CHAPTER XV

Structure of Bilinear Forms

There are three major types of bilinear forms: hermitian (or symmetric), unitary, and alternating (skew-symmetric). In this chapter, we give structure theorems giving normalized expressions for these forms with respect to suitable bases. The chapter also follows the standard pattern of decomposing an object into a direct sum of simple objects, insofar as possible.

§1. PRELIMINARIES, ORTHOGONAL SUMS

The purpose of this chapter is to go somewhat deeper into the structure theory for our three types of forms. To do this we shall assume most of the time that our ground ring is a field, and in fact a field of characteristic $\neq 2$ in the symmetric case.

We recall our three definitions. Let E be a module over a commutative ring R. Let $g : E \times E \to R$ be a map. If g is bilinear, we call g a **symmetric** form if $g(x, y) = g(y, x)$ for all $x, y \in E$. We call g **alternating** if $g(x, x) = 0$, and hence $g(x, y) = -g(y, x)$ for all x, $y \in E$. If R has an automorphism of order 2, written $a \mapsto \bar{a}$, we say that g is a **hermitian** form if it is linear in its first variable, antilinear in its second, and

$$g(x, y) = \overline{g(y, x)}.$$

We shall write $g(x, y) = \langle x, y \rangle$ if the reference to g is clear. We also occasionally write $g(x, y) = x \cdot y$ or $g(x, x) = x^2$. We sometimes call g a **scalar product**.

S. Lang, *Algebra*

If $v_1, \dots, v_m \in E$, we denote by $(v_1, \dots, v_m)$ the submodule of E generated by $v_1, \dots, v_m$.

Let g be symmetric, alternating, or hermitian. Then it is clear that the left kernel of g is equal to its right kernel, and it will simply be called the **kernel** of g.

In any one of these cases, we say that g is **non-degenerate** if its kernel is 0. Assume that E is finite dimensional over the field k. The form is non-degenerate if and only if it is non-singular, i.e., induces an isomorphism of E with its dual space (anti-dual in the case of hermitian forms).

Except for the few remarks on the anti-linearity made in the previous chapter, we don't use the results of the duality in that chapter. We need only the duality over fields, given in Chapter III. Furthermore, we don't essentially meet matrices again, except for the remarks on the pfaffian in §10.

We introduce one more notation. In the study of forms on vector spaces, we shall frequently decompose the vector space into direct sums of orthogonal subspaces. If E is a vector space with a form g as above, and F, F' are subspaces, we shall write

$$E = F \perp F'$$

to mean that E is the direct sum of F and F', and that F is orthogonal (or perpendicular) to F', in other words, $x \perp y$ (or $\langle x, y \rangle = 0$) for all $x \in F$ and $y \in F'$. We then say that E is the **orthogonal sum** of F and F'. There will be no confusion with the use of the symbol $\perp$ when we write $F \perp F'$ to mean simply that F is perpendicular to F'. The context always makes our meaning clear.

Most of this chapter is devoted to giving certain orthogonal decompositions of a vector space with one of our three types of forms, so that each factor in the sum is an easily recognizable type.

In the symmetric and hermitian case, we shall be especially concerned with direct sum decompositions into factors which are 1-dimensional. Thus if $\langle\ ,\ \rangle$ is symmetric or hermitian, we shall say that $\{v_1, \dots, v_n\}$ is an **orthogonal basis** (with respect to the form) if $\langle v_i, v_j \rangle = 0$ whenever $i \neq j$. We see that an orthogonal basis gives such a decomposition. If the form is nondegenerate, and if $\{v_1, \dots, v_n\}$ is an orthogonal basis, then we see at once that $\langle v_i, v_i \rangle \neq 0$ for all i.

Proposition 1.1. *Let E be a vector space over the field k, and let g be a form of one of the three above types. Suppose that E is expressed as an orthogonal sum,*

$$E = E_1 \perp \cdots \perp E_m.$$

Then g is non-degenerate on E if and only if it is non-degenerate on each E_i. If E_i^0 is the kernel of the restriction of g to E_i, then the kernel of g in E is the orthogonal sum

$$E^0 = E_1^0 \perp \cdots \perp E_m^0.$$

Proof. Elements v, w of E can be written uniquely

$$v = \sum_{i=1}^{m} v_i, \qquad w = \sum_{i=1}^{m} w_i$$

with $v_i, w_i \in E_i$. Then

$$v \cdot w = \sum_{i=1}^{m} v_i \cdot w_i,$$

and $v \cdot w = 0$ if $v_i \cdot w_i = 0$ for each $i = 1, \dots, m$. From this our assertion is obvious.

Observe that if $E_1, \dots, E_m$ are vector spaces over k, and $g_1, \dots, g_m$ are forms on these spaces respectively, then we can define a form $g = g_1 \oplus \cdots \oplus g_m$ on the direct sum $E = E_1 \oplus \cdots \oplus E_m$; namely if v, w are written as above, then we let

$$g(v, w) = \sum_{i=1}^{m} g_i(v_i, w_i).$$

It is then clear that, in fact, we have $E = E_1 \perp \cdots \perp E_m$. We could also write $g = g_1 \perp \cdots \perp g_m$.

Proposition 1.2. *Let E be a finite-dimensional space over the field k, and let g be a form of the preceding type on E. Assume that g is non-degenerate. Let F be a subspace of E. The form is non-degenerate on F if and only if $F + F^\perp = E$, and also if and only if it is non-degenerate on $F^\perp$.*

Proof. We have (as a trivial consequence of Chapter III, §5)

$$\dim F + \dim F^\perp = \dim E = \dim(F + F^\perp) + \dim(F \cap F^\perp).$$

Hence $F + F^\perp = E$ if and only if $\dim(F \cap F^\perp) = 0$. Our first assertion follows at once. Since F, $F^\perp$ enter symmetrically in the dimension condition, our second assertion also follows.

Instead of saying that a form is non-degenerate on E, we shall sometimes say, by abuse of language, that E is non-degenerate.

Let E be a finite-dimensional space over the field k, and let g be a form of the preceding type. Let E_0 be the kernel of the form. Then we get an induced form of the same type

$$g_0 : E/E_0 \times E/E_0 \to k,$$

because $g(x, y)$ depends only on the coset of x and the coset of y modulo E_0. Furthermore, g_0 is non-degenerate since its kernel on both sides is 0.

Let E, E' be finite-dimensional vector spaces, with forms g, g' as above, respectively. A linear map $\sigma : E \to E'$ is said to be **metric** if

$$g'(\sigma x, \sigma y) = g(x, y)$$

or in the dot notation, $\sigma x \cdot \sigma y = x \cdot y$ for all $x, y \in E$. If σ is a linear isomorphism, and is metric, then we say that σ is an **isometry**.

Let E, E_0 be as above. Then we have an induced form on the factor space E/E_0. If W is a complementary subspace of E_0, in other words, $E = E_0 \oplus W$, and if we let $\sigma : E \to E/E_0$ be the canonical map, then σ is metric, and induces an isometry of W on E/E_0. This assertion is obvious, and shows that if

$$E = E_0 \oplus W'$$

is another direct sum decomposition of E, then W' is isometric to W. We know that $W \approx E/E_0$ is nondegenerate. Hence our form determines a unique nondegenerate form, up to isometry, on complementary subspaces of the kernel.

§2. QUADRATIC MAPS

Let R be a commutative ring and let E, F be R-modules. We suppress the prefix R- as usual. We recall that a bilinear map $f : E \times E \to F$ is said to be symmetric if $f(x, y) = f(y, x)$ for all $x, y \in E$.

We say that F is **without 2-torsion** if for all $y \in F$ such that $2y = 0$ we have $y = 0$. (This holds if 2 is invertible in R.)

Let $f : E \to F$ be a mapping. We shall say that f is **quadratic** (i.e. R-quadratic) if there exists a symmetric bilinear map $g : E \times E \to F$ and a linear map $h : E \to F$ such that for all $x \in E$ we have

$$f(x) = g(x, x) + h(x).$$

Proposition 2.1. *Assume that F is without 2-torsion. Let $f : E \to F$ be quadratic, expressed as above in terms of a symmetric bilinear map and a linear map. Then g, h are uniquely determined by f. For all $x, y \in E$ we have*

$$2g(x, y) = f(x + y) - f(x) - f(y).$$

Proof. If we compute $f(x + y) - f(x) - f(y)$, then we obtain $2g(x, y)$. If g_1 is symmetric bilinear, h_1 is linear, and $f(x) = g_1(x, x) + h_1(x)$, then $2g(x, y) = 2g_1(x, y)$. Since F is assumed to be without 2-torsion, it follows that $g(x, y) = g_1(x, y)$ for all $x, y \in E$, and thus that g is uniquely determined. But then h is determined by the relation

$$h(x) = f(x) - g(x, x).$$

We call g, h the bilinear and linear maps **associated** with f.

If $f : E \to F$ is a map, we define

$$\Delta f : E \times E \to F$$

by

$$\Delta f(x, y) = f(x + y) - f(x) - f(y).$$

We say that f is **homogeneous quadratic** if it is quadratic, and if its associated linear map is 0. We shall say that F is **uniquely divisible** by 2 if for each $z \in F$ there exists a unique $u \in F$ such that $2u = z$. (Again this holds if 2 is invertible in R.)

Proposition 2.2. *Let $f: E \to F$ be a map such that Δf is bilinear. Assume that F is uniquely divisible by 2. Then the map $x \mapsto f(x) - \frac{1}{2}\Delta f(x, x)$ is $\mathbf{Z}$-linear. If f satisfies the condition $f(2x) = 4f(x)$, then f is homogeneous quadratic.*

Proof. Obvious.

By a **quadratic form** on E, one means a homogeneous quadratic map $f: E \to R$, with values in R.

In what follows, we are principally concerned with symmetric bilinear forms. The quadratic forms play a secondary role.

§3. SYMMETRIC FORMS, ORTHOGONAL BASES

Let k be a field of characteristic $\neq 2$.

Let E be a vector space over k, with the symmetric form g. We say that g is a **null** form or that E is a **null** space if $\langle x, y\rangle = 0$ for all $x, y \in E$. Since we assumed that the characteristic of k is $\neq 2$, the condition $x^2 = 0$ for all $x \in E$ implies that g is a null form. Indeed,

$$4x \cdot y = (x + y)^2 - (x - y)^2.$$

Theorem 3.1. *Let E be $\neq 0$ and finite dimensional over k. Let g be a symmetric form on E. Then there exists an orthogonal basis.*

Proof. We assume first that g is non-degenerate, and prove our assertion by induction in that case. If the dimension n is 1, then our assertion is obvious.

Assume $n > 1$. Let $v_1 \in E$ be such that $v_1^2 \neq 0$ (such an element exists since g is assumed non-degenerate). Let $F = (v_1)$ be the subspace generated by v_1. Then F is non-degenerate, and by Proposition 1.2, we have

$$E = F + F^{\perp}.$$

Furthermore, $\dim F^{\perp} = n - 1$. Let $\{v_2, \ldots, v_n\}$ be an orthogonal basis of $F^{\perp}$.

Then $\{v_1, \dots, v_n\}$ are pairwise orthogonal. Furthermore, they are linearly independent, for if

$$a_1 v_1 + \cdots + a_n v_n = 0$$

with $a_i \in k$ then we take the scalar product with v_i to get $a_i v_i^2 = 0$ whence $a_i = 0$ for all i.

Remark. We have shown in fact that if g is non-degenerate, and $v \in E$ is such that $v^2 \neq 0$ then we can complete v to an orthogonal basis of E.

Suppose that the form g is degenerate. Let E_0 be its kernel. We can write E as a direct sum

$$E = E_0 \oplus W$$

for some subspace W. The restriction of g to W is non-degenerate; otherwise there would be an element of W which is in the kernel of E, and $\neq 0$. Hence if $\{v_1, \dots, v_r\}$ is a basis of E_0, and $\{w_1, \dots, w_{n-r}\}$ is an orthogonal basis of W, then

$$\{v_1, \dots, v_r, w_1, \dots, w_{n-r}\}$$

is an orthogonal basis of E, as was to be shown.

Corollary 3.2. *Let $\{v_1, \dots, v_n\}$ be an orthogonal basis of E. Assume that $v_i^2 \neq 0$ for $i \leqq r$ and $v_i^2 = 0$ for $i > r$. Then the kernel of E is equal to $(v_{r+1}, \dots, v_n)$.*

Proof. Obvious.

If $\{v_1, \dots, v_n\}$ is an orthogonal basis of E and if we write

$$X = x_1 v_1 + \cdots + x_n v_n$$

with $x_i \in k$, then

$$X^2 = a_1 x_1^2 + \cdots + a_n x_n^2$$

where $a_i = \langle v_i, v_i \rangle$. In this representation of the form, we say that it is **diagonalized**. With respect to an orthogonal basis, we see at once that the associated matrix of the form is a diagonal matrix, namely

$$\begin{pmatrix} a_1 & & & & & \\ & a_2 & & 0 & & \\ & & \ddots & & & \\ & & & a_r & & \\ & 0 & & & 0 & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}.$$

Example. Note that Exercise 33 of Chapter XIII gave an interesting example of an orthogonal decomposition involving harmonic polynomials.

§4. SYMMETRIC FORMS OVER ORDERED FIELDS

Theorem 4.1. (Sylvester) *Let k be an ordered field and let E be a finite dimensional vector space over k, with a non-degenerate symmetric form g. There exists an integer $r \geqq 0$ such that, if $\{v_1, \dots, v_n\}$ is an orthogonal basis of E, then precisely r among the n elements $v_1^2, \dots, v_n^2$ are > 0, and $n - r$ among these elements are < 0.*

Proof. Let $a_i = v_i^2$, for $i = 1, \dots, n$. After renumbering the basis elements, say $a_1, \dots, a_r > 0$ and $a_i < 0$ for $i > r$. Let $\{w_1, \dots, w_n\}$ be any orthogonal basis, and let $b_i = w_i^2$. Say $b_1, \dots, b_s > 0$ and $b_j < 0$ for $j > s$. We shall prove that $r = s$. Indeed, it will suffice to prove that

$$v_1, \dots, v_r, w_{s+1}, \dots, w_n$$

are linearly independent, for then we get $r + n - s \leqq n$, whence $r \leqq s$, and $r = s$ by symmetry. Suppose that

$$x_1 v_1 + \cdots + x_r v_r + y_{s+1} w_{s+1} + \cdots + y_n w_n = 0.$$

Then

$$x_1 v_1 + \cdots + x_r v_r = -y_{s+1} w_{s+1} - \cdots - y_n w_n.$$

Squaring both sides yields

$$a_1 x_1^2 + \cdots + a_r x_r^2 = b_{s+1} y_{s+1}^2 + \cdots + b_n y_n^2.$$

The left-hand side is $\geqq 0$, and the right-hand side is $\leqq 0$. Hence both sides are equal to 0, and it follows that $x_i = y_j = 0$, in other words that our vectors are linearly independent.

Corollary 4.2. *Assume that every positive element of k is a square. Then there exists an orthogonal basis $\{v_1, \dots, v_n\}$ of E such that $v_i^2 = 1$ for $i \leqq r$ and $v_i^2 = -1$ for $i > r$, and r is uniquely determined.*

Proof. We divide each vector in an orthogonal basis by the square root of the absolute value of its square.

A basis having the property of the corollary is called **orthonormal**. If X is an element of E having coordinates $(x_1, \dots, x_n)$ with respect to this basis, then

$$X^2 = x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_n^2.$$

We say that a symmetric form g is **positive definite** if $X^2 > 0$ for all $X \in E$, $X \neq 0$. This is the case if and only if $r = n$ in Theorem 4.1. We say that g is **negative definite** if $X^2 < 0$ for all $X \in E$, $X \neq 0$.

Corollary 4.3. *The vector space E admits an orthogonal decomposition $E = E^+ \perp E^-$ such that g is positive definite on E^+ and negative definite on E^-. The dimension of E^+ (or E^-) is the same in all such decompositions.*

Let us now assume that the form g is positive definite and that every positive element of k is a square.

We define the **norm** of an element $v \in E$ by

$$|v| = \sqrt{v \cdot v}.$$

Then we have $|v| > 0$ if $v \neq 0$. We also have the **Schwarz inequality**

$$|v \cdot w| \leqq |v|\,|w|$$

for all $v, w \in E$. This is proved in the usual way, expanding

$$0 \leqq (av \pm bw)^2 = (av \pm bw) \cdot (av \pm bw)$$

by bilinearity, and letting $b = |v|$ and $a = |w|$. One then gets

$$\mp 2ab\, v \cdot w \leqq 2|v|^2|w|^2.$$

If $|v|$ or $|w| = 0$ our inequality is trivial. If neither is 0 we divide by $|v|\,|w|$ to get what we want.

From the Schwarz inequality, we deduce the triangle inequality

$$|v + w| \leqq |v| + |w|.$$

We leave it to the reader as a routine exercise.

When we have a positive definite form, there is a canonical way of getting an orthonormal basis, starting with an arbitrary basis $\{v_1, \ldots, v_n\}$ and proceeding inductively. Let

$$v_1' = \frac{1}{|v_1|} v_1.$$

Then v_1 has norm 1. Let

$$w_2 = v_2 - (v_2 \cdot v_1')v_1',$$

and then

$$v_2' = \frac{1}{|w_2|} w_2.$$

Inductively, we let

$$w_r = v_r - (v_r \cdot v'_1)v'_1 - \cdots - (v_r \cdot v'_{r-1})v'_{r-1}$$

and then

$$v'_r = \frac{1}{|w_r|} w_r.$$

The $\{v'_1, \ldots, v'_n\}$ is an orthonormal basis. The inductive process just described is known as the **Gram-Schmidt orthogonalization.**

§5. HERMITIAN FORMS

Let k_0 be an ordered field (a subfield of the reals, if you wish) and let $k = k_0(i)$, where $i = \sqrt{-1}$. Then k has an automorphism of order 2, whose fixed field is k_0.

Let E be a finite-dimensional vector space over k. We shall deal with a hermitian form on E, i.e. a map

$$E \times E \to k$$

written

$$(x, y) \mapsto \langle x, y \rangle$$

which is k-linear in its first variable, k-anti-linear in its second variable, and such that

$$\langle x, y \rangle = \overline{\langle y, x \rangle}$$

for all $x, y \in E$.

We observe that $\langle x, x \rangle \in k_0$ for all $x \in E$. This is essentially the reason why the proofs of statements concerning symmetric forms hold essentially without change in the hermitian case. We shall now make the list of the properties which apply to this case.

Theorem 5.1. *There exists an orthogonal basis. If the form is non-degenerate, there exists an integer r having the following property. If $\{v_1, \ldots, v_n\}$ is an orthogonal basis, then precisely r among the n elements*

$$\langle v_1, v_1 \rangle, \ldots, \langle v_n, v_n \rangle$$

are > 0 and $n - r$ among these elements are < 0.

An orthogonal basis $\{v_1, \dots, v_n\}$ such that $\langle v_i, v_i \rangle = 1$ or -1 is called an **orthonormal** basis.

Corollary 5.2. *Assume that the form is non-degenerate, and that every positive element of k_0 is a square. Then there exists an orthonormal basis.*

We say that the hermitian form is **positive definite** if $\langle x, x \rangle > 0$ for all $x \in E$. We say that it is **negative definite** if $\langle x, x \rangle < 0$ for all $x \in E$, $x \neq 0$.

Corollary 5.3. *Assume that the form is non-degenerate. Then E admits an orthogonal decomposition $E = E^+ \perp E^-$ such that the form is positive definite on E^+ and negative definite on E^-. The dimension of E^+ (or E^-) is the same in all such decompositions.*

The proofs of Theorem 5.1 and its corollaries are identical with those of the analogous results for symmetric forms, and will be left to the reader.

We have the **polarization identity**, for any k-linear map $A : E \to E$, namely

$$\langle A(x + y), (x + y) \rangle - \langle A(x - y), (x - y) \rangle = 2[\langle Ax, y \rangle + \langle Ay, x \rangle].$$

If $\langle Ax, x \rangle = 0$ for all x, we replace x by ix and get

$$\langle Ax, y \rangle + \langle Ay, x \rangle = 0,$$

$$i\langle Ax, y \rangle - i\langle Ay, x \rangle = 0.$$

From this we conclude:

If $\langle Ax, x \rangle = 0$, for all x, then $A = 0$.

This is the only statement which has no analogue in the case of symmetric forms. The presence of i in one of the above linear equations is essential to the conclusion. In practice, one uses the statement in the complex case, and one meets an analogous situation in the real case when A is symmetric. Then the statement for symmetric maps is obvious.

Assume that the hermitian form is positive definite, and that every positive element of k_0 is a square.

We have the **Schwarz inequality**, namely

$$|\langle x, y \rangle|^2 \leqq \langle x, x \rangle \langle y, y \rangle$$

whose proof comes again by expanding

$$0 \leqq \langle \alpha x + \beta y, \alpha x + \beta y \rangle$$

and setting $\alpha = \langle y, y \rangle$ and $\beta = -\langle x, y \rangle$.

We define the norm of $|x|$ to be

$$|x| = \sqrt{\langle x, x \rangle}.$$

Then we get at once the triangle inequality

$$|x + y| \leqq |x| + |y|,$$

and for $\alpha \in k$,

$$|\alpha x| = |\alpha|\,|x|.$$

Just as in the symmetric case, given a basis, one can find an orthonormal basis by the inductive procedure of subtracting successive projections. We leave this to the reader.

§6. THE SPECTRAL THEOREM (HERMITIAN CASE)

Throughout this section, we let E be a finite dimensional space over $\mathbf{C}$, *of dimension* $\geqq 1$, *and we endow E with a positive definite hermitian form.*

Let $A : E \to E$ be a linear map (*i.e.* $\mathbf{C}$-linear map) of E into itself. For fixed $y \in E$, the map $x \mapsto \langle Ax, y \rangle$ is a linear functional, and hence there exists a unique element $y^* \in E$ such that

$$\langle Ax, y \rangle = \langle x, y^* \rangle$$

for all $x \in E$. We define the map $A^* : E \to E$ by $A^*y = y^*$. It is immediately clear that A^* is linear, and we shall call A^* the **adjoint** of A with respect to our hermitian form.

The following formulas are trivially verified, for any linear maps A, B of E into itself:

$$(A + B)^* = A^* + B^*, \qquad A^{**} = A,$$

$$(\alpha A)^* = \bar{\alpha} A^*, \qquad (AB)^* = B^* A^*.$$

A linear map A is called **self-adjoint** (or **hermitian**) if $A^* = A$.

Proposition 6.1. *A is hermitian if and only if $\langle Ax, x \rangle$ is real for all $x \in E$.*

Proof. Let A be hermitian. Then

$$\overline{\langle Ax, x \rangle} = \overline{\langle x, Ax \rangle} = \langle Ax, x \rangle,$$

whence $\langle Ax, x \rangle$ is real. Conversely, assume $\langle Ax, x \rangle$ is real for all x. Then

$$\langle Ax, x \rangle = \overline{\langle Ax, x \rangle} = \langle x, Ax \rangle = \langle A^*x, x \rangle,$$

and consequently $\langle (A - A^*)x, x \rangle = 0$ for all x. Hence $A = A^*$ by polarization.

Let $A : E \to E$ be a linear map. An element $\xi \in E$ is called an **eigenvector** of A if there exists $\lambda \in \mathbf{C}$ such that $A\xi = \lambda\xi$. If $\xi \neq 0$, then we say that λ is an **eigenvalue** of A, belonging to ξ.

Proposition 6.2. *Let A be hermitian. Then all eigenvalues belonging to nonzero eigenvectors of A are real. If ξ, ξ' are eigenvectors $\neq 0$ having eigenvalues λ, λ' respectively, and if $\lambda \neq \lambda'$, then $\xi \perp \xi'$.*

Proof. Let λ be an eigenvalue, belonging to the eigenvector $\xi \neq 0$. Then $\langle A\xi, \xi \rangle = \langle \xi, A\xi \rangle$, and these two numbers are equal respectively to $\lambda\langle \xi, \xi \rangle$ and $\bar{\lambda}\langle \xi, \xi \rangle$. Since $\xi \neq 0$, it follows that $\lambda = \bar{\lambda}$, i.e. that λ is real. Secondly, assume that ξ, ξ' and λ, λ' are as described above. Then

$$\langle A\xi, \xi' \rangle = \lambda\langle \xi, \xi' \rangle = \langle \xi, A\xi' \rangle = \lambda'\langle \xi, \xi' \rangle,$$

from which it follows that $\langle \xi, \xi' \rangle = 0$.

Lemma 6.3. *Let $A : E \to E$ be a linear map, and* $\dim E \geqq 1$. *Then there exists at least one non-zero eigenvector of A.*

Proof. We consider $\mathbf{C}[A]$, i.e. the ring generated by A over $\mathbf{C}$. As a vector space over $\mathbf{C}$, it is contained in the ring of endomorphisms of E, which is finite dimensional, the dimension being the same as for the ring of all $n \times n$ matrices if $n = \dim E$. Hence there exists a non-zero polynomial P with coefficients in $\mathbf{C}$ such that $P(A) = 0$. We can factor P into a product of linear factors,

$$P(X) = (X - \lambda_1) \cdots (X - \lambda_m)$$

with $\lambda_j \in \mathbf{C}$. Then $(A - \lambda_1 I) \cdots (A - \lambda_m I) = 0$. Hence not all factors $A - \lambda_j I$ can be isomorphisms, and there exists $\lambda \in \mathbf{C}$ such that $A - \lambda I$ is not an isomorphism. Hence it has an element $\xi \neq 0$ in its kernel, and we get $A\xi - \lambda\xi = 0$. This shows that ξ is a non-zero eigenvector, as desired.

Theorem 6.4. (Spectral Theorem, Hermitian Case). *Let E be a non-zero finite dimensional vector space over the complex numbers, with a positive definite hermitian form. Let $A : E \to E$ be a hermitian linear map. Then E has an orthogonal basis consisting of eigenvectors of A.*

Proof. Let ξ_1 be a non-zero eigenvector, with eigenvalue λ_1, and let E_1 be the subspace generated by ξ_1. Then A maps $E_1^\perp$ into itself, because

$$\langle AE_1^\perp, \xi_1 \rangle = \langle E_1^\perp, A\xi_1 \rangle = \langle E_1^\perp, \lambda_1\xi_1 \rangle = \lambda_1\langle E_1^\perp, \xi_1 \rangle = 0,$$

whence $AE_1^\perp$ is perpendicular to ξ_1.

Since $\xi_1 \neq 0$ we have $\langle \xi_1, \xi_1 \rangle > 0$ and hence, since our hermitian form is non-degenerate (being positive definite), we have

$$E = E_1 \oplus E_1^\perp.$$

The restriction of our form to $E_1^\perp$ is positive definite (if $\dim E > 1$). From Proposition 6.1, we see at once that the restriction of A to $E_1^\perp$ is hermitian. Hence we can complete the proof by induction.

Corollary 6.5. *Hypotheses being as in the theorem, there exists an orthonormal basis consisting of eigenvectors of A.*

Proof. Divide each vector in an orthogonal basis by its norm.

Corollary 6.6. *Let E be a non-zero finite dimensional vector space over the complex numbers, with a positive definite hermitian form f. Let g be another hermitian form on E. Then there exists a basis of E which is orthogonal for both f and g.*

Proof. We write $f(x, y) = \langle x, y\rangle$. Since f is non-singular, being positive definite, there exists a unique hermitian linear map A such that $g(x, y) = \langle Ax, y\rangle$ for all $x, y \in E$. We apply the theorem to A, and find a basis as in the theorem, say $\{v_1, \dots, v_n\}$. Let λ_i be the eigenvalue such that $Av_i = \lambda_i v_i$. Then

$$g(v_i, v_j) = \langle Av_i, v_j\rangle = \lambda_i\langle v_i, v_j\rangle,$$

and therefore our basis is also orthogonal for g, as was to be shown.

We recall that a linear map $U : E \to E$ is **unitary** if and only if $U^* = U^{-1}$. This condition is equivalent to the property that $\langle Ux, Uy\rangle = \langle x, y\rangle$ for all elements $x, y \in E$. In other words, U is an automorphism of the form f.

Theorem 6.7. (Spectral Theorem, Unitary Case). *Let E be a non-zero finite dimensional vector space over the complex numbers, with a positive definite hermitian form. Let $U : E \to E$ be a unitary linear map. Then E has an orthogonal basis consisting of eigenvectors of U.*

Proof. Let $\xi_1 \neq 0$ be an eigenvector of U. It is immediately verified that the subspace of E orthogonal to ξ_1 is mapped into itself by U, using the relation $U^* = U^{-1}$, because if η is perpendicular to ξ_1, then

$$\langle U\eta, \xi_1\rangle = \langle \eta, U^*\xi_1\rangle = \langle \eta, U^{-1}\xi_1\rangle = \langle \eta, \lambda^{-1}\xi_1\rangle = 0.$$

Thus we can finish the proof by induction as before.

Remark. If λ is an eigenvalue of the unitary map U, then λ has necessarily absolute value 1 (because U preserves length), whence λ can be written in the form $e^{i\theta}$ with θ real, and we may view U as a rotation.

Let $A : E \to E$ be an invertible linear map. Just as one writes a non-zero complex number $z = re^{i\theta}$ with $r > 0$, there exists a decomposition of A as a product called its polar decomposition. Let $P : E \to E$ be linear. We say that P is **semipositive** if P is hermitian and we have $\langle Px, x\rangle \geqq 0$ for all $x \in E$. If we have $\langle Px, x\rangle > 0$ for all $x \neq 0$ in E then we say that P is **positive definite**. For example, if we let $P = A^*A$ then we see that P is positive definite, because

$$\langle A^*Ax, x\rangle = \langle Ax, Ax\rangle > 0 \text{ if } x \neq 0.$$

Proposition 6.8. *Let P be semipositive. Then P has a unique semipositive square root $B : E \to E$, i.e. a semipositive linear map such that $B^2 = P$.*

Proof. For simplicity, we assume that P is positive definite. By the spectral theorem, there exists a basis of E consisting of eigenvectors. The eigenvalues must be > 0 (immediate from the condition of positivity). The linear map defined by sending each eigenvector to its multiple by the square root of the corresponding eigenvalue satisfies the required conditions. As for uniqueness, since B commutes with P because $B^2 = P$, it follows that if $\{v_1, \ldots, v_n\}$ is a basis consisting of eigenvectors for P, then each v_i is also an eigenvector for B. (Cf. Chapter XIV, Exercises 12 and 13(d).) Since a positive number has a unique positive square root, it follows that B is uniquely determined as the unique linear map whose effect on v_i is multiplication by the square root of the corresponding eigenvalue for P.

Theorem 6.9. *Let $A : E \to E$ be an invertible linear map. Then A can be written in a unique way as a product $A = UP$, where U is unitary and P is positive definite.*

Proof. Let $P = (A^*A)^{1/2}$, and let $U = AP^{-1}$. Using the defiitions, it is immediately verified that U is unitary, so we get the existence of the decomposition. As for uniqueness, suppose $A = U_1P_1$. Let

$$U_2 = PP_1^{-1} = U^{-1}U_1.$$

Then U_2 is unitary, so $U_2^*U_2 = I$. From the fact that $P^* = P$ and $P_1^* = P_1$, we conclude that $P^2 = P_1^2$. Since P, P_1 are Hermitian positive definite, it follows as in Proposition 6.8 that $P = P_1$, thus proving the theorem.

Remark. The arguments used to prove Theorem 6.9 apply in the case of Hilbert space in analysis. *Cf.* my *Real Analysis*. However, for the uniqueness, since there may not be "eigenvalues", one has to use another technique from analysis, described in that book.

As a matter of terminology, the expression $A = UP$ in Theorem 6.9 is called the **polar decomposition** of A. Of course, it does matter in what order we write the decomposition. There is also a unique decomposition $A = P_1U_1$ with P_1 positive definite and U_1 unitary (apply Theorem 6.9 to A^{-1}, and then take inverses).

§7. THE SPECTRAL THEOREM (SYMMETRIC CASE)

Let E be a finite dimensional vector space over the real numbers, and let g be a symmetric positive definite form on E. If $A : E \to E$ is a linear map, then we know that its transpose, relative to g, is defined by the condition

$$\langle Ax, y\rangle = \langle x, {}^tAy\rangle$$

for all $x, y \in E$. We say that A is **symmetric** if $A = {}^tA$. As before, an element $\xi \in E$ is called an eigenvector of A if there exists $\lambda \in R$ such that $A\xi = \lambda\xi$, and λ is called an eigenvalue if $\xi \neq 0$.

Theorem 7.1. (Spectral Theorem, Symmetric Case). *Let $E \neq 0$. Let $A : E \to E$ be a symmetric linear map. Then E has an orthogonal basis consisting of eigenvectors of A.*

Proof. If we select an orthogonal basis for the positive definite form, then the matrix of A with respect to this basis is a real symmetric matrix, and we are reduced to considering the case when $E = \mathbf{R}^n$. Let M be the matrix representing A. We may view M as operating on $\mathbf{C}^n$, and then M represents a hermitian linear map. Let $z \neq 0$ be a complex eigenvector for M, and write

$$z = x + iy,$$

with $x, y \in \mathbf{R}^n$. By Proposition 6.2, we know that an eigenvalue λ for M, belonging to z, is real, and we have $Mz = \lambda z$. Hence $Mx = \lambda x$ and $My = \lambda y$. But we must have $x \neq 0$ or $y \neq 0$. Thus we have found a nonzero eigenvector for M, namely, A, in E. We can now proceed as before. The orthogonal complement of this eigenvector in E has dimension $(n - 1)$, and is mapped into itself by A. We can therefore finish the proof by induction.

Remarks. The spectral theorems are valid over a real closed field; our proofs don't need any change. Furthermore, the proofs are reasonably close to those which would be given in analysis for Hilbert spaces, and compact operators. The existence of eigenvalues and eigenvectors must however be proved differently, for instance using the Gelfand-Mazur theorem which we have actually proved in Chapter XII, or using a variational principle (i.e. finding a maximum or minimum for the quadratic function depending on the operator).

Corollary 7.2. *Hypotheses being as in the theorem, there exists an orthonormal basis consisting of eigenvectors of A.*

Proof. Divide each vector in an orthogonal basis by its norm.

Corollary 7.3. *Let E be a non-zero finite dimensional vector space over the reals, with a positive definite symmetric form f. Let g be another symmetric form on E. Then there exists a basis of E which is orthogonal for both f and g.*

Proof. We write $f(x, y) = \langle x, y\rangle$. Since f is non-singular, being positive definite, there exists a unique symmetric linear map A such that

$$g(x, y) = \langle Ax, y\rangle$$

for all $x, y \in E$. We apply the theorem to A, and find a basis as in the theorem. It is clearly an orthogonal basis for g (cf. the same proof in the hermitian case).

The analogues of Proposition 6.8 and the polar decomposition also hold in the present case, with the same proofs. See Exercise 9.

§8. ALTERNATING FORMS

Let E be a vector space over the field k, on which we now make no restriction. We let f be an alternating form on E, i.e. a bilinear map $f : E \times E \to k$ such that $f(x, x) = x^2 = 0$ for all $x \in E$. Then

$$x \cdot y = -y \cdot x$$

for all $x, y \in E$, as one sees by substituting $(x + y)$ for x in $x^2 = 0$.

We define a **hyperbolic plane** (for the alternating form) to be a 2-dimensional space which is non-degenerate. We get automatically an element w such that $w^2 = 0$, $w \neq 0$. If P is a hyperbolic plane, and $w \in P$, $w \neq 0$, then there exists an element $y \neq 0$ in P such that $w \cdot y \neq 0$. After dividing y by some constant, we may assume that $w \cdot y = 1$. Then $y \cdot w = -1$. Hence the matrix of the form with respect to the basis $\{w, y\}$ is

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

The pair w, y is called a **hyperbolic pair** as before. Given a 2-dimensional vector space over k with a bilinear form, and a pair of elements $\{w, y\}$ satisfying the relations

$$w^2 = y^2 = 0, \qquad y \cdot w = -1, \qquad w \cdot y = 1,$$

then we see that the form is alternating, and that (w, y) is a hyperbolic plane for the form.

Given an alternating form f on E, we say that E (or f) is **hyperbolic** if E is an orthogonal sum of hyperbolic planes. We say that E (or f) is **null** if $x \cdot y = 0$ for all $x, y \in E$.

Theorem 8.1. *Let f be an alternating form on the finite dimensional vector space E over k. Then E is an orthogonal sum of its kernel and a hyperbolic subspace. If E is non-degenerate, then E is a hyperbolic space, and its dimension is even.*

Proof. A complementary subspace to the kernel is non-degenerate, and hence we may assume that E is non-degenerate. Let $w \in E$, $w \neq 0$. There exists $y \in E$ such that $w \cdot y \neq 0$ and $y \neq 0$. Then (w, y) is non-degenerate, hence is a hyperbolic plane P. We have $E = P \oplus P^{\perp}$ and $P^{\perp}$ is non-degenerate. We complete the proof by induction.

Corollary 8.2. *All alternating non-degenerate forms of a given dimension over a field k are isometric.*

We see from Theorem 8.1 that there exists a basis of E such that relative to this basis, the matrix of the alternating form is

$$\begin{pmatrix} 0 & 1 & & & & & & & & \\ -1 & 0 & & & & & & & & \\ & & 0 & 1 & & & & & & \\ & & -1 & 0 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & 0 & 1 & & & \\ & & & & & -1 & 0 & & & \\ & & & & & & & 0 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & 0 \end{pmatrix}.$$

For convenience of writing, we reorder the basis elements of our orthogonal sum of hyperbolic planes in such a way that the matrix of the form is

$$\begin{pmatrix} 0 & I_r & 0 \\ -I_r & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

where I_r is the unit $r \times r$ matrix. The matrix

$$\begin{pmatrix} 0 & I_r \\ -I_r & 0 \end{pmatrix}$$

is called the **standard alternating** matrix.

Corollary 8.3. *Let E be a finite dimensional vector space over k, with a non-degenerate symmetric form denoted by $\langle\ ,\ \rangle$. Let Ω be a non-degenerate alternating form on E. Then there exists a direct sum decomposition $E = E_1 \oplus E_2$ and a symmetric automorphism A of E (with respect to $\langle\ ,\ \rangle$) having the following property. If $x, y \in E$ are written*

$$x = (x_1, x_2) \quad \text{with} \quad x_1 \in E_1 \quad \text{and} \quad x_2 \in E_2,$$

$$y = (y_1, y_2) \quad \text{with} \quad y_1 \in E_1 \quad \text{and} \quad y_2 \in E_2,$$

then

$$\Omega(x, y) = \langle Ax_1, y_2\rangle - \langle Ax_2, y_1\rangle.$$

Proof. Take a basis of E such that the matrix of Ω with respect to this basis is the standard alternating matrix. Let f be the symmetric non-degenerate form on E given by the dot product with respect to this basis. Then we obtain a direct sum decomposition of E into subspaces E_1, E_2 (corresponding to the first n, resp. the last n coordinates), such that

$$\Omega(x, y) = f(x_1, y_2) - f(x_2, y_1).$$

Since $\langle\ ,\ \rangle$ is assumed non-degenerate, we can find an automorphism A having the desired effect, and A is symmetric because f is symmetric.

§9. THE PFAFFIAN

An alternating matrix is a matrix G such that ${}^tG = -G$ and the diagonal elements are equal to 0. As we saw in Chapter XIII, §6, it is the matrix of an alternating form. We let G be an $n \times n$ matrix, and assume n is *even.* (For odd n, cf. exercises.)

We start over a field of characteristic 0. By Corollary 8.2, there exists a non-singular matrix C such that ${}^tCGC$ is the matrix

$$\begin{pmatrix} 0 & I_r & 0 \\ -I_r & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and hence

$$\det(C)^2 \det(G) = 1 \quad \text{or} \quad 0$$

according as the kernel of the alternating form is trivial or non-trivial. Thus in any case, we see that $\det(G)$ is a square in the field.

Now we move over to the integers $\mathbf{Z}$. Let t_{ij} $(1 \leqq i < j \leqq n)$ be $n(n-1)/2$ algebraically independent elements over $\mathbf{Q}$, let $t_{ii} = 0$ for $i = 1, \ldots, n$, and let $t_{ij} = -t_{ji}$ for $i > j$. Then the matrix $T = (t_{ij})$ is alternating, and hence $\det(T)$ is a square in the field $\mathbf{Q}(t)$ obtained from $\mathbf{Q}$ by adjoining all the variables t_{ij}. However, $\det(T)$ is a polynomial in $\mathbf{Z}[t]$, and since we have unique factorization in $\mathbf{Z}[t]$, it follows that $\det(T)$ is the square of a polynomial in $\mathbf{Z}[t]$. We can write

$$\det(T) = P(t)^2.$$

The polynomial P is uniquely determined up to a factor of ± 1. If we substitute values for the t_{ij} so that the matrix T specializes to

$$\begin{pmatrix} 0 & I_{n/2} \\ -I_{n/2} & 0 \end{pmatrix},$$

then we see that there exists a unique polynomial P with integer coefficients taking the value 1 for this specialized set of values of (t). We call P the **generic Pfaffian** of size n, and write it Pf.

Let R be a commutative ring. We have a homomorphism

$$\mathbf{Z}[t] \to R[t]$$

induced by the unique homomorphism of $\mathbf{Z}$ into R. The image of the generic Pfaffian of size n in $R[t]$ is a polynomial with coefficients in R, which we still denote by Pf. If G is an alternating matrix with coefficients in R, then we write $\mathrm{Pf}(G)$ for the value of $\mathrm{Pf}(t)$ when we substitute g_{ij} for t_{ij} in Pf. Since the determinant commutes with homomorphisms, we have:

Theorem 9.1. *Let R be a commutative ring. Let $(g_{ij}) = G$ be an alternating matrix with $g_{ij} \in R$. Then*

$$\det(G) = (\mathrm{Pf}(G))^2.$$

Furthermore, if C is an $n \times n$ matrix in R, then

$$\mathrm{Pf}(CG{}^tC) = \det(C)\,\mathrm{Pf}(G).$$

Proof. The first statement has been proved above. The second statement will follow if we can prove it over $\mathbf{Z}$. Let u_{ij} $(i, j = 1, \ldots, n)$ be algebraically independent over $\mathbf{Q}$, and such that u_{ij}, t_{ij} are algebraically independent over $\mathbf{Q}$. Let U be the matrix (u_{ij}). Then

$$\mathrm{Pf}(UT{}^tU) = \pm\det(U)\,\mathrm{Pf}(T),$$

as follows immediately from taking the square of both sides. Substitute values for U and T such that U becomes the unit matrix and T becomes the standard alternating matrix. We conclude that we must have a $+$ sign on the right-hand side. Our assertion now follows as usual for any substitution of U to a matrix in R, and any substitution of T to an alternating matrix in R, as was to be shown.

§10. WITT'S THEOREM

We go back to symmetric forms and we let k be a field of characteristic $\neq 2$.

Let E be a vector space over k, with a symmetric form. We say that E is a **hyperbolic plane** if the form is non-degenerate, if E has dimension 2, and if there exists an element $w \neq 0$ in E such that $w^2 = 0$. We say that E is a **hyperbolic space** if it is an orthogonal sum of hyperbolic planes. We also say that the form on E is hyperbolic.

Suppose that E is a hyperbolic plane, with an element $w \neq 0$ such that $w^2 = 0$. Let $u \in E$ be such that $E = (w, u)$. Then $u \cdot w \neq 0$; otherwise w would be a non-zero element in the kernel. Let $b \in k$ be such that $w \cdot bu = bw \cdot u = 1$.

Then select $a \in k$ such that

$$(aw + bu)^2 = 2abw \cdot u + b^2u^2 = 0.$$

(This can be done since we deal with a linear equation in a.) Put $v = aw + bu$. Then we have found a basis for E, namely $E = (w, v)$ such that

$$w^2 = v^2 = 0 \quad \text{and} \quad w \cdot v = 1.$$

Relative to this basis, the matrix of our form is therefore

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

We observe that, conversely, a space E having a basis $\{w, v\}$ satisfying $w^2 = v^2 = 0$ and $w \cdot v = 1$ is non-degenerate, and thus is a hyperbolic plane. A basis $\{w, v\}$ satisfying these relations will be called a **hyperbolic pair**.

An orthogonal sum of non-degenerate spaces is non-degenerate and hence a hyperbolic space is non-degenerate. We note that a hyperbolic space always has even dimension.

Lemma 10.1. *Let E be a finite dimensional vector space over k, with a non-degenerate symmetric form g. Let F be a subspace, F_0 the kernel of F, and suppose we have an orthogonal decomposition*

$$F = F_0 \perp U.$$

Let $\{w_1, \ldots, w_s\}$ be a basis of F_0. Then there exist elements $v_1, \ldots, v_s$ in E perpendicular to U, such that each pair $\{w_i, v_i\}$ is a hyperbolic pair generating a hyperbolic plane P_i, and such that we have an orthogonal decomposition

$$U \perp P_1 \perp \cdots \perp P_s.$$

Proof. Let

$$U_1 = (w_2, \ldots, w_s) \oplus U.$$

Then U_1 is contained in $F_0 \oplus U$ properly, and consequently $(F_0 \oplus U)^\perp$ is contained in $U_1^\perp$ properly. Hence there exists an element $u_1 \in U_1^\perp$ but

$$u_1 \notin (F_0 \oplus U)^\perp.$$

We have $w_1 \cdot u_1 \neq 0$, and hence (w_1, u_1) is a hyperbolic plane P_1. We have seen previously that we can find $v_1 \in P_1$ such that $\{w_1, v_1\}$ is a hyperbolic pair. Furthermore, we obtain an orthogonal sum decomposition

$$F_1 = (w_2, \ldots, w_s) \perp P_1 \perp U.$$

Then it is clear that $(w_2, \ldots, w_s)$ is the kernel of F_1, and we can complete the proof by induction.

Theorem 10.2 *Let E be a finite dimensional vector space over k, and let g be a non-degenerate symmetric form on E. Let F, F' be subspaces of E, and let $\sigma : F \to F'$ be an isometry. Then σ can be extended to an isometry of E onto itself.*

Proof. We shall first reduce the proof to the case when F is non-degenerate.

We can write $F = F_0 \perp U$ as in the lemma of the preceding section, and then $\sigma F = F' = \sigma F_0 \perp \sigma U$. Furthermore, $\sigma F_0 = F'_0$ is the kernel of F'. Now we can enlarge both F and F' as in the lemma to orthogonal sums

$$U \perp P_1 \perp \cdots \perp P_s \quad \text{and} \quad \sigma U \perp P'_1 \perp \cdots \perp P'_s$$

corresponding to a choice of basis in F_0 and its corresponding image in F'_0. Thus we can extend σ to an isometry of these extended spaces, which are non-degenerate. This gives us the desired reduction.

We assume that F, F' are non-degenerate, and proceed stepwise.

Suppose first that $F' = F$, i.e. that σ is an isometry of F onto itself. We can extend σ to E simply by leaving every element of $F^\perp$ fixed.

Next, assume that $\dim F = \dim F' = 1$ and that $F \neq F'$. Say $F = (v)$ and $F' = (v')$. Then $v^2 = v'^2$. Furthermore, (v, v') has dimension 2.

If (v, v') is non-degenerate, it has an isometry extending σ, which maps v on v' and v' on v. We can apply the preceding step to conclude the proof.

If (v, v') is degenerate, its kernel has dimension 1. Let w be a basis for this kernel. There exist $a, b \in k$ such that $v' = av + bw$. Then $v'^2 = a^2v^2$ and hence $a = \pm 1$. Replacing v' by $-v'$ if necessary, we may assume $a = 1$. Replacing w by bw, we may assume $v' = v + w$. Let $z = v + v'$. We apply Lemma 10.1 to the space

$$(w, z) = (w) \perp (z).$$

We can find an element $y \in E$ such that

$$y \cdot z = 0, \qquad y^2 = 0, \quad \text{and} \quad w \cdot y = 1.$$

The space $(z, w, y) = (z) \perp (w, y)$ is non-degenerate, being an orthogonal sum of (z) and the hyperbolic plane (w, y). It has an isometry such that

$$z \leftrightarrow z, \qquad w \leftrightarrow -w, \qquad y \leftrightarrow -y.$$

But $v = \frac{1}{2}(z - w)$ is mapped on $v' = \frac{1}{2}(z + w)$ by this isometry. We have settled the present case.

We finish the proof by induction. By the existence of an orthogonal basis (Theorem 3.1), every subspace F of dimension > 1 has an orthogonal decomposition into a sum of subspaces of smaller dimension. Let $F = F_1 \perp F_2$ with dim F_1 and dim $F_2 \geqq 1$. Then

$$\sigma F = \sigma F_1 \perp \sigma F_2.$$

Let $\sigma_1 = \sigma | F_1$ be the restriction of σ to F_1. By induction, we can extend σ_1 to an isometry

$$\bar{\sigma}_1 : E \to E.$$

Then $\bar{\sigma}_1(F_1^\perp) = (\sigma_1 F_1)^\perp$. Since σF_2 is perpendicular to $\sigma F_1 = \sigma_1 F_1$, it follows that σF_2 is contained in $\bar{\sigma}_1(\mathrm{F}_1^\perp)$. Let $\sigma_2 = \sigma | F_2$. Then the isometry

$$\sigma_2 : F_2 \to \sigma_2 F_2 = \sigma F_2$$

extends by induction to an isometry

$$\bar{\sigma}_2 : F_1^\perp \to \bar{\sigma}_1(F_1^\perp).$$

The pair $(\sigma_1, \bar{\sigma}_2)$ gives us an isometry of $F_1 \perp F_1^\perp = E$ onto itself, as desired.

Corollary 10.3. *Let E, E' be finite dimensional vector spaces with non-degenerate symmetric forms, and assume that they are isometric. Let F, F' be subspaces, and let $\sigma : F \to F'$ be an isometry. Then σ can be extended to an isometry of E onto E'.*

Proof. Clear.

Let E be a space with a symmetric form g, and let F be a null subspace. Then by Lemma 10.1, we can embed F in a hyperbolic subspace H whose dimension is 2 dim F.

As applications of Theorem 10.2, we get several corollaries.

Corollary 10.4. *Let E be a finite dimensional vector space with a non-degenerate symmetric form. Let W be a maximal null subspace, and let W' be some null subspace. Then* dim $W' \leqq$ dim W, *and W' is contained in some maximal null subspace, whose dimension is the same as* dim W.

Proof. That W' is contained in a maximal null subspace follows by Zorn's lemma. Suppose $\dim W' \geqq \dim W$. We have an isometry of W onto a subspace of W' which we can extend to an isometry of E onto itself. Then $\sigma^{-1}(W')$ is a null subspace containing W, hence is equal to W, whence $\dim W = \dim W'$. Our assertions follow by symmetry.

Let E be a vector space with a non-degenerate symmetric form. Let W be a null subspace. By Lemma 10.1 we can embed W in a hyperbolic subspace H of E such that W is the maximal null subspace of H, and H is non-degenerate. Any such H will be called a **hyperbolic enlargement** of W.

Corollary 10.5. *Let E be a finite dimensional vector space with a non-degenerate symmetric form. Let W and W' be maximal null subspaces. Let H, H' be hyperbolic enlargements of W, W' respectively. Then H, H' are isometric and so are $H^{\perp}$ and $H'^{\perp}$.*

Proof. We have obviously an isometry of H on H', which can be extended to an isometry of E onto itself. This isometry maps $H^{\perp}$ on $H'^{\perp}$, as desired.

Corollary 10.6. *Let g_1, g_2, h be symmetric forms on finite dimensional vector spaces over the field of k. If $g_1 \oplus h$ is isometric to $g_2 \oplus h$, and if g_1, g_2 are non-degenerate, then g_1 is isometric to g_2.*

Proof. Let g_1 be a form on E_1 and g_2 a form on E_2. Let h be a form on F. Then we have an isometry between $F \oplus E_1$ and $F \oplus E_2$. Extend the identity $\mathrm{id}: F \to F$ to an isometry σ of $F \oplus E_1$ to $F \oplus E_2$ by Corollary 10.3. Since E_1 and E_2 are the respective orthogonal complements of F in their two spaces, we must have $\sigma(E_1) = E_2$, which proves what we wanted.

If g is a symmetric form on E, we shall say that g is **definite** if $g(x, x) \neq 0$ for any $x \in E$, $x \neq 0$ (i.e. $x^2 \neq 0$ if $x \neq 0$).

Corollary 10.7. *Let g be a symmetric form on E. Then g has a decomposition as an orthogonal sum*

$$g = g_0 \oplus g_{\mathrm{hyp}} \oplus g_{\mathrm{def}}$$

where g_0 is a null form, g_{hyp} is hyperbolic, and g_{def} is definite. The form $g_{\mathrm{hyp}} \oplus g_{\mathrm{def}}$ is non-degenerate. The forms g_0, g_{hyp}, and g_{def} are uniquely determined up to isometries.

Proof. The decomposition $g = g_0 \oplus g_1$ where g_0 is a null form and g_1 is non-degenerate is unique up to an isometry, since g_0 corresponds to the kernel of g.

We may therefore assume that g is non-degenerate. If

$$g = g_h \oplus g_d$$

where g_h is hyperbolic and g_d is definite, then g_h corresponds to the hyperbolic enlargement of a maximal null subspace, and by Corollary 10.5 it follows that g_h is uniquely determined. Hence g_d is uniquely determined as the orthogonal complement of g_h. (By uniquely determined, we mean of course up to an isometry.)

We shall abbreviate g_{hyp} by g_h and g_{def} by g_d.

§11. THE WITT GROUP

Let g, φ by symmetric forms on finite dimensional vector spaces over k. We shall say that they are **equivalent** if g_d is isometric to φ_d. The reader will verify at once that this is an equivalence relation. Furthermore the (orthogonal) sum of two null forms is a null form, and the sum of two hyperbolic forms is hyperbolic. However, the sum of two definite forms need not be definite. We write our equivalence $g \sim \varphi$. Equivalence is preserved under orthogonal sums, and hence equivalence classes of symmetric forms constitute a monoid.

Theorem 11.1. *The monoid of equivalence classes of symmetric forms* (*over the field* k) *is a group.*

Proof. We have to show that every element has an additive inverse. Let g be a symmetric form, which we may assume definite. We let $-g$ be the form such that $(-g)(x, y) = -g(x, y)$. We contend that $g \oplus -g$ is equivalent to 0. Let E be the space on which g is defined. Then $g \oplus -g$ is defined on $E \oplus E$. Let W be the subspace consisting of all pairs (x, x) with $x \in E$. Then W is a null space for $g \oplus -g$. Since $\dim(E \oplus E) = 2 \dim W$, it follows that W is a maximal null space, and that $g \oplus -g$ is hyperbolic, as was to be shown.

The group of Theorem 11.1 will be called the **Witt group** of k, and will be denoted by $W(k)$. It is of importance in the study of representations of elements of k by the quadratic form f arising from g [i.e. $f(x) = g(x, x)$], for instance when one wants to classify the definite forms f.

We shall now define another group, which is of importance in more functorial studies of symmetric forms, for instance in studying the quadratic forms arising from manifolds in topology.

We observe that isometry classes of non-degenerate symmetric forms (over k) constitute a monoid $M(k)$, the law of composition being the orthogonal sum. Furthermore, the cancellation law holds (Corollary 10.6). We let

$$\mathrm{cl} : M(k) \to WG(k)$$

be the canonical map of $M(k)$ into the Grothendieck group of this monoid, which we shall call the **Witt-Grothendieck** group over k. As we know, the cancellation law implies that cl is injective.

If g is a symmetric non-degenerate form over k, we define its dimension dim g to be the dimension of the space E on which it is defined. Then it is clear that

$$\dim(g \oplus g') = \dim g + \dim g'.$$

Hence dim factors through a homomorphism

$$\dim : WG(k) \to \mathbf{Z}.$$

This homomorphism splits since we have a non-degenerate symmetric form of dimension 1.

Let $WG_0(k)$ be the kernel of our homomorphism dim. If g is a symmetric non-degenerate form we can define its determinant $\det(g)$ to be the determinant of a matrix G representing g relative to a basis, modulo squares. This is well defined as an element of k^*/k^{*2}. We define det of the 0-form to be 1. Then det is a homomorphism

$$\det : M(k) \to k^*/k^{*2},$$

and can therefore be factored through a homomorphism, again denoted by det, of the Witt-Grothendieck group, $\det : WG(k) \to k^*/k^{*2}$.

Other properties of the Witt-Grothendieck group will be given in the exercises.

EXERCISES

1. (a) Let E be a finite dimensional space over the complex numbers, and let

$$h : E \times E \to \mathbf{C}$$

be a hermitian form. Write

$$h(x, y) = g(x, y) + if(x, y)$$

where g, f are real valued. Show that g, f are $\mathbf{R}$-bilinear, g is symmetric, f is alternating.

(b) Let E be finite dimensional over $\mathbf{C}$. Let $g : E \times E \to \mathbf{C}$ be $\mathbf{R}$-bilinear. Assume that for all $x \in E$, the map $y \mapsto g(x, y)$ is $\mathbf{C}$-linear, and that the R-bilinear form

$$f(x, y) = g(x, y) - g(y, x)$$

is real-valued on $E \times E$. Show that there exists a hermitian form h on E and a symmetric **C**-bilinear form ψ on E such that $2ig = h + \psi$. Show that h and ψ are uniquely determined.

2. Prove the real case of the unitary spectral theorem: If E is a non-zero finite dimensional space over **R**, with a positive definite symmetric form, and $U : E \to E$ is a unitary linear map, then E has an orthogonal decomposition into subspaces of dimension 1 or 2, invariant under U. If $\dim E = 2$, then the matrix of U with respect to any orthonormal basis is of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

depending on whether $\det(U) = 1$ or -1. Thus U is a rotation, or a rotation followed by a reflection.

3. Let E be a finite-dimensional, non-zero vector space over the reals, with a positive definite scalar product. Let $T : E \to E$ be a unitary automorphism of E. Show that E is an orthogonal sum of subspaces

$$E = E_1 \perp \cdots \perp E_m$$

such that each E_i is T-invariant, and has dimension 1 or 2. If E has dimension 2, show that one can find a basis such that the matrix associated with T with respect to this basis is

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

according as $\det T = 1$ or $\det T = -1$.

4. Let E be a finite dimensional non-zero vector space over **C**, with a positive definite hermitian product. Let $A, B : E \to E$ be a hermitian endomorphism. Assume that $AB = BA$. Prove that there exists a basis of E consisting of common eigenvectors for A and B.

5. Let E be a finite-dimensional space over the complex, with a positive definite hermitian form. Let S be a set of (**C**-linear) endomorphisms of E having no invariant subspace except 0 and E. (This means that if F is a subspace of E and $BF \subset F$ for all $B \in S$, then $F = 0$ or $F = E$.) Let A be a hermitian map of E into itself such that $AB = BA$ for all $B \in S$. Show that $A = \lambda I$ for some real number λ. [*Hint*: Show that there exists exactly one eigenvalue of A. If there were two eigenvalues, say $\lambda_1 \neq \lambda_2$, one could find two polynomials f and g with real coefficients such that $f(A) \neq 0$, $g(A) \neq 0$ but $f(A)g(A) = 0$. Let F be the kernel of $g(A)$ and get a contradiction.]

6. Let E be as in Exercise 5. Let T be a **C**-linear map of E into itself. Let

$$A = \tfrac{1}{2}(T + T^*).$$

Show that A is hermitian. Show that T can be written in the form $A + iB$ where A, B are hermitian, and are uniquely determined.

7. Let S be a commutative set of **C**-linear endomorphisms of E having no invariant subspace unequal to 0 or E. Assume in addition that if $B \in S$, then $B^* \in S$. Show that each

element of S is of type αI for some complex number α. [*Hint*: Let $B_0 \in S$. Let

$$A = \tfrac{1}{2}(B_0 + B_0^*).$$

Show that $A = \lambda I$ for some real λ.]

8. An endomorphism B of E is said to be **normal** if B commutes with B^*. State and prove a spectral theorem for normal endomorphisms.

Symmetric endomorphisms

For Exercises 9, 10 and 11 we let E be a non-zero finite dimensional vector space over **R**, *with a symmetric positive definite scalar product g, which gives rise to a norm* $|\ |$ *on E.*

Let $A : E \to E$ be a symmetric endomorphism of E with respect to g. Define $A \geqq 0$ to mean $\langle Ax, x\rangle \geqq 0$ for all $x \in E$.

9. (a) Show that $A \geqq 0$ if and only if all eigenvalues of A belonging to non-zero eigenvectors are $\geqq 0$. Both in the hermitian case and the symmetric case, one says that A is **semipositive** if $A \geqq 0$, and **positive definite** if $\langle Ax, x\rangle > 0$ for all $x \neq 0$.

   (b) Show that an automorphism A of E can be written in a unique way as a product $A = UP$ where U is real unitary (that is, ${}^tUU = I$), and P is symmetric positive definite. For two hermitian or symmetric endomorphisms A, B, define $A \geqq B$ to mean $A - B \geqq 0$, and similarly for $A > B$. Suppose $A > 0$. Show that there are two real numbers $\alpha > 0$ and $\beta > 0$ such that $\alpha I \leqq A \leqq \beta I$.

10. If A is an endomorphism of E, define its norm $|A|$ to be the greatest lower bound of all numbers C such that $|Ax| \leqq C|x|$ for all $x \in E$.

    (a) Show that this norm satisfies the triangle inequality.

    (b) Show that the series

$$\exp(A) = I + A + \frac{A^2}{2!} + \cdots$$

converges, and if A commutes with B, then $\exp(A + B) = \exp(A)\exp(B)$. If A is sufficiently close to I, show that the series

$$\log(A) = \frac{(A - I)}{1} - \frac{(A - I)^2}{2} + \cdots$$

converges, and if A commutes with B, then

$$\log(AB) = \log A + \log B.$$

(c) Using the spectral theorem, show how to define $\log P$ for arbitrary positive definite endomorphisms P.

11. Again, let E be non-zero finite dimensional over **R**, and with a positive definite symmetric form. Let $A : E \to E$ be a linear map. Prove:

    (a) If A is symmetric (resp. alternating), then $\exp(A)$ is symmetric positive definite (resp. real unitary).

    (b) If A is a linear automorphism of E sufficiently close to I, and is symmetric positive definite (resp. real unitary), then log A is symmetric (resp. alternating).

(c) More generally, if A is positive definite, then log A is symmetric.

12. Let R be a commutative ring, let E, F be R-modules, and let $f: E \to F$ be a mapping. Assume that multiplication by 2 in F is an invertible map. Show that f is homogeneous quadratic if and only if f satisfies the **parallelogram law**:

$$f(x + y) + f(x - y) = 2f(x) + 2f(y)$$

for all $x, y \in E$.

13. (Tate) Let E, F be complete normed vector spaces over the real numbers. Let $f: E \to F$ be a map having the following property. There exists a number $C > 0$ such that for all $x, y \in E$ we have

$$|f(x + y) - f(x) - f(y)| \leqq C.$$

Show that there exists a unique additive map $g : E \to F$ such that $|g - f|$ is bounded (i.e. $|g(x) - f(x)|$ is bounded as a function of x). Generalize to the bilinear case. [*Hint*: Let

$$g(x) = \lim_{n \to \infty} \frac{f(2^n x)}{2^n}.]$$

14. (Tate) Let S be a set and $f: S \to S$ a map of S into itself. Let $h: S \to \mathbf{R}$ be a real valued function. Assume that there exists a real number $d > 1$ such that $h \circ f - df$ is bounded. Show that there exists a unique function h_f such that $h_f - h$ is bounded, and $h_f \circ f = dh_f$. [*Hint*: Let $h_f(x) = \lim h(f^n(x))/d^n$.]

15. Define maps of degree > 2, from one module into another. [*Hint*: For degree 3, consider the expression

$$f(x + y + z) - f(x + y) - f(x + z) - f(y + z) + f(x) + f(y) + f(z).]$$

Generalize the statement proved for quadratic maps to these higher-degree maps, i.e. the uniqueness of the various multilinear maps entering into their definitions.

Alternating forms

16. Let E be a vector space over a field k and let g be a bilinear form on E. Assume that whenever $x, y \in E$ are such that $g(x, y) = 0$, then $g(y, x) = 0$. Show that g is symmetric or alternating.

17. Let E be a module over $\mathbf{Z}$. Assume that E is free, of dimension $n \geqq 1$, and let f be a bilinear alternating form on E. Show that there exists a basis $\{e_i\}$ $(i = 1, \dots, n)$ and an integer r such that $2r \leqq n$,

$$e_1 \cdot e_2 = a_1, \qquad e_3 \cdot e_4 = a_2, \ \dots, e_{2r-1} \cdot e_{2r} = a_r$$

where $a_1, \dots, a_r \in \mathbf{Z}$, $a_i \neq 0$, and a_i divides a_{i+1} for $i = 1, \dots, r - 1$ and finally $e_i \cdot e_j = 0$ for all other pairs of indices $i \leqq j$. Show that the ideals $\mathbf{Z}a_i$ are uniquely determined. [*Hint*: Consider the injective homomorphism $\varphi_f : E \to E^\vee$ of E into the dual space over $\mathbf{Z}$, viewing $\varphi_f(E)$ as a free submodule of $E^\vee$.]. Generalize to principal rings when you know the basis theorem for modules over these rings.

Remark. A basis as in Exercise 18 is called a **symplectic basis**. For one use of such a basis, see the theory of theta functions, as in my *Introduction to Algebraic and Abelian Functions* (Second Edition, Springer Verlag), Chapter VI, §3.

18. Let E be a finite-dimensional vector space over the reals, and let $\langle\ ,\ \rangle$ be a symmetric positive definite form. Let Ω be a non-degenerate alternating form on E. Show that there exists a direct sum decomposition

$$E = E_1 \oplus E_2$$

having the following property. If $x, y \in E$ are written

$$x = (x_1, x_2) \quad \text{with} \quad x_1 \in E_1 \quad \text{and} \quad x_2 \in E_2,$$

$$y = (y_1, y_2) \quad \text{with} \quad y_1 \in E_1 \quad \text{and} \quad y_2 \in E_2,$$

then $\Omega(x, y) = \langle x_1, y_2\rangle - \langle x_2, y_1\rangle$. [*Hint*: Use Corollary 8.3, show that A is positive definite, and take its square root to transform the direct sum decomposition obtained in that corollary.]

19. Show that the pfaffian of an alternating $n \times n$ matrix is 0 when n is odd.

20. Prove all the properties for the pfaffian stated in Artin's *Geometric Algebra* (*Interscience*, 1957), p. 142.

The Witt group

21. Show explicitly how $W(k)$ is a homomorphic image of $WG(k)$.

22. Show that $WG(k)$ can be expressed as a homomorphic image of $\mathbf{Z}[k^*/k^{*2}]$ [*Hint*: Use the existence of orthogonal bases.]

23. Witt's theorem is still true for alternating forms. Prove it or look it up in Artin (ref. in Exercise 20).

$SL_n(\mathbf{R})$

There is a whole area of linear algebraic groups, giving rise to an extensive algebraic theory as well as the possibility of doing Fourier analysis on such groups. The group $SL_n(\mathbf{R})$ (or $SL_n(\mathbf{C})$) can serve as a prototype, and a number of basic facts can be easily verified. Some of them are listed below as exercises. Readers wanting to see solutions can look them up in [JoL 01], *Spherical Inversion on* $SL_n(\mathbf{R})$, Chapter I.

24. **Iwasawa decomposition**. We start with $GL_n(\mathbf{R})$. Let:

$G = GL_n(\mathbf{R})$;

$K =$ subgroup of real unitary $n \times n$ matrices;

$U =$ group of real unipotent upper triangular matrices, that is having components 1 on the diagonal, arbitrary above the diagonal, and 0 below the diagonal;

A = group of diagonal matrices with positive diagonal components.

Prove that the product map $U \times A \times K \to UAK \subset G$ is actually a bijection. This amounts to Gram–Schmidt orthogonalization. Prove the similar statement in the complex case, that is, for $G(\mathbf{C}) = GL_n(\mathbf{C})$, $K(\mathbf{C})$ = complex unitary group, $U(\mathbf{C})$ = complex unipotent upper triangular group, and A the same group of positive diagonal matrices as in the real case.

25. Let now $G = SL_n(\mathbf{R})$, and let K, A be the corresponding subgroups having determinant 1. Show that the product $U \times A \times K \to UAK$ again gives a bijection with G.

26. Let $\mathfrak{a}$ be the $\mathbf{R}$-vector space of real diagonal matrices with trace 0. Let $\mathfrak{a}^\vee$ be the dual space. Let α_i $(i = 1, \dots, n-1)$ be the functional defined on an element $H = \operatorname{diag}(h_1, \dots, h_n)$ by $\alpha_i(H) = h_i - h_{i+1}$. (a) Show that $\{\alpha_1, \dots, \alpha_{n-1}\}$ is a basis of $\mathfrak{a}^\vee$ over $\mathbf{R}$. (b) Let $H_{i,i+1}$ be the diagonal matrix with $h_i = 1$, $h_{i+1} = -1$, and $h_j = 0$ for $j \neq i, i+1$. Show that $\{H_{1,2}, \dots, H_{n-1,n}\}$ is a basis of $\mathfrak{a}$. (c) Abbreviate $H_{i,i+1} = H_i$ $(i = 1, \dots, n-1)$. Let $\alpha'_i \in \mathfrak{a}^\vee$ be the functional such that $\alpha'_i(H_j) = \delta_{ij}$ $(= 1$ if $i = j$ and 0 otherwise). Thus $\{\alpha'_1, \dots, \alpha'_{n-1}\}$ is the dual basis of $\{H_1, \dots, H_{n-1}\}$. Show that
$$\alpha'_i(H) = h_1 + \cdots + h_i.$$

27. **The trace form.** Let $\operatorname{Mat}_n(\mathbf{R})$ be the vector space of real $n \times n$ matrices. Define the **twisted trace form** on this space by
$$B_t(X, Y) = \operatorname{tr}(X\,{}^tY) = \langle X, Y \rangle_t.$$
As usual, ${}^tY$ is the transpose of a matrix Y. Show that B_t is a symmetric positive definite bilinear form on $\operatorname{Mat}_n(\mathbf{R})$. What is the analogous positive definite hermitian form on $\operatorname{Mat}_n(\mathbf{C})$?

28. **Positivity.** On $\mathfrak{a}$ (real diagonal matrices with trace 0) the form of Exercise 27 can be defined by $\operatorname{tr}(XY)$, since elements $X, Y \in \mathfrak{a}$ are symmetric. Let $\mathscr{A} = \{\alpha_1, \dots, \alpha_{n-1}\}$ denote the basis of Exercise 26. Define an element $H \in \mathfrak{a}$ to be **semipositive** (writen $H \geqq 0$) if $\alpha_i(H) \geqq 0$ for all $i = 1, \dots, n-1$. For each $\alpha \in \mathfrak{a}^\vee$, let $H_\alpha \in \mathfrak{a}$ represent α with respect to B_t, that is $\langle H_\alpha, H \rangle = \alpha(H)$ for all $H \in \mathfrak{a}$. Show that $H \geqq 0$ if and only if
$$H = \sum_{i=1}^{n-1} s_i H_{\alpha'_i} \qquad \text{with } s_i \geqq 0.$$
Similarly, define H to be **positive** and formulate the similar condition with $s_i > 0$.

29. Show that the elements $n\alpha'_i$ $(i = 1, \dots, n-1)$ can be expressed as linear combinations of $\alpha_1, \dots, \alpha_{n-1}$ with positive coefficients in $\mathbf{Z}$.

30. Let W be the group of permutations of the diagonal elements in the vector space $\mathfrak{a}$ of diagonal matrices. Show that $\mathfrak{a}_{\geqq 0}$ is a fundamental domain for the action of W on $\mathfrak{a}$ (i.e., given $H \in \mathfrak{a}$, there exists a unique $H^+ \geqq 0$ such that $H^+ = wH$ for some $w \in W$.