Math 551 – Algebra – Fall 2000

Richard Lyons
Rutgers University
New Brunswick, New Jersey, USA

B. Abelian Groups and Modules over Principal Ideal Domains

3. Bilinear forms; duality.

3a. Definitions

An important function of matrices is that they are the concrete versions not only of linear transformations, but also, in the square case, of bilinear forms. This leads to a rather different equivalence relation on the set of all $n \times n$ matrices over a field F. We briefly indicate the flavor here.

Definition. *Let V be a finite-dimensional vector space over a field F. A bilinear form on V is a function $\mathcal{B} : V \times V \to F$ such that*

$$\mathcal{B}(v, \alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 \mathcal{B}(v, w_1) + \alpha_2 \mathcal{B}(v, w_2) \text{ and}$$
$$\mathcal{B}(\alpha_1 v_1 + \alpha_2 v_2, w) = \alpha_1 \mathcal{B}(v_1, w) + \alpha_2 \mathcal{B}(v_2, w)$$

for all $v, v_1, v_2, w, w_1, w_2 \in V$.

Given a bilinear form $\mathcal{B}$, if we choose a basis $C = \{v_1, \ldots, v_n\}$ of V, we obtain the matrix of $\mathcal{B}$ with respect to the basis, namely

$$[\mathcal{B}]_C = [\mathcal{B}(v_i, v_j)]_{i,j=1}^n .$$

This matrix completely determines $\mathcal{B}$, since $\mathcal{B}(\sum_i \alpha_i v_i, \sum_j \beta_j v_j) = \sum_{i,j} \alpha_i \beta_j (v_i, v_j)$ by the bilinearity of $\mathcal{B}$. Furthermore, starting with any matrix A, one can define

$$\mathcal{B}(\sum_i \alpha_i v_i, \sum_j \beta_j v_j) = \sum_{i,j} \alpha_i \beta_j A_{ij}$$

and obtain a bilinear form with matrix A. Thus as with linear transformations, if we fix a basis C of V, then $\mathcal{B} \mapsto [\mathcal{B}]_C$ is a bijective correspondence between the set of bilinear forms on V and the set of square matrices of size $\dim V$.

However, if we want "equivalence" of matrices then to mean "same bilinear form, different basis", we do not come out with similarity but rather the following equivalence relation on matrices:

$$A \sim A' \iff \exists \text{ an invertible } D \text{ such that } A' = D^T A D,$$

the "T" indicating transpose.

For if $w_j = \sum_i d_{ij} v_i$, $j = 1, \ldots, n$, then

$$(w_i, w_j) = (\sum_k d_{ki} v_k, \sum_\ell d_{\ell j} v_\ell) = \sum_{k,\ell} d_{ki}(v_k, v_\ell) d_{\ell j}.$$

If V' is another vector space over F, and B' is a bilinear form on V', then an isometric isomorphism from V to V' (with respect to $\mathcal{B}$ and $\mathcal{B}'$) is an isomorphism $\phi : V \to V'$ of vector spaces such that $\mathcal{B}'(\phi(v), \phi(w)) = \mathcal{B}(v, w)$ for all $v, w \in V$. This last condition holds as long as it holds for all v, w in a basis of V, because of the bilinearity of $\mathcal{B}$ and $\mathcal{B}'$. Thus:

Proposition. *Let V and V' be vector spaces with bases C and C', and bilinear forms $\mathcal{B}$ and $\mathcal{B}'$, respectively. Then there exists an isometric isomorphism from V to V' (with respect to $\mathcal{B}$ and $\mathcal{B}'$) if and only if there is an invertible matrix D such that*

$$[\mathcal{B}']_{C'} = D^T [\mathcal{B}]_C D.$$

The "canonical form" problem for bilinear forms — i.e., the problem of classifying them — is thus completely different from the classification of linear transformations we have developed in the previous sections.

There is some connection, which can be formulated in terms of the dual space V^*. Let us fix V and a bilinear form $\mathcal{B}$; the ensuing discussion applies to V and $\mathcal{B}$, unless explicitly stated otherwise.

Definition. *If V is a finite-dimensional vector space over F, then $V^* = \mathrm{Hom}_F(V, F)$.*

The elements of $\mathrm{Hom}_F(V, F)$ are called linear functionals on V. When V is infinite-dimensional, V^* can get out of hand and it is normally defined to be a subspace of $\mathrm{Hom}_F(V, F)$, e.g. consisting in the case of Banach spaces, for example, of all bounded linear functionals.

For finite dimensions, however, the fact that direct sums are coproducts means that

$$\underset{F}{\mathrm{Hom}}(V_1 \oplus V_2, F) \cong_\phi \underset{F}{\mathrm{Hom}}(V_1, F) \times \underset{F}{\mathrm{Hom}}(V_2, F)$$

under the mapping $\phi(f) = (f|V_1, f|V_2)$. Moreover, if V is 1-dimensional, and we choose any $0 \neq v_1 \in V$, then $\mathrm{Hom}_F(V, F) \cong F$, via $f \mapsto f(v_1)$. Note however this isomorphism is not "natural" but depends on the choice of v_1. The displayed isomorphism is "natural", however.

Choosing a basis $v_1, \ldots, v_n$ of V, in general, we therefore get an isomorphism

$$V^* = \mathrm{Hom}(V, F) = \mathrm{Hom}(Fv_1 \oplus \cdots \oplus Fv_n, F) \cong \mathrm{Hom}(Fv_1, F) \times \cdots \times \mathrm{Hom}(Fv_n, F) \cong F \oplus \cdots \oplus F \cong V. \blacksquare$$

However the isomorphism we get here depends on the choice of basis. Following the progress of v_i from right to left, we see that under our isomorphism v_i corresponds to the functional v_i^* defined by

$$v_i^*(v_j) = \delta_{ij} \quad (= 1 \text{ or } 0 \text{ according as } i = j \text{ or } i \neq j).$$

The basis $v_1^*, \ldots, v_n^*$ of V^* is called the dual basis to $v_1, \ldots, v_n$. We have

$$v = \sum_j v_j^*(v) v_j$$

for each $v \in V$; this is obvious for $v = v_i$ and then extends to all v by linearity.

Now from the bilinear form $\mathcal{B}$ on V, we obtain for each $v \in V$ the following two elements $\rho_v, \lambda_v \in V^*$:

$$\rho_v(w) = \mathcal{B}(w, v) \text{ and } \lambda_v(w) = \mathcal{B}(v, w).$$

The bilinearity of $\mathcal{B}$ implies not only that these lie in V^*, but also that the mappings

$$\rho = \rho_{\mathcal{B}} : V \to V^* \text{ taking } v \mapsto \rho_v \text{ and } \lambda = \lambda_{\mathcal{B}} : V \to V^* \text{ taking } v \mapsto \lambda_v$$

are linear transformations.

(Conversely a linear transformation $\beta : V \to V^*$ determines a bilinear form such that $\rho = \beta$.)

Then for a basis $C = \{v_1, \ldots, v_n\}$ and its dual basis $C^* = \{v_1^*, \ldots, v_n^*\}$, we have $\mathcal{B}(v_i, v_j) = \lambda_{\mathcal{B}}(v_i)(v_j)$, so

$$\lambda_{\mathcal{B}}(v_i) = \sum \beta_{\mathcal{B}}(v_i)(v_j) v_j^* = \sum \mathcal{B}(v_i, v_j) v_j^*.$$

Thus

$$[\mathcal{B}]_C = [\lambda_{\mathcal{B}}]_{C^*}^C, \text{ and similarly } [\mathcal{B}]_C^T = [\rho_{\mathcal{B}}]_{C^*}^C.$$

This implies:

Lemma. *$\lambda_{\mathcal{B}}$ is an isomorphism if and only if $\rho_{\mathcal{B}}$ is an isomorphism.*

Definition. *$\mathcal{B}$ is nondegenerate if and only if $\lambda_{\mathcal{B}}$ is an isomorphism.*

There are many equivalent formulations of nondegeneracy, because of the lemma and the fact that $\dim V = \dim V^*$. For example: $\ker \lambda_{\mathcal{B}} = 0$. Or $\ker \rho_{\mathcal{B}} = 0$. Or: the only $v \in V$ such that $\mathcal{B}(v, w) = 0$ for all $w \in V$ is $v = 0$, (i.e. $\ker \lambda = 0$). Or: for every $\omega \in V^*$ there is $w \in V$ such that $\omega(v) = \mathcal{B}(v, w)$ for all $v \in V$ (i.e., ρ is surjective). And so on.

Instead of saying that $\mathcal{B}$ is nondegenerate, it is customary to say that V is nondegenerate. Then if W is a subspace of V, the restriction of $\mathcal{B}$ to W is a bilinear form on W, and we say that W is nondegenerate if and only if $\mathcal{B}|W$ is nondegenerate.

3b. Orthogonal complements

For simplicity let us write (v, w) for $\mathcal{B}(v, w)$. In order for orthogonality to be an easy concept, we shall assume from now on that our bilinear forms $\mathcal{B}$ satisfy the condition:

$$\text{if } (v, w) = 0, \text{ then } (w, v) = 0. \qquad 3A$$

This condition is satisfied, for example, if either

a) $\mathcal{B}$ is **symmetric**: $(v, w) = (w, v)$ for all $v, w \in V$, or

b) $\mathcal{B}$ is **alternating**: $(v, v) = 0$ for all $v \in V$; this implies that $(v, w) + (w, v) = (v + w, v + w) - (v, v) - (w, w) = 0$ so $(v, w) = -(w, v)$ for all $v, w \in V$, i.e., $\mathcal{B}$ is antisymmetric. (Conversely, antisymmetry implies the alternating condition as long as $1 \neq -1$, i.e, as long as the underlying field is not $\mathbf{Z}_2$ and is not any other field of "characteristic 2".)

Under the hypothesis $(3A)$, we can then define $v \perp w$, for $v, w \in V$, to mean $(v, w) = 0$; then $\perp$ is a symmetric relation. Then for any subset (usually a subspace) $W \subseteq V$, we define

$$W^\perp = \{v \in V \mid v \perp w \text{ for all } w \in W\}.$$

The bilinearity of $\mathcal{B}$ implies immediately that $W^\perp$ is a subspace of V.

Exercise. *If $W_1 \subseteq W_2$, then $W_2^\perp \subseteq W_1^\perp$.*

Proposition. *Suppose that $\mathcal{B}$ is a symmetric or alternating bilinear form. Then For any subspace $W \subseteq V$,*

a) $W \subseteq W^{\perp\perp}$.

b) $\dim W + \dim W^\perp \geq \dim V$. *Moreover, if V is nondegenerate, then equality holds in (a) and (b).*

Proof. a) is trivial from the definitions.

Restriction to W gives a homomorphism

$$\phi : V^* \to W^*, \quad \omega \mapsto \omega|W,$$

which is surjective; choosing any complement X to W in V (i.e. writing $V = W \oplus X$) we may extend any linear functional $W \to V$ to a linear functional on V by prescribing that X go to 0. (Each complement X gives a different extension of the functional, but we only need one extension anyway.) Let

$$\gamma = \phi \circ \rho_{\mathcal{B}} : V \to V^* \to W^*.$$

Then $v \in \ker \gamma$ if and only if $\rho_v|W = 0$, i.e., $\mathcal{B}(w, v) = 0$ for all $w \in W$, i.e., $v \in W^\perp$. Therefore

$$V/W^\perp \cong \operatorname{im} \gamma \subseteq W^*,$$

and counting dimensions gives (b). Moreover, if V is nondegenerate, then ρ is an isomorphism, and since ϕ is surjective, so is γ. Hence $V/W^\perp \cong W^*$ and equality holds in (b). This in turn implies that $\dim W^{\perp\perp} = \dim V - \dim W^\perp = \dim W$ so equality holds in (a). QED

Corollary. *If $\mathcal{B}$ is a nondegenerate symmetric or alternating bilinear form on the finite-dimensional vector space V, then for any subspace $W \subseteq V$, we have*

$$\dim V = \dim W + \dim W^\perp.$$

Corollary. *If $\mathcal{B}$ is a nondegenerate symmetric or alternating bilinear form on the finite-dimensional vector space V, then for any nondegenerate subspace $W \subseteq V$, we have*

$$V = W \oplus W^{\perp}$$

and $W^{\perp}$ is nondegenerate.

Proof. Nondegeneracy of W means that $W \cap W^{\perp} = 0$. The dimensions add up by the proposition, so $V = W \oplus W^{\perp}$. Also $W^{\perp\perp} \cap W^{\perp} = W \cap W^{\perp} = 0$ so $W^{\perp}$ is nondegenerate.

3c. Symmetric and alternating forms

We give some examples in this section.

Ex. A. The usual Euclidean inner product $\mathcal{E}$ on $\mathbf{R}^n$ is a bilinear form; it is symmetric and nondegenerate. It is also positive definite ($(v,v) > 0$ for all $v \neq 0$), a condition which implies (but is not equivalent to) nondegeneracy. There exist orthonormal bases C (i.e., $[\mathcal{E}]_C = I$). Any positive definite symmetric bilinear form $\mathcal{B}$ on $\mathbf{R}^n$ is equivalent to $\mathcal{E}$. This assertion is equivalent to the statement that there exists an orthonormal basis for $\mathcal{B}$, by the proposition in section 3a. To see that such a basis exists, choose any $v \in V$, $v \neq 0$. Then $\mathcal{B}(v,v) > 0$ and we let $c = \mathcal{B}(v,v)^{-1/2}$ and $w_1 = cv$. Then $\mathcal{B}(w_1,w_1) = 1$, and in particular $\mathbf{R}w_1$ is nondegenerate. Let $V' = \mathbf{R}w_1^{\perp}$. Then $V = \mathbf{R}w_1 \perp V'$ by the proposition, $\mathcal{B}|V'$ is still positive definite, and our assertion follows by induction.

Ex. B. Let us consider an arbitrary bilinear symmetric form $\mathcal{B}$ on $V = \mathbf{R}^n$, dropping the positive definiteness and nondegeneracy conditions. First choose any vector-space complement W to $V^{\perp}$ and write

$$V = V^{\perp} \oplus W = V^{\perp} \perp W;$$

observe also that W is nondegenerate, because $W^{\perp}$ is orthogonal to both $V^{\perp}$ and W, hence $W^{\perp} \subseteq V^{\perp}$, whence $W^{\perp} \cap W \subseteq V^{\perp} \cap W = 0$. This means that to obtain our form $\mathcal{B}$, we first may study the nondegenerate case and then attach an arbitrary number of basis vectors orthogonal to everything.

In the nondegenerate case, we must have $(v,v) \neq 0$ for some v. For otherwise,

$$2(v,w) = (v+w, v+w) - (v,v) - (w,w) = 0$$

for all $v, w \in V$, by bilinearity and symmetry. Choosing any v with $(v,v) \neq 0$ and setting $w_1 = |(v,v)|^{-1/2}v$, we have $(w_1,w_1) = \pm 1$. Then $V' = \mathbf{R}w_1$ is nondegenerate, so $V = \mathbf{R}w_1 \perp V'$ and we may continue. In this way we have proved the existence part of Sylvester's Theorem:

Theorem. *Let $\mathcal{B}$ be a symmetric bilinear form on the real finite-dimensional vector space V. Then V has an orthogonal basis consisting of vectors v for which $(v,v) = 0$ or ± 1.*

Moreover, the number of basis vectors for which (v,v) takes on each of these three values is uniquely determined by $\mathcal{B}$, i.e., independent of the choice of basis.

Proof. It remains to prove the uniqueness statement. Choose a basis C consisting of mutually orthogonal vectors $c_1,\dots,c_m$, $d_1,\dots,d_n$ and $e_1,\dots,e_p$, with $(c_i,c_i)=1$, $(d_i,d_i)=0$ and $(e_i,e_i)=-1$ for all i. Let $V_C^{+,0}$ be the span of the c's and d's; then $(v,v)\ge 0$ for all $v\in C^{+,0}$. Likewise if V_C^- is the span of the $e's$, then $(v,v)<0$ for all $0\ne v\in V_C^-$. We can make the same construction for any basis as in the theorem. We must prove that the dimensions m,n,p are unique; but they are determined by the four dimensions $V_C^{\pm,0}$ and $V_C^{\pm}$, so it is enough to prove the equality of these. As $V_C^{\pm,0}$ and $V_C^{\mp}$ are complementary, it is enough to show the uniqueness of $\dim V_C^{\pm,0}$. Let D be another such basis. Then $V_C^{+,0}\cap V_D^-=0$, since its nonzero vectors v satisfy $(v,v)\ge 0$ and $(v,v)<0$ simultaneously. Therefore $\dim V_C^{+,0}\le \dim V-\dim V_D^-=\dim V_D^{+,0}$. By symmetry, we have equality. QED

Ex. C. Suppose that $\mathcal{B}$ is a nondegenerate alternating form (here F can be any field). Let $v\in V$, $v\ne 0$. Since $\mathcal{B}$ is nondegenerate, $(v,v')\ne 0$ for some $v'\in V$. Replacing v' by a scalar multiple we may arrange that $(v,v')=1$. Such a pair v,v' is called a hyperbolic pair, and satisfies

$$(v,v)=(v',v')=0,\ (v,v')=1,\ (v',v)=-1.$$

Obviously $v'\notin Fv$. Let $W=Fv+Fv'$. Then $\dim W=2$ (W is a "hyperbolic plane"), and the matrix of the form $\mathcal{B}$ with respect to $\{v,v'\}$ is

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

In particular $\mathcal{B}|W$ is nondegenerate, so $\mathcal{B}|W^\perp$ is too by the corollary in the previous section. By induction we have proved:

Theorem. *Let $\mathcal{B}$ be a nondegenerate alternating bilinear form on a f.d.v.s. V. Then V is the orthogonal sum of hyperbolic planes for $\mathcal{B}$. In particular,* $\dim V$ *is even.*

Corollary. *Any two nondegenerate vector spaces of the same dimension with alternating bilinear forms are isometrically isomorphic.*

3d. Sesquilinear and Hermitian Forms

The results of this section are specific to vector spaces over $\mathbf{C}$. For them, since $z\overline{z}\ge 0$ for all $z\in\mathbf{C}$, it is more favorable to consider "hermitian" forms. These are an example of sesquilinear forms ("$\frac{11}{2}$-linear").

A mapping $\phi: V\to W$ of complex vector spaces is called conjugate-linear if and only if $\phi(v+v')=\phi(v)+\phi(v')$ (ϕ is "bi-additive") and $\phi(\alpha v)=\overline{\alpha}\phi(v)$ for all $v,v'\in v$ and $\alpha\in\mathbf{C}$. Here $\overline{\alpha}$ is the complex conjugate of α.

Exercise. *For a conjugate-linear mapping $\phi : V \to W$, $\ker\phi$ and $\operatorname{im}\phi$ are subspaces and $\dim V = \dim\ker\phi + \dim\operatorname{im}\phi$.*

Definition. *A hermitian form on a complex vector space V is a form $\mathcal{B} : V \times V \to V$ which is linear in the first argument (for each fixed value of the second) and conjugate linear in the second (for each fixed value of the first), and which is hermitian-symmetric:*

$$(w, v) = \overline{(v, w)}$$

for all $v, w \in V$.

Of these three conditions, either of the first two follows from the others. Without the hermitian-symmetry, the form still could be called sesquilinear.

For a sesquilinear form, each λ_v is conjugate-linear and the mapping λ is a linear transformation from V to the vector space $V^{*?}$ of all conjugate linear transformations $V \to F$. On the other hand, each ρ_v is linear and the mapping ρ is a conjugate linear transformation $V \to V^*$.

For a hermitian form, $(v, w) = 0 \iff (w, v) = 0$, and our general results on orthogonality from the previous section still hold; the fact that ρ is conjugate-linear rather than linear does not change the arguments. (Check this!) In particular, $\ker\lambda = \ker\rho$ by the hermitian symmetry, and we define if V to be nondegenerate if and only if these kernels are trivial. Moreover $V = W \perp W^\perp$ for any nondegenerate subspace $W \subseteq V$, and $W^\perp$ is then nondegenerate too. Also $W^{\perp\perp} = W$.

We have $(v, v) = \overline{(v, v)} \in \mathbf{R}$ by the hermitian-symmetry, for all $v \in V$. We again call the form positive-definite if and only if

$$(v, v) > 0 \text{ for all } 0 \neq v \in V.$$

If a basis $C = \{v_1, \ldots, v_n\}$ of V is given, then a hermitian form on V is determined by its values (v_i, v_j):

$$(\sum_i \alpha_i v_i, \sum_j \beta_j v_j) = \sum_{i,j} \alpha_i \overline{\beta}_j (v_i, v_j).$$

So we can again speak of the matrix $[\mathcal{B}]_C$ of $\mathcal{B}$ with respect to C.

Proposition. *The matrix of a hermitian form is hermitian-symmetric.*

Proof. By definition A is hermitian-symmetric if and only if $A^T = \overline{A}$, the matrix obtained by replacing each entry by its complex conjugate. In our situation this is just $(v_i, v_j) = \overline{(v_j, v_i)}$. QED

Exercise. *If V is a f.d.v.s over $\mathbf{C}$ with a positive-definite hermitian form, then V has an orthonormal basis, i.e., its matrix with respect to some basis is I. Any two positive-definite hermitian forms on V are equivalent.*

Notice that if V_0 is a real vector space with a symmetric form, with matrix A (which is then symmetric) relative to a basis $C = \{v_1, \ldots, v_n\}$, then we may form a complex vector

space V with the same basis, containing V_0 as an $\mathbf{R}$-subspace, and extend the given form on V_0 to a hermitian form on V, defined with respect to C by the same matrix (which is hermitian-symmetric, being real symmetric).

3e. Adjoints

We explore duality further in this section. Here V, W, X,etc. are vector spaces over a fixed field F. We investigate not only how to form the dual V^* of the object (vector space) V, but also how to dualize mappings. For each linear transformation of vector spaces

$$T : V \to W,$$

we define the adjoint

$$T^* : W^* \to V^*$$

by

$$T^*(\omega) = \omega \circ T.$$

Thus $T^*(\omega)$ is the composite of linear transformations from V to W to F, so lies in V^*. The reader should check that T^* is a linear transformation.

The reversal of arrows is the only way to get a "functorial" definition of T^*, one that does not depend on choice of basis. Of course if we choose isomorphisms $V \cong V^*$ and $W \cong W^*$, we can obtain from T a linear transformation $V^* \to W^*$, but as it depends heavily on the isomorphisms, and there is no "natural" isomorphism between a vector space and its dual, this would be a lousy way to define the dual of T.

The following properties are immediate:

a) If $T : V \to W$ and $U : W \to X$, then $(U \circ T)^* = T^* \circ U^*$.

b) The adjoint of $id_V : V \to V$ is $(id_V)^* = id_{V^*}$.

c) If B and C are bases of V and W respectively, and B^* and C^* are the dual bases of V^* and W^*, then

$$[T]_C^B \text{ and } [T^*]_{B^*}^{C^*} \text{ are transposes of each other.}$$

In c), the equation $T(v_j) = \sum_i a_{ij} w_i$ leads to the equation $T^*(w_i^*)(v_k) = w_i^*(T(v_k)) = a_{ik}$, so $T^*(w_i^*) = \sum_k a_{ik} v_k^*$ holds when applied to any v_k and hence is valid.

Repeating the dualizing process gives us V^{**}, W^{**} and $T^{**} : V^{**} \to W^{**}$.

Proposition. *The mapping $\nu_V : V \to V^{**}$ given by $\nu(v)(\omega) = \omega(v)$ for all $\omega \in V^*$ is an isomorphism.*

Proof. The expression $\omega(v)$ is linear in each of ω and v when the other is held fixed. Therefore ν_V is well-defined and linear, and $\nu_V(v) = 0$ if and only if $v \in \ker \omega$ for all $\omega \in V^*$. This implies $v = 0$, for otherwise v would be part of a basis of V and an element of the dual basis would not annihilate v. Thus ν_V is injective, hence an isomorphism as V is finite-dimensional.

Proposition. *(Naturality of ν) If $T : V \to W$ is a linear transformation then the following diagram commutes:*

$$\begin{array}{rccc} T: & V & \longrightarrow & W \\ & \nu_V \downarrow & & \nu_W \downarrow \\ T^{**}: & V^{**} & \longrightarrow & W^{**} \end{array}$$

Proof. You must check it yourself to believe it. For any $v \in V$ and $\omega \in W^*$, $\nu_W(T(v))(\omega) = \omega(T(v)) = T^*(\omega)(v)$ while $T^{**}(\nu_V(v))(\omega) = \nu_V(v)[T^*(\omega)] = T^*(\omega(v))$. QED

It is for this reason that the relation between V and V^* is said to be one of "duality".

When V has a nondegenerate bilinear or sesquilinear form $\mathcal{B}$, then $\mathcal{B}$ determines a way to identify V and V^*, namely by the mapping $\rho_{\mathcal{B}}$. In this case given $T : V \to V$ we can use this identification to interpret T^* as a mapping from $V \to V$, also called (by abuse of notation) the adjoint of T.

Definition. *Let V have the nondegenerate bilinear form $\mathcal{B}$. Let $T : V \to V$ be a linear transformation. Then the adjoint $T^* : V \to V$ is the unique linear transformation such that*

$$\mathcal{B}(Tv, w) = \mathcal{B}(v, T^*w) \text{ for all } v, w \in V.$$

One can see that this makes sense in two ways. First, for each fixed $w \in W$, $\mathcal{B}(Tv, w)$ is linear in v, so by the nondegeneracy of $\mathcal{B}$ there is a unique element (which we call T^*w) of V such that the above equation holds. Then $\mathcal{B}(v, T^*(w_1 + w_2)) = \mathcal{B}(Tv, w_1 + w_2) = \mathcal{B}(Tv, w_1) + \mathcal{B}(Tv, w_2) = \mathcal{B}(v, T^*w_1) + \mathcal{B}(v, T^*w_2) = \mathcal{B}(v, T^*w_1 + T^*w_2)$, so $T^*(w_1 + w_2) = T^*(w_1) + T^*(w_2)$ by the uniqueness just invoked. Similarly scalar multiplication is preserved by T^*.

On the other hand, one can check that with our two definitions of T^* (one on V, one on V^*) the following diagram commutes:

$$\begin{array}{rccc} T^*: & V & \longrightarrow & V \\ & \rho_{\mathcal{B}} \downarrow & & \rho_{\mathcal{B}} \downarrow \\ T^*: & V^* & \longrightarrow & V^* \end{array}$$

namely, $\rho(T^*w)(v) = (v, T^*w) = (Tv, w) = \rho(w)(Tv) = (\rho(w) \circ T)(v) = (T^*\rho(w))(v)$ for all $v, w \in V$, so $\rho \circ T^* = T^* \circ \rho$.

This works just as well if $\mathcal{B}$ is nondegenerate hermitian.

Theorem. *Let $T : V \to V$ be a linear transformation on the vector space V, and let $T^* : V \to V$ be its adjoint with respect to a nondegenerate bilinear or sesquilinear form $\mathcal{B}$. Then*

$$\ker T = (\operatorname{im} T^*)^\perp \text{ and } \ker T^* = (\operatorname{im} T)^\perp.$$

Proof. $v \in \ker T \iff Tv = 0 \forall v \iff (Tv, w) = 0 \forall v, w \iff (v, T^*w) = 0 \forall v, w \iff v \in (\operatorname{im} T^*)^\perp$. The second statement is proved similarly. QED

Note that finite-dimensionality is not used here. But to obtain the corollary $(\ker T)^\perp = \operatorname{im} T^*$ by taking perps of both sides, one needs finite-dimensionality (or some other sufficient condition for $\operatorname{im} T^* = (\operatorname{im} T^*)^{\perp\perp}$).

Exercise. $(T_1 + T_2)^* = T_1^* + T + 2^*$, $(\alpha T)^* = \alpha T^*$ *or* $\overline{\alpha} T^*$ *(the latter if* $\mathcal{B}$ *is sesquilinear). Moreover* $T^{**} = T$.

Exercise. *In the sesquilinear case, with respect to an orthonormal basis* C *of* V*, the matrices of* T *and* T^* *satisfy*

$$[T^*]_C^C = \overline{[T]_C^C}^T.$$

3f. The Finite-Dimensional Spectral Theorem

We have earlier derived an algebraic necessary and sufficient condition for a matrix, or a linear transformation $V \to V$, to be diagonalizable: its minimal polynomial is square-free. When V comes equipped with a bilinear form (i.e., some geometry), there may be geometric sufficient conditions for diagonalizability. We give an important one here: the spectral theorem.

Suppose then that $T : V \to V$ is a linear transformation, and V is a complex finite-dimensional vector space with a nondegenerate hermitian form $\mathcal{B}$.

Definition. T *is called*

a) *self-adjoint if and only if* $T^* = T$;

b) *unitary if and only if* $T^* = T^{-1}$;

c) *normal if and only if* $TT^* = T^*T$.

Thus self-adjoint and unitary transformations are normal.

Exercise. *If* $A = [T]_C^C$*, with* C *an orthonormal basis, then* T *is self-adjoint, unitary or normal according as* $A^T = \overline{A}$, $A^T\overline{A} = I$, *or* $A^T\overline{A} = \overline{A}A^T$.

Spectral Theorem. *Let* V *be a finite-dimensional complex vector space with a positive definite hermitian bilinear form. Let* $T : V \to V$ *be a linear transformation. Then* T *is normal if and only if* V *has an orthonormal basis consisting of eigenvectors of* T*. Moreover, if* T *is self-adjoint, all its eigenvalues are real; if* T *is unitary, then all its eigenvalues are on the unit circle.*

We prove this as the culmination of several simple remarks.

1 . For any $v \in V$, $Tv = 0$ if and only if $T^*Tv = 0$.

Proof. If $T^*Tv = 0$, then $0 = (T^*Tv, v) = (Tv, T^{**}v) = (Tv, Tv)$. By positive-definiteness, $Tv = 0$. The converse is obvious.

2 . If T is normal, then $\ker T = \ker T^*$.

Proof. $Tv = 0 \iff T^*Tv = 0 \iff TT^*v = 0 \iff T^{**}T^*v = 0 \iff T^*v = 0$.

3 . If T is normal, then so is $T - \lambda id_V$ for any $\lambda \in \mathbf{C}$.

Proof. $(T - \lambda id_V)^* = T^* - \overline{\lambda} id_V$ commutes with T and hence with $T - \lambda id_V$.

4 . If T is normal and $\lambda \in \mathbf{C}$, then $\ker(T - \lambda) = \ker(T^* - \overline{\lambda}) = \operatorname{im}(T - \lambda)^\perp$.

Proof. By 3, we may apply 2 and the theorem of the previous section.

5 . If T is normal and $\lambda \in \mathbf{C}$, then $\ker(T - \lambda)$ and $\operatorname{im}(T - \lambda)$ are T^*-invariant.

Proof. Let $U = T - \lambda$. Since T^* commutes with T, it commutes with U. The equation $T(Uv) = U(Tv)$, with v varying over V, shows $T(\operatorname{im} U) \subseteq \operatorname{im} U$. With $v \in \ker U$ it shows that $UTv = 0$, so $Tv \in \ker U$, as required.

6 . If T is normal, and $W \subseteq V$ is a subspace invariant under both T and T^*, then $T|W$ is normal.

Proof. $(Tv, w) = (v, T^*w)$ for all $v, w \in V$, hence all $v, w \in W$. Since $T^*w \in W$ for all $w \in W$, this equation and the definition of T^* (and the uniqueness of $(T|W)^*w$) implies that $(T|W)^*w = T^*w$. That is, $(T|W)^* = T^*|W$. Now restrict the equation $TT^* = T^*T$ to W to get the normality of $T|W$.

Now we prove the "hard" direction of the spectral theorem, by induction on $\dim V$. Suppose that T is normal. If there exists a proper subspace W such that $V = W \oplus W^\perp$, and both W and $W^\perp$ are T-invariant and T^*-invariant, then $T|W$ and $T|W^\perp$ are both normal by 6, and we are done by induction applied to W and $W^\perp$. So we may assume that no such decomposition exists.

Using the fundamental theorem of algebra, take an eigenvalue λ of T. By 4 and 5,

$$V = \ker(T - \lambda id_V) \oplus \operatorname{im}(T - \lambda id_V)$$

is an orthogonal decomposition of V with both subspaces invariant under T and T^*. Since no such proper decomposition exists and $\ker(T - \lambda) \neq 0$ by choice of λ, we must have $V = \ker(T - \lambda id_V)$, so $T = \lambda id_V$ and the conclusion is obvious.

If T is self-adjoint, then for any eigenvector v and corresponding eigenvalue λ,

$$\lambda(v, v) = (Tv, v) = (v, Tv) = \overline{(Tv, v)}$$

is real. Also $(v, v) = \overline{(v, v)}$ is real, so λ is as well. Likewise if T is unitary, then T is invertible by definition, and the equation $Tv = \lambda v$ implies that $T^{-1}v = \lambda^{-1}v$. Now $\lambda(v, v) = (Tv, v) = (v, T^{-1}v) = \overline{\lambda}^{-1}(v, v)$, so $\lambda = \overline{\lambda}^{-1}$ is on the unit circle.

Conversely, if there exists an orthonormal basis $v_1, \ldots, v_n$ of eigenvectors of T, say $Tv_i = \lambda_i v_i$, then we put $Uv_i = \overline{\lambda}_i v_i$ for each i and extend U to a linear transformation on V; then

$$(Tv_i, v_j) = \delta_{ij}\lambda_i = \delta_{ij}\lambda_j = (v_i, Uv_j)$$

for all i and j, and so $U = T^*$. But obviously $TU = UT$ since the matrices of both T and U with respect to C are diagonal. Hence T is normal. QED