CHAPTER XIII

Matrices and Linear Maps

Presumably readers of this chapter will have had some basic acquaintance with linear algebra in elementary courses. We go beyond such courses by pointing out that a lot of results hold for free modules over a commutative ring. This is useful when one wants to deal with families of linear maps, and reduction modulo an ideal.

Note that §8 and §9 give examples of group theory in the context of linear groups.

Throughout this chapter, we let R be a commutative ring, and we let E, F be R-modules. We suppress the prefix R in front of linear maps and modules.

§1. MATRICES

By an $m \times n$ **matrix** in R one means a doubly indexed family of elements of R, (a_{ij}), ($i = 1, \dots, m$ and $j = 1, \dots, n$), usually written in the form

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

We call the elements a_{ij} the **coefficients** or **components** of the **matrix**. A $1 \times n$ matrix is called a **row vector** (of dimension, or size, n) and a $m \times 1$ matrix is called a **column vector** (of dimension, or size, m). In general, we say that (m, n) is the **size** of the matrix, or also $m \times n$.

We define addition for matrices of the same size by components. If $A = (a_{ij})$ and $B = (b_{ij})$ are matrices of the same size, we define $A + B$ to be the matrix whose ij-component is $a_{ij} + b_{ij}$. Addition is obviously associative. We define the multiplication of a matrix A by an element $c \in R$ to be the matrix (ca_{ij}),

S. Lang, *Algebra*

whose ij-component is ca_{ij}. Then the set of $m \times n$ matrices in R is a module (i.e. an R-module).

We define the product AB of two matrices only under certain conditions. Namely, when A has size (m, n) and B has size (n, r), i.e. only when the size of the rows of A is the same as the size of the columns of B. If that is the case, let $A = (a_{ij})$ and let $B = (b_{jk})$. We define AB to be the $m \times r$ matrix whose ik-component is

$$\sum_{j=1}^{n} a_{ij} b_{jk}.$$

If A, B, C are matrices such that AB is defined and BC is defined, then so is $(AB)C$ and $A(BC)$ and we have

$$(AB)C = A(BC).$$

This is trivial to prove. If $C = (c_{kl})$, then the reader will see at once that the il-component of either of the above products is equal to

$$\sum_j \sum_k a_{ij} b_{jk} c_{kl}.$$

An $m \times n$ matrix is said to be a **square matrix** if $m = n$. For example, a 1×1 matrix is a square matrix, and will sometimes be identified with the element of R occurring as its single component.

For a given integer $n \geqq 1$ the set of square $n \times n$ matrices forms a ring.

This is again trivially verified and will be left to the reader.

The unit element of the ring of $n \times n$ matrices is the matrix

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

whose components are equal to 0 except on the diagonal, in which case they are equal to 1. We sometimes write I instead of I_n.

If $A = (a_{ij})$ is a square matrix, we define in general its **diagonal components** to be the elements a_{ii}.

We have a natural ring-homomorphism of R into the ring of $n \times n$ matrices, given by

$$c \mapsto cI_n.$$

Thus cI_n is the square $n \times n$ matrix having all its components equal to 0 except the diagonal components, which are equal to c. Let us denote the ring of $n \times n$

matrices in R by $\mathrm{Mat}_n(R)$. Then $\mathrm{Mat}_n(R)$ is an algebra over R (with respect to the above homomorphism).

Let $A = (a_{ij})$ be an $m \times n$ matrix. We define its **transpose** ${}^tA$ to be the matrix (a_{ji}) ($j = 1, \ldots, n$ and $i = 1, \ldots, m$). Then ${}^tA$ is an $n \times m$ matrix. The reader will verify at once that if A, B are of the same size, then

$$ {}^t(A + B) = {}^tA + {}^tB. $$

If $c \in R$ then ${}^t(cA) = c\,{}^tA$. If A, B can be multiplied, then ${}^tB\,{}^tA$ is defined and we have

$$ {}^t(AB) = {}^tB\,{}^tA. $$

We note the operations on matrices commute with homomorphisms. More precisely, let $\varphi : R \to R'$ be a ring-homomorphism. If A, B are matrices in R, we define φA to be the matrix obtained by applying φ to all the components of A. Then

$$ \varphi(A + B) = \varphi A + \varphi B, \qquad \varphi(AB) = (\varphi A)(\varphi B), \qquad \varphi(cA) = \varphi(c)\varphi A, $$
$$ \varphi({}^tA) = {}^t\varphi(A). $$

A similar remark will hold throughout our discussion of matrices (for instance in the next section).

Let $A = (a_{ij})$ be a square $n \times n$ matrix in a commutative ring R. We define the **trace** of A to be

$$ \mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}; $$

in other words, the trace is the sum of the diagonal elements.

If A, B are $n \times n$ matrices, then

$$ \mathrm{tr}(AB) = \mathrm{tr}(BA). $$

Indeed, if $A = (a_{ij})$ and $B = (b_{ij})$ then

$$ \mathrm{tr}(AB) = \sum_{i} \sum_{\nu} a_{i\nu} b_{\nu i} = \mathrm{tr}(BA). $$

As an application, we observe that if B is an invertible $n \times n$ matrix, then

$$ \mathrm{tr}(B^{-1}AB) = \mathrm{tr}(A). $$

Indeed, $\mathrm{tr}(B^{-1}AB) = \mathrm{tr}(ABB^{-1}) = \mathrm{tr}(A)$.

§2. THE RANK OF A MATRIX

Let k be a field and let A be an $m \times n$ matrix in k. By the **row rank** of A we shall mean the maximum number of linearly independent rows of A, and by the **column rank** of A we shall mean the maximum number of linearly independent columns of A. Thus these ranks are the dimensions of the vector spaces generated respectively by the rows of A and the columns of A. We contend that these ranks are equal to the same number, and we define the **rank** of A to be that number.

Let $A^1, \ldots, A^n$ be the columns of A, and let $A_1, \ldots, A_m$ be the rows of A. Let ${}^tX = (x_1, \ldots, x_m)$ have components $x_i \in k$. We have a linear map

$$X \mapsto x_1A_1 + \cdots + x_mA_m$$

of $k^{(m)}$ onto the space generated by the row vectors. Let W be its kernel. Then W is a subspace of $k^{(m)}$ and

$$\dim W + \text{row rank} = m.$$

If Y is a column vector of dimension m, then the map

$$(X, Y) \mapsto {}^tXY = X \cdot Y$$

is a bilinear map into k, if we view the 1×1 matrix ${}^tXY$ as an element of k. We observe that W is the orthogonal space to the column vectors $A^1, \ldots, A^n$, i.e. it is the space of all X such that $X \cdot A^j = 0$ for all $j = 1, \ldots, n$. By the duality theorem of Chapter III, we know that $k^{(m)}$ is its own dual under the pairing

$$(X, Y) \mapsto X \cdot Y$$

and that $k^{(m)}/W$ is dual to the space generated by $A^1, \ldots, A^n$. Hence

$$\dim k^{(m)}/W = \text{column rank},$$

or

$$\dim W + \text{column rank} = m.$$

From this we conclude that

$$\text{column rank} = \text{row rank},$$

as desired.

We note that W may be viewed as the space of solutions of the system of n linear equations

$$x_1A_1 + \cdots + x_mA_m = 0,$$

in m unknowns $x_1, \dots, x_m$. Indeed, if we write out the preceding vector equation in terms of all the coordinates, we get the usual system of n linear equations. We let the reader do this if he or she wishes.

§3. MATRICES AND LINEAR MAPS

Let E be a module, and assume that there exists a basis $\mathcal{B} = \{\xi_1, \dots, \xi_n\}$ for E over R. This means that every element of E has a unique expression as a linear combination

$$x = x_1\xi_1 + \cdots + x_n\xi_n$$

with $x_i \in R$. We call $(x_1, \dots, x_n)$ the **components** of x with respect to the basis. We may view this n-tuple as a row vector. We shall denote by X the transpose of the row vector $(x_1, \dots, x_n)$. We call X the **column vector of x with respect to the basis**.

We observe that if $\{\xi'_1, \dots, \xi'_m\}$ is another basis of E over R, then $m = n$. Indeed, let $\mathfrak{p}$ be a maximal ideal of R. Then $E/\mathfrak{p}E$ is a vector space over the field $R/\mathfrak{p}R$, and it is immediately clear that if we denote by $\bar{\xi}_i$ the residue class of ξ_i mod $\mathfrak{p}E$, then $\{\bar{\xi}_1, \dots, \bar{\xi}_n\}$ is a basis for $E/\mathfrak{p}E$ over $R/\mathfrak{p}R$. Hence n is also the dimension of this vector space, and we know the invariance of the cardinality for bases of vector spaces over fields. Thus $m = n$. We shall call n the **dimension** of the module E over R.

We shall view $R^{(n)}$ as the module of column vectors of size n. It is a free module of dimension n over R. It has a basis consisting of the unit vectors $e^1, \dots, e^n$ such that

$${}^t e^i = (0, \dots, 0, 1, 0, \dots, 0)$$

has components 0 except for its i-th component, which is equal to 1.

An $m \times n$ matrix A gives rise to a linear map

$$L_A : R^{(n)} \to R^{(m)}$$

by the rule

$$X \mapsto AX.$$

Namely, we have $A(X + Y) = AX + AY$ and $A(cX) = cAX$ for column vectors X, Y and $c \in R$.

The above considerations can be extended to a slightly more general context, which can be very useful. Let E be an abelian group and assume that R is a commutative subring of

$$\mathrm{End}_{\mathbf{Z}}(E) = \mathrm{Hom}_{\mathbf{Z}}(E, E).$$

Then E is an R-module. Furthermore, if A is an $m \times n$ matrix in R, then we get a linear map

$$L_A : E^{(n)} \to E^{(m)}$$

defined by a rule similar to the above, namely $X \mapsto AX$. However, this has to be interpreted in the obvious way. If $A = (a_{ij})$ and X is a column vector of elements of E, then

$$AX = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix},$$

where $y_i = \sum_{j=1}^{n} a_{ij}x_j$.

If A, B are matrices in R whose product is defined, then for any $c \in R$ we have

$$L_{AB} = L_A L_B \quad \text{and} \quad L_{cA} = cL_A.$$

Thus we have associativity, namely

$$A(BX) = (AB)X.$$

An arbitrary commutative ring R may be viewed as a module over itself. In this way we recover the special case of our map from $R^{(n)}$ into $R^{(m)}$. Furthermore, if E is a module over R, then R may be viewed as a ring of endomorphisms of E.

Proposition 3.1. *Let E be a free module over R, and let $\{x_1, \ldots, x_n\}$ be a basis. Let $y_1, \ldots, y_n$ be elements of E. Let A be the matrix in R such that*

$$A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Then $\{y_1, \ldots, y_n\}$ is a basis of E if and only if A is invertible.

Proof. Let X, Y be the column vectors of our elements. Then $AX = Y$. Suppose Y is a basis. Then there exists a matrix C in R such that $CY = X$.

Then $CAX = X$, whence $CA = I$ and A is invertible. Conversely, assume that A is invertible. Then $X = A^{-1}Y$ and hence $x_1, \dots, x_n$ are in the module generated by $y_1, \dots, y_n$. Suppose that we have a relation

$$b_1 y_1 + \cdots + b_n y_n = 0$$

with $b_i \in R$. Let B be the row vector $(b_1, \dots, b_n)$. Then

$$BY = 0$$

and hence $BAX = 0$. But $\{x_1, \dots, x_n\}$ is a basis. Hence $BA = 0$, and hence $BAA^{-1} = B = 0$. This proves that the components of Y are linearly independent over R, and proves our proposition.

We return to our situation of modules over an arbitrary commutative ring R.

Let E, F be modules. We shall see how we can associate a matrix with a linear map whenever bases of E and F are given. We assume that E, F are free. We let $\mathfrak{B} = \{\xi_1, \dots, \xi_n\}$ and $\mathfrak{B}' = \{\xi'_1, \dots, \xi'_m\}$ be bases of E and F respectively. Let

$$f : E \to F$$

be a linear map. There exist unique elements $a_{ij} \in R$ such that

$$f(\xi_1) = a_{11}\xi'_1 + \cdots + a_{m1}\xi'_m,$$
$$\cdots$$
$$f(\xi_n) = a_{1n}\xi'_1 + \cdots + a_{mn}\xi'_m,$$

or in other words,

$$f(\xi_j) = \sum_{i=1}^{m} a_{ij}\xi'_i$$

(Observe that the sum is over the *first* index.) We define

$$M^{\mathfrak{B}}_{\mathfrak{B}'}(f) = (a_{ij}).$$

If $x = x_1\xi_1 + \cdots + x_n\xi_n$ is expressed in terms of the basis, let us denote the column vector X of components of x by $M_{\mathfrak{B}}(x)$. We see that

$$M_{\mathfrak{B}'}(f(x)) = M^{\mathfrak{B}}_{\mathfrak{B}'}(f)M_{\mathfrak{B}}(x).$$

In other words, if X' is the column vector of $f(x)$, and M is the matrix associated with f then $X' = MX$. Thus the operation of the linear map is reflected by the matrix multiplication, and we have $f = L_M$.

Proposition 3.2. *Let E, F, D be modules, and let $\mathcal{B}$, $\mathcal{B}'$, $\mathcal{B}''$ be finite bases of E, F, D, respectively. Let*

$$E \xrightarrow{f} F \xrightarrow{g} D$$

be linear maps. Then

$$M^{\mathcal{B}}_{\mathcal{B}''}(g \circ f) = M^{\mathcal{B}'}_{\mathcal{B}''}(g) M^{\mathcal{B}}_{\mathcal{B}'}(f).$$

Proof. Let A and B be the matrices associated with the maps f, g respectively, with respect to our given bases. If X is the column vector associated with $x \in E$, the vector associated with $g(f(x))$ is $B(AX) = (BA)X$. Hence BA is the matrix associated with $g \circ f$. This proves what we wanted.

Corollary 3.3. *Let $E = F$. Then*

$$M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}) M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}) = M^{\mathcal{B}'}_{\mathcal{B}'}(\mathrm{id}) = I.$$

Each matrix $M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id})$ is invertible (i.e. is a unit in the ring of matrices).

Proof. Obvious.

Corollary 3.4. *Let $N = M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id})$. Then*

$$M^{\mathcal{B}'}_{\mathcal{B}'}(f) = M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}) M^{\mathcal{B}}_{\mathcal{B}}(f) M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}) = N M^{\mathcal{B}}_{\mathcal{B}}(f) N^{-1}.$$

Proof. Obvious

Corollary 3.5. *Let E be a free module of dimension n over R. Let $\mathcal{B}$ be a basis of E over R. The map*

$$f \mapsto M^{\mathcal{B}}_{\mathcal{B}}(f)$$

is a ring-isomorphism of the ring of endomorphisms of E onto the ring of $n \times n$ matrices in R. In fact, the isomorphism is one of algebras over R.

We shall call the matrix $M^{\mathcal{B}}_{\mathcal{B}}(f)$ the **matrix associated with f with respect to the basis $\mathcal{B}$**.

Let E be a free module of dimension n over R. By $GL(E)$ or $\mathrm{Aut}_R(E)$ one means the group of linear automorphisms of E. It is the group of units in $\mathrm{End}_R(E)$. By $GL_n(R)$ one means the group of invertible $n \times n$ matrices in R. Once a basis is selected for E over R, we have a group-isomorphism

$$GL(E) \leftrightarrow GL_n(R)$$

with respect to this basis.

Let E be as above. If

$$f: E \to E$$

is a linear map, we select a basis $\mathcal{B}$ and let M be the matrix associated with f relative to $\mathcal{B}$. We define the **trace** of f to be the trace of M, thus

$$\operatorname{tr}(f) = \operatorname{tr}(M).$$

If M' is the matrix of f with respect to another basis, then there exists an invertible matrix N such that $M' = N^{-1}MN$, and hence the trace is independent of the choice of basis.

§4. DETERMINANTS

Let $E_1, \dots, E_n, F$ be modules. A map

$$f: E_1 \times \cdots \times E_n \to F$$

is said to be ***R*-multilinear** (or simply multilinear) if it is linear in each variable, i.e. if for every index i and elements $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n, x_j \in E_j$, the map

$$x \mapsto f(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)$$

is a linear map of E_i into F.

A multilinear map defined on an n-fold product is also called n-multilinear. If $E_1 = \cdots = E_n = E$, we also say that f is a **multilinear map on *E***, instead of saying that it is multilinear on $E^{(n)}$.

Let f be an n-multilinear map. If we take two indices i, j and $i \neq j$ then fixing all the variables except the i-th and j-th variable, we can view f as a bilinear map on $E_i \times E_j$.

Assume that $E_1 = \cdots = E_n = E$. We say that the multilinear map f is **alternating** if $f(x_1, \dots, x_n) = 0$ whenever there exists an index i, $1 \leqq i \leqq n - 1$, such that $x_i = x_{i+1}$ (in other words, when two adjacent elements are equal).

Proposition 4.1. *Let f be an n-multilinear alternating map on E. Let $x_1, \dots, x_n \in E$. Then*

$$f(\dots, x_i, x_{i+1}, \dots) = -f(\dots, x_{i+1}, x_i, \dots).$$

In other words, when we interchange two adjacent arguments of f, the value of f changes by a sign. If $x_i = x_j$ for $i \neq j$ then $f(x_1, \dots, x_n) = 0$.

Proof. Restricting our attention to the factors in the i-th and j-th place, with $j = i + 1$, we may assume f is bilinear for the first statement. Then for all x, $y \in E$ we have

$$0 = f(x + y, x + y) = f(x, y) + f(y, x).$$

This proves what we want, namely $f(y, x) = -f(x, y)$. For the second assertion, we can interchange successively adjacent arguments of f until we obtain an n-tuple of elements of E having two equal adjacent arguments. This shows that when $x_i = x_j$, $i \neq j$, then $f(x_1, \dots, x_n) = 0$.

Corollary 4.2. *Let f be an n-multilinear alternating map on E. Let $x_1, \dots, x_n \in E$. Let $i \neq j$ and let $a \in R$. Then the value of f on $(x_1, \dots, x_n)$ does not change if we replace x_i by $x_i + ax_j$ and leave all other components fixed.*

Proof. Obvious.

A multilinear alternating map taking its value in R is called a multilinear alternating **form**.

On repeated occasions we shall evaluate multilinear alternating maps on linear combinations of elements of E. Let

$$w_1 = a_{11}v_1 + \cdots + a_{1n}v_n,$$
$$\dots$$
$$w_n = a_{n1}v_1 + \cdots + a_{nn}v_n.$$

Let f be n-multilinear alternating on E. Then

$$f(w_1, \dots, w_n) = f(a_{11}v_1 + \cdots + a_{1n}v_n, \dots, a_{n1}v_1 + \cdots + a_{nn}v_n).$$

We expand this by multilinearity, and get a sum of terms of type

$$a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} f(v_{\sigma(1)}, \dots, v_{\sigma(n)}),$$

where σ ranges over arbitrary maps of $\{1, \dots, n\}$ into itself. If σ is not a bijection (i.e. a permutation), then two arguments $v_{\sigma(i)}$ and $v_{\sigma(j)}$ are equal for $i \neq j$, and the term is equal to 0. Hence we may restrict our sum to permutations σ. Shuffling back the elements $(v_{\sigma(1)}, \dots, v_{\sigma(n)})$ to their standard ordering and using Proposition 4.1, we see that we have obtained the following expansion:

Lemma 4.3. *If $w_1, \dots, w_n$ are as above, then*

$$f(w_1, \dots, w_n) = \sum_{\sigma} \epsilon(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} f(v_1, \dots, v_n)$$

where the sum is taken over all permutations σ of $\{1, \dots, n\}$ and $\epsilon(\sigma)$ is the sign of the permutation.

For determinants, I shall follow Artin's treatment in *Galois Theory*.

By an $n \times n$ **determinant** we shall mean a mapping

$$\det : \mathrm{Mat}_n(R) \to R$$

also written

$$D : \mathrm{Mat}_n(R) \to R$$

which, when viewed as a function of the column vectors $A^1, \ldots, A^n$ of a matrix A, is multilinear alternating, and such that $D(I) = 1$. In this chapter, we use mostly the letter D to denote determinants.

We shall prove later that determinants exist. For the moment, we derive properties.

Theorem 4.4. **(Cramer's Rule).** *Let $A^1, \ldots, A^n$ be column vectors of dimension n. Let $x_1, \ldots, x_n \in R$ be such that*

$$x_1A^1 + \cdots + x_nA^n = B$$

for some column vector B. Then for each i we have

$$x_iD(A^1, \ldots, A^n) = D(A^1, \ldots, B, \ldots, A^n),$$

where B in this last line occurs in the i-th place.

Proof. Say $i = 1$. We expand

$$D(B, A^2, \ldots, A^n) = \sum_{j=1}^{n} x_jD(A^j, A^2, \ldots, A^n),$$

and use Proposition 4.1 to get what we want (all terms on the right are equal to 0 except the one having x_1 in it).

Corollary 4.5. *Assume that R is a field. Then $A^1, \ldots, A^n$ are linearly dependent if and only if $D(A^1, \ldots, A^n) = 0$.*

Proof. Assume we have a relation

$$x_1A^1 + \cdots + x_nA^n = 0$$

with $x_i \in R$. Then $x_iD(A) = 0$ for all i. If some $x_i \neq 0$ then $D(A) = 0$. Conversely, assume that $A^1, \ldots, A^n$ are linearly independent. Then we can express the unit vectors $e^1, \ldots, e^n$ as linear combinations

$$e^1 = b_{11}A^1 + \cdots + b_{1n}A^n,$$

$$\cdots$$

$$e^n = b_{n1}A^1 + \cdots + b_{nn}A^n$$

with $b_{ij} \in R$. But

$$1 = D(e^1, \dots, e^n).$$

Using a previous lemma, we know that this can be expanded into a sum of terms involving $D(A^1, \dots, A^n)$, and hence $D(A)$ cannot be 0.

Proposition 4.6. *If determinants exist, they are unique. If $A^1, \dots, A^n$ are the column vectors of dimension n, of the matrix $A = (a_{ij})$, then*

$$D(A^1, \dots, A^n) = \sum_{\sigma} \epsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n},$$

where the sum is taken over all permutations σ of $\{1, \dots, n\}$, and $\epsilon(\sigma)$ is the sign of the permutation.

Proof. Let $e^1, \dots, e^n$ be the unit vectors as usual. We can write

$$A^1 = a_{11}e^1 + \cdots + a_{n1}e^n,$$
$$\cdots$$
$$A^n = a_{1n}e^n + \cdots + a_{nn}e^n.$$

Therefore

$$D(A^1, \dots, A^n) = \sum_{\sigma} \epsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}$$

by the lemma. This proves that the value of the determinant is uniquely determined and is given by the expected formula.

Corollary 4.7. *Let $\varphi : R \to R'$ be a ring-homomorphism into a commutative ring. If A is a square matrix in R, define φA to be the matrix obtained by applying φ to each component of A. Then*

$$\varphi(D(A)) = D(\varphi A).$$

Proof. Apply φ to the expression of Proposition 4.6.

Proposition 4.8. *If A is a square matrix in R then*

$$D(A) = D({}^tA).$$

Proof. In a product

$$a_{\sigma(1),1} \cdots a_{\sigma(n),n}$$

each integer k from 1 to n occurs precisely once among the integers $\sigma(1), \dots, \sigma(n)$. Hence we can rewrite this product in the form

$$a_{1,\sigma^{-1}(1)} \cdots a_{n,\sigma^{-1}(n)}.$$

Since $\epsilon(\sigma) = \epsilon(\sigma^{-1})$, we can rewrite the sum in Proposition 4.6 in the form

$$\sum_{\sigma} \epsilon(\sigma^{-1})a_{1,\sigma^{-1}(1)} \cdots a_{n,\sigma^{-1}(n)}.$$

In this sum, each term corresponds to a permutation σ. However, as σ ranges over all permutations, so does σ^{-1}. Hence our sum is equal to

$$\sum_{\sigma} \epsilon(\sigma)a_{1,\sigma(1)} \cdots a_{n,\sigma(n)},$$

which is none other than $D({}^tA)$, as was to be shown.

Corollary 4.9. *The determinant is multilinear and alternating with respect to the rows of a matrix.*

We shall now prove existence, and prove simultaneously one additional important property of determinants.

When $n = 1$, we define $D(a) = a$ for any $a \in R$.

Assume that we have proved the existence of determinants for all integers $< n$ $(n \geqq 2)$. Let A be an $n \times n$ matrix in R, $A = (a_{ij})$. We let A_{ij} be the $(n-1) \times (n-1)$ matrix obtained from A by deleting the i-th row and j-th column. Let i be a fixed integer, $1 \leqq i \leqq n$. We define inductively

$$D(A) = (-1)^{i+1}a_{i1}D(A_{i1}) + \cdots + (-1)^{i+n}a_{in}D(A_{in}).$$

(This is known as the **expansion of *D* according to the *i*-th row**.) We shall prove that D satisfies the definition of a determinant.

Consider D as a function of the k-th column, and consider any term

$$(-1)^{i+j}a_{ij}D(A_{ij}).$$

If $j \neq k$ then a_{ij} does not depend on the k-th column, and $D(A_{ij})$ depends linearly on the k-th column. If $j = k$, then a_{ij} depends linearly on the k-th column, and $D(A_{ij})$ does not depend on the k-th column. In any case our term depends linearly on the k-th column. Since $D(A)$ is a sum of such terms, it depends linearly on the k-th column, and thus D is multilinear.

Next, suppose that two adjacent columns of A are equal, say $A^k = A^{k+1}$. Let j be an index $\neq k$ and $\neq k+1$. Then the matrix A_{ij} has two adjacent equal columns, and hence its determinant is equal to 0. Thus the term corresponding to an index $j \neq k$ or $k+1$ gives a zero contribution to $D(A)$. The other two terms can be written

$$(-1)^{i+k}a_{ik}D(A_{ik}) + (-1)^{i+k+1}a_{i,k+1}D(A_{i,k+1}).$$

The two matrices A_{ik} and $A_{i,k+1}$ are equal because of our assumption that the k-th column of A is equal to the $(k+1)$-th column. Similarly, $a_{ik} = a_{i,k+1}$.

Hence these two terms cancel since they occur with opposite signs. This proves that our form is alternating, and gives:

Proposition 4.10. *Determinants exist and satisfy the rule of expansion according to rows and columns.*

(For columns, we use the fact that $D(A) = D({}^tA)$.)

Example. We mention explicity one of the most important determinants. Let $x_1, \dots, x_n$ be elements of a commutative ring. The **Vandermonde determinant** $V = V(x_1, \dots, x_n)$ of these elements is defined to be

$$V = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix},$$

whose value can be determined explicitly to be

$$V = \prod_{i<j} (x_j - x_i).$$

If the ring is entire and $x_i \neq x_j$ for $i \neq j$, it follows that $V \neq 0$. The proof for the stated value is done by multiplying the next to the last row by x_1 and subtracting from the last row. Then repeat this step going up the rows, thus making the elements of the first column equal to 0, except for 1 in the upper left-hand corner. One can then expand according to the first column, and use the homogeneity property and induction to conclude the proof of the evaluation of V.

Theorem 4.11. *Let E be a module over R, and let $v_1, \dots, v_n$ be elements of E. Let $A = (a_{ij})$ be a matrix in R, and let*

$$A\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}.$$

Let Δ be an n-multilinear alternating map on E. Then

$$\Delta(w_1, \dots, w_n) = D(A)\,\Delta(v_1, \dots, v_n).$$

Proof. We expand

$$\Delta(a_{11}v_1 + \cdots + a_{1n}v_n, \dots, a_{n1}v_1 + \cdots + a_{nn}v_n),$$

and find precisely what we want, taking into account $D(A) = D({}^tA)$.

Let E, F be modules, and let $L_a^n(E, F)$ denote the set of n-multilinear alternating maps of E into F. If $F = R$, we also write $L_a^n(E, R) = L_a^n(E)$. It is clear that $L_a^n(E, F)$ is a module over R, i.e. is closed under addition and multiplication by elements of R.

Corollary 4.12. *Let E be a free module over R, and let $\{v_1, \dots, v_n\}$ be a basis. Let F be any module, and let $w \in F$. There exists a unique n-multilinear alternating map*

$$\Delta_w : E \times \cdots \times E \to F$$

such that $\Delta_w(v_1, \dots, v_n) = w$.

Proof. Without loss of generality, we may assume that $E = R^{(n)}$, and then, if $A^1, \dots, A^n$ are column vectors, we define

$$\Delta_w(A^1, \dots, A^n) = D(A)w.$$

Then Δ_w obviously has the required properties.

Corollary 4.13. *If E is free over R, and has a basis consisting of n elements, then $L_a^n(E)$ is free over R, and has a basis consisting of 1 element.*

Proof. We let Δ_1 be the multilinear alternating map taking the value 1 on a basis $\{v_1, \dots, v_n\}$. Any element $\varphi \in L_a^n(E)$ can then be written in a unique way as $c\Delta_1$, with some $c \in R$, namely $c = \varphi(v_1, \dots, v_n)$. This proves what we wanted.

Any two bases of $L_a^n(E)$ in the preceding corollary differ by a unit in R. In other words, if Δ is a basis of $L_a^n(E)$, then $\Delta = c\Delta_1 = \Delta_c$ for some $c \in R$, and c must be a unit. Our Δ_1 depends of course on the choice of a basis for E. When we consider $R^{(n)}$, our determinant D is precisely Δ_1, relative to the standard basis consisting of the unit vectors $e^1, \dots, e^n$.

It is sometimes convenient terminology to say that any basis of $L_a^n(E)$ is a **determinant** on E. In that case, the corollary to Cramer's rule can be stated as follows.

Corollary 4.14. *Let R be a field. Let E be a vector space of dimension n. Let Δ be any determinant on E. Let $v_1, \dots, v_n \in E$. In order that $\{v_1, \dots, v_n\}$ be a basis of E it is necessary and sufficient that*

$$\Delta(v_1, \dots, v_n) \neq 0.$$

Proposition 4.15. *Let A, B be $n \times n$ matrices in R. Then*

$$D(AB) = D(A)D(B).$$

Proof. This is actually a corollary of Theorem 4.11. We take $v_1, \dots, v_n$ to be the unit vectors $e^1, \dots, e^n$, and consider

$$AB\begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}.$$

We obtain

$$D(w_1, \dots, w_n) = D(AB)D(e^1, \dots, e^n).$$

On the other hand, by associativity, applying Theorem 4.11 twice,

$$D(w_1, \dots, w_n) = D(A)D(B)D(e^1, \dots, e^n).$$

Since $D(e^1, \dots, e^n) = 1$, our proposition follows.

Let $A = (a_{ij})$ be an $n \times n$ matrix in R. We let

$$\tilde{A} = (b_{ij})$$

be the matrix such that

$$b_{ij} = (-1)^{i+j}D(A_{ji}).$$

(Note the reversal of indices!)

Proposition 4.16. *Let $d = D(A)$. Then $A\tilde{A} = \tilde{A}A = dI$. The determinant $D(A)$ is invertible in R if and only if A is invertible, and then*

$$A^{-1} = \frac{1}{d}\tilde{A}.$$

Proof. For any pair of indices i, k the ik-component of $A\tilde{A}$ is

$$a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk} = a_{i1}(-1)^{k+1}D(A_{k1}) + \cdots + a_{in}(-1)^{k+n}D(A_{kn}).$$

If $i = k$, then this sum is simply the expansion of the determinant according to the i-th row, and hence this sum is equal to d. If $i \neq k$, let $\bar{A}$ be the matrix obtained from A by replacing the k-th row by the i-th row, and leaving all other rows unchanged. If we delete the k-th row and the j-th column from $\bar{A}$, we obtain the same matrix as by deleting the k-th row and j-th column from A. Thus

$$\bar{A}_{kj} = A_{kj},$$

and hence our sum above can be written

$$a_{i1}(-1)^{k+1}D(\bar{A}_{k1}) + \cdots + a_{in}(-1)^{k+n}D(\bar{A}_{kn}).$$

This is the expansion of the determinant of $\bar{A}$ according to the i-th row. Hence $D(\bar{A}) = 0$, and our sum is 0. We have therefore proved that the ik-component of $A\tilde{A}$ is equal to d if $i = k$ (i.e. if it is a diagonal component), and is equal to 0 otherwise. This proves that $A\tilde{A} = dI$. On the other hand, we see at once from the definitions that ${}^t\tilde{A} = \widetilde{{}^tA}$. Then

$$ {}^t(\tilde{A}A) = {}^tA\,{}^t\tilde{A} = {}^tA\,\widetilde{{}^tA} = dI, $$

and consequently, $\tilde{A}A = dI$ also, since ${}^t(dI) = dI$. When d is a unit in R, then A is invertible, its inverse being $d^{-1}\tilde{A}$. Conversely, if A is invertible, and $AA^{-1} = I$, then $D(A)D(A^{-1}) = 1$, and hence $D(A)$ is invertible, as was to be shown.

Corollary 4.17. *Let F be any R-module, and let $w_1, \dots, w_n$ be elements of F. Let $A = (a_{ij})$ be an $n \times n$ matrix in R. Let*

$$ a_{11}w_1 + \cdots + a_{1n}w_n = v_1 $$
$$ \cdots $$
$$ a_{n1}w_1 + \cdots + a_{nn}w_n = v_n. $$

Then one can solve explicitly

$$ \begin{pmatrix} D(A)w_1 \\ \vdots \\ D(A)w_n \end{pmatrix} = D(A)\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \tilde{A}\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}. $$

In particular, if $v_i = 0$ for all i, then $D(A)w_i = 0$ for all i. If $v_i = 0$ for all i and F is generated by $w_1, \dots, w_n$, then $D(A)F = 0$.

Proof. This is immediate from the relation $\tilde{A}A = D(A)I$, using the remarks in §3 about applying matrices to column vectors whose components lie in the module.

Proposition 4.18. *Let E, F be free modules of dimension n over R. Let $f: E \to F$ be a linear map. Let $\mathfrak{B}$, $\mathfrak{B}'$ be bases of E, F respectively over R. Then f is an isomorphism if and only if the determinant of its associated matrix $M^{\mathfrak{B}}_{\mathfrak{B}'}(f)$ is a unit in R.*

Proof. Let $A = M^{\mathfrak{B}}_{\mathfrak{B}'}(f)$. By definition, f is an isomorphism if and only if there exists a linear map $g : F \to E$ such that $g \circ f = \text{id}$ and $f \circ g = \text{id}$. If f is an isomorphism, and $B = M^{\mathfrak{B}'}_{\mathfrak{B}}(g)$, then $AB = BA = I$. Taking the determinant of the product, we conclude that $D(A)$ is invertible in R. Conversely, if $D(A)$ is a unit, then we can define A^{-1} by Proposition 4.16. This A^{-1} is the associated matrix of a linear map $g: F \to E$ which is an inverse for f, as desired.

Finally, we shall define the determinant of an endomorphism.

Let E be a free module over R, and let $\mathcal{B}$ be a basis. Let $f: E \to E$ be an endomorphism of E. Let

$$M = M^{\mathcal{B}}_{\mathcal{B}}(f).$$

If $\mathcal{B}'$ is another basis of E, and $M' = M^{\mathcal{B}'}_{\mathcal{B}'}(f)$, then there exists an invertible matrix N such that

$$M' = NMN^{-1}.$$

Taking the determinant, we see that $D(M') = D(M)$. Hence the determinant does not depend on the choice of basis, and will be called the **determinant of the linear map** f. We shall give below a characterization of this determinant which does not depend on the choice of a basis.

Let E be any module. Then we can view $L^n_a(E)$ as a functor in the variable E (contravariant). In fact, we can view $L^n_a(E, F)$ as a functor of two variables, contravariant in the first, and covariant in the second. Indeed, suppose that

$$E' \xrightarrow{f} E$$

is a linear map. To each multilinear map $\varphi : E^{(n)} \to F$ we can associate the composite map $\varphi \circ f^{(n)}$,

$$E' \times \cdots \times E' \xrightarrow{f^{(n)}} E \times \cdots \times E \xrightarrow{\varphi} F$$

where $f^{(n)}$ is the product of f with itself n times. The map

$$L^n_a(f) : L^n_a(E, F) \to L^n_a(E', F)$$

given by

$$\varphi \mapsto \varphi \circ f^{(n)},$$

is obviously a linear map, which defines our functor. We shall sometimes write f^* instead of $L^n_a(f)$.

In particular, consider the case when $E = E'$ and $F = R$. We get an induced map

$$f^* : L^n_a(E) \to L^n_a(E).$$

Proposition 4.19. *Let E be a free module over R, of dimension n. Let $\{\Delta\}$ be a basis of $L^n_a(E)$. Let $f: E \to E$ be an endomorphism of E. Then*

$$f^*\Delta = D(f)\Delta.$$

Proof. This is an immediate consequence of Theorem 4.11. Namely, we let $\{v_1, \dots, v_n\}$ be a basis of E, and then take A (or ${}^tA$) to be a matrix of f relative to this basis. By definition,

$$f^*\Delta(v_1, \dots, v_n) = \Delta(f(v_1), \dots, f(v_n)),$$

and by Theorem 4.11, this is equal to

$$D(A)\,\Delta(v_1, \ldots, v_n).$$

By Corollary 4.12, we conclude that $f^*\Delta = D(A)\Delta$ since both of these forms take on the same value on $(v_1, \ldots, v_n)$.

The above considerations have dealt with the determinant as a function on all endomorphisms of a free module. One can also view it multiplicatively, as a homomorphism.

$$\det : GL_n(R) \to R^*$$

from the group of invertible $n \times n$ matrices over R into the group of units of R. The kernel of this homomorphism, consisting of those matrices with determinant 1, is called the **special linear group**, and is denoted by $SL_n(R)$.

We now give an application of determinants to the situation of a free module and a submodule considered in Chapter III, Theorem 7.8.

Proposition 4.20. *Let R be a principal entire ring. Let F be a free module over R and let M be a finitely generated submodule. Let $\{e_1, \ldots, e_m, \ldots\}$ be a basis of F such that there exist non-zero elements $a_1, \ldots, a_m \in R$ such that:*

(i) *The elements $a_1e_1, \ldots, a_me_m$ form a basis of M over R.*
(ii) *We have $a_i \mid a_{i+1}$ for $i = 1, \ldots, m - 1$.*

Let L_a^s be the set of all s-multilinear alternating forms on F. Let J_s be the ideal generated by all elements $f(y_1, \ldots, y_s)$, with $f \in L_a^s$ and $y_1, \ldots, y_s \in M$. Then

$$J_s = (a_1 \cdots a_s).$$

Proof. We first show that $J_s \subset (a_1 \cdots a_s)$. Indeed, an element $y \in M$ can be written in the form

$$y = c_1a_1e_1 + \cdots + c_ra_re_r.$$

Hence if $y_1, \ldots, y_s \in M$, and f is multilinear alternating on F, then $f(y_1, \ldots, y_s)$ is equal to a sum in terms of type

$$c_{i_1} \cdots c_{i_s} a_{i_1} \cdots a_{i_s} f(e_{i_1}, \ldots, e_{i_s}).$$

This is non-zero only when $e_{i_1}, \ldots, e_{i_s}$ are distinct, in which case the product $a_1 \cdots a_s$ divides this term, and hence J_s is contained in the stated ideal.

Conversely, we show that there exists an s-multilinear alternating form which gives precisely this product. We deduce this from determinants. We can write F as a direct sum

$$F = (e_1, \ldots, e_r) \oplus F_r$$

with some submodule F_r. Let f_i $(i = 1, \dots, r)$ be the linear map $F \to R$ such that $f_i(e_j) = \delta_{ij}$, and such that f_i has value 0 on F_r. For $v_1, \dots, v_s \in F$ we define

$$f(v_1, \dots, v_s) = \det(f_i(v_j)).$$

Then f is multilinear alternating and takes on the value

$$f(e_2, \dots, e_s) = 1,$$

as well as the value

$$f(a_1e_1, \dots, a_se_s) = a_1 \cdots a_s.$$

This proves the proposition.

The uniqueness of Chapter III, Theorem 7.8 is now obvious, since first (a_1) is unique, then (a_1a_2) is unique and the quotient (a_2) is unique, and so forth by induction.

Remark. Compare the above theorem with Theorem 2.9 of Chapter XIX, in the theory of Fitting ideals, which gives a fancier context for the result.

§5. DUALITY

Let R be a commutative ring, and let E, F be modules over R. An ***R*-bilinear form** on $E \times F$ is a map

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

having the following properties: For each $x \in E$, the map

$$y \mapsto f(x, y)$$

is R-linear, and for each $y \in F$, the map

$$x \mapsto f(x, y)$$

is R-linear. We shall omit the prefix R- in the rest of this section, and write $\langle x, y\rangle_f$ or $\langle x, y\rangle$ instead of $f(x, y)$. If $x \in F$, we write $x \perp y$ if $\langle x, y\rangle = 0$. Similarly, if S is a subset of F, we define $x \perp S$ if $x \perp y$ for all $y \in S$. We then say that x is **perpendicular** to S. We let $S^\perp$ consist of all elements of E which are perpendicular to S. It is obviously a submodule of E. We define perpendicularity on the other side in the same way. We define the **kernel** of f on the left to be $F^\perp$ and the kernel on the right to be $E^\perp$. We say that f is **non-degenerate** on the left if its kernel on the left is 0. We say that f is **non-degenerate** on the right if its kernel on the right is 0. If E_0 is the kernel of f on the left, then we get an induced bilinear map

$$E/E_0 \times F \to R$$

which is non-degenerate on the left, as one verifies trivially from the definitions. Similarly, if F_0 is the kernel of f on the right, we get an induced bilinear map

$$E/E_0 \times F/F_0 \to R$$

which is non-degenerate on either side. This map arises from the fact that the value $\langle x, y\rangle$ depends only on the coset of x modulo E_0 and the coset of y modulo F_0.

We shall denote by $L^2(E, F; R)$ the set of all bilinear maps of $E \times F$ into R. It is clear that this set is a module (i.e. an R-module), addition of maps being the usual one, and also multiplication of maps by elements of R.

The form f gives rise to a homomorphism

$$\varphi_f : E \to \mathrm{Hom}_R(F, R)$$

such that

$$\varphi_f(x)(y) = f(x, y) = \langle x, y\rangle,$$

for all $x \in E$ and $y \in F$. We shall call $\mathrm{Hom}_R(F, R)$ the **dual module** of F, and denote it by $F^\vee$. We have an *isomorphism*

$$\boxed{L^2(E, F; R) \leftrightarrow \mathrm{Hom}_R(E, \mathrm{Hom}_R(F, R))}$$

given by $f \mapsto \varphi_f$, its inverse being defined in the obvious way: If

$$\varphi : E \to \mathrm{Hom}_R(F, R)$$

is a homomorphism, we let f be such that

$$f(x, y) = \varphi(x)(y).$$

We shall say that f is **non-singular on the left** if φ_f is an isomorphism, in other words if our form can be used to identify E with the dual module of F. We define **non-singular on the right** in a similar way, and say that f is **non-singular** if it is non-singular on the left and on the right.

Warning: Non-degeneracy does not necessarily imply non-singularity.

We shall now obtain an *isomorphism*

$$\boxed{\mathrm{End}_R(E) \mapsto L^2(E, F; R)}$$

depending on a fixed non-singular bilinear map $f : E \times F \to R$.

Let $A \in \mathrm{End}_R(E)$ be a linear map of E into itself. Then the map

$$(x, y) \mapsto \langle Ax, y\rangle = \langle Ax, y\rangle_f$$

is bilinear, and in this way, we associate linearly with each $A \in \mathrm{End}_R(E)$ a bilinear map in $L^2(E, F; R)$.

Conversely, let $h : E \times F \to R$ be bilinear. Given $x \in E$, the map $h_x : F \to R$ such that $h_x(y) = h(x, y)$ is linear, and is in the dual space $F^\vee$. By assumption, there exists a unique element $x' \in E$ such that for all $y \in F$ we have

$$h(x, y) = \langle x', y\rangle.$$

It is clear that the association $x \mapsto x'$ is a linear map of E into itself. Thus with each bilinear map $E \times F \to R$ we have associated a linear map $E \to E$.

It is immediate that the mappings described in the last two paragraphs are inverse isomorphisms between $\mathrm{End}_R(E)$ and $L^2(E, F; R)$. We emphasize of course that they depend on our form f.

Of course, we could also have worked on the right, and thus we have a similar *isomorphism*

$$\boxed{L^2(E, F; R) \leftrightarrow \mathrm{End}_R(F)}$$

depending also on our fixed non-singular form f.

As an application, let $A : E \to E$ be linear, and let $(x, y) \mapsto \langle Ax, y\rangle$ be its associated bilinear map. There exists a unique linear map

$${}^tA : F \to F$$

such that

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

for all $x \in E$ and $y \in F$. We call ${}^tA$ the **transpose of** A **with respect to** f.

It is immediately clear that if, A, B are linear maps of E into itself, then for $c \in R$,

$${}^t(cA) = c\,{}^tA, \qquad {}^t(A + B) = {}^tA + {}^tB, \quad \text{and} \quad {}^t(AB) = {}^tB\,{}^tA.$$

More generally, let E, F be modules with non-singular bilinear forms denoted by $\langle\ ,\ \rangle_E$ and $\langle\ ,\ \rangle_F$ respectively. Let $A: E \to F$ be a linear map. Then by the non-singularity of $\langle\ ,\ \rangle_E$ there exists a unique linear map ${}^tA: F \to E$ such that

$$\langle Ax, y\rangle_F = \langle x, {}^tAy\rangle_E \text{ for all } x \in E \text{ and } y \in F.$$

We also call ${}^tA$ the **transpose** with respect to these forms.

Examples. For a nice classical example of a transpose, see Exercise 33. For the systematic study when a linear map is equal to its transpose, see the spectral theorems of Chapter XV. Next I give another example of a transpose from analysis as follows. Let E be the (infinite dimensional) vector space of C^∞ functions on $\mathbf{R}$, having compact support, i.e. equal to 0 outside some finite interval. We define the scalar product

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)dx.$$

Let $D: E \to E$ be the derivative. Then one has the formula

$$\langle Df, g\rangle = -\langle f, Dg\rangle.$$

Thus one says that ${}^tD = -D$, even though the scalar product is not "non-singular", but much of the formalism of non-singular forms goes over. Also in analysis, one puts various norms on the spaces and one extends the bilinear form by continuity to the completions, thus leaving the domain of algebra to enter the domain of estimates (analysis). Then the spectral theorems become more complicated in such analytic contexts.

Let us assume that $E = F$. Let $f: E \times E \to R$ be bilinear. By an **automorphism of the pair (E,f)**, or simply of f, we shall mean a linear automorphism $A: E \to E$ such that

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

for all $x, y \in E$. The group of automorphisms of f is denoted by $\mathrm{Aut}(f)$.

Proposition 5.1. *Let $f: E \times E \to R$ be a non-singular bilinear form. Let $A: E \to E$ be a linear map. Then A is an automorphism of f if and only if ${}^tAA = \mathrm{id}$, and A is invertible.*

Proof. From the equality

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

holding for all $x, y \in E$, we conclude that ${}^tAA = \mathrm{id}$ if A is an automorphism of f. The converse is equally clear.

Note. If E is free and finite dimensional, then the condition ${}^tAA = \mathrm{id}$ implies that A is invertible.

Let $f: E \times E \to R$ be a bilinear form. We say that f is **symmetric** if $f(x, y) = f(y, x)$ for all $x, y \in E$. The set of symmetric bilinear forms on E will be denoted by $L_s^2(E)$. Let us take a fixed symmetric non-singular bilinear form f on E, denoted by $(x, y) \mapsto \langle x, y\rangle$. An endomorphism $A: E \to E$ will be said to be **symmetric with respect** to f if ${}^tA = A$. It is clear that the set of symmetric endomorphisms of E is a module, which we shall denote by $\mathrm{Sym}(E)$.

Depending on our fixed symmetric non-singular f, we have an isomorphism

$$\boxed{L_s^2(E) \leftrightarrow \mathrm{Sym}(E)}$$

which we describe as follows. If g is symmetric bilinear on E, then there exists a unique linear map A such that

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

for all $x, y \in E$. Using the fact that both f, g are symmetric, we obtain

$$\langle Ax, y\rangle = \langle Ay, x\rangle = \langle y, {}^tAx\rangle = \langle {}^tAx, y\rangle.$$

Hence $A = {}^tA$. The association $g \mapsto A$ gives us a homomorphism from $L_s^2(E)$ into $\mathrm{Sym}(E)$. Conversely, given a symmetric endomorphism A of E, we can define a symmetric form by the rule $(x, y) \mapsto \langle Ax, y\rangle$, and the association of this form to A clearly gives a homomorphism of $\mathrm{Sym}(E)$ into $L_s^2(E)$ which is inverse to the preceding homomorphism. Hence $\mathrm{Sym}(E)$ and $L_s^2(E)$ are isomorphic.

We recall that a bilinear form $g : E \times E \to R$ is said to be **alternating** if $g(x, x) = 0$ for all $x \in E$, and consequently $g(x, y) = -g(y, x)$ for all $x, y \in E$. The set of bilinear alternating forms on E is a module, denoted by $L_a^2(E)$.

Let f be a fixed *symmetric* non-singular bilinear form on E. An endomorphism $A : E \to E$ will be said to be **skew-symmetric** or **alternating** with respect to f if ${}^tA = -A$, and also $\langle Ax, x\rangle = 0$ for all $x \in E$. If for all $a \in R$, $2a = 0$ implies $a = 0$, then this second condition $\langle Ax, x\rangle = 0$ is redundant, because $\langle Ax, x\rangle = -\langle Ax, x\rangle$ implies $\langle Ax, x\rangle = 0$. It is clear that the set of alternating endomorphisms of E is a module, denoted by $\mathrm{Alt}(E)$. *Depending on our fixed symmetric non-singular form f, we have an isomorphism*

$$\boxed{L_a^2(E) \leftrightarrow \mathrm{Alt}(E)}$$

described as usual. If g is an alternating bilinear form on E, its corresponding linear map A is the one such that

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

for all $x, y \in E$. One verifies trivially in a manner similar to the one used in the symmetric case that the correspondence $g \leftrightarrow A$ gives us our desired isomorphism.

Examples. Let k be a field and let E be a finite-dimensional vector space over k. Let $f : E \times E \to E$ be a bilinear map, denoted by $(x, y) \mapsto xy$. To each

$x \in E$, we associate the linear map $\lambda_x : E \mapsto E$ such that

$$\lambda_x(y) = xy.$$

Then the map obtained by taking the trace, namely

$$(x, y) \mapsto \operatorname{tr}(\lambda_{xy})$$

is a bilinear form on E. If $xy = yx$, then this bilinear form is symmetric.

Next, let E be the space of continuous functions on the interval $[0, 1]$. Let $K(s, t)$ be a continuous function of two real variables defined on the square $0 \leqq s \leqq 1$ and $0 \leqq t \leqq 1$. For $\varphi, \psi \in E$ we define

$$\langle \varphi, \psi \rangle = \iint \varphi(s)K(s, t)\psi(t)\, ds\, dt,$$

the double integral being taken on the square. Then we obtain a bilinear form on E. If $K(s, t) = K(t, s)$, then the bilinear form is symmetric. When we discuss matrices and bilinear forms in the next section, the reader will note the similarity between the preceding formula and the bilinear form defined by a matrix.

Thirdly, let U be an open subset of a real Banach space E (or a finite-dimensional Euclidean space, if the reader insists), and let $f : U \to \mathbf{R}$ be a map which is twice continuously differentiable. For each $x \in U$, the derivative $Df(x) : E \to \mathbf{R}$ is a continuous linear map, and the second derivative $D^2f(x)$ can be viewed as a continuous symmetric bilinear map of $E \times E$ into $\mathbf{R}$.

§6. MATRICES AND BILINEAR FORMS

We shall investigate the relation between the concepts introduced above and matrices. Let $f : E \times F \to R$ be bilinear. Assume that E, F are free over R. Let $\mathcal{B} = \{v_1, \ldots, v_m\}$ be a basis for E over R, and let $\mathcal{B}' = \{w_1, \ldots, w_n\}$ be a basis for F over R. Let $g_{ij} = \langle v_i, w_j \rangle$. If

$$x = x_1v_1 + \cdots + x_mv_m$$

and

$$y = y_1w_1 + \cdots + y_nw_n$$

are elements of E and F respectively, with coordinates $x_i, y_j \in R$, then

$$\langle x, y \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} g_{ij}x_iy_j.$$

Let X, Y be the column vectors of coordinates for x, y respectively, with respect to our bases. Then

$$\langle x, y\rangle = {}^tXGY$$

where G is the matrix (g_{ij}). We could write $G = M^{\mathcal{B}}_{\mathcal{B}'}(f)$. We call G the **matrix associated with the form f relative to the bases $\mathcal{B}$, $\mathcal{B}'$**.

Conversely, given a matrix G (of size $m \times n$), we get a bilinear form from the map

$$(X, Y) \mapsto {}^tXGY.$$

In this way, we get a correspondence from bilinear forms to matrices and back, and it is clear that this correspondence induces an *isomorphism* (of R-modules)

$$\boxed{L^2(E, F; R) \leftrightarrow \mathrm{Mat}_{m\times n}(R)}$$

given by

$$f \mapsto M^{\mathcal{B}}_{\mathcal{B}'}(f).$$

The two maps between these two modules which we described above are clearly inverse to each other.

If we have bases $\mathcal{B} = \{v_1, \dots, v_n\}$ and $\mathcal{B}' = \{w_1, \dots, w_n\}$ such that $\langle v_i, w_j\rangle = \delta_{ij}$, then we say that these bases are **dual** to each other. In that case, if X is the coordinate vector of an element of E, and Y the coordinate vector of an element of F, then the bilinear map on X, Y has the value

$$X \cdot Y = x_1y_1 + \cdots + x_ny_n$$

given by the usual dot product.

It is easy to derive in general how the matrix G changes when we change bases in E and F. However, we shall write down the explicit formula only when $E = F$ and $\mathcal{B} = \mathcal{B}'$. Thus we have a bilinear form $f: E \times E \to R$. Let $\mathcal{C}$ be another basis of E and write $X_{\mathcal{B}}$ and $X_{\mathcal{C}}$ for the column vectors belonging to an element x of E, relative to the two bases. *Let C be the invertible matrix* $M^{\mathcal{C}}_{\mathcal{B}}(\mathrm{id})$, so that

$$X_{\mathcal{B}} = CX_{\mathcal{C}}.$$

Then our form is given by

$$\langle x, y\rangle = {}^tX_{\mathcal{C}}{}^tCGCY_{\mathcal{C}}.$$

We see that

$$M^{\mathcal{C}}_{\mathcal{C}}(f) = {}^tCM^{\mathcal{B}}_{\mathcal{B}}(f)C. \tag{1}$$

In other words, the matrix of the bilinear form changes by the *transpose*.

If F is free over R, with a basis $\{\eta_1, \dots, \eta_n\}$, then $\mathrm{Hom}_R(F, R)$ *is also free, and we have a dual basis $\{\eta'_1, \dots, \eta'_n\}$ such that*

$$\eta'_i(\eta_j) = \delta_{ij}.$$

This has already been mentioned in Chapter III, Theorem 6.1.

Proposition 6.1. *Let E, F be free modules of dimension n over R and let $f : E \times F \to R$ be a bilinear form. Then the following conditions are equivalent:*

f is non-singular on the left.
f is non-singular on the right.
f is non-singular.
The determinant of the matrix of f relative to any bases is invertible in R.

Proof. Assume that f is non-singular on the left. Fix bases of E and F relative to which we write elements of these modules as column vectors, and giving rise to the matrix G for f. Then our form is given by

$$(X, Y) \mapsto {}^tXGY$$

where X, Y are column vectors with coefficients in R. By assumption the map

$$X \mapsto {}^tXG$$

gives an isomorphism between the module of column vectors, and the module of row vectors of length n over R. Hence G is invertible, and hence its determinant is a unit in R. The converse is equally clear, and if $\det(G)$ is a unit, we see that the map

$$Y \to GY$$

must also be an isomorphism between the module of column vectors and itself. This proves our assertion.

We shall now investigate how the transpose behaves in terms of matrices. Let E, F be free over R, of dimension n.

Let $f : E \times F \to R$ be a non-singular bilinear form, and assume given a basis $\mathcal{B}$ of E and $\mathcal{B}'$ of F. Let G be the matrix of f relative to these bases. Let $A : E \to E$ be a linear map. If $x \in E$, $y \in F$, let X, Y be their column vectors relative to $\mathcal{B}$, $\mathcal{B}'$. Let M be the matrix of A relative to $\mathcal{B}$. Then for $x \in E$ and $y \in F$ we have

$$\langle Ax, y \rangle = {}^t(MX)GY = {}^tX{}^tMGY.$$

Let N be the matrix of ${}^tA$ relative to the basis $\mathcal{B}'$. Then NY is the column vector of ${}^tAy$ relative to $\mathcal{B}'$. Hence

$$\langle x, {}^tAy \rangle = {}^tXGNY.$$

From this we conclude that ${}^tMG = GN$, and since G is invertible, we can solve for N in terms of M. We get:

Proposition 6.2. *Let E, F be free over R, of dimension n. Let $f : E \times F \to R$ be a non-singular bilinear form. Let $\mathcal{B}$, $\mathcal{B}'$ be bases of E and F respectively over R, and let G be the matrix of f relative to these bases. Let $A : E \to E$ be a linear map, and let M be its matrix relative to $\mathcal{B}$. Then the matrix of ${}^tA$ relative to $\mathcal{B}'$ is*

$$(G^{-1})^t MG.$$

Corollary 6.3. *If G is the unit matrix, then the matrix of the transpose is equal to the transpose of the matrix.*

In terms of matrices and bases, we obtain the following characterization for a matrix to induce an automorphism of the form.

Corollary 6.4. *Let the notation be as in Proposition 6.2, and let $E = F$, $\mathcal{B} = \mathcal{B}'$. An $n \times n$ matrix M is the matrix of an automorphism of the form f (relative to our basis) if and only if*

$${}^tMGM = G.$$

If this condition is satisfied, then in particular, M is invertible.

Proof. We use the definitions, together with the formula given in Proposition 6.2. We note that M is invertible, for instance because its determinant is a unit in R.

A matrix M is said to be **symmetric** (resp. **alternating**) if ${}^tM = M$ (resp. ${}^tM = -M$ and the diagonal elements of M are 0).

Let $f : E \times E \to R$ be a bilinear form. We say that f is **symmetric** if $f(x, y) = f(y, x)$ for all $x, y \in E$. We say that f is **alternating** if $f(x, x) = 0$ for all $x \in E$.

Proposition 6.5. *Let E be a free module of dimension n over R, and let $\mathcal{B}$ be a fixed basis. The map*

$$f \mapsto M_{\mathcal{B}}^{\mathcal{B}}(f)$$

induces an isomorphism between the module of symmetric bilinear forms on $E \times E$ (resp. the module of alternating forms on $E \times E$) and the module of symmetric $n \times n$ matrices over R (resp. the module of alternating $n \times n$ matrices over R).

Proof. Consider first the symmetric case. Assume that f is symmetric. In terms of coordinates, let $G = M^{\mathcal{B}}_{\mathcal{B}}(f)$. Our form is given by ${}^tXGY$ which must be equal to ${}^tYGX$ by symmetry. However, ${}^tXGY$ may be viewed as a 1×1 matrix, and is equal to its transpose, namely ${}^tY{}^tGX$. Thus

$${}^tYGX = {}^tY{}^tGX$$

for all vectors X, Y. It follows that $G = {}^tG$. Conversely, it is clear that any symmetric matrix defines a symmetric form.

As for the alternating case, replacing x by $x + y$ in the relation $\langle x, x\rangle = 0$ we obtain

$$\langle x, y\rangle + \langle y, x\rangle = 0.$$

In terms of the coordinate vectors X, Y and the matrix G, this yields

$${}^tXGY + {}^tYGX = 0.$$

Taking the transpose of, say, the second of the 1×1 matrices entering in this relation, yields (for all X, Y):

$${}^tXGY + {}^tX{}^tGY = 0.$$

Hence $G + {}^tG = 0$. Furthermore, letting X be any one of the unit vectors

$${}^t(0, \dots, 0, 1, 0, \dots, 0)$$

and using the relation ${}^tXGX = 0$, we see that the diagonal elements of G must be equal to 0. Conversely, if G is an $n \times n$ matrix such that ${}^tG + G = 0$, and such that $g_{ii} = 0$ for $i = 1, \dots, n$ then one verifies immediately that the map

$$(X, Y) \mapsto {}^tXGY$$

defines an alternating form. This proves our proposition.

Of course, if as is usually the case, 2 is invertible in R, then our condition ${}^tM = -M$ implies that the diagonal elements of M must be 0. Thus in that case, showing that $G + {}^tG = 0$ implies that G is alternating.

§7. SESQUILINEAR DUALITY

There exist forms which are not quite bilinear, and for which the results described above hold almost without change, but which must be handled separately for the sake of clarity in the notation involved.

Let R have an automorphism of period 2. We write this automorphism as $a \mapsto \bar{a}$ (and think of complex conjugation).

Following Bourbaki, we say that a map

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

is a **sesquilinear form** if it is **Z**-bilinear, and if for $x \in E$, $y \in F$, and $a \in R$ we have

$$f(ax, y) = af(x, y)$$

and

$$f(x, ay) = \bar{a}f(x, y).$$

(**Sesquilinear** means $1\frac{1}{2}$ times linear, so the terminology is rather good.)

Let E, E' be modules. A map $\varphi : E \to E'$ is said to be **anti-linear** (or **semi-linear**) if it is **Z**-linear, and $\varphi(ax) = \bar{a}\varphi(x)$ for all $x \in E$. Thus we may say that a sesquilinear form is linear in its first variable, and anti-linear in its second variable. We let $\overline{\mathrm{Hom}}_R(E, E')$ denote the module of anti-linear maps of E into E'.

We shall now go systematically through the same remarks that we made previously for bilinear forms.

We define perpendicularity as before, and also the kernel on the right and on the left for any sesquilinear form f. These kernels are submodules, say E_0 and F_0, and we get an induced sesquilinear form

$$E/E_0 \times F/F_0 \to R,$$

which is non-degenerate on either side.

Let F be an R-module. We define its **anti-module** $\bar{F}$ to be the module whose additive group is the same as F, and such that the operation $R \times \bar{F} \to \bar{F}$ is given by

$$(a, y) \mapsto \bar{a}y.$$

Then $\bar{F}$ is a module. We have a natural isomorphism

$$\mathrm{Hom}_R(\bar{F}, R) \leftrightarrow \overline{\mathrm{Hom}}_R(F, R),$$

as R-modules.

The sesquilinear form $f : E \times F \to R$ induces a linear map

$$\varphi_f : E \to \mathrm{Hom}_R(\bar{F}, R).$$

We say that f is **non-singular on the left** if φ_f is an isomorphism. Similarly, we have a corresponding linear map

$$\varphi'_f : \bar{F} \to \mathrm{Hom}_R(E, R)$$

from $\bar{F}$ into the dual space of E, and we say that f is **non-singular on the right** if φ'_f is an isomorphism. We say that f is **non-singular** if it is non-singular on the left and on the right.

We observe that our sesquilinear form f can be viewed as a **bilinear** form

$$f : E \times \bar{F} \to R,$$

and that our notions of non-singularity are then compatible with those defined previously for bilinear forms.

If we have a fixed non-singular sesquilinear form on $E \times F$, then depending on this form, we obtain an isomorphism between the module of sesquilinear forms on $E \times F$ and the module of endomorphisms of E. We also obtain an anti-isomorphism between these modules and the module of endomorphisms of F. In particular, we can define the analogue of the transpose, which in the present case we shall call the adjoint. Thus, let $f : E \times F \to R$ be a non-singular sesquilinear form. Let $A : E \to E$ be a linear map. There exists a unique linear map

$$A^* : F \to F$$

such that

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

for all $x \in E$ and $y \in F$. Note that A^* is linear, not anti-linear. We call A^* the **adjoint** of A with respect to our form f. We have the rules

$$(cA)^* = \bar{c}A^*, \qquad (A + B)^* = A^* + B^*, \qquad (AB)^* = B^*A^*$$

for all linear maps A, B of E into itself, and $c \in R$.

Let us assume that $E = F$. Let $f : E \times E \to R$ be sesquilinear. By an **automorphism** of f we shall mean a linear automorphism $A : E \to E$ such that

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

just as we did for bilinear forms.

Proposition 7.1. *Let $f : E \times E \to R$ be a non-singular sesquilinear form. Let $A : E \to E$ be a linear map. Then A is an automorphism of f if and only if $A^*A = \mathrm{id}$, and A is invertible.*

The proof, and also the proofs of subsequent propositions, which are completely similar to those of the bilinear case, will be omitted.

A sesquilinear form $g : E \times E \to R$ is said to be **hermitian** if

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

for all $x, y \in E$. The set of hermitian forms on E will be denoted by $L^2_h(E)$. Let R_0 be the subring of R consisting of all elements fixed under our automorphism

$a \to \bar{a}$ (i.e. consisting of all elements $a \in R$ such that $a = \bar{a}$). Then $L^2_h(E)$ is an R_0-module.

Let us take a fixed hermitian non-singular form f on E, denoted by $(x, y) \mapsto \langle x, y \rangle$. An endomorphism $A : E \to E$ will be said to be **hermitian** with respect to f if $A^* = A$. It is clear that the set of hermitian endomorphisms is an R_0-module, which we shall denote by Herm(E). *Depending on our fixed hermitian non-singular form f, we have an R_0-isomorphism*

$$\boxed{L^2_h(E) \leftrightarrow \mathrm{Herm}(E)}$$

described in the usual way. A hermitian form g corresponds to a hermitian map A if and only if

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

for all $x, y \in E$.

We can now describe the relation between our concepts and matrices, just as we did with bilinear forms.

We start with a sesquilinear form $f : E \times F \to R$.

If E, F are free, and we have selected bases as before, then we can again associate a matrix G with the form, and in terms of coordinate vectors X, Y our sesquilinear form is given by

$$(X, Y) \mapsto {}^tXG\bar{Y},$$

where $\bar{Y}$ is obtained from Y by applying the automorphism to each component of Y.

If $E = F$ and we use the same basis on the right and on the left, then with the same notation as that used in formula (1), if f is sesquilinear, the formula now reads

$$\text{(1S)} \qquad M^{\mathscr{C}}_{\mathscr{C}}(f) = {}^tCM^{\mathscr{B}}_{\mathscr{B}}(f)\bar{C}.$$

The automorphism appears.

Proposition 7.2. *Let E, F be free modules of dimension n over R, and let $f : E \times F \to R$ be a sesquilinear form. Then the following conditions are equivalent.*

f is non-singular on the left.
f is non-singular on the right.
f is non-singular.

The determinant of the matrix of f relative to any bases is invertible in R.

Proposition 7.3. *Let E, F be free over R, of dimension n. Let $f : E \times F \to R$ be a non-singular sesquilinear form. Let $\mathcal{B}$, $\mathcal{B}'$ be bases of E and F respectively over R, and let G be the matrix of f relative to these bases. Let $A : E \to E$ be a linear map, and let M be its matrix relative to $\mathcal{B}$. Then the matrix of A^* relative to $\mathcal{B}'$ is*

$$(\bar{G}^{-1})^t \bar{M} \bar{G}.$$

Corollary 7.4. *If G is the unit matrix, then the matrix of A^* is equal to ${}^t\bar{M}$.*

Corollary 7.5. *Let the notation be as in the proposition, and let $\mathcal{B} = \mathcal{B}'$ be a basis of E. An $n \times n$ matrix M is the matrix of an automorphism of f (relative to our basis) if and only if*

$${}^tMG\bar{M} = G.$$

A matrix M is said to be **hermitian** if ${}^tM = \bar{M}$.

Let R_0 be as before the subring of R consisting of all elements fixed under our automorphism $a \mapsto \bar{a}$ (i.e. consisting of all elements $a \in R$ such that $a = \bar{a}$).

Proposition 7.6. *Let E be a free module of dimension n over R, and let $\mathcal{B}$ be a basis. The map*

$$f \mapsto M_{\mathcal{B}}^{\mathcal{B}}(f)$$

induces an R_0-isomorphism between the R_0-module of hermitian forms on E and the R_0-module of $n \times n$ hermitian matrices in R.

Remark. If we had assumed at the beginning that our automorphism $a \mapsto \bar{a}$ has period 2 or 1 (i.e. if we allow it to be the identity), then the results on bilinear and symmetric forms become special cases of the results of this section. However, the notational differences are sufficiently disturbing to warrant a repetition of the results as we have done.

Terminology

For some confusing reason, the group of automorphisms of a symmetric (resp. alternating, resp. hermitian) form on a vector space is called the **orthogonal** (resp. **symplectic**, resp. **unitary**) group of the form. The word orthogonal is especially unfortunate, because an orthogonal map preserves more than orthogonality: It also preserves the scalar product, i.e. length. Furthermore, the word symplectic is also unfortunate. It turns out that one can carry out a discussion of hermitian forms over certain division rings (having automorphisms of order 2), and their group of automorphisms have also been called symplectic, thereby creating genuine confusion with the use of the word relative to alternating forms.

In order to unify and improve the terminology, I have discussed the matter with several persons, and it seems that one could adopt the following conventions.

As said in the text, the group of automorphisms of any form f is denoted by $\mathrm{Aut}(f)$.

On the other hand, there is a standard form, described over the real numbers in terms of coordinates by

$$f(x, x) = x_1^2 + \cdots + x_n^2,$$

over the complex numbers by

$$f(x, x) = x_1\bar{x}_1 + \cdots + x_n\bar{x}_n,$$

and over the quaternions by the same formula as in the complex case. The group of automorphisms of this form would be called the **unitary group**, and be denoted by U_n. The points of this group in the reals (resp. complex, resp. quaternions) would be denoted by

$$U_n(\mathbf{R}), \qquad U_n(\mathbf{C}), \qquad U_n(\mathbf{K}),$$

and these three groups would be called the **real unitary group** (resp. **complex unitary group**, resp. **quaternion unitary group**). Similarly, the group of points of U_n in any subfield or subring k of the quaternions would be denoted by $U_n(k)$.

Finally, if f is the standard alternating form, whose matrix is

$$\begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix},$$

one would denote its group of automorphisms by A_{2n}, and call it the **alternating form group**, or simply the alternating group, if there is no danger of confusion with the permutation group. The group of points of the alternating form group in a field k would then be denoted by $A_{2n}(k)$.

As usual, the subgroup of $\mathrm{Aut}(f)$ consisting of those elements whose determinant is 1 would be denoted by adding the letter S in front, and would still be called the **special group**. In the four standard cases, this yields

$$SU_n(\mathbf{R}), \qquad SU_n(\mathbf{C}), \qquad SU_n(\mathbf{K}), \qquad SA_{2n}(k).$$

§8. THE SIMPLICITY OF $SL_2(F)/\pm 1$

Let F be a field. Let n be a positive integer. By $GL_n(F)$ we mean the group of $n \times n$ invertible matrices over F. By $SL_n(F)$ we mean the subgroup of those matrices whose determinant is equal to 1. By $PGL_n(F)$ we mean the factor group of $GL_n(F)$ by the subgroup of scalar matrices (which are in the center).

Similarly for $PSL_n(F)$. In this section, we are interested in giving an application of matrices to the group theoretic structure of SL_2. The analogous statements for SL_n with $n \geqq 3$ will be proved in the next section.

The **standard Borel subgroup** B of GL_2 is the group of all matrices

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

with $a, b, d \in F$ and $ad \neq 0$. For the Borel subgroup of SL_2, we require in addition that $ad = 1$. By a **Borel subgroup** we mean a subgroup which is conjugate to the standard Borel subgroup (whether in GL_2 or SL_2). We let U be the group of matrices

$$u(b) = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \quad \text{with } b \in F.$$

We let A be the group of diagonal matrices

$$\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}, \quad \text{with } a, d \in F^*.$$

Let

$$s(a) = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \qquad \text{with } a \in F^*$$

and

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

For the rest of this section, we let

$$G = GL_2(F) \quad \text{or} \quad SL_2(F).$$

Lemma 8.1. *The matrices*

$$X(b) = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad Y(c) = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$$

generate $SL_2(F)$.

Proof. Multiplying an arbitrary element of $SL_2(F)$ by matrices of the above type on the right and on the left corresponds to elementary row and column operations, that is adding a scalar multiple of a row to the other, etc. Thus a given matrix can always be brought into a form

$$\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$$

by such multiplications. We want to express this matrix with $a \neq 1$ in the form

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ d & 1 \end{pmatrix}.$$

Matrix multiplication will show that we can solve this equation, by selecting x arbitrarily $\neq 0$, then solving for b, c, and d successively so that

$$1 + bx = a, \quad c = \frac{-x}{1 + bx}, \quad d = \frac{-b}{1 + bc}.$$

Then one finds $1 + bc = (1 + xb)^{-1}$ and the two symmetric conditions

$$b + bcd + d = 0$$
$$c + bcx + x = 0,$$

so we get what we want, and thereby prove the lemma.

Let $\overline{U}$ be the group of lower matrices

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

Then we see that

$$wUw^{-1} = \overline{U}.$$

Also note the commutation relation

$$w\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}w^{-1} = \begin{pmatrix} d & 0 \\ 0 & a \end{pmatrix},$$

so w normalizes A. Similarly,

$$wBw^{-1} = \overline{B}$$

is the group of lower triangular matrices.

We note that

$$B = AU = UA,$$

and also that A normalizes U.

There is a decomposition of G into disjoint subsets

$$G = B \cup BwB.$$

Indeed, view G as operating on the left of column vectors. The isotropy group of

$$e^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

is obviously U. The orbit Be^1 consists of all column vectors whose second component is 0. On the other hand,

$$we^1 = \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$

and therefore the orbit Bwe^1 consists of all vectors whose second component is $\neq 0$, and whose first component is arbitrary. Since these two orbits of B and BwB cover the orbit Ge^1, it follows that the union of B and BwB is equal to G (because the isotropy group U is contained in B), and they are obviously disjoint. This decomposition is called the **Bruhat decomposition**.

Proposition 8.2. *The Borel subgroup B is a maximal proper subgroup.*

Proof. By the Bruhat decomposition, any element not in B lies in BwB, so the assertion follows since B, BwB cover G.

Theorem 8.3. *If F has at least four elements, then $SL_2(F)$ is equal to its own commutator group.*

Proof. We have the commutator relation (by matrix multiplication)

$$s(a)u(b)s(a)^{-1}u(b)^{-1} = u(ba^2 - b) = u(b(a^2 - 1)).$$

Let $G = SL_2(F)$ for this proof. We let G' be the commutator subgroup, and similarly let B' be the commutator subgroup of B. We prove the first assertion that $G = G'$. From the hypothesis that F has at least four elements, we can find an element $a \neq 0$ in F such that $a^2 \neq 1$, whence the commutator relation shows that $B' = U$. It follows that $G' \supset U$, and since G' is normal, we get

$$G' \supset wUw^{-1}.$$

From Lemma 8.1, we conclude that $G' = G$.

Let Z denote the center of G. It consists of $\pm I$, that is $\pm$ the identity 2×2 matrix if $G = SL_2(F)$; and Z is the subgroup of scalar matrices if $G = GL_2(F)$.

Theorem 8.4. *If F has at least four elements, then $SL_2(F)/Z$ is simple.*

The proof will result from two lemmas.

Lemma 8.5. *The intersection of all conjugates of B in G is equal to Z.*

Proof. We leave this to the reader, as a simple fact using conjugation with w.

Lemma 8.6. *Let $G = SL_2(F)$. If H is normal in G, then either $H \subset Z$ or $H \supset G'$.*

Proof. By the maximality of B we must have

$$HB = B \quad \text{or} \quad HB = G.$$

If $HB = B$ then $H \subset B$. Since H is normal, we conclude that H is contained in every conjugate of B, whence in the center by Lemma 8.5. On the other hand, suppose that $HB = G$. Write

$$w = hb$$

with $h \in H$ and $b \in B$. Then

$$wUw^{-1} = \bar{U} = hbUb^{-1}h^{-1} = hUh^{-1} \subset HU$$

because H is normal. Since $U \subset HU$ and $U, \bar{U}$ generate $SL_2(F)$, it follows that $HU = G$. Hence

$$G/H = HU/H \approx U/(U \cap H)$$

is abelian, whence $H \supset G'$, as was to be shown.

The simplicity of Theorem 8.4 is an immediate consequence of Lemma 8.6.

§9. THE GROUP $SL_n(F)$, $n \geqq 3$.

In this section we look at the case with $n \geqq 3$, and follow parts of Artin's *Geometric Algebra*, Chapter IV. (Artin even treats the case of a non-commutative division algebra as the group ring, but we omit this for simplicity.)

For $i, j = 1, \ldots, n$ and $i \neq j$ and $c \in F$, we let

$$E_{ij}(c) = \begin{pmatrix} 1 & & & \\ & 1 & & 0 \\ 0 & & \ddots & \\ & c_{ij} & & \\ 0 & & 0 & 1 \end{pmatrix}$$

be the matrix which differs from the unit matrix by having c in the ij-component instead of 0. We call such $E_{ij}(c)$ an **elementary matrix**. Note that

$$\det E_{ij}(c) = 1.$$

If A is any $n \times n$ matrix, then multiplication $E_{ij}(c)A$ on the left adds c times the j-th row to the i-th row of A. Multiplication $AE_{ij}(c)$ on the right adds c times the i-th column to the j-th column. We shall mostly multiply on the left.

For fixed $i \neq j$ the map

$$c \mapsto E_{ij}(c)$$

is a homomorphism of F into the multiplicative group of $n \times n$ matrices $GL_n(F)$.

Proposition 9.1. *The group $SL_n(F)$ is generated by the elementary matrices. If $A \in GL_n(F)$, then A can be written in the form*

$$A = SD,$$

where $S \in SL_n(F)$ and D is a diagonal matrix of the form

$$D = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d \end{pmatrix}$$

so D has 1 *on the diagonal except on the lower right corner, where the component is $d = \det(A)$.*

Proof. Let $A \in GL_n(F)$. Since A is non-singular, the first component of some row is not zero, and by an elementary row operation, we can make $a_{11} \neq 0$. Adding a suitable multiple of the first row to the second row, we make $a_{21} \neq 0$, and then adding a suitable multiple of the second row to the first we make $a_{11} = 1$. Then we subtract multiples of the first row from the others to make $a_{i1} = 0$ for $i \neq 1$.

We now repeat the procedure with the second row and column, to make $a_{22} = 1$ and $a_{i2} = 0$ if $i > 2$. But then we can also make $a_{12} = 0$ by subtracting a suitable multiple of the second row from the first, so we can get $a_{i2} = 0$ for $i \neq 2$.

We repeat this procedure until we are stopped at $a_{nn} = d \neq 0$, and $a_{nj} = 0$ for $j \neq n$. Subtracting a suitable multiple of the last row from the preceding ones yields a matrix D of the form indicated in the statement of the theorem, and concludes the proof.

Theorem 9.2. *For $n \geqq 3$, $SL_n(F)$ is equal to its own commutator group.*

Proof. It suffices to prove that $E_{ij}(c)$ is a commutator. Using $n \geqq 3$, let $k \neq i, j$. Then by direct computation,

$$E_{ij}(c) = E_{ik}(c)E_{kj}(1)E_{ik}(-c)E_{kj}(-1)$$

expresses $E_{ij}(c)$ as a commutator. This proves the theorem.

We note that if a matrix M commutes with every element of $SL_n(F)$, then it must be a scalar matrix. Indeed, just the commutation with the elementary matrices

$$E_{ij}(1) = I + 1_{ij}$$

shows that M commutes with all matrices 1_{ij} (having 1 in the ij-component, 0 otherwise), so M commutes with all matrices, and is a scalar matrix. Taking the determinant shows that the center consists of $\boldsymbol{\mu}_n(F)I$, where $\boldsymbol{\mu}_n(F)$ is the group of n-th roots of unity in F.

We let Z be the center of $SL_n(F)$, so we have just seen that Z is the group of scalar matrices such that the scalar is an n-th root of unity. Then we define

$$PSL_n(F) = SL_n(F)/Z.$$

Theorem 9.3. *For $n \geqq 3$, $PSL_n(F)$ is simple.*

The rest of this section is devoted to the proof. We view $GL_n(F)$ as operating on the vector space $E = F^n$. If λ is a non-zero functional on E, we let

$$H_\lambda = \operatorname{Ker} \lambda,$$

and call H_λ (or simply H) the **hyperplane associated with** λ. Then $\dim H = n - 1$, and conversely, if H is a subspace of codimension 1, then E/H has dimension 1, and is the kernel of a functional.

An element $T \in GL_n(F)$ is called a **transvection** if it keeps every element of some hyperplane H fixed, and for all $x \in E$, we have

$$Tx = x + h \qquad \text{for some } h \in H.$$

Given any element $u \in H_\lambda$ we define a transvection T_u by

$$T_u x = x + \lambda(x)u.$$

Every transvection is of this type. If $u, v \in H_\lambda$, it is immediate that

$$T_{u+v} = T_u \circ T_v.$$

If T is a transvection and $A \in GL_n(F)$, then the conjugate ATA^{-1} is obviously a transvection.

The elementary matrices $E_{ij}(c)$ are transvections, and it will be useful to use them with this geometric interpretations, rather than formally as we did before. Indeed, let $e_1, \dots, e_n$ be the standard unit vectors which form a basis of $F^{(n)}$. Then $E_{ij}(c)$ leaves e_k fixed if $k \neq j$, and the remaining vector e_j is moved by a multiple of e_i. We let H be the hyperplane generated by e_k with $k \neq j$, and thus see that $E_{ij}(c)$ is a transvection.

Lemma 9.4. *For $n \geqq 3$, the transvections $\neq I$ form a single conjugacy class in $SL_n(F)$.*

Proof. First, by picking a basis of a hyperplane $H = H_\lambda$ and using one more element to form a basis of $F^{(n)}$, one sees from the matrix of a transvection T that $\det T = 1$, i.e. transvections are in $SL_n(F)$.

Let T' be another transvection relative to a hyperplane H'. Say

$$Tx = x + \lambda(x)u \quad \text{and} \quad T'x = x + \lambda'(x)u'$$

with $u \in H$ and $u' \in H'$. Let z and z' be vectors such that $\lambda(z) = 1$ and $\lambda'(z') = 1$. Since a basis for H together with z is a basis for $F^{(n)}$, and similarly a basis for H' together with z' is a basis for $F^{(n)}$, there exists an element $A \in GL_n(F)$ such that

$$Au = u', \qquad AH = H', \qquad Az = z'.$$

It is then immediately verified that

$$ATA^{-1} = T',$$

so T, T' are conjugate in $GL_n(F)$. But in fact, using $n \geqq 3$, the hyperplanes H, H' contain vectors which are independent. We can change the image of a basis vector in H' which is independent of u' by some factor in F so as to make $\det A = 1$, so $A \in SL_n(F)$. This proves the lemma.

We now want to show that certain subgroups of $GL_n(F)$ are either contained in the center, or contain $SL_n(F)$. Let G be a subgroup of $GL_n(F)$. We say that G is **SL_n-invariant** if

$$AGA^{-1} \subset G \quad \text{for all } A \in SL_n(F).$$

Lemma 9.5. *Let $n \geqq 3$. Let G be SL_n-invariant, and suppose that G contains a transvection $T \neq I$. Then $SL_n(F) \subset G$.*

Proof. By Lemma 9.4, all transvections are conjugate, and the set of transvections contains the elementary matrices which generate $SL_n(F)$ by Proposition 9.1, so the lemma follows.

Theorem 9.6. *Let $n \geqq 3$. If G is a subgroup of $GL_n(F)$ which is SL_n-invariant and which is not contained in the center of $GL_n(F)$, then $SL_n(F) \subset G$.*

Proof. By the preceding lemma, it suffices to prove that G contains a transvection, and this is the key step in the proof of Theorem 9.3.

We start with an element $A \in G$ which moves some line. This is possible since G is not contained in the center. So there exists a vector $u \neq 0$ such that Au is not a scalar multiple of u, say $Au = v$. Then u, v are contained in some hyperplane $H = \operatorname{Ker} \lambda$. Let $T = T_u$ and let

$$B = ATA^{-1}T^{-1}.$$

Then

$$ATA^{-1} \neq T \quad \text{and} \quad B = ATA^{-1}T^{-1} \neq I.$$

This is easily seen by applying say B to an arbitrary vector x, and using the definition of T_u. In each case, for some x the left-hand side cannot equal the right-hand side.

For any vector $x \in F^{(n)}$ we have

$$Bx - x \in (u, v),$$

where (u, v) is the plane generated by u, v. It follows that $BH \subset H$, so

$$BH = H \quad \text{and} \quad Bx - x \in H.$$

We now distinguish two cases to conclude the proof. First assume that B commutes with all transvections with respect to H. Let $w \in H$. Then from the definitions, we find for any vector x:

$$BT_w x = Bx + \lambda(x)Bw$$

$$T_w Bx = Bx + \lambda(Bx)w = Bx + \lambda(x)w.$$

Since we are in the case $BT_w = T_w B$, it follows that $Bw = w$. Therefore B leaves every vector of H fixed. Since we have seen that $Bx - x \in H$ for all x, it follows that B is a transvection and is in G, thus proving the theorem in this case.

Second, suppose there is a transvection T_w with $w \in H$ such that B does not commute with T_w. Let

$$C = BT_w B^{-1} T_w^{-1}.$$

Then $C \neq I$ and $C \in G$. Furthermore C is a product of T_w^{-1} and $BT_w B^{-1}$ whose hyperplanes are H and BH, which is also H by what we have already proved. Therefore C is a transvection, since it is a product of transvections with the same hyperplane. And $C \in G$. This concludes the proof in the second case, and also concludes the proof of Theorem 9.6.

We now return to the main theorem, that $PSL_n(F)$ is simple. Let $\bar{G}$ be a normal subgroup of $PSL_n(F)$, and let G be its inverse image in $SL_n(F)$. Then G is SL_n-invariant, and if $\bar{G} \neq 1$, then G is not equal to the center of $SL_n(F)$. Therefore G contains $SL_n(F)$ by Theorem 9.6, and therefore $\bar{G} = PSL_n(F)$, thus proving that $PSL_n(F)$ is simple.

Example. By Exercise 41 of Chapter I, or whatever other means, one sees that $PSL_2(\mathbf{F}_5) \approx A_5$ (where $\mathbf{F}_5$ is the finite field with 5 elements). While you are in the mood, show also that

$$PGL_2(\mathbf{F}_3) \approx S_4 \quad \text{but} \quad SL_2(\mathbf{F}_3) \not\approx S_4; \quad PSL_2(\mathbf{F}_3) \approx A_4.$$

EXERCISES

1. Interpret the rank of a matrix A in terms of the dimensions of the image and kernel of the linear map L_A.

2. (a) Let A be an invertible matrix in a commutative ring R. Show that $({}^tA)^{-1} = {}^t(A^{-1})$.
   (b) Let f be a non-singular bilinear form on the module E over R. Let A be an R-automorphism of E. Show that $({}^tA)^{-1} = {}^t(A^{-1})$. Prove the same thing in the hermitian case, i.e. $(A^*)^{-1} = (A^{-1})^*$.

3. Let V, W be finite dimensional vector spaces over a field k. Suppose given non-degenerate bilinear forms on V and W respectively, denoted both by $\langle\ ,\ \rangle$. Let $L\colon V \to W$ be a surjective linear map and let ${}^tL$ be its transpose; that is, $\langle Lv, w\rangle = \langle v, {}^tLw\rangle$ for $v \in V$ and $w \in W$.
   (a) Show that ${}^tL$ is injective.
   (b) Assume in addition that if $v \in V$, $v \neq 0$ then $\langle v, v\rangle \neq 0$. Show that

$$V = \operatorname{Ker} L \oplus \operatorname{Im} {}^tL,$$

   and that the two summands are orthogonal. (Cf. Exercise 33 for an example.)

4. Let $A_1, \dots, A_r$ be row vectors of dimension n, over a field k. Let $X = (x_1, \dots, x_n)$. Let $b_1, \dots, b_r \in k$. By a system of linear equations in k one means a system of type

$$A_1 \cdot X = b_1, \dots, A_r \cdot X = b_r.$$

   If $b_1 = \cdots = b_r = 0$, one says the system is homogeneous. We call n the number of variables, and r the number of equations. A solution X of the homogeneous system is called **trivial** if $x_i = 0$, $i = 1, \dots, n$.
   (a) Show that a homogeneous system of r linear equations in n unknowns with $n > r$ always has a non-trivial solution.
   (b) Let L be a system of homogeneous linear equations over a field k. Let k be a subfield of k'. If L has a non-trivial solution in k', show that it has a non-trivial solution in k.

5. Let M be an $n \times n$ matrix over a field k. Assume that $\operatorname{tr}(MX) = 0$ for all $n \times n$ matrices X in k. Show that $M = O$.

6. Let S be a set of $n \times n$ matrices over a field k. Show that there exists a column vector $X \neq 0$ of dimension n in k, such that $MX = X$ for all $M \in S$ if and only if there exists such a vector in some extension field k' of k.

7. Let $\mathbf{H}$ be the division ring over the reals generated by elements i, j, k such that $i^2 = j^2 = k^2 = -1$, and

$$ij = -ji = k, \qquad jk = -kj = i, \qquad ki = -ik = j.$$

   Then $\mathbf{H}$ has an automorphism of order 2, given by

$$a_0 + a_1 i + a_2 j + a_3 k \mapsto a_0 - a_1 i - a_2 j - a_3 k.$$

   Denote this automorphism by $\alpha \mapsto \bar{\alpha}$. What is $\alpha\bar{\alpha}$? Show that the theory of hermitian forms can be carried out over $\mathbf{H}$, which is called the division ring of **quaternions** (or by abuse of language, the non-commutative field of quaternions).

8. Let N be a strictly upper triangular $n \times n$ matrix, that is $N = (a_{ij})$ and $a_{ij} = 0$ if $i \geqq j$. Show that $N^n = 0$.

9. Let E be a vector space over k, of dimension n. Let $T: E \to E$ be a linear map such that T is nilpotent, that is $T^m = 0$ for some positive integer m. Show that there exists a basis of E over k such that the matrix of T with respect to this basis is strictly upper triangular.

10. If N is a nilpotent $n \times n$ matrix, show that $I + N$ is invertible.

11. Let R be the set of all upper triangular $n \times n$ matrices (a_{ij}) with a_{ij} in some field k, so $a_{ij} = 0$ if $i > j$. Let J be the set of all strictly upper triangular matrices. Show that J is a two-sided ideal in R. How would you describe the factor ring R/J?

12. Let G be the group of upper triangular matrices with non-zero diagonal elements. Let H be the subgroup consisting of those matrices whose diagonal element is 1. (Actually prove that H is a subgroup). How would you describe the factor group G/H?

13. Let R be the ring of $n \times n$ matrices over a field k. Let L be the subset of matrices which are 0 except on the first column.
    (a) Show that L is a left ideal.
    (b) Show that L is a minimal left ideal; that is, if $L' \subset L$ is a left ideal and $L' \neq 0$, then $L' = L$. (For more on this situation, see Chapter VII, §5.)

14. Let F be any field. Let D be the subgroup of diagonal matrices in $GL_n(F)$. Let N be the normalizer of D in $GL_n(F)$. Show that N/D is isomorphic to the symmetric group on n elements.

15. Let F be a finite field with q elements. Show that the order of $GL_n(F)$ is

$$(q^n - 1)(q^n - q) \cdots (q^n - q^{n-1}) = q^{n(n-1)/2} \prod_{i=1}^{n} (q^i - 1).$$

[*Hint*: Let $x_1, \ldots, x_n$ be a basis of F^n. Any element of $GL_n(F)$ is uniquely determined by its effect on this basis, and thus the order of $GL_n(F)$ is equal to the number of all possible bases. If $A \in GL_n(F)$, let $Ax_i = y_i$. For y_1 we can select any of the $q^n - 1$ non-zero vectors in F^n. Suppose inductively that we have already chosen $y_1, \ldots, y_r$ with $r < n$. These vectors span a subspace of dimension r which contains q^r elements. For y_{i+1} we can select any of the $q^n - q^r$ elements outside of this subspace. The formula drops out.]

16. Again let F be a finite field with q elements. Show that the order of $SL_n(F)$ is

$$q^{n(n-1)/2} \prod_{i=2}^{n} (q^i - 1);$$

and that the order of $PSL_n(F)$ is

$$\frac{1}{d} q^{n(n-1)/2} \prod_{i=2}^{n-1} (q^i - 1),$$

where d is the greatest common divisor of n and $q - 1$.

17. Let F be a finite field with q elements. Show that the group of all upper triangular matrices with 1 on the diagonal is a Sylow subgroup of $GL_n(F)$ and of $SL_n(F)$.

18. The reduction map $\mathbf{Z} \to \mathbf{Z}/N\mathbf{Z}$, where N is a positive integer defines a homomorphism
$$SL_2(\mathbf{Z}) \to SL_2(\mathbf{Z}/N\mathbf{Z}).$$
Show that this homomorphism is surjective. [*Hint*: Use elementary divisors, i.e. the structure of submodules of rank 2 over the principal ring $\mathbf{Z}$.]

19. Show that the order of $SL_2(\mathbf{Z}/N\mathbf{Z})$ is equal to
$$N^3 \prod_{p \mid N} \left(1 - \frac{1}{p^2}\right),$$
where the product is taken over all primes dividing N.

20. Show that one has an exact sequence
$$1 \to SL_2(\mathbf{Z}/N\mathbf{Z}) \to GL_2(\mathbf{Z}/N\mathbf{Z}) \xrightarrow{\det} (\mathbf{Z}/N\mathbf{Z})^* \to 1.$$
In fact, show that
$$GL_2(\mathbf{Z}/N\mathbf{Z}) = SL_2(\mathbf{Z}/N\mathbf{Z})G_N,$$
where G_N is the group of matrices
$$\begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix} \quad \text{with} \quad d \in (\mathbf{Z}/N\mathbf{Z})^*.$$

21. Show that $SL_2(\mathbf{Z})$ is generated by the matrices
$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

22. Let p be a prime $\geqq 5$. Let G be a subgroup of $SL_2(\mathbf{Z}/p^n\mathbf{Z})$ with $n \geqq 1$. Assume that the image of G in $SL_2(\mathbf{Z}/p\mathbf{Z})$ under the natural homomorphism is all of $SL_2(\mathbf{Z}/p\mathbf{Z})$. Prove that $G = SL_2(\mathbf{Z}/p^n\mathbf{Z})$.

Note. Exercise 22 is a generalization by Serre of a result of Shimura; see Serre's *Abelian ℓ-adic Representations and elliptic curves*, Benjamin, 1968, IV, §3, Lemma 3. See also my exposition in *Elliptic Functions*, Springer Verlag, reprinted from Addison-Wesley, 1973, Chapter 17, §4.

23. Let k be a field in which every quadratic polynomial has a root. Let B be the Borel subgroup of $GL_2(k)$. Show that G is the union of all the conjugates of B. (This cannot happen for finite groups!)

24. Let A, B be square matrices of the same size over a field k. Assume that B is non-singular. If t is a variable, show that $\det(A + tB)$ is a polynomial in t, whose leading coefficient is $\det(B)$, and whose constant term is $\det(A)$.

25. Let $a_{11}, \ldots, a_{1n}$ be elements from a principal ideal ring, and assume that they generate the unit ideal. Suppose $n > 1$. Show that there exists a matrix (a_{ij}) with this given first row, and whose determinant is equal to 1.

26. Let A be a commutative ring, and $I = (x_1, \dots, x_r)$ an ideal. Let $c_{ij} \in A$ and let

$$y_i = \sum_{j=1}^{r} c_{ij}x_j.$$

Let $I' = (y_1, \dots, y_r)$. Let $D = \det(c_{ij})$. Show that $DI \subset I'$.

27. Let L be a free module over $\mathbf{Z}$ with basis $e_1, \dots, e_n$. Let M be a free submodule of the same rank, with basis $u_1, \dots, u_n$. Let $u_i = \sum c_{ij}e_j$. Show that the index $(L:M)$ is given by the determinant:

$$(L:M) = |\det(c_{ij})|.$$

28. **(The Dedekind determinant).** Let G be a finite commutative group and let F be the vector space of functions of G into $\mathbf{C}$. Show that the characters of G (homomorphisms of G into the roots of unity) form a basis for this space. If $f: G \to \mathbf{C}$ is a function, show that for $a, b \in G$.

$$\det(f(ab^{-1})) = \prod_{\chi} \sum_{a \in G} \chi(a)f(a),$$

where the product is taken over all characters. [*Hint*: Use both the characters and the characteristic functions of elements of G as bases for F, and consider the linear map

$$T = \sum f(a)T_a,$$

where T_a is translation by a.] Also show that

$$\det(f(ab^{-1})) = \left(\sum_{a \in G} f(a)\right) \det(f(ab^{-1}) - f(b^{-1})),$$

where the determinant on the left is taken for all $a, b \in G$, and the determinant on the right is taken only for $a, b \neq 1$.

29. Let $\mathfrak{g}$ be a module over the commutative ring R. A bilinear map $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, written $(x, y) \mapsto [x, y]$, is said to make $\mathfrak{g}$ a **Lie algebra** if it is anti-symmetric, i.e. $[x, y] = -[y, x]$, and if the map $D_x: g \to g$ defined by $D_x(y) = [x, y]$ is a derivation of g into itself, that is

$$D([y, z]) = [Dy, z] + [y, Dz] \qquad \text{and} \qquad D(cy) = cD(y)$$

for all $x, y, z \in g$ and $c \in R$.

(a) Let A be an associative algebra over R. For $x, y \in A$, define $[x, y] = xy - yx$. Show that this makes A into a Lie algebra. Example: the algebra of R-endomorphisms of a module M, especially the algebra of matrices $\mathrm{Mat}_n(R)$.

(b) Let M be a module over R. For two derivations D_1, D_2 of M, define $[D_1, D_2] = D_1D_2 - D_2D_1$. Show that the set of derivations of M is a Lie subalgebra of $\mathrm{End}_R(M)$.

(c) Show that the map $x \mapsto \bar{E}_x$ is a Lie homomorphism of $\mathfrak{g}$ into the Lie algebra of derivations of $\mathfrak{g}$ into itself.

30. Given a set of polynomials $\{P_\nu(X_{ij})\}$ in the polynomial ring $R[X_{ij}]$ $(1 \leqq i, j \leqq n)$, a zero of this set in R is a matrix $x = (x_{ij})$ such that $x_{ij} \in R$ and $P_\nu(x_{ij}) = 0$ for all ν. We use vector notation, and write $(X) = (X_{ij})$. We let $G(R)$ denote the set of zeros

of our set of polynomials $\{P_\nu\}$. Thus $G(R) \subset M_n(R)$, and if R' is any commutative associative R-algebra we have $G(R') \subset M_n(R')$. We say that the set $\{P_\nu\}$ defines an **algebraic group over** R if $G(R')$ is a subgroup of the group $GL_n(R')$ for all R' (where $GL_n(R')$ is the multiplicative group of invertible matrices in R').

As an example, the group of matrices satisfying the equation ${}^tXX = I_n$ is an algebraic group.

Let R' be the R-algebra which is free, with a basis $\{1, t\}$ such that $t^2 = 0$. Thus $R' = R[t]$. Let $\mathfrak{g}$ be the set of matrices $x \in M_n(R)$ such that $I_n + tx \in G(R[t])$. Show that $\mathfrak{g}$ is a Lie algebra. [*Hint*: Note that

$$P_\nu(I_n + tX) = P_\nu(I_n) + \operatorname{grad} P_\nu(I_n)tX.$$

Use the algebra $R[t, u]$ where $t^2 = u^2 = 0$ to show that if $I_n + tx \in G(R[t])$ and $I_n + uy \in G(R[u])$ then $[x, y] \in \mathfrak{g}$.]

(I have taken the above from the first four pages of [Se 65]. For more information on Lie algebras and Lie Groups, see [Bo 82] and [Ja 79].

[Bo 82] N. Bourbaki, *Lie Algebras and Lie Groups*, Masson, 1982

[Ja 79] N. Jacobson, *Lie Algebras*, Dover, 1979 (reprinted from Interscience, 1962)

[Se 65] J. P. Serre, *Lie Algebras and Lie Groups*, Benjamin, 1965. Reprinted Springer Lecture Notes 1500. Springer/Verlag 1992

Non-commutative cocycles

Let K be a finite Galois extension of a field k. Let $\Gamma = GL_n(K)$, and $G = \operatorname{Gal}(K/k)$. Then G operates on Γ. By a **cocycle** of G in Γ we mean a family of elements $\{A(\sigma)\}$ satisfying the relation

$$A(\sigma)\sigma A(\tau) = A(\sigma\tau).$$

We say that the cocycle **splits** if there exists $B \in \Gamma$ such that

$$A(\sigma) = B^{-1}\sigma B \qquad \text{for all } \sigma \in G.$$

In this non-commutative case, cocycles do not form a group, but one could define an equivalence relation to define cohomology classes. For our purposes here, we care only whether a cocycle splits or not. When every cocycle splits, we also say that $H^1(G, \Gamma) = 0$ (or 1).

31. Prove that $H^1(G, GL_n(K)) = 1$. [*Hint*: Let $\{e_1, \dots, e_N\}$ be a basis of $\operatorname{Mat}_n(k)$ over k, say the matrices with 1 in some component and 0 elsewhere. Let

$$x = \sum_{i=1}^{N} x_i e_i$$

with variables x_i. There exists a polynomial $P(X)$ such that x is invertible if and only if $P(x_1, \dots, x_N) \neq 0$. Instead of $P(x_1, \dots, x_N)$ we also write $P(x)$. Let $\{A(\sigma)\}$ be a cocycle. Let $\{t_\sigma\}$ be algebraically independent variables over k. Then

$$P\left(\sum_{\gamma \in G} t_\gamma A(\gamma)\right) \neq 0$$

because the polynomial does not vanish when one t_γ is replaced by 1 and the others are replaced by 0. By the algebraic independence of automorphisms from Galois theory, there exists an element $y \in K$ such that if we put

$$B = \sum_\gamma (\gamma y)A(\gamma)$$

then $P(B) \neq 0$, so B is invertible. It is then immediately verified that $A(\sigma) = B\sigma B^{-1}$. But when k is finite, cf. my *Algebraic Groups over Finite Fields*, Am. J. Vol 78 No. 3, 1956.]

32. **Invariant bases.** (A. Speiser, Zahlentheoretische Sätze aus der Gruppentheorie, *Math. Z.* **5** (1919) pp. 1–6. See also Kolchin-Lang, *Proc. AMS* Vol. 11 No. 1, 1960). Let K be a finite Galois extension of k, $G = \mathrm{Gal}(K/k)$ as in the preceding exercise. Let V be a finite-dimensional vector space over K, and suppose G operates on V in such a way that $\sigma(av) = \sigma(a)\sigma(v)$ for $a \in K$ and $v \in V$. Prove that there exists a basis $\{w_1, \dots, w_n\}$ such that $\sigma w_i = w_i$ for all $i = 1, \dots, n$ and all $\sigma \in G$ (an invariant basis). *Hint*: Let $\{v_1, \dots, v_n\}$ be any basis, and let

$$\sigma\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = A(\sigma)\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

where $A(\sigma)$ is a matrix in $GL_n(K)$. Solve for B in the equation $(\sigma B)A(\sigma) = B$, and let

$$\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = B\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

The next exercises on harmonic polynomials have their source in Whittaker, *Math. Ann.* 1902; see also Whittaker and Watson, *Modern Analysis*, Chapter XIII.

33. **Harmonic polynomials.** Let $\mathrm{Pol}(n, d)$ denote the vector space of homogeneous polynomials of degree d in n variables $X_1, \dots, X_n$ over a field k of characteristic 0. For an n-tuple of integers $(\nu_1, \dots, \nu_n)$ with $\nu_i \geqq 0$ we denote by $M_{(\nu)}$ as usual the monomial

$$M_{(\nu)}(X) = X_1^{\nu_1} \cdots X_n^{\nu_n}.$$

Prove:

(a) The number of monomials of degree d is $\binom{n-1+d}{n-1}$, so this number is the dimension of $\mathrm{Pol}(n, d)$.

(b) Let $(D) = (D_1, \dots, D_n)$ where D_i is the partial derivative with respect to the i-th variable. Then we can define $P(D)$ as usual. For $P, Q \in \mathrm{Pol}(n, d)$, define

$$\langle P, Q \rangle = P(D)Q(0).$$

Prove that this defines a symmetric non-degenerate scalar product on $\mathrm{Pol}(n, d)$. If k is not real, it may happen that $P \neq 0$ but $\langle P, P \rangle = 0$. However, if the ground field is real, then $\langle P, P \rangle > 0$ for $P \neq 0$. Show also that the monomials of degree d form an orthogonal basis. What is $\langle M_{(\nu)}, M_{(\nu)} \rangle$?

(c) The map $P \mapsto P(D)$ is an isomorphism of $\mathrm{Pol}(n, d)$ onto its dual.

(d) Let $\Delta = D_1^2 + \cdots + D_n^2$. Note that $\Delta\colon \mathrm{Pol}(n, d) \to \mathrm{Pol}(n, d - 2)$ is a linear map. Prove that Δ is surjective.

(e) Define $\mathrm{Har}(n, d) = \mathrm{Ker}\Delta$ = vector space of **harmonic homogeneous polynomials** of degree d. Prove that

$$\dim \mathrm{Har}(n, d) = (n + d - 3)!(n + 2d - 2)/(n - 2)!d!.$$

In particular, if $n = 3$, then $\dim \mathrm{Har}(3, d) = 2d + 1$.

(f) Let $r^2 = X_1^2 + \cdots + X_n^2$. Let S denote multiplication by r^2. Show that

$$\langle \Delta P, Q\rangle = \langle P, SQ\rangle \text{ for } P \in \mathrm{Pol}(n, d) \text{ and } Q \in \mathrm{Pol}(n, d - 2),$$

so ${}^t\Delta = S$. More generally, for $R \in \mathrm{Pol}(n, m)$ and $Q \in \mathrm{Pol}(n, d - m)$ we have

$$\langle R(D)P, Q\rangle = \langle P, RQ\rangle.$$

(g) Show that $[\Delta, S] = 4d + 2n$ on $\mathrm{Pol}(n, d)$. Here $[\Delta, S] = \Delta \circ S - S \circ \Delta$. Actually, $[\Delta, S] = 4E + 2n$, where E is the Euler operator $E = \sum X_i D_i$, which is, however, the degree operator on homogeneous polynomials.

(h) Prove that $\mathrm{Pol}(n, d) = \mathrm{Har}(n, d) \oplus r^2\mathrm{Pol}(n, d - 2)$ and that the two summands are orthogonal. This is a classical theorem used in the theory of the Laplace operator.

(i) Let $(c_1, \ldots, c_n) \in k^n$ be such that $\sum c_i^2 = 0$. Let

$$H_c^d(X) = (c_1X_1 + \cdots + c_nX_n)^d.$$

Show that H_c^d is harmonic, i.e. lies in $\mathrm{Har}(n, d)$.

(j) For any $Q \in \mathrm{Pol}(n, d)$, and a positive integer m, show that

$$Q(D)H_c^m(X) = m(m - 1) \cdots (m - d + 1)Q(c)H_c^{m-d}(X).$$

34. (Continuation of Exercise 33). Prove:

Theorem. *Let k be algebraically closed of characteristic* 0. *Let $n \geqq 3$. Then* $\mathrm{Har}(n, d)$ *as a vector space over k is generated by all polynomials H_c^d with $(c) \in k^n$ such that $\sum c_i^2 = 0$.*

[*Hint*: Let $Q \in \mathrm{Har}(n, d)$ be orthogonal to all polynomials H_c^d with $(c) \in k^n$. By Exercise 33(h), it suffices to prove that $r^2 | Q$. But if $\sum c_i^2 = 0$, then by Exercise 33(j) we conclude that $Q(c) = 0$. By the Hilbert Nullstellensatz, it follows that there exists a polynomial $F(X)$ such that

$$Q(X)^s = r^2(X)F(X) \text{ for some positive integer } s.$$

But $n \geqq 3$ implies that $r^2(X)$ is irreducible, so $r^2(X)$ divides $Q(X)$.]

35. (Continuation of Exercise 34). Prove that the representation of $O(n) = U_n(\mathbf{R})$ on $\mathrm{Har}(n, d)$ is irreducible.
Readers will find a proof in the following:

S. Helgason, *Topics in Harmonic Analysis on Homogeneous Spaces*, Birkhäuser, 1981 (see especially §3, Theorem 3.1(ii))

N. Vilenkin, *Special Functions and the Theory of Group Representations*, AMS Translations of mathematical monographs **Vol. 22**, 1968 (Russian original, 1965), Chapter IX, §2.

R. HOWE and E. C. TAN, *Non-Abelian Harmonic Analysis*, Universitext, Springer Verlag, New York, 1992.

The Howe-Tan proof runs as follows. We now use the hermitian product

$$\langle P, Q\rangle = \int_{\mathbf{S}^{n-1}} P(x)\,\overline{Q(x)}\,d\sigma(x),$$

where σ is the rotation invariant measure on the $(n-1)$-sphere $\mathbf{S}^{n-1}$. Let $e_1, \dots, e_n$ be the unit vectors in $\mathbf{R}^n$. We can identify $O(n-1)$ as the subgroup of $O(n)$ leaving e_n fixed. Observe that $O(n)$ operates on $\text{Har}(n, d)$, say on the right by composition $P \mapsto P \circ A$, $A \in O(n)$, and this operation commutes with Δ. Let

$$\lambda\colon \text{Har}(n, d) \to \mathbf{C}$$

be the functional such that $\lambda(P) = P(e_n)$. Then λ is $O(n-1)$-invariant, and since the hermitian product is non-degenerate, there exists a harmonic polynomial Q_n such that

$$\lambda(P) = \langle P, Q_n\rangle \quad \text{for all } P \in \text{Har}(n, d).$$

Let $M \subset \text{Har}(n, d)$ be an $O(n)$-submodule. Then the restriction λ_M of λ to M is nontrivial because $O(n)$ acts transitively on $\mathbf{S}^{n-1}$. Let Q_n^M be the orthogonal projection of Q_n on M. Then Q_n^M is $O(n-1)$-invariant, and so is a linear combination

$$Q_n^M(x) = \sum_{j+2k=d} c_j\, x_n^j\, r_{n-1}^{2k}.$$

Furthermore Q_n^H is harmonic. From this you can show that Q_n^H is uniquely determined, by showing the existence of recursive relations among the coefficients c_j. Thus the submodule M is uniquely determined, and must be all of $\text{Har}(n, d)$.

Irreducibility of $\mathfrak{sl}_n(F)$.

36. Let F be a field of characteristic 0. Let $\mathfrak{g} = \mathfrak{sl}_n(F)$ be the vector space of matrices with trace 0, with its Lie algebra structure $[X, Y] = XY - YX$. Let E_{ij} be the matrix having (i, j)-component 1 and all other components 0. Let $G = SL_n(F)$. Let A be the multiplicative group of diagonal matrices over F.
    (a) Let $H_i = E_{ii} - E_{i+1,i+1}$ for $i = 1, \dots, n-1$. Show that the elements E_{ij} $(i \neq j)$, $H_1, \dots, H_{n-1}$ form a basis of $\mathfrak{g}$ over F.
    (b) For $g \in G$ let $\mathbf{c}(g)$ be the conjugation action on $\mathfrak{g}$, that is $\mathbf{c}(g)X = gXg^{-1}$. Show that each E_{ij} is an eigenvector for this action restricted to the group A.
    (c) Show that the conjugation representation of G on $\mathfrak{g}$ is irreducible, that is, if $V \neq 0$ is a subspace of $\mathfrak{g}$ which is $\mathbf{c}(G)$-stable, then $V = \mathfrak{g}$. *Hint:* Look up the sketch of the proof in [JoL 01], Chapter VII, Theorem 1.5, and put in all the details. Note that for $i \neq j$ the matrix E_{ij} is nilpotent, so for variable t, the exponential series $\exp(tE_{ij})$ is actually a polynomial. The derivative with respect to t can be taken in the formal power series $F[[t]]$, not using limits. If X is a matrix, and $x(t) = \exp(tX)$, show that

$$\frac{d}{dt}x(t)\,Yx(t)^{-1}\Big|_{t=0} = XY - YX = [X, Y].$$