogonal basis of V with respect to \( \Phi \) (which exists by (e)). If \( c_i := \Phi(x_i, x_i) \neq 0 \) for \( i \), then replacing the vector \( x_i \) by \( x'_i := x_i/\sqrt{|c_i|} \), we get \( \Phi(x'_i, x'_i) = c_i/|c_i| = \text{Sign} c_i \). Therefore, we may assume that there exists an orthogonal basis \( x_1, \ldots, x_n \) such that (renumbering the \( x_i \) if necessary) the Gram’s matrix of \( \Phi \) with respect to \( x_1, \ldots, x_n \) is a diagonal matrix \( C_n := \text{Diag}(1, 1, \ldots, 1; -1, -1, \ldots, -1; 0, 0, \ldots, 0) \).

3.3 Definition. Let \( \Phi : V \times V \to K \) be a Hermitian form on the finite dimensional \( K \)-vector space V. Then \( \Phi \) is called:

1. Positive definite if \( \Phi(x, x) > 0 \) for all \( x \in V \), \( x \neq 0 \).
2. Negative definite if \( \Phi(x, x) < 0 \) for all \( x \in V \), \( x \neq 0 \).
3. Positive semi-definite if \( \Phi(x, x) \geq 0 \) for all \( x \in V \).
4. Negative semi-definite if \( \Phi(x, x) \leq 0 \) for all \( x \in V \).
5. Indefinite if there are vectors \( x, y \in V \) with \( \Phi(x, x) > 0 \) and \( \Phi(y, y) < 0 \).
Therefore, the sign of the Gram’s determinant has an orthonormal basis with respect to such that the Gram’s matrix of \( \mathbf{V} \)

\[ \Phi \begin{pmatrix} a_{ij} \end{pmatrix} \]

is Hermitian matrix. The pair \((\Phi, K)\) is called the type of a Hermitian form \( q \) is called the Morse-index of \( \Phi \). We denote the rank, signature and type of a Hermitian form \( \Phi \) by rank \( \Phi \), sign \( \Phi \) and type \( \Phi \), respectively. The type of a Hermitian matrix \( \mathbf{C} \in M_n(K) \) is by definition the type a form with \( \mathbf{C} \) as the Gram’s matrix with respect to an (arbitrary) \( K \)-basis of \( K^n \). The matrix analog of the Sylvester’s Law of Inertia is the following:

3.4 Sylvester’s Law of Inertia\(^6\) Let \( \Phi \) be a Hermitian form on the finite dimensional \( K \)-vector space \( V \) and \( x_1, \ldots, x_n \), \( n := \dim K V \) be an (orthogonal) basis of \( V \) with respect to \( \Phi \) and the Gram’s matrix of \( \Phi \) is the diagonal matrix of the form

\[ \Phi_n^{p,q} := \text{Diag} \left( 1, \ldots, 1, -1, \ldots, -1, \underbrace{0, \ldots, 0}_{q\text{-times}}, \underbrace{n-p-q\text{-times}}_{p\text{-times}} \right) \]

Then \( p \) is the maximum of the dimensions of subspaces of \( V \) on which \( \Phi \) positive definite, and \( q \) the maximum of the dimensions of subspaces of \( V \) on which \( \Phi \) negative definite. — In particular, \( p \) and \( q \) do not depend on the special choice of the orthogonal basis \( x_1, \ldots, x_n \) of \( V \).

**Proof** Let \( p' \) is the maximum of the dimensions of subspaces of \( V \) on which \( \Phi \) positive definite. Since \( \Phi \) is positive definite on the subspace \( Kx_1 + \cdots + Kx_p \), it follows \( p \leq p' \). Let \( W \) be an arbitrary subspace on which \( \Phi \) is positive definite. Since \( \Phi \) is negative semi-definite on \( U := Kx_{p+1} + \cdots + Kx_n \), we have \( W \cap U = 0 \). Therefore, \( \dim W \leq n - \dim U = p \). Analogously one can prove the characterization of \( q \).

The pair \((p, q)\) as in the Sylvester’s Law of Inertia is called the type of the form \( \Phi \). The natural number \( p \) is called the inertia index of the form \( \Phi \), the integer \( p - q \) is called the signature and the natural number \( q \) is called the Morse-index of \( \Phi \). We denote the rank, signature and type of a Hermitian form \( \Phi \) by rank \( \Phi \), sign \( \Phi \) and type \( \Phi \), respectively.

The type of a Hermitian matrix \( \mathbf{C} \in M_n(K) \) is by definition the type a form with \( \mathbf{C} \) as the Gram’s matrix with respect to an (arbitrary) \( K \)-basis of \( K^n \).

3.5 Corollary Let \( \Phi \) be a Hermitian form on the \( n \)-dimensional \( K \)-vector space \( V \). Let \( x_1, \ldots, x_n \) be a \( K \)-basis of \( V \) and \( \Phi(x) = (\Phi(x_i, x_j)) \in M_n(K) \) be the Gram’s matrix of \( \Phi \) with respect to the basis \( x_1, \ldots, x_n \). Then \( \Phi \) is of type \((p, q)\) if and only if \( \Phi(x) \) is congruent to the matrix \( \Phi_n^{p,q} \), i.e. there exists an invertible matrix \( \mathbf{A} \in \text{GL}_n(K) \) such that \( \Phi(x) = \mathbf{A} \cdot \Phi_n^{p,q} \cdot \mathbf{A}^{-1} \). Two Hermitian matrices \( \mathbf{C}, \mathbf{C}' \in M_n(K) \) have the same type if and only if they are congruent. In particular, a Hermitian matrix \( \mathbf{C} \in M_n(K) \) have type \((p, q)\) if and only if \( \mathbf{C} \) is congruent to the matrix \( \Phi_n^{p,q} \). A symmetric matrix \( \mathbf{C} \in M_n(K) \) is of type \((p, q)\) if and only if it is congruent to the matrix \( \Phi_n^{p,q} \).

In the case \( K = \mathbb{R} \) (real closed), one can choose \( \mathbf{A} \in \text{GL}_n^+(\mathbb{R}) \). In the situation of Corollary 3.5, if \( \Phi \) is non-degenerate, i.e. if \( p + q = n \), then \( \det \Phi(x) = (-1)^q \det \mathbf{A}^2 \), i.e. sign \( (\det \Phi(x) = (-1)^q \). Therefore, the sign of the Gram’s determinant \( \det \Phi(x) \) determines the parity of \( q \). From this one can deduce the following useful criterion for the determination of the type:

\(^6\) Sylvester, James Joseph (1814-1897) was an English mathematician who made fundamental contributions to matrix theory, invariant theory, number theory, partition theory, and combinatorics. He played a leadership role in American mathematics in the later half of the 19th century as a professor at the Johns Hopkins University and in 1878 he founded the American Journal of Mathematics, the first mathematical journal in the United States.

\(^7\) Use the following observation: Let \( V \) be an oriented vector space over a real-closed field \( K \) of dimension \( n \in \mathbb{N}^* \) and \( \Phi \) be a Hermitian form of type \((p, q)\) on \( V \). Then there exists an orientation of \( V \) represented by a basis \( x_1, \ldots, x_n \) of \( V \) such that the Gram’s matrix of \( \Phi \) is equal to the matrix \( \Phi_n^{p,q} \).
3.6 Hurwitz’s Criterion[8] Let $\Phi$ be a Hermitian form on the $n$-dimensional $K$-vector space $V$. Let $x_1, \ldots, x_n$ be a basis of $V$ and $\mathbf{G}_\Phi (x) = (\Phi(x_i, x_j))_{i,j\leq n} \in M_n (K)$ be the Gram’s matrix of $\Phi$ with respect to the basis $x_1, \ldots, x_n$. Suppose that the principal minors

\[
D_i := \begin{vmatrix}
\Phi(x_1, x_1) & \cdots & \Phi(x_1, x_i) \\
\vdots & \ddots & \vdots \\
\Phi(x_i, x_1) & \cdots & \Phi(x_i, x_i)
\end{vmatrix}, \quad i = 0, \ldots, n,
\]

are all $\neq 0$. Then $\Phi$ is of type $(n - q, q)$, where $q$ is the number of sign changes in the sequence $1 = D_0, D_1, \ldots, D_n = \det \mathbf{G}_\Phi (x)$.

**Proof** Recall that we say that in a sequence $a_0, \ldots, a_n$ of non-zero real numbers, changing the $i$-th place if $0 \leq i < n$ and $a_i a_{i+1} < 0$.

Note that since $D_n \neq 0$, $\Phi$ is non-degenerate. Let $(p, q) = (n - q, q)$ be the type of $\Phi$. We prove the assertion by induction on $n$. The case $n = 0$ is trivial (and so is the case $n = 1$). For the inductive step from $n - 1$ to $n$, by induction hypothesis, the type of $\Phi$ on $V' := \mathbb{K} x_1 + \cdots + \mathbb{K} x_{n-1}$ is of type $(n - 1 - q', q')$, where $q'$ is the number of sign changes in the sequence $D_0, \ldots, D_{n-1}$. It follows $q' \leq q \leq q' + 1$, since by Sylvester’s Law of Inertia 3.4. Then $\Phi$ is of type $(n - 1 - q' \leq n - q)$ and $q' \leq q'$. Now, since $\text{Sign} D_{n-1} = (-1)^q$ and $\text{Sign} D_n = (-1)^q$, the assertion follows.

The Criterion of Hurwitz yields the following lemma:

3.7 Lemma Let $F_{ij} \in K[T]$, $1 \leq i, j \leq n$, be polynomials such that $F_{ij} = F_{ji}$. Let $s \in K$ be such that the bilinear form defined by the symmetric matrix $(F_{ij}(s))_{1 \leq i, j \leq n}$ is non-degenerate, then there exists an $\varepsilon > 0$ such that the type of the symmetric matrices $(F_{ij}(t))_{1 \leq i, j \leq n}$ is the same for all $t \in ]s - \varepsilon, s + \varepsilon[$. In particular, “being of type $(p, q)$” is an open property for non-degenerate symmetric bilinear forms over $K$.

**Proof** Note that for $t \in K$, the principal minors $D_m(t) = (F_{ij}(t))_{1 \leq i, j \leq m}$, $m = 1, \ldots, n$, of the symmetric matrix $(F_{ij}(t))_{1 \leq i, j \leq n}$, are polynomial functions in $t$ over $K$ and hence (see 3.1(c)) $D_m(t)$, $m = 1, \ldots, n$, are continuous functions in $t$. Therefore there exists $\varepsilon > 0$ such that $D_m(t)$ has the same sign as $D_m(s)$ for all $t \in ]s - \varepsilon, s + \varepsilon[$ and all $m = 0, \ldots, n$. Now, the assertion follows from the Hurwitz’s Criterion 3.6.

3.8 Remark If in Lemma 3.7 the Gramian matrix $(F_{ij}(s))$ is degenerate of type $(p, q)$, then for some $\varepsilon > 0$ the Gramian matrix $(F_{ij}(t))$ is of type $(p', q')$ with $p' \geq p$, $q' \geq q$ for all $t \in ]s - \varepsilon, s + \varepsilon[$.

3.9 Corollary Let $\Phi$ be a Hermitian form on the $n$-dimensional $K$-vector space $V$ and let $x_1, \ldots, x_n$ be an arbitrary basis of $V$. Let

\[
D_i := \begin{vmatrix}
\Phi(x_1, x_1) & \cdots & \Phi(x_1, x_i) \\
\vdots & \ddots & \vdots \\
\Phi(x_i, x_1) & \cdots & \Phi(x_i, x_i)
\end{vmatrix}, \quad i = 0, \ldots, n,
\]

be the principal minors of the Gram’s matrix $\mathbf{G}^\Phi (\Phi)$. Then:

1. $\Phi$ is positive definite if and only $D_i > 0$ for all $i = 1, \ldots, n$.
2. $\Phi$ is negative definite if and only $(-1)^n D_i > 0$ for all $i = 1, \ldots, n$, i.e. at every position in the sequence $D_0, D_1, \ldots, D_n$ there is a sign change.

**Proof** (1) If all $D_i > 0$, then $\Phi$ is positive definite by the Hurwitz’s Criterion 3.6 Conversely, if $\Phi$ is positive definite, then all $D_i > 0$, since the restriction of $\Phi$ on $V_i := \mathbb{K} v_i + \cdots + \mathbb{K} v_i$ are positive definite and hence of the type $(i, 0)$ and $\text{Sign} D_i = (-1)^0 = 1$ for all $i = 1, \ldots, n$.

---

[8] Hurwitz, Adolf (1859-1919) was a German mathematician who worked on algebra, analysis and number theory. He was a doctoral student of Felix Klein finishing dissertation in 1881 on elliptic modular functions. He used Riemann surface theory to prove many basic results on algebraic curves, for instance Hurwitz’s automorphism theorem.

[9] For an arbitrary sequence of real numbers $b_0, \ldots, b_m$ by a change of signs means a change of signs in the sequence obtained by removing the zeros from the original sequence.
(2) Replace $\Phi$ by $-\Phi$ and use (1).

### 3.10 Example

Let $v_1, v_2$ be a basis of the 2-dimensional $K$-vector space $V$. For the symmetric bilinear form $\Phi = \langle - , - \rangle$ on $V$. Let $D_1 = \langle v_1, v_1 \rangle$ and $D_2 = \text{Det} \left( \begin{pmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle \end{pmatrix} \right) = \langle v_1, v_1 \rangle \langle v_2, v_2 \rangle - |\langle v_1, v_2 \rangle|^2$. Then the following table shows the dependence of the sign $D_1$, sign $D_2$, and the type of $\Phi$:

\[
\begin{array}{c|cccccccc}
D_1 & + & + & - & - & + & 0 & 0 & 0 \\
D_2 & + & - & + & - & 0 & 0 & 0 & 0 \\
\langle x_1, x_2 \rangle & (1,0) & (1,1) & (0,2) & (1,1) & (1,0) & (0,1) & (1,1) & (0,0) \\
\text{Type} & (2,0) & (1,1) & (0,2) & (1,1) & (1,0) & (0,1) & (1,1) & (0,0) \\
\end{array}
\]

Note that the case $D_1 = 0, D_2 > 0$ is not possible.

### 3.11 Example

Let $z \in C_K \setminus K, \pi := (X-z)(X-z) \in K[X], A := K[X]/(\pi) := K[\pi]$, where $x$ is the image of $X$ modulo $\langle \pi \rangle$. Further, let $H \in K[X], H \not\subseteq \langle \pi \rangle, h = h(x) \in A$ be the image of $H$ in $A$ and let $\Phi_h : A \times A \rightarrow K$ be the symmetric bilinear form defined by $\Phi_h (f, g) = \text{tr}_K (hf g)$, $f, g \in A$. Then the Gram’s matrix

\[
\Theta_{\Phi_h} (1, x) = \left( \begin{array}{cc}
h(z) + h(\bar{z}) & h(z) \cdot \bar{z} + h(\bar{z}) \cdot z \\
h(z) \cdot z + h(\bar{z}) \cdot \bar{z} & h(z) \cdot z^2 + h(\bar{z}) \cdot \bar{z}^2
\end{array} \right) \in M_2(K)
\]

is a symmetric matrix with $D_1 = h(z) + h(\bar{z}) = 2 \Re(h(z))$ and $D_2 = \text{Det} \Theta_{\Phi_h} (1, x) = h(z) (h(\bar{z}) (z - \bar{z})^2 = -4|h(z)|^2 (\Re(\bar{z}))^2 < 0$. Therefore, by the table in Example 3.10, the type of $\Phi_h$ is $(1, 1)$.

The type of a Hermitian form on a finite dimensional vector space $V$ over $C_K$ can also be determined by using the eigenvalues of the Gram’s matrix, see Theorem 3.13 below. The standard text book proof of this fact uses the Principal Axis Theorem. We give here a direct proof using the following interesting Lemma 3.12.

In the following, as in 3.2(d), let $K$ be a real closed field, $C_K = K[i]$ with $i^2 = -1$ be the algebraic closure of $K$ and $\sigma = \iota : C_K \rightarrow C_K, i \mapsto -i$, be the conjugation of $C_K$. The $\sigma$-Hermitian standard form on $C_K^n$ is the form $\langle - , - \rangle : C_K^n \times C_K^n \rightarrow C_K, ((a_i), (b_i)) \mapsto \sum_{j=1}^n a_i \bar{b}_j$. The Gram’s matrix of the $\sigma$-Hermitian standard form with respect to the standard basis of $C_K^n$ is the unit matrix $\Theta_C$. Further, this form is positive definite. Let $x_1, \ldots, x_n$ be an orthonormal basis of $C_K^n$ and $A \in \text{GL}_n (C_K)$ be the matrix whose columns are $x_1, \ldots, x_n$. Since $\langle x_i, x_j \rangle = \iota x_i \bar{x}_j = \delta_{ij}$ for all $1 \leq i, j \leq n, A^{-1} = \overline{A}$. Such a matrix $A$ is called $\sigma$-unitary.

### 3.12 Lemma

With the notation as in 3.2(d). Let $K$ be a real closed field and $C_K$ be the algebraic closure of $K$. Further, let $f: V \rightarrow V$ be a $C_K$-linear operator on the $n$-dimensional $C_K$-vector space $V$ and $\langle - , - \rangle$ be a positive definite $C_K$-Hermitian form on $V$. Then there exists an orthonormal basis $\vec{x} = (x_1, \ldots, x_n)$ of $V$ such that the matrix $M_f^\sigma(f)$ of $f$ with respect to $\vec{x}$ is an upper triangular matrix.

**Proof**

First, we prove that there exists a basis $\eta = (y_1, \ldots, y_n)$ of $V$ such that the matrix $M_\eta^\sigma(f)$ of $f$ with respect to the basis $\eta$ is an upper triangular matrix. It is enough to prove that there exists a chain of $f$-invariant subspaces $0 = V_0 \subseteq V_2 \subseteq V_2 \subseteq \cdots \subseteq V_{n-1} \subseteq V_n = V$. We prove this by induction on $n$. For a proof, let $z = (z_1, \ldots, z_n) \in C_K^n$ be arbitrary $C_K$-basis of $V$ and $M_\eta^\sigma(f)$ be the matrix of $f$ with respect to the basis $\eta$. Note that since $C_K$ is an algebraically closed field (see FTA), it follows that the characteristic polynomial $\chi_f := \text{Det} (X \Theta_C - M_\eta^\sigma(f)) \in C_K[X]$ factors into linear polynomials in $C_K[X]$. In particular, there exists an eigenvector $y_1 \in V$, i.e. the subspace $V_1 := Ky_1$ is $f$-invariant. Now, apply induction hypothesis to the operator $\overline{f} : V/V_1 \rightarrow V/V_1$ induced by $f$ to construct the chain of $f$-invariant subspaces $0 = V_0 \subseteq V_2 \subseteq V_2 \subseteq \cdots \subseteq V_{n-1} \subseteq V_n = V$. Now, if $\eta = (y_1, \ldots, y_n)$ is a basis of $V$ with the matrix $M_\eta^\sigma(f)$ of $f$ with respect to the basis $\eta$ is an upper triangular matrix. Then the basis obtained from the basis $\eta$ by applying the Schmidt’s orthonormalization process is also an upper triangular matrix.

Now we prove the following important theorem:

---

pv2018-archiv.tex January 25, 2019 : 2:37 a.m.
3.13 Theorem With the notation as in \[\text{(d)}\]. Let \(K\) be a real closed field and \(C_K\) be the algebraic closure of \(K\). Further, let \(\mathcal{C} \in M_n(C_K)\) be a Hermitian matrix. Then all the eigenvalues of \(\mathcal{C}\) are in \(K\) and \(\mathcal{C}\) is of type \((p,q)\) if \(p\) is the number of positive eigenvalues and \(q\) is the number of negative eigenvalues of \(\mathcal{C}\) counted with their multiplicities in the characteristic polynomial of \(\mathcal{C}\).

Proof Let \(f : C^n_K \to C^n_K\) be the operator defined by the matrix \(\mathcal{C}\) (i.e. the matrix of \(f\) with respect to the standard basis of \(C^n_K\) is the matrix \(\mathcal{C}\)). Use Lemma 3.12 to choose an orthonormal basis \(\mathbf{r} = (x_1, \ldots, x_n)\) of \(C^n_K\) (with respect to the standard form on \(C^n_K\)) such that the matrix \(\mathcal{D} := \mathcal{M}_{ij}(f)\) of \(f\) is an upper triangular matrix. If \(\mathcal{A}\) is the (unitary) matrix with columns \(x_1, \ldots, x_n\), i.e. \(\mathcal{A}\) is the transition matrix from the standard basis \(\mathbf{e} = (e_1, \ldots, e_n)\) to the basis \(\mathbf{r}\), then \(\mathcal{A}^{-1} = \overline{\mathcal{A}}\) and \(\mathcal{D} = \mathcal{M}_{ij}(f) = \mathcal{M}_{ij}(\operatorname{id}_{C_K}) \mathcal{M}_{ij}(f) \mathcal{M}_{ij}(\operatorname{id}_{C_K}) = \mathcal{A}^{-1}\mathcal{C}\mathcal{A}, \) i.e. \(\mathcal{D}\) and \(\mathcal{C}\) are congruent. Therefore, since \(\mathcal{C}\) is Hermitian, \(\mathcal{D}\) is also Hermitian. Further, since \(\mathcal{D}\) is an upper triangular matrix, \(\mathcal{D}\) is a diagonal matrix (with diagonal elements in \(K\), see \[\text{(b)}\]). Now, since \(\chi_D = \chi_C\), it follows that multiplicities of eigenvalues of \(\mathcal{C}\) and \(\mathcal{D}\) are same and hence \(\mathcal{C}\) is also of type \((p,q)\).

3.14 Remark The proof of the above Theorem 3.13 shows that every Hermitian matrix in \(M_n(C_K)\) is diagonalizable (even with respect to an orthonormal basis of \(C^n_K\)).

3.15 Corollary With the notation as in \[\text{(d)}\]. Let \(K\) be a real closed field and \(C_K\) be the algebraic closure of \(K\). Further, let \(\mathcal{C} \in M_n(C_K)\) be a Hermitian matrix and \(\chi_C = c_0 + c_1X + \cdots + c_{n-1}X^{n-1} + X^n \in K[X]\) be the characteristic polynomial of \(\mathcal{C}\). Then \(\mathcal{C}\) is of type \((p,q)\) where \(p\) is the number of sign changes in the sequence \(c_0, c_1, \ldots, c_{n-1}\), \(c_n = 1\) and \(q\) is the number of sign changes in the sequence \(c_0, -c_1, \ldots, (-1)^{n-1}c_{n-1}, c_n = 1\). If \(c_0 = c_1 = \cdots = c_{r-1} = 0\) and \(c_r \neq 0\), then \(p + q = n - r\).

Proof Note that, since all the eigenvalues of \(\mathcal{C}\) are real by Theorem 3.13, indeed \(\chi_A \in K[X]\). The assertion is immediate from Theorem 3.13 and the following classical theorem of Descartes.\(\text{\cite{10}}\)

3.16 Theorem (Descartes’ Rule of Signs) With the notation as in \[\text{(d)}\]. Let \(K\) be a real closed field and \(C_K\) be the algebraic closure of \(K\). Further, let \(f = a_0 + a_1X + \cdots + a_{n-1}X^{n-1} + a_nX^n \in K[X]\), \(a_n \neq 0\) be a polynomial of degree \(n\) over a real closed field \(K\). Let \(V_+\), resp. \(V_-\) denote the number of sign changes in the sequence \(a_0, a_1, \ldots, a_{n-1}, a_n\) (resp. in the sequence \(a_0, -a_1, \ldots, (-1)^{n-1}a_{n-1}, a_n\)). Further, let \(N_+\) (resp. \(N_-\)) denote the number of positive (resp. negative) zeros of \(f\) (each zero \(f\) is counted with its multiplicity). Then there exist \(\rho_+\) and \(\rho_-\) such that \(N_+ = V_+ - 2\rho_+\) and \(N_- = V_- - 2\rho_-\). Moreover, if all zeros of \(f\) belong to \(K\), i.e. if \(f\) splits into linear factors in \(K[X]\), then \(N_+ = V_+\) and \(N_- = V_-\).

Proof See \[\text{12} \text{Theorem 57.15}\].

We now consider family of non-degenerate symmetric bilinear forms on \(n\)-dimensional vector space over a real closed field \(K\) and their Gram’s matrices \(\mathcal{H}(t) = (R_{ij}(t))_{1 \leq i, j \leq n}\), where \(R_{ij}(t) = R_{ij}(t_1, \ldots, t_n)\) are rational functions on a subset \(U \subseteq K^N\). With this we prove the following important Rigidity Theorem. More precisely:

3.17 Rigidity Theorem for Quadratic Forms Let \(K\) be a real closed field, \(R_{ij}, 1 \leq i, j \leq n\), be rational functions on a line connected subset \(U \subseteq K^N\) such that \(R_{ij} = R_{ji}\) for all \(1 \leq i, j \leq n\) and \(\mathcal{H}(t) := (R_{ij}(t))_{1 \leq i, j \leq n} \in M_n(K), t \in U\) with \(\mathcal{H}(t) \neq 0\) for all \(t \in U\). Then all the matrices \(\mathcal{H}(t) \in M_n(K), t \in U\), have the same type \((p,q)\), or equivalently, the same signature \(p - q\).

Proof By \[\text{3.1 (f)}\], it is enough to prove that the signature of \(\mathcal{H}(t)\) is constant on a line segment \(L \subseteq U\). To prove this, we parametrize the points on \(L\) by the unit interval \([0,1] \subseteq K\) and assume that \(R_{ij} = F_{ij}/G_{ij}, 1 \leq i, j \leq n\), be rational functions defined on \([0,1]\) and \(G_{ij}\) do not vanish on any point in \([0,1]\) for all \(1 \leq i, j \leq n\). Further, by Intermediate Value Theorem \[\text{3.1 (g)}\], the values of \(G := \operatorname{lcm}\{G_{ij} \mid 1 \leq i, j \leq n\}\) have the constant sign on \([0,1]\) which we may assume positive.

\(\text{\cite{10}}\) René Descartes (1596-1650) was a French philosopher, mathematician, and scientist. He is generally considered one of the most notable intellectual representatives of the Dutch Golden Age. Descartes` work, La Géométrie, is a technical work written for expert mathematicians published as an appendix to Discours de la méthode (1637). It includes application of algebra to geometry from which we now have Cartesian geometry (geometry with coordinates).
Then for every $t \in [0, 1]$, the symmetric matrices $R(t)$ and $(F_{ij})_{1 \leq i, j \leq n}$ have the same type $(p, q)$. Therefore the matrix $(F_{ij}(0))_{1 \leq i, j \leq n}$ is congruent to a diagonal matrix $\text{Diag}(a_1, \ldots, a_p, b_1, \ldots, b_q)$ such that $a_i > 0$; $i = 1, \ldots, p$, $b_j < 0$; $j = 1, \ldots, q$ and $p + q = n$. In particular, by Hurwitz’s Criterion, the principal minors $D_m(t) := \text{Det} (F_{ij}(t))_{1 \leq i, j \leq m}$ are non-zero polynomials for all $m = 0, \ldots, n$, and the sequence $D_0(0), D_1(0), \ldots, D_m(0)$ has precisely $q$ sign changes. Therefore, there are points $0 = t_0 < t_1 < \cdots < t_l = 1$ such that all $D_m(t)$ are $\neq 0$ on the intervals $[t_i, t_{i+1}]$, $i = 0, \ldots, l - 1$. On each of these intervals, by Intermediate Value Theorem, all minors $D_m(t)$ have the same sign and by Hurwitz’s Criterion, the matrices $(F_{ij}(t))_{1 \leq i, j \leq n}, t \in [0, 1]$, have the same type. Now, it follows from [3.1(f)] that the type is constant even on $[0, 1]$.

### §4 The Trace Form and its Generalizations

#### 4.1 Preliminaries
In this subsection, we recall the concepts from elementary commutative algebra for the sake of completeness. For details the reader is recommended to see [1], [9] and [10].

(a) **Prime spectrum, Maximal spectrum and Radicals** Let $A$ be an arbitrary commutative ring (with unity). The set Spec($A$) (resp. Spm $A$) of prime (resp. maximal) ideals in $A$ is called the prime (resp. maximal) **spectrum** of $A$. Then Spm $A \subseteq$ Spec $A$ and a well-known theorem of Krull asserts that if $A$ $\neq 0$ then Spm $A \neq 0$. For example, Spm $\mathbb{Z} = \{ (p) | p \in \mathbb{P} \}$, where $\mathbb{P}$ is the set of prime numbers and Spec $\mathbb{Z} = \{ 0 \} \cup$ Spm $\mathbb{Z}$. The ring $R$ is a field if and only if only $\{ 0 \} \in$ Spec $A$. For an ideal $a$ in $R$, the ideal $\sqrt{a} := \{ fR | f' \in a \text{ for some integer } f \geq 1 \}$ is called the **radical** of $a$. Clearly $a \subseteq \sqrt{a}$. If $\sqrt{a} = a$, then $a$ is called a radical **ideal**. Obviously, $\sqrt{\sqrt{a}} = \sqrt{a}$. Therefore the radical of an ideal is a radical ideal. Prime ideals are radical ideals. An ideal $a$ in $\mathbb{Z}$ is a radical ideal if and only if if $a = 0$ or $a$ is generated by a square-free integer.

The radical $n_{A} := \sqrt{0}$ of the zero ideal is the ideal of nilpotent elements and is called the **nilradical** of $A$. The **nilradical** $n_{A} = \cap \in \text{Spec} A \ p$ is the intersection of all prime ideals in $A$. More generally, for a ring $A$, the intersection $n_{A} = \cap \in \text{Spec} A \ p$ for every ideal $a$ in $A$, see [1], [9].

The intersection $m_{A} := \cap \in \text{Spec} A \ m$ of maximal ideals in $A$ is called the **Jacobson radical** of $A$. Clearly, $m_{A} \subseteq m_{A}$. The Jacobson radical of $\mathbb{Z}$ (resp. the polynomial algebra $K[X_{1}, \ldots, X_{n}]$ over a field $K$) is 0.

(b) **The K-Spectrum and the K-rational of a K-algebra** (see [10]) Let $K$ be a field. Then using the universal property of the polynomial algebra $K[X_{1}, \ldots, X_{n}]$, the affine space $K^n$ can be identified with the set of $K$-algebra homomorphisms $\text{Hom}_{K-alg}(K[X_{1}, \ldots, X_{n}], K)$ by identifying $a = (a_{1}, \ldots, a_{n}) \in K^{n}$ with the substitution homomorphism $\xi_{a} : K[X_{1}, \ldots, X_{n}] \rightarrow K, X_{i} \mapsto a_{i}$. The kernel of $\xi_{a}$ is the maximal ideal $m_{a} = (X_{i} - a_{i}, \ldots, X_{n} - a_{n})$ in $K[X_{1}, \ldots, X_{n}]$. Moreover, every maximal ideal $m$ in $K[X_{1}, \ldots, X_{n}]$ with $K[X_{1}, \ldots, X_{n}]/m = K$ is of the type $m_{a}$ for a unique $a = (a_{1}, \ldots, a_{n}) \in K^{n}$; the component $a_{i}$ is determined by the congruence $X_{i} \equiv a_{i} \text{ mod } m$.

The subset $K$-$\text{Spec} K[X_{1}, \ldots, X_{n}] := \{ m_{a} | a \in K^{n} \}$ of Spm $K[X_{1}, \ldots, X_{n}]$ is called the **$K$-spectrum of $K[X_{1}, \ldots, X_{n}]$**. We have the identifications:

\[
\begin{align*}
K^n & \xleftarrow{\text{Hom}_{K-alg}(K[X_{1}, \ldots, X_{n}], K)} \xrightarrow{\xi_{a}} \text{Spec} K[X_{1}, \ldots, X_{n}], \\
\text{a} & \xleftarrow{} \text{Hom}_{K-alg}(A, K) = \{ m \in \text{Spm} A | A/m = K \}, \xi_{a} \xrightarrow{} \text{Ker} \xi_{a}, \text{is bijective.}
\end{align*}
\]

More generally, for any $K$-algebra $A$, the map $\text{Hom}_{K-alg}(A, K) \rightarrow \{ m \in \text{Spm} A | A/m = K \}$, $\xi_{a} \rightarrow \text{Ker} \xi_{a}$, is bijective. Therefore we make the following definition:

For any $K$-algebra $A$ of finite type, the subset $K$-$\text{Spec} A := \{ m \in \text{Spm} A | A/m = K \}$ is called the $K$-$\text{spec}$ of $A$ and is denoted by Spec $A$. Further, if $A \rightarrow K[X_{1}, \ldots, X_{n}]/\mathfrak{a}$ is a representation of the finite type $K$-algebra $A$, then the $K$-algebraic set $V_{K}(\mathfrak{a}) := \{ a \in K^{n} | F(a) = 0 \text{ for all } F \in \mathfrak{a} \}$ is called the set of $K$-$\text{rational points of } A$. Under the above bijective maps, we have the identification $V_{K}(\mathfrak{a}) = \text{Hom}_{K-alg}(A, K) = \text{Spec} A$. For example, since $C$ is an algebraically closed field, Spm $C[X] = C$-$\text{Spec} C[X]$, but $\text{Spec} K[X] \subseteq \text{Spm} K[X]$. In fact, the maximal ideal $m := \langle x^{2} + 1 \rangle$ does not belong to $\text{Spec} K[X]$. More generally, a field $K$ is algebraically closed if and only if Spm $K[X] = \text{Spec} K[X]$, see [1], [9] or [4, Theorem 2.10, HNS 3].

(c) **Local components of a finite algebra** Let $A$ be a finite algebra over a field $K$, i.e. A finite dimensional as a $K$-vector space of dimension $\dim_{K} A$. Then Spm $A = \text{Spec} A$ (since any finite $K$-algebra which is an integral domain is already a field). Moreover, Spm $A$ is a finite set. This follows immediately from the Chinese Remainder Theorem: If $m_{i}, i \in I$, is a finite family of pairwise distinct maximal ideals of $A$, then the
Let $A$ be a finite $K$-algebra over an arbitrary field. Suppose that $A$ has a primitive element $x \in A$, i.e., a generator of $A$ as a $K$-algebra. Then $A = K[x] \cong K[X]/(\mu_i)$, where $\mu_i$ is the minimal polynomial $x \in K[X]$ with pairwise distinct prime polynomials $\pi_1, \ldots, \pi_n$ and (positive) exponents $\alpha_1, \ldots, \alpha_n \in \mathbb{N}^*$, then $A_1 := K[X]/(\pi_i^\alpha_i)$ are the local components of $A$ and $K_i := K[X]/(\pi_i)$ its residue class fields, $i = 1, \ldots, n$. The $K$-rational points of $A$ correspond to the local prime factors of $\mu_i$, i.e., to the zeros of $\mu_i$ in $K$. The radical of $A$ is generated by $\pi_1(x) \cdots \pi_n(x) \in A$ and is reduced if and only if $\alpha_1 = \cdots = \alpha_n = 1$. Moreover, if all prime factors $\pi_1, \ldots, \pi_n$ of $\mu_i$ are separable prime polynomials (equivalently, $\gcd(\mu_i, \mu_i') = 1$, where $\mu_i'$ is the derivative of $\mu_i$), then $A$ is a separable $K$-algebra.

A primitive element of a finite $K$-algebra $A$ is often called a resolvent of $A$ and its minimal polynomial a resolvent polynomial (or a resolvent equation) for $A$.

**Primitive Element Theorem:** A finite separable algebra over an infinite field $K$ has a resolvent.

**Sketch of a proof:** There exists a finite field extension $L/K$ such that all residue fields of the $L$-algebra $A(L) = L \otimes_K A$ coincide with $L$ (such a field $L$ is called a splitting field of $A$ over $K$). Then, since $A(L)$ is separable over $L$ (by Lemma 4.2 below), it is isomorphic to the product algebra $L^n$, $n := \dim_L A(L) = \dim_K A$, and therefore has a primitive element. (Any $n$-tuple in $L^n$ with $n$ pairwise distinct components is a primitive element!) Then the $K$-algebra $A$ also has a primitive element, since $K$ is infinite.

A resolvent of a finite Galois field extension $L/K$ (which exists by the Primitive Element Theorem) is called a Galois resolvent of $L/K$ and its minimal polynomial is called a Galois resolvent polynomial of $L/K$.

**4.3 The trace form** Let $A$ be a finite algebra over the field $K$. A classical tool for studying $A$ is the trace form, which is the symmetric $K$-bilinear form $tr : (f, g) \mapsto tr_K^A(fg)$ on $A$. Usually, we denote it by $tr = tr_K$. The decomposition of $A = A_1 \times \cdots \times A_r$ into its local components (cf. [4.1]) yields the orthogonal decomposition (see Decomposition Theorem [2.24])

$$tr_K^A = tr_K^{A_1} \oplus \cdots \oplus tr_K^{A_r}$$

of the trace form. The degeneration space $A^\perp = A^\perp_0 = \{ f \in A \mid tr(Af) = 0 \}$ is an ideal in $A$.

**Lemma** Let $A$ be a finite algebra over an arbitrary field $K$ and let $A^\perp$ be the degeneration space of the trace form $tr_K^A$. Then radical $m_A := n_A \subseteq A^\perp$. Moreover, equality holds if and only if all the residue class fields of $A$ are separable over $K$, i.e., if and only if the reduction $A_{red} = A/n_A$ is a separable $K$-algebra. — In particular, the trace form is non-degenerate if and only if $A$ is a separable $K$-algebra.

**Proof** If $f \in n_A$, then $gf \in n_A$ for every $g \in A$. Therefore, multiplication by $gf$ is a nilpotent operator on $A$ and $tr(gf) = 0$ and hence $n_A \subseteq A^\perp$. For the remaining assertions, it is enough to prove
that the trace form of a reduced finite $K$-algebra $A = K_1 \times \cdots \times K_r$ is non-degenerate if and only if $A$ is a separable $K$-algebra. Since $\text{tr}_A^K = \text{tr}_K^{K_1} \oplus \cdots \oplus \text{tr}_K^{K_r}$, we have to show that for a finite field extension $K \subseteq L$, the trace form $\text{tr}_K^L$ is non-degenerate if and only if $L$ is separable over $K$. This follows from the well-known fact that the trace form $\text{tr}_K^L : L \to K$ is non-zero if and only if $L$ is separable over $K$, see [12, Ch. XI, §94, Corollary 94.3].

Now, if $K$ is a real closed field, then the characteristic of $K$ is zero, and hence $A_{\text{red}}$ is always separable over $K$ and $A^\perp = n_A$. The inclusion $A^\perp \subseteq n_A$ may be deduced in this case in the following simple way: If $f \in A^\perp$, then, in particular, $\text{tr}_K^A(f^n) = 0$ for all $n \in \mathbb{N}^\ast$. Now, a well-known result in linear algebra states that in characteristic zero a linear operator $f$ on a finite dimensional vector space $V$ with $tr(f^n) = 0$ for all $n \in \mathbb{N}^\ast$ (or at least for all $n = 1, \ldots, \dim_K V$) is nilpotent.

From the Lemma 4.4 we get:

4.5 Corollary Let $A$ be a finite separable algebra over an arbitrary field $K$. Then

$$\text{rank} \text{tr}_A^K = \dim_K (A/m_A) = \sum_{i=1}^r [K_i : K].$$

Moreover, if $K$ is an ordered field, then $\text{type} \text{tr}_A^K = \sum_{i=1}^r \text{type} \text{tr}_K^{K_i}$ and $\text{sign} \text{tr}_A^K = \sum_{i=1}^r \text{sign} \text{tr}_K^{K_i}$.

Now, we prove the following important and classical criterion for the existence of $K$-rational points for real closed fields.

4.6 Theorem Let $A$ be a finite commutative algebra over a real closed field $K$. Then

$$\text{sign} \text{tr}_A^K = \# K - \text{Spec} A.$$

In particular, $K$ is a residue class field of $A$ if and only if $\text{sign} \text{tr}_A^K \neq 0$.

Proof Since $\text{sign} \text{tr}_K^K = 1$ and by the formula above, it suffices to show that $\text{sign} \text{tr}_K^K = 0$ for every finite field extension $L \neq K$ of $K$. To prove this, we consider for every $x \in L, x \neq 0$, the symmetric bilinear forms $\Phi_x : (f, g) \mapsto \text{tr}_K^K(xfg)$ on $L$ which are non-degenerate. We have $\Phi_1 = \text{tr}_K^K$ and $\Phi_{-1} = - \text{tr}_K^K$. Since $\dim_K L \geq 2$, the punctured space $L^\times = L \setminus \{0\}$ is line connected and, by the Rigidity Theorem for Quadratic Forms [3, 11], the signature $\text{sign} \Phi_x$ is constant on $L^\times$. In particular, $\text{sign} \text{tr}_K^K = \text{sign} (- \text{tr}_K^K) = - \text{sign} \text{tr}_K^K$, and therefore $\text{sign} \text{tr}_K^K = 0$.

4.7 Example Note that the only non-trivial field extension $L$ of a real closed field $K$ is, up to isomorphism, given by the quadratic field $L = \mathbb{C}_K = K[i]$ with $i^2 = -1$, of complex numbers over $K$. The Gram’s matrix of $\text{tr}_K^\mathbb{C}$ with respect to the basis $1, i$ is given by

$$\begin{pmatrix}
\text{tr}(1) & \text{tr}(i) \\
\text{tr}(i) & \text{tr}(-1)
\end{pmatrix} = \begin{pmatrix} 2 & 0 \\
0 & -2
\end{pmatrix}.$$ 

Thus, $\text{type} \text{tr}_K^\mathbb{C} = (1, 1)$ and $\text{sign} \text{tr}_K^\mathbb{C} = 0$.

4.8 Remark For an ordered field $K$, there exists a real closed field $\mathbb{K}$ and an order preserving embedding $K \hookrightarrow \mathbb{K}$ (see [17, Ch. 11, Theorem 11.4]). Further, if $A$ is a finite $K$-algebra and if $\hat{A} := \mathbb{K} \otimes_K A$, then $\text{sign} \text{tr}_K^A = \text{sign} \text{tr}_K^\hat{A}$ and so by Theorem 4.6, $\text{sign} \text{tr}_K^A = \# \text{Hom}_{\hat{A}}(\hat{A}, \mathbb{K}) = \# \text{Hom}_{\hat{A}}(A, \mathbb{K})$ is the number of $\hat{A}$-rational points of $A$.

4.9 Corollary Let $A$ be a finite algebra over a real closed field $K$. Then the trace form $\text{tr}_K^A$ is positive definite if and only if $A$ is separable over $K$ and $A$ splits over $K$, i.e. $A \cong K^\dim_K A$ (as $K$-algebras).

4.10 Corollary Let $K$ be a real closed field and $f \in K[X]$ be a monic polynomial. Then all zeroes (in $K$) belong to $K$ and are simple if and only if the trace form $\text{tr}_K^A$ of the $K$-algebra $A = K[X]/(f)$ is positive definite.
4.11 Generalized trace forms The statement of Theorem 4.6 may be generalized to some extent. For this, let $\text{Sym}_K(V,K) := \{ \Phi \in \text{Mult}_K(V,K) \mid \Phi \text{ is symmetric} \}$ denote the $K$-vector space of all symmetric bilinear forms on $V$ and consider a $K$-linear embedding

$$E_{A/K} := \text{Hom}_K(A,K) \longrightarrow \text{Sym}_K(V,K), \quad \alpha \longmapsto \Phi_\alpha : A \times A \rightarrow K, \quad (f,g) \longmapsto \alpha(fg).$$

The elements of the image of this map are called generalized trace forms on $A$. Note that, $E_{A/K}$ is also an $A$-module, called the d u a l i z i n g m o d u l e of $A$, with the scalar multiplication defined by $(g\alpha)(f) = \alpha(fg)$ for $\alpha \in E$, $g \in A$. Therefore, $\Phi_\alpha(f,g) = (g\alpha)(f) = (f\alpha)(g)$. The degeneration space $A^{1\alpha}$ of $\Phi_\alpha$ is the largest ideal of $A$ contained in $\text{Ker} \alpha$. If $\overline{\alpha} : A/A^{1\alpha} \rightarrow K$ denotes the linear form on $\overline{A} := A/A^{1\alpha}$ induced by $\alpha$, then $\text{rank} \Phi_\alpha = \text{rank} \overline{\alpha}$.

In general, $\Phi_\alpha$ is non-degenerate if and only if $f\alpha \neq 0$ for all $f \in A \setminus \{0\}$, that is, if $A\alpha \subseteq E$ is a free $A$-submodule of rank one. Since $\text{Dim}_K E = \text{Dim}_K A$, even the equality $A\alpha = E$ holds, i.e. $\alpha$ is an $A$-basis of $E$. A finite $K$-algebra is called a Frobenius algebra if it possesses such a $K$-linear form $\alpha$ for which $\Phi_\alpha$ is non-degenerate or, equivalently, if $E \cong A$ as $A$-modules. For instance, $A/A^{1\alpha}$ is a Frobenius algebra for every $\alpha \in E$.

With these notions, we can formulate a partial generalization of Theorem 4.6:

4.12 Theorem Let $\alpha$ be a $K$-linear form on a finite commutative algebra $A$ over a real closed field $K$. If $\text{sign} \Phi_\alpha \neq 0$, then $A$ has a $K$-rational point, i.e. $K\text{-Spec} \ A \neq \emptyset$.

Proof Since $\text{sign} \Phi_\alpha = \text{sign} \overline{\alpha}$ for the induced $\overline{\alpha}$ on $\overline{A} := A/A^{1\alpha}$, we may assume that $\Phi_\overline{\alpha}$ is non-degenerate. We consider the family of non-degenerate symmetric bilinear forms $\Phi_f \alpha, f \in A^\times = A \setminus (m_1 \cup \cdots \cup m_r)$, where $m_1, \ldots, m_r$ denote the maximal ideals of $A$. If $A/m_i \neq K$ for every $i = 1, \ldots, r$, then all $m_i$ have at least codimension 2 in $A$, and therefore $A^\times$ is line connected by 3.1(f). By the Rigidity Theorem for Quadratic Forms 3.16 then all $\Phi_f \alpha, f \in A^\times$, have the same signature. In particular, $\text{sign} \Phi_\alpha = \text{sign} \Phi_{-\alpha} = - \text{sign} \Phi_\alpha$. Therefore $\text{sign} \Phi_\alpha = 0$. Contradiction!  

4.13 Corollary If $L$ is a non-trivial finite field extension of a real closed field $K$ and $\alpha : L \rightarrow K$ is a $K$-linear form on $L$, then $\text{sign} \Phi_\alpha = 0$.

Proof Note that, since $K$ is a real closed field, the only possibility is $L = C_K = K[i]$ with $i^2 = -1$ and hence the assertion follows from Example 4.7 and Theorem 4.12.

§ 5 Counting rational points of 0-dimensional affine varieties

In this section we will apply generalized trace forms to count the rational points of 0-(Krull) dimensional affine varieties over real closed fields. Our method is a modern version of old results of Hermite and Sylvester who had used signatures of quadratic forms to count real zeros of polynomials in one variable, see [5], [6] and [14]. We use elementary commutative algebra to treat multivariate versions of these problems.

5.1 Notation, Assumptions and Consequences Throughout this section, we use the following notation and assumptions and their consequences: Let $K$ be a real closed field and $\mathbb{K} := C_K = K[i]$, $i^2 = -1$, be the algebraic closure of $K$ (see 3.1(g)), $\mathfrak{A} \subseteq K[X_1, \ldots, X_n]$ an ideal in the polynomial ring $K[X_1, \ldots, X_n]$ over $K$ and let $V_K(\mathfrak{A}) := \{ a \in K^n \mid F(a) = 0 \text{ for all } F \in \mathfrak{A} \}$, (resp. $V_K(\mathfrak{A}) = \{ a \in K^n \mid F(a) = 0 \text{ for all } F \in \mathfrak{A} \}$) be the affine algebraic set in $K^n$ (resp. in $K^n$) defined by $\mathfrak{A}$. Polynomials in $K[X_1, \ldots, X_n]$ are denoted by capital letters $F, G, H, \ldots$ and their images in the residue class algebra $A := K[X_1, \ldots, X_n]/\mathfrak{A}$ are denoted by small letters $f, g, h, \ldots$. An element $f \in A$ defines a (regular or polynomial) function on $V_K(\mathfrak{A})$, namely $f : V_K(\mathfrak{A}) \rightarrow K, a \mapsto f(a)$. For $f, g \in A$, clearly, $f = g$ on $V_K(\mathfrak{A}) \iff f = g$ in $A \iff F \equiv G \pmod{\mathfrak{A}}$, i.e. $F - G \in \mathfrak{A}$.
We assume that the residue class algebra $A := K[X_1, \ldots, X_n]/\mathfrak{a}$ is finite dimensional $K$-vector space, or equivalently, the $K$-algebra $K \otimes_K A = A_K = K[X_1, \ldots, X_n]/(\mathfrak{a})$ is finite dimensional over $K$. Further, equivalently, $V_K(\mathfrak{a})$ (resp. $V_K(\mathfrak{a}) = \{a_1, \ldots, a_n\}$) is a finite set. Since $\mathfrak{a} \subseteq K[X_1, \ldots, X_n]$, it follows that if $a \in V_K(\mathfrak{a})$, then its conjugate $\bar{a} \in V_K(\mathfrak{a})$. Renumbering we may assume that $V_K(\mathfrak{a}) = \{a_1, \ldots, a_r\} \subseteq V_K(\mathfrak{a}) = \{a_1, \ldots, a_r, a_{r+1}, \bar{a}_{r+1}, \ldots, a_{r+s}, \bar{a}_{r+s}\}$, where $r := \# V_K(\mathfrak{a})$, $r+s = \# \text{Spm} A$ and $r+2s = m = \dim_k A = \dim_k A_K = \# V_K(A)$.

Since $K$ is a real closed field, Char $K = 0$, in particular, $K$ is infinite and hence by a linear change of coordinates (over $K$) (for instance, $Y_i = X_i$ for all $i = 1, \ldots, n-1$ and $Y_n = X_n + \sum_{i=1}^{n-1} X_i t^i$ for suitable $t \in K$ avoiding finitely many), we may assume that $V_K(\mathfrak{a})$ is in general $X_n$-position, or the ideal $\mathfrak{a}$ in general $X_n$-position (The intention is to separate all zeros in an algebraic closure of $K$ by their last coordinate), i.e.:

5.1.a The $n$-th coordinates $a_{in}$ of the points $a_i = (a_{i1}, \ldots, a_{in}) \in K^n$, $i = 1, \ldots, m$ are all distinct.

Note that $V_K(\mathfrak{a}) = V_K(\mathfrak{a}) \cap K^n$ is the set of $K$-rational points of $V_K(\mathfrak{a}) \cong K$-Spec $A_K = \text{Spec} A_K$ (the first equality follows from Hilbert’s Nullstellensatz, see [9] or [4] Theorem 2.10, HNS 3]) and $V_K(\mathfrak{a}) \cong K$-Spec $A \subseteq \text{Spm} A = \text{Spec} A$, see [4,1(b)]. Further, since $A$ and $A_K$ are reduced, the local components (see [4,1(c)]) of $A$ corresponding to the $K$-rational points $a_i \in V_K(\mathfrak{a})$, $i = 1, \ldots, m$, are isomorphic to $K$ and corresponding to $m \in \text{Spm} A \setminus \text{Spec} A$ are isomorphic to $K$, but local components of $A_K$ corresponding to all the points $a \in V_K(\mathfrak{a})$ are all isomorphic to $K$. Therefore the explicit structures of the $K$-algebra $A$ and the $K$-algebra $A_K$ are determined by the algebra isomorphisms which are defined by the substitutions:

5.1.b $A \cong K^r \times K^s, h \mapsto (h(\text{mod } m))_{m \in \text{Spm} A}$ and $A_K \cong K^m, f \mapsto (f(\bar{a}))_{a \in V_K(\mathfrak{a})}$, where $r := \# V_K(\mathfrak{a})$, $r+s = \# \text{Spm} A$ and $r+2s = m = \dim_k A = \dim_k A_K = \# V_K(A)$.

Furthermore we note the following eigenvector theorem (see [2] Ch. 2, §4, Theorem 4.5) which follows directly from 5.1.b:

5.1.c For every $h \in A$, the eigenvalues of the $K$-linear map $\lambda_h : A \rightarrow A, f \mapsto hf$, are the values $h(a_1), h(a_r), h(a_{r+1}), h(a_{r+s})$ of the function $h : V_K(\mathfrak{a}) \rightarrow K$.

For more accessible determination of the signature of the trace form $\text{tr}_K^A$, we need a nice basis of $A$ over $K$. The following crucial key observation so-called Shape Lemma guarantees a distinguished generating set for a radical ideal $\mathfrak{a}$ in $K[X_1, \ldots, X_n]$.

5.2 Lemma (Shape Lemma) Let $K$ be an arbitrary infinite perfect field and let $\mathfrak{a} \subseteq K[X_1, \ldots, X_n]$ be a 0-(Krull) dimensional radical ideal, i.e. the residue class $K$-algebra $A = K[X_1, \ldots, X_n]/\mathfrak{a}$ is a finite dimensional $K$-vector space. With further notation and assumptions as in 5.1. There exist polynomials $g_1, \ldots, g_{n-1}, g_n \in K[T]$ with $g_n \neq 0$ square free of degree $m$, such that $\mathfrak{a}$ is generated by $X_1 - g_1(X_n), \ldots, X_{n-1} - g_{n-1}(X_n), g_n(X_n)$. In particular, $\mathfrak{e} = \{1, x_1, \ldots, x_n^{m-1}\}$ is a $K$-basis of $A$, where $x_n$ is the image of $X_n$ in $A$.

Proof Let $\overline{K}$ be an algebraic closure of $K$. Since $K$ is perfect $\overline{K}|K$ is a Galois extension with Galois group $\text{Gal}(\overline{K}|K)$. Let $V_{\overline{K}}(\mathfrak{a}) := \{a \in \overline{K}^n \mid f(a) = 0 \text{ for all } F \in \mathfrak{a}\}$. Then $V_{\overline{K}}(\mathfrak{a})$ is finite by assumption on $\mathfrak{a}$ and the projection map $q : V_{\overline{K}}(\mathfrak{a}) \rightarrow \overline{K}, (a_1, \ldots, a_n) \mapsto a_n$, which is injective, i.e. $q$ separates points in $V_{\overline{K}}(\mathfrak{a})$ by assumption (see 5.1.a). The Galois group $\text{Gal}(\overline{K}|K)$ operates naturally on $V_{\overline{K}}(\mathfrak{a})$: $\text{Gal}(\overline{K}|K) \times V_{\overline{K}}(\mathfrak{a}) \rightarrow V_{\overline{K}}(\mathfrak{a}), (\sigma, (a_1, \ldots, a_n)) \mapsto (\sigma(a_1), \ldots, \sigma(a_n))$. Obviously, the image $q(V_{\overline{K}}(\mathfrak{a})) = W_1 \cup W_2 \cup \cdots \cup W_r$ is the union of orbits of this operation and each orbit $W_\rho = V_{\overline{K}}(\pi_\rho)$ is the zero set of the irreducible polynomial $\pi_\rho \in K[T], \rho = 1, \ldots, \ell$, see [7] or [12] Ch. XI, §93, 93.2. Therefore, since $K$ is perfect, the polynomial $g_n := \pi_1 \cdots \pi_\ell \in K[T]$ is square free and $q(V_{\overline{K}}(\mathfrak{a})) = V_{\overline{K}}(g_n), \deg g_n = \# V_{\overline{K}}(m) = m$.

5.2.a For all $a_n \in q(V_{\overline{K}}(\mathfrak{a}))$, there exist polynomials $g_1, \ldots, g_{n-1} \in K[T]$ with $\deg g_i < \deg g = m$ such that $(g_1(a_n), \ldots, g_{n-1}(a_n), a_n)$ is the unique point lying over $a_n$.

To prove 5.2.a, let $a_n \in q(V_{\overline{K}}(\mathfrak{a}))$ and $(a_1, \ldots, a_{n-1}, a_n)$ be the unique point lying over $a_n$. We
may assume that \( a_n \in W_1 = \{ \sigma_j(a_n) \mid j = 1, \ldots, d_1, \sigma_1 = \text{id}_K \} \) with \( d_1 = \#W_1 \). Let \( W'_i \) denote the orbit of \( a_i \). Then, since \( q \) is injective, \( \#W'_i \leq \#W_1 = d_1 \). Moreover, for all \( i = 1, \ldots, n - 1 \), \( W'_i = \{ \sigma_j(a_i) \mid j = 1, \ldots, d_1, \sigma_1 = \text{id}_K \} \), but all \( \sigma_j(a_i), j = 1, \ldots, d_1 \), may not be distinct.

Now, since \( \sigma_j(a_n), j = 1, \ldots, d_1 \), are distinct elements in \( K \), by Lagrange’s Interpolation Formula\(^{[1]}\) for each \( i = 1, \ldots, n - 1 \), there exists a polynomial \( g_i \in K[X], \deg g_i < d_1 < \deg g_n \), such that \( g_i(\sigma_j(a_n)) = \sigma_j(a_i) \) for all \( j = 1, \ldots, d_1 \). Moreover, \( g_1, \ldots, g_{n-1} \in K[X] \).

Finally we claim the equality \( \mathcal{A}' := \langle X_1 - g_1(X_n), \ldots, X_{n-1} - g_{n-1}(X_n), g_n(X_n) \rangle = \mathcal{A} \). For this first note that \( K[X_1, \ldots, X_n]/\mathcal{A}' \sim K[X_n]/(g) \) is reduced, since \( g_n \) is separable over \( K \) and hence \( \mathcal{A}' \) is a radical ideal. Further, from 5.3.a it follows that \( V_{K}(\mathcal{A}') = V_{K}(\mathcal{A}) \). Now, use Hilbert’s Nullstellensatz see \([11, 9] \) or \([10 \text{ Theorem 2.10, HNS 2}] \) to conclude the equality \( \mathcal{A}' = \mathcal{A} \).

5.3 Remark
The Shape Lemma\(^{[2]}\) appeared first time in \([3] \) which may be regarded as a natural generalization of the Primitive Element Theorem, see Example 3.2. Further, it gives a very useful presentation of the radical ideal \( \mathcal{A} \) which allows to find the solution space \( V_{K}(\mathcal{A}) \) immediately, namely:
\[
V_{K}(\mathcal{A}) = \{ (g_1(a), \ldots, g_{n-1}(a), a) \in K^n \mid g_n(a) = 0 \}.
\]
In other words the last coordinates are zeros of \( g_n \) and for a fixed last coordinate \( a_n \), the other coordinates are determined by evaluation of polynomials \( g_1, \ldots, g_{n-1} \) at \( a_n \). This simple shape of the solution space \( V_{K}(\mathcal{A}) \) is quite convenient to work with. The primary decomposition of \( \mathcal{A} \) is given by the prime factorization of the polynomial \( g_n \). Under the conditions on the polynomials \( g_1, \ldots, g_{n-1}, g_n \in K[X] \) as in the proof of the Shape Lemma\(^{[5,2]}\) one can easily verify that \( X_1 - g_1(X_n), \ldots, X_{n-1} - g_{n-1}(X_n), g_n(X_n) \) form a reduced (=minimal) Gröbner basis of the radical ideal \( \mathcal{A} \) relative to the lexicographic order \( X_1 > X_2 > \cdots > X_n \). For a different proof of the Shape Lemma\(^{[5,2]}\) see \([8] \) Theorem 3.7.25] and a detailed recipe for solving systems of polynomial equations efficiently using the Shape Lemma\(^{[5,2]}\) is also given in \([8] \) Theorem 3.7.26]. The Shape Lemma\(^{[5,2]}\) also appeared in \([2 \text{ Exercise 16, § 4, Chapter 2}] \).

5.4 Theorem
Let \( K \) be a real closed field and \( K := \mathbb{C}_K = K[i], i^2 = -1, \) be the algebraic closure of \( K \)(see 3.1(g)), \( \mathfrak{A} \subset K[X_1, \ldots, X_n] \) a 0-(Krull) dimensional radical ideal. With the notation and assumptions as in 5.1 further, let \( H \in K[X_1, \ldots, X_n] \), \( H \neq 0 \), \( h \) be the image of \( H \) in \( A \) and let \( \Phi_h : A \times A \rightarrow K, \quad (f, f') \mapsto tr^K_{K}(hf f') \), be the generalized trace form associated with \( h \) in \( A \).

(a) There exists an element \( x \in A \) such that \( x = \{ 1, x, \ldots, x^{m-1} \} \) is a \( K \)-basis of \( A \) and the Gram's matrix \( \Phi_{\Phi_h}(x) \) of \( \Phi_h \) with respect to the \( K \)-basis \( x \) is a symmetric matrix in \( M_m(K) \).

(b) \( \Phi_{\Phi_h}(x) = \mathcal{D} \otimes \mathcal{D} \), where \( \mathcal{D} := \text{Diag} \{(h(z) \mid z \in V_{K}(\mathcal{A})) \} \) is the diagonal matrix and \( \mathcal{D} = V(z \mid z \in V_{K}(\mathcal{A})) \subset GL_m(K) \) is the Vandermonde’s matrix of the elements \( z \in V_{K}(\mathcal{A}) \).

(c) Let \( p_{H} := \# \{ a \in V_{K}(\mathcal{A}) \mid H(a) > 0 \} \) and \( q_{H} := \# \{ a \in V_{K}(\mathcal{A}) \mid H(a) < 0 \} \). Then type \( \Phi_h = (p_{H} + s, q_{H} + s) \), where \( s = \# \{ V_{K}(\mathcal{A}) \setminus V_{K}(\mathcal{A}) \} \) and rank \( \Phi_h = \# \{ a \in V_{K}(\mathcal{A}) \mid H(a) \neq 0 \} \). In particular, sign \( \Phi_h = p_{H} - q_{H} \).

Proof
Remember the notation and assumptions from 5.1 that \( V_{K}(\mathcal{A}) = \{ a_1, \ldots, a_r \} \subset V_{K}(\mathcal{A}) = \{ a_1, \ldots, a_r, a_{r+1}, \ldots, a_{r+s}, \ldots, a_{r+s}, \ldots, \} \), where \( r := \#V_{K}(\mathcal{A}), \quad r + s = \# \text{Sp} A \) and \( A = \text{Dim}_{K} A = \text{Dim}_{K} A_{K} = \#V_{K}(A) \) and that \( V_{K}(\mathcal{A}) \) is in general \( X_n \)-position, see 5.1.a. Therefore by using the Shape Lemma\(^{[5,2]}\) and notation in its proof, we reduce to the one variable case, i. e. we may assume that: There exists a square free polynomial \( g \in K[X] \) such that \( \mathcal{A} = \langle g \rangle \subset K[X] \) and a \( K \)-algebra isomorphism \( A \rightarrow K[X]/\langle g \rangle \). Further, since \( K \) is a real closed field (see 3.1(g)), \( g = (X - a_1) \cdots (X - a_r) \pi_1 \cdots \pi_s \), where \( a_i \in K, i = 1, \ldots, r \) and \( \pi_j = (X - z_j)(X - z_j) \in K[X], z_j \in K \setminus K, j = 1, \ldots, s \). Therefore identifying \( A \) with \( K[X]/\langle g \rangle \), we have a \( K \)-basis \( 1, x, \ldots, x^{m-1} \) of \( A \), where \( x \) is the image of \( X \) modulo \( \langle g \rangle \), \( m = \deg g \). Further, \( V_{K}(\mathcal{A}) = V_{K}(g) = \{ a_1, \ldots, a_r \} \subset V_{K}(\mathcal{A}) = V_{K}(g) = \{ a_1, \ldots, a_r, z_1, \ldots, z_s, z_s \} \). With these identifications, for \( \mathcal{H} \in 1, \ldots, n \). For a proof consider the polynomial \( g := \sum \frac{y_i}{z_i} \prod_{j \neq i} (X - x_j), \) where \( z_i := \prod_{j \neq i} (x_i - x_j) \).
With the notation and assumptions as in 5.1, our main goal is to relate the cardinality \( \#V \) of the \( k \)-dimensional affine algebraic set defined by \( h \) on \( A \). Then:

\[ \Phi_h : A \times A \to K, (f, g) \mapsto \text{tr}^A(hfg), \]

is the Vandermonde’s matrix of the distinct elements \( a_1, \ldots, a_r, z_1, \ldots, z_s \in K \) which belongs to \( \text{GL}_m(K) \). This proves (a) and (b).

(c) The equality for the rank follows from the factorization \( \mathcal{G}_\Phi(x) = \mathcal{G}_\Phi^{(1)} \mathcal{G}_\Phi^{(2)} \) of \( \mathcal{G}_\Phi(x) \) (in b), since \( \mathcal{G}_\Phi \) is invertible. Note that the local decomposition \( A \sim \to K^r \times K^s \) yields the orthogonal decompositions

\[ \Phi_h = (\Phi_h)^K \oplus \cdots \oplus (\Phi_h)^K \oplus (\Phi_h)^K \oplus \cdots (\Phi_h)^K, \]

where \( (\Phi_h)^K = \Phi_h|K \) is the restrictions of \( \Phi_h \) to the real component at \( a_i \in K \) with Gram’s matrix \( \mathcal{G}_{\Phi_h}^{(1)}(1, 1) = (h(a_i)) \in M_1(K), i = 1, \ldots, r \) and \( (\Phi_h)^K = \Phi_h|K \), is the restrictions of \( \Phi_h \) to the non-real component \( K[X]/(\pi_j) \sim \to K \) at \( m_j = (\pi_j) \in \text{Spec} (A) \) for all \( j = 1, \ldots, s \), see 5.1.b, Lemma 5.2 and 4.3. Further, obviously, type \( (\Phi_h)_i^K = \text{sign}(\Phi_h)_i^K = \text{sign} h(a_i) = \text{sign} H(a_i) \) for all \( i = 1, \ldots, r \) and by Example 5.1.11 (since \( \pi_j = (X - z_j)(X - \bar{z}_j), z_j \in K \setminus K \)), we have \( \text{sign}(\Phi_h)_j^K = (1, -1) \) for all \( j = 1, \ldots, s \). Therefore type \( \Phi_h = \sum_i \text{type}(\Phi_h)_i^K + \sum_{j=r+1}^{s} \text{type}(\Phi_h)_j^K = (pH + s, qH + s) \) and so \( \Phi_h = pH - qH \). 

5.5 Corollary (Hermite) Let \( K \) be an arbitrary real closed field and let \( g = b_0 + b_1X + \cdots + b_{m-1}X^{m-1} + X^m \in K[X], A := K[X]/(g) \). Then type \( \text{tr}^A(h) = (p + s, s) \), where \( \text{tr}^A : A \times A \to K, (f, f') \mapsto \text{tr}_K(ff') = \text{the trace of the linear map } \lambda_{ff'} : A \to A, \text{ the trace form on } A, \) is the number of zeros \( \#V_K(g) \) of \( g \) in \( K \) and \( s \) is the half of the number of zeros of \( g \) in the algebraic closure \( K \) of \( K \) which are not in \( K \). In particular, \( \text{tr}^A_k = p = \#V_K(g) \).

With the notation and assumptions as in 5.1, our main goal is to relate the cardinality \( \#V_K(\mathfrak{A}) \) with the signatures of the generalized trace form on the finite \( K \)-algebra \( A \).

5.6 With the notation and assumptions as in 5.1 and as in the proof of Theorem 5.4. Further, let \( H \in K[X_1, \ldots, X_n], H \neq 0 \) and \( V_K(H) := \{ a \in K^n \mid H(a) = 0 \} \) be the hypersurface (an \( (n - 1) \)-dimensional affine algebraic set in \( K^n \)) defined by \( H \). Then the complement of \( V_K(H) \) in \( K^n \) is the union of line-connected subsets (in the strong topology on \( K^n \) (see 5.1.d)) on which \( H \) takes either all positive values or all negative values, i.e. \( K^n \setminus V_K(H) = H^+ \cup H^-, \) where \( H^+ := \{ a \in K^n \mid H(a) > 0 \} \) and \( H^- := \{ a \in K^n \mid H(a) < 0 \} \).

Further, since \( V_K(\mathfrak{A}) = (V_K(\mathfrak{A}) \cap H^+) \cup (V_K(\mathfrak{A}) \cap H^-) \cup (V_K((\mathfrak{A}, H))) \), we have:

\[ \#V_K(\mathfrak{A}) = \#(V_K(\mathfrak{A}) \cap H^+) + \#(V_K(\mathfrak{A}) \cap H^-) + \#(V_K((\mathfrak{A}, H))), \]

and hence to compute \( \#V_K(\mathfrak{A}) \), we can use arbitrary polynomial \( H \in K[X_1, \ldots, X_n] \) and compute the cardinalities \( \#V_K(\mathfrak{A}) \cap H^+, \#V_K(\mathfrak{A}) \cap H^- \) and \( \#V_K((\mathfrak{A}, H)) \).

More precisely, we have:

5.7 Theorem With the notation and assumptions as in 5.1. For \( H \in K[X_1, \ldots, X_n], H \neq 0 \), let \( p_H := \#V_K(\mathfrak{A}) \cap H^+ \) and \( d_H := \#V_K(\mathfrak{A}) \cap H^- \). Further, let \( h \) denote the image of \( H \) in \( A = K[X_1, \ldots, X_n]/\mathfrak{A} \) and \( \Phi_h : A \times A \to K, (f, g) \mapsto \text{tr}^A(hfg) \), be the generalized trace form defined by \( h \) on \( A \). Then:

(a) (Pederson-Roy-Szpirglas [11, Theorem 2.1])
\[
\text{sign } \Phi_h = p_H - q_H \quad \text{and} \quad \text{rank } \Phi_h = \#(V_K(\mathfrak{A}) \setminus V_K(H)).
\]

(b) \sign \Phi_h^2 = p_H + q_H \quad \text{and} \quad \text{rank } \Phi_h^2 = \#(V_K(\mathfrak{A}) \setminus V_K(H)).

(c) Let \mathfrak{B} := (\mathfrak{A}, H) be the ideal (in \( K[X_1, \ldots, X_n] \)) generated by \( \mathfrak{A} \) and \( H \). Then the \( K \)-algebra \( B := K[X_1, \ldots, X_n]/\mathfrak{B} \) is finite over \( K \) and \( \text{sign tr}^B_K = \#V_K(\mathfrak{B}) \).

(d) The three signatures \( \text{sign } \Phi_h \), \( \text{sign } \Phi_h^2 \) and \( \text{rank } \Phi_h^2 \) uniquely determine the natural numbers \( p_H, q_H \) and \( \#V_K(\mathfrak{B}) = V_K(\mathfrak{A}) \cap V_K(H) \). In particular, they determine the cardinality \( \#V_K(\mathfrak{A}) = p_h + q_h + \#V_K(\mathfrak{B}) \).

**Proof** (a) : Proved in Theorem 5.4(b).

(b) : Since \( H^2(a) = H(a) H(a) > 0 \) for every \( a \in H^+ \cup H^- \) and \( V_K(H^2) = V_K(H) \), from Theorem 5.4(b) it follows that \( \text{sign } \Phi_h^2 = p_H + q_H \quad \text{and} \quad \text{rank } \Phi_h^2 = \#(V_K(\mathfrak{A}) \setminus V_K(H)) \).

(c) : Since the \( K \)-algebra \( B \) is a homomorphic image of the \( K \)-algebra \( A, B \) is also finite over \( K \). The equality \( \text{sign } \Phi_h^2 = \#V_K(\mathfrak{B}) \) is immediate from Theorem 5.4(a) \( (H = 1) \) or Theorem 4.6.

(d) : Immediate from the formula 5.6.a for \#V_K(\mathfrak{A}) in 5.6 and the parts (a), (b) and (c).

\section*{References}

[1] Atiyah, M. F., Macdonald, I. G. : *Introduction to commutative algebra*. Addison-Wesley, Reading, Mass. (1969), x+128 pp.

[2] Cox, David, A., Little, John and O’Shea, Donald : *Using Algebraic Geometry*, Graduate Texts in Mathematics, 185, (2nd edition), Springer-Verlag, New York (2005).

[3] Gianni, P. and Mora, T. : Algebraic solutions of systems of polynomial equations using Gröbner bases, *Proc. AAECC 5, LNCS 356* (1989), 247-257.

[4] Goel, Kriti, Patil, Dilip P. and Verma, Jugal : Nullstellensätze and Applications, *Preprint, IIT Bombay* 2018.

[5] Hermite, C. : Sur l’Extension du Théorème de M. Sturm a un Système D’Équations Simultanées, *C. R. Acad. Sci., Paris*, 35 (1852).

[6] Hermite, C. : Sur l’Extension du Théorème de M. Sturm a un Système D’Équations Simultanées, *Oeuvres de Charles Hermite*, Tome III., 1-34.

[7] Jacobson, N. : *Basic Algebra*, Vol. 2, (2nd edition), W. H. Freeman and Company (1989), xv+pp. 686.

[8] Kreuzer, K. and Robbiano, L. : *Computational Commutative Algebra I*, Springer-Verlag Berlin Heidelberg (2000), x+324 pp.

[9] Kunz, E. : *Introduction to commutative algebra and algebraic geometry*, Translated from the 1980 German original, Modern Birkhäuser Classics. Birkhäuser/Springer, New York (2013), xiv+238 pp.

[10] Patil, Dilip P. and Storch, U. : *Introduction to Algebraic Geometry and Commutative Algebra*. IISc Lecture Notes Series, No. 1, IISc Press / World Scientific Publications Singapore / Chennai, 2010, pp. x + 207. Indian Edition Published by Cambridge University press India Pvt. Ltd. 2012, pp. x+209.

[11] Pedersen, P., Roy, M.-F. and Sziprgas, A. : Counting real zeros in the multivariate case, *Computational Algebraic Geometry*, Editors : F. Eyssette and A. Galligo, Progr. Math., 109, Birkhäuser, Boston (1993), 203-224.

[12] Scheja, G. and Storch, U. : *Lehrbuch der Algebra*, Teil 2, B. G. Teubner, Stuttgart (1988), 816 pp.

[13] Storch, U. and Wiebe, H. : *Lehrbuch der Mathematik*, Bd 4, (2. Aufl. 1999, Korr. Nachdruck 2010 edition) Spektrum Akademischer Verlag, Heidelberg (2011), 782 pp.

[14] Sylvester, J. On a Theory of the Syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm’s Functions, and that of the greatest Algebraical Common Measure, *Philos. Trans. Roy. Soc. London* 143 (1853), 407-548.

\footnote{1} Department of Mathematics, Indian Institute of Science Bangalore  
E-mail address: patil@iisc.ac.in

\footnote{2} Department of Mathematics, Indian Institute of Technology Bombay  
E-mail address: jkv@math.iitb.ac.in