Chapter 4
Tensor and Exterior Products

4.1 Tensor Product of Vector Spaces

4.1.1 Construction of the Tensor Product

Let E and F be K-vector spaces whose dimensions are finite. We construct their *tensor product* $E \otimes_K F$ as follows.

We start with their Cartesian product. We warn the reader that we do not equip $E \times F$ with the usual addition. Thus we do not think of it as a vector space. The first step is to consider the set G of formal linear combinations of elements of $E \times F$

$$\sum_{j=1}^{r} \lambda_j (x_j, y_j),$$

where r is an arbitrary natural integer, λ_j are scalars, and $(x_j, y_j) \in E \times F$.

The set G has a natural structure of K-vector space, where $E \times F$ is a basis. The zero element is the empty sum ($r = 0$). We warn the reader that $(x + x', y) - (x, y) - (x', y)$ and $\lambda(x, y) - (\lambda x, y)$ cannot be simplified, and $(0_E, 0_F)$ is not equal to 0_G. Actually, G is infinite-dimensional whenever K is infinite!

We now consider the subspace G_0 generated by all elements of the form $(x + x', y) - (x, y) - (x', y)$, $(x, y + y') - (x, y) - (x, y')$, $\lambda(x, y) - (\lambda x, y)$ or $\lambda(x, y) - (x, \lambda y)$. The quotient space G/G_0 is what we call the tensor product of E and F (in this order) and denote by $E \otimes_K F$. When there is no ambiguity about the scalars, we simply write $E \otimes F$.

By construction, $E \otimes_K F$ is a vector space. The class of an elementary pair (x, y) is denoted $x \otimes y$. The elements of G_0 can be viewed as the *simplification rules* in $E \otimes_K F$:

$$(\lambda x + x') \otimes y = \lambda (x \otimes y) + x' \otimes y, \qquad x \otimes (\lambda y + y') = \lambda (x \otimes y) + x \otimes y'.$$

In particular, we have $x \otimes 0_F = 0_E \otimes y = 0$ for every x and y.

D. Serre, *Matrices,* Graduate Texts in Mathematics 216,
DOI 10.1007/978-1-4419-7683-3_4, © Springer Science+Business Media, LLC 2010

Theorem 4.1 *Let E,F,H be K-vector spaces. Then $\mathbf{Bil}(E\times F;H)$ is isomorphic to $\mathscr{L}(E\otimes F;H)$ through the formula*

$$b(x,y)=u(x\otimes y).$$

Proof. If $u\in\mathscr{L}(E\otimes F;H)$ is given, then $b(x,y):=u(x\otimes y)$ clearly defines a bilinear map.

Conversely, let $b\in\mathbf{Bil}(E\times F;H)$ be given. Then $\beta:G\mapsto H$, defined by

$$\beta\left(\sum_{j=1}^{r}\lambda_j(x_j,y_j)\right):=\sum_{j=1}^{r}\lambda_j b(x_j,y_j),$$

is linear. Because of the bilinearity, β vanishes identically over G_0, thus passes to the quotient as a linear map u. □

Corollary 4.1 *Let $x\in E$ and $y\in F$ be given vectors. If $x\neq 0$ and $y\neq 0$, then $x\otimes y\neq 0$.*

Proof. There exist linear forms $\ell\in E'$ and $m\in F'$ such that $\ell(x)=m(y)=1$. Then the map

$$(w,z)\mapsto b(w,z):=\ell(w)m(z)$$

is bilinear over $E\times F$ and is such that $b(x,y)=1$. Theorem 4.1 provides a linear form u over $E\otimes F$ such that $u(x\otimes y)=1$. Thus $x\otimes y\neq 0$. □

We say that an element $x\otimes y$ is *rank one* if $x\neq 0$ and $y\neq 0$. More generally, the *rank* of an element of $E\otimes_K F$ is the minimal length r among all of its representations of the form

$$x_1\otimes y_1+\cdots+x_r\otimes y_r.$$

Given a basis $\mathscr{B}_E$ of E and a basis $\mathscr{B}_F$ of F, one may form a basis of $E\otimes_K F$ by taking all the products $\mathbf{e}^i\otimes\mathbf{f}^j$ where $\mathbf{e}^i\in\mathscr{B}_E$ and $\mathbf{f}^j\in\mathscr{B}_F$. This is obviously a generating family. To see that it is free, let us consider an identity

$$\sum_{i,j}\lambda_{ij}\mathbf{e}^i\otimes\mathbf{f}^j=0.$$

Let ℓ_p and m_q be the elements of the dual bases such that $\ell_p(\mathbf{e}^i)=\delta_p^i$ and $m_q(\mathbf{f}^j)=\delta_q^j$. The bilinear map $b(x,y):=\ell_p(x)m_q(y)$ extends as a linear map u over $E\otimes F$, according to Theorem 4.1. We have

$$0=u\left(\sum_{i,j}\lambda_{ij}\mathbf{e}^i\otimes\mathbf{f}^j\right)=\sum_{i,j}\lambda_{ij}\ell_p(\mathbf{e}^i)m_q(\mathbf{f}^j)=\lambda_{pq}.$$

The dimension of $E\otimes_K F$ is thus equal to the product of $\dim E$ and $\dim F$:

$$\dim E\otimes_K F=\dim E\cdot\dim F. \tag{4.1}$$

This contrasts with the formula

$$\dim E\times F=\dim E+\dim F.$$

4.1.2 *Linearity versus Bilinearity*

4.1.2.1 The Dual of a Tensor Product

If $\ell\in E'$ and $m\in F'$ are linear forms, then the pair (m,ℓ) defines a bilinear form over $E\times F$, by

$$(x,y)\mapsto \ell(x)\cdot m(y).$$

By Theorem 4.1, there corresponds a linear form $T_{(m,\ell)}$ over $E\otimes F$. We notice that the map $(m,\ell)\mapsto T_{(m,\ell)}$ is bilinear too. Invoking the theorem again, we infer a linear map

$$T:F'\otimes E'\to (E\otimes F)',$$

defined by

$$[T(m\otimes\ell)](x\otimes y)=\ell(x)\cdot m(y).$$

We verify easily that T is an isomorphism, a canonical one.

4.1.2.2 $\mathscr{L}(\mathbf{E};\mathbf{F})$ as a Tensor Product

Given a linear map $f:E\mapsto F$, we may construct a bilinear form over $E\times F'$ by

$$(x,m)\mapsto m(f(x)).$$

By Theorem 4.1, it extends as a linear form over $E\otimes F'$, satisfying

$$x\otimes m\mapsto m(f(x)).$$

By the previous paragraph, f can be identified as an element of $(E\otimes F')'=(F'')\otimes E'=F\otimes E'$. This provides a homomorphism from $\mathscr{L}(E;F)$ into $F\otimes E'$. This morphism is obviously one-to-one: if $m(f(x))\equiv 0$ for every m and x, then $f(x)=0$ for every x; that is, $f=0$. It is also onto: an element

$$\sum_{j=1}^{r} v^j\otimes m^j$$

is the image of the linear map

$$x\mapsto f(x):=\sum_j m^j(x)v^j.$$

Therefore $\mathscr{L}(E;F)$ identifies canonically with $F \otimes E'$.

4.1.3 Iterating the Tensor Product

Given three vector spaces E, F, and G over K, with bases $(\mathbf{e}^i)_{1\le i\le m}$, $(\mathbf{f}^j)_{1\le j\le n}$, $(\mathbf{g}^k)_{1\le k\le p}$, the spaces $(E\otimes F)\otimes G$ and $E\otimes(F\otimes G)$ are isomorphic through $(\mathbf{e}^i\otimes \mathbf{f}^j)\otimes \mathbf{g}^k \longleftrightarrow \mathbf{e}^i\otimes(\mathbf{f}^j\otimes \mathbf{g}^k)$. We thus identify both spaces, and denote them simply by $E\otimes F\otimes G$.

This rule allows us to define the tensor product of an arbitrary finite number of vector spaces $E^1,\dots,E^r$, denoted as $E^1\otimes\cdots\otimes E^r$. The following generalization of Theorem 4.1 is immediate.

Theorem 4.2 *Let $E^1,\dots,E^r,F$ be K-vector spaces. Then the vector space of r-linear maps from $E^1\times\cdots\times E^r$ into F is isomorphic to $\mathscr{L}(E^1\otimes\cdots\otimes E^r;F)$ through the formula*

$$\psi(x^1,\dots,x^r) = u(x^1\otimes\cdots\otimes x^r).$$

4.2 Exterior Calculus

4.2.1 Tensors of Degree Two

We assume temporarily that the characteristic of the field K is not 2, even though in general we do not need this hypothesis. Let E be a finite-dimensional K-vector space, with basis $\{\mathbf{e}^1,\dots,\mathbf{e}^n\}$. The linear map over G

$$\sum_i \lambda_i(x_i,y_i) \mapsto \sum_i \lambda_i(y_i,x_i)$$

sends G_0 into G_0. It thus passes to the quotient, defining a linear map $\sigma : x\otimes y \mapsto y\otimes x$. Because σ is an involution, the group $\mathbb{Z}/2\mathbb{Z} \sim \{1,\sigma\}$ operates over $E\otimes E$ by

$$1\cdot(x\otimes y) = (x\otimes y), \qquad \sigma\cdot(x\otimes y) = (y\otimes x).$$

$$\mathbf{Sym}^2(E) := \{w\in E\otimes E \mid \sigma(w) = w\}$$

the set of *symmetric* tensors, and

$$\Lambda^2(E) := \{w\in E\otimes E \mid \sigma(w) = -w\}$$

the set of *skew-symmetric* tensors. These are subspaces, with obviously $\mathbf{Sym}^2(E)\cap \Lambda^2(E) = \{0\}$. If $w\in E\otimes E$, we have

$$\frac{1}{2}(w+\sigma(w)) \in \mathbf{Sym}^2(E), \qquad \frac{1}{2}(w-\sigma(w)) \in \Lambda^2(E),$$

and their sum equals w. We deduce

$$E \otimes E = \mathbf{Sym}^2(E) \oplus \Lambda^2(E).$$

Thus we can view as well $\Lambda^2(E)$ as the quotient of $E \otimes E$ by the subspace spanned by the tensors of the form $x \otimes x$ (which is nothing but $\mathbf{Sym}^2(E)$).

Proposition 4.1 *A basis of* $\mathbf{Sym}^2(E)$ *is provided by the tensors*

$$\mathbf{e}^j \mathbf{e}^k := \frac{1}{2}(\mathbf{e}^j \otimes \mathbf{e}^k + \mathbf{e}^k \otimes \mathbf{e}^j), \qquad 1 \le j \le k \le n,$$

and a basis of $\Lambda^2(E)$ *is provided by the tensors*

$$\mathbf{e}^j \wedge \mathbf{e}^k := \frac{1}{2}(\mathbf{e}^j \otimes \mathbf{e}^k - \mathbf{e}^k \otimes \mathbf{e}^j), \qquad 1 \le j < k \le n.$$

Proof. Clearly, the elements $\mathbf{e}^j \mathbf{e}^k$ span a subspace S of $\mathbf{Sym}^2(E)$, whereas the elements $\mathbf{e}^j \wedge \mathbf{e}^k$ span a subspace A of $\Lambda^2(E)$. We thus have $S \cap A = \{0\}$. Also, these elements together span $E \otimes E$ because

$$\mathbf{e}^j \otimes \mathbf{e}^k = \mathbf{e}^j \mathbf{e}^k + \mathbf{e}^j \wedge \mathbf{e}^k,$$

and $\mathbf{e}^j \otimes \mathbf{e}^k$ form a basis of $E \otimes E$. Therefore $E \otimes E = S \oplus A$, which implies $S = \mathbf{Sym}^2(E)$ and $A = \Lambda^2(E)$. Because the family made of elements $\mathbf{e}^j \mathbf{e}^k$ and $\mathbf{e}^j \wedge \mathbf{e}^k$ has cardinal n^2 and is generating, it is a basis of $E \otimes E$. Therefore the elements $\mathbf{e}^j \mathbf{e}^k$ form a basis of $\mathbf{Sym}^2(E)$ and the elements $\mathbf{e}^j \wedge \mathbf{e}^k$ form a basis of $\Lambda^2(E)$. □

As a by-product, we have

$$\dim \mathbf{Sym}^2(E) = \frac{n(n+1)}{2}, \qquad \dim \Lambda^2(E) = \frac{n(n-1)}{2}.$$

We point out that

$$\mathbf{e}^j \mathbf{e}^k = \mathbf{e}^k \mathbf{e}^j, \qquad \mathbf{e}^j \wedge \mathbf{e}^k = -\mathbf{e}^k \wedge \mathbf{e}^j.$$

4.2.2 Exterior Products

Let $k \ge 1$ be an integer. We denote $T^k(E) = E \otimes \cdots \otimes E$ the tensor product of k copies of E. We also write $T^k(E) = V^{\otimes k}$. When $k = 1$, $T^1(E)$ is nothing but E. By convention, we set $T^0(E) = K$.

We extend the definition of $\Lambda^2(E)$ to other integers $k \ge 0$ as follows. We first consider the subspace L_k of $T^k(E)$ spanned by the elementary products $x^1 \otimes \cdots \otimes x^k$ in which at least two vectors x^i and x^j are equal. Notice that $L_0 = \{0\}$ and $L_1 = \{0\}$. Then we define the quotient space

$$\Lambda^k(E) := T^k(E)/L_k.$$

In particular, $\Lambda^0(E) = K$ and $\Lambda^1(E) = E$. The class of $w^1 \otimes \cdots \otimes w^k$ is denoted $w^1 \wedge \cdots \wedge w^k$. The operation $\wedge$ is called the *exterior product* of tensors. It is extended by multilinearity.

By definition, and because

$$x \otimes y + y \otimes x = (x+y) \otimes (x+y) - x \otimes x - y \otimes y \in L_k,$$

the exchange of two consecutive factors w^s and w^{s+1} just flips the sign:

$$\cdots \wedge w^{s+1} \wedge w^s \wedge \cdots = -(\cdots \wedge w^s \wedge w^{s+1} \wedge \cdots).$$

By induction, we infer the next lemma.

Lemma 3.

$$w^{\sigma(1)} \wedge \cdots \wedge w^{\sigma(k)} = \varepsilon(\sigma)\, w^1 \wedge \cdots \wedge w^k, \qquad \forall \sigma \in S_k.$$

Theorem 4.3 *If* $\dim E = n$*, then*

$$\dim \Lambda^k(E) = \binom{n}{k}.$$

A basis of $\Lambda^k(E)$ *is given by the set of tensors* $\mathbf{e}^{i_1} \wedge \cdots \wedge \mathbf{e}^{i_k}$ *with* $i_1 < \cdots < i_k$.

In particular, $\Lambda^k(E) = \{0\}$ *if* $k > n$.

Proof. Because $T^k(E)$ is spanned by the vectors of the form $\mathbf{e}^{i_1} \otimes \cdots \otimes \mathbf{e}^{i_k}$, $\Lambda^k(E)$ is spanned by elements of the form $\mathbf{e}^{i_1} \wedge \cdots \wedge \mathbf{e}^{i_k}$. However, this vector vanishes if two indices i_r are equal, by construction. Thus $\Lambda^k(E)$ is spanned by those for which the indices $i_1, \dots, i_k$ are pairwise distinct. And because of Lemma 3, those that satisfy $i_1 < \cdots < i_k$ form a generating family. In particular, $\Lambda^k(E) = \{0\}$ if $k > n$.

There remains to prove that this generating family is free. It suffices to treat the case $k \le n$.

Case $k = n$. We may assume that $E = K^n$. With $\mathbf{w} = (w^1, \dots, w^n) \in V^n$, we associate the matrix W whose jth column is w^j for each j. The map $\mathbf{w} \mapsto \det W$ is n-linear and thus corresponds to a linear form D over $T^n(E)$, according to Theorem 4.2. Obviously, D vanishes over L_n, and thus defines a linear form Δ over the quotient $\Lambda^n(E)$. This form is nontrivial, because $\Delta(\mathbf{e}^1, \dots, \mathbf{e}^n) = \det I_n = 1 \neq 0$. This shows that $\dim \Lambda^n(E) \ge 1$. This space is spanned by at most one element $\mathbf{e}^1 \wedge \cdots \wedge \mathbf{e}^n$, therefore we deduce $\dim \Lambda^n(E) = 1$, and $\mathbf{e}^1 \wedge \cdots \wedge \mathbf{e}^n \neq 0$ too.

Case $k < n$. Let us assume that

$$\sum_J \mu_J \mathbf{e}^{j_1} \wedge \cdots \wedge \mathbf{e}^{j_k} = 0, \tag{4.2}$$

where the sum runs over increasing lists $J = (j_1 < \cdots < j_k)$ and the μ_J are scalars. Let us choose an increasing list I of length k. We denote by I^c the complement of I in $\{1, \dots, n\}$, arranged in increasing order. To be specific,

$$I = (i_1, \dots, i_k), \qquad I^c = (i_{k+1}, \dots, i_n).$$

We define a linear map $S : T^k(E) \to T^n(E)$ by

$$S(\mathbf{e}^{j_1} \otimes \cdots \otimes \mathbf{e}^{j_n}) := \mathbf{e}^{j_1} \otimes \cdots \otimes \mathbf{e}^{j_n} \otimes \mathbf{e}^{i_{k+1}} \otimes \cdots \otimes \mathbf{e}^{i_n}.$$

Obviously, we have $S(L_k) \subset L_n$ and thus S passes to the quotient. This yields a linear map $s : \Lambda^k(E) \to \Lambda^n(E)$ satisfying

$$s(\mathbf{e}^{j_1} \wedge \cdots \wedge \mathbf{e}^{j_n}) := \mathbf{e}^{j_1} \wedge \cdots \wedge \mathbf{e}^{j_n} \wedge \mathbf{e}^{i_{k+1}} \wedge \cdots \wedge \mathbf{e}^{i_n}.$$

Applying s to (4.2), and remembering that $\mathbf{e}^{j_1} \wedge \cdots \wedge \mathbf{e}^{j_n} \wedge \mathbf{e}^{i_{k+1}} \wedge \cdots \wedge \mathbf{e}^{i_n}$ vanishes if an index occurs twice, we obtain

$$\mu_I \mathbf{e}^1 \wedge \cdots \wedge \mathbf{e}^n = 0.$$

Because $\mathbf{e}^1 \wedge \cdots \wedge \mathbf{e}^n \neq 0$ (from the case $k = n$ above), we deduce $\mu_I = 0$. This proves that our generating set is free.
□

4.2.3 *The Tensor and Exterior Algebras*

The spaces $T^k(E)$ may be summed up so as to form the *tensor algebra* of E, denoted $T(E)$:

$$T(E) = K \oplus E \oplus T^2(E) \oplus \cdots \oplus T^k(E) \oplus \cdots.$$

We recall that an element of $T(E)$ is a sequence $(x^0, x^1, \dots, x^k, \dots)$ whose kth element is in $T^k(E)$ and only finitely many of them are nonzero. It is thus a *finite* sum $x^0 \oplus \cdots \oplus x^k + \cdots$.

The word *algebra* is justified by the following bilinear operation, defined from $T^k(E) \times T^\ell(E)$ into $T^{k+\ell}(E)$ and then extended to $T(E)$ by linearity. It makes $T(E)$ a *graded algebra*. We need only to define a product over elements of the bases:

$$(\mathbf{e}^{i_1} \otimes \cdots \otimes \mathbf{e}^{i_k}) \cdot (\mathbf{e}^{j_1} \otimes \cdots \otimes \mathbf{e}^{j_\ell}) = \mathbf{e}^{i_1} \otimes \cdots \otimes \mathbf{e}^{i_k} \otimes \mathbf{e}^{j_1} \otimes \cdots \otimes \mathbf{e}^{j_\ell}.$$

When $k = 0$, we simply have

$$\lambda \cdot (\mathbf{e}^{j_1} \otimes \cdots \otimes \mathbf{e}^{j_\ell}) = (\lambda \mathbf{e}^{j_1}) \otimes \mathbf{e}^{j_2} \otimes \cdots \otimes \mathbf{e}^{j_\ell}.$$

Because of the identities above, it is natural to denote this product with the same tensor symbol $\otimes$. We thus have

$$(\mathbf{e}^{i_1} \otimes \cdots \otimes \mathbf{e}^{i_k}) \otimes (\mathbf{e}^{j_1} \otimes \cdots \otimes \mathbf{e}^{j_\ell}) = \mathbf{e}^{i_1} \otimes \cdots \otimes \mathbf{e}^{i_k} \otimes \mathbf{e}^{j_1} \otimes \cdots \otimes \mathbf{e}^{j_\ell}.$$

If $E \neq \{0\}$, the vector space $T(E)$ is infinite dimensional. As an algebra, it is associative, inasmuch as associativity holds true for elements of the canonical basis.

The bilinear map $B : (x,y) \mapsto x \otimes y$, from $T^k(E) \times T^\ell(E)$ into $T^{k+\ell}(E)$, obviously satisfies

$$B(L_k \times T^\ell(E)) \subset L_{k+\ell}, \qquad B(T^k(E) \times L_\ell) \subset L_{k+\ell}.$$

It therefore corresponds to a bilinear map $b : \Lambda^k(E) \times \Lambda^\ell(E) \to \Lambda^{k+\ell}(E)$ verifying

$$b(x^1 \wedge \cdots \wedge x^k, x^{k+1} \wedge \cdots \wedge x^{k+\ell}) = x^1 \wedge \cdots \wedge x^{k+\ell}.$$

This operation is called the *exterior product*. From its definition, it is natural to denote it again with the same wedge symbol $\wedge$. We thus have

$$(x^1 \wedge \cdots \wedge x^k) \wedge (x^{k+1} \wedge \cdots \wedge x^{k+\ell}) = x^1 \wedge \cdots \wedge x^{k+\ell}.$$

Again, the exterior product allows us to define the *graded algebra*

$$\Lambda(E) = K \oplus E \oplus \Lambda^2(E) \oplus \cdots \oplus \Lambda^n(E).$$

We point out that because of Theorem 4.3, this sum involves only $n+1$ terms. Again, the *exterior algebra* $\Lambda(E)$ is associative. Its dimension equals 2^n, thanks to the identity

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

4.2.3.1 Rules

Let $x, y \in E$ be given. Because $x \otimes y + y \otimes x = (x+y) \otimes (x+y) - x \otimes x - y \otimes y \in L_2$, we have

$$y \wedge x = -x \wedge y, \qquad \forall x, y \in E. \tag{4.3}$$

When dealing with the exterior product of higher order, the situation is slightly different. For instance, if x, y, z belong to E, then (4.3) together with associativity give

$$x \wedge (y \wedge z) = (x \wedge y) \wedge z = -(y \wedge x) \wedge z = -y \wedge (x \wedge z) = y \wedge (z \wedge x) = (y \wedge z) \wedge x.$$

By linearity, we deduce that if $x \in E$ and $w \in \Lambda^2(E)$, then $x \wedge w = w \wedge x$. More generally, we prove the following by induction over k and ℓ.

Proposition 4.2 *If $w \in \Lambda^k(E)$ and $z \in \Lambda^\ell(E)$, then*

$$z \wedge w = (-1)^{k\ell} w \wedge z.$$

4.2.3.2 A Commutative Subalgebra

The sum

$$\Lambda_{\text{even}}(E) = K \oplus \Lambda^2(E) \oplus \Lambda^4(E) \oplus \cdots$$

is obviously a subalgebra of $\Lambda(E)$, of dimension 2^{n-1} because of

$$\sum_{0\leq k\leq n/2}\binom{n}{2k}=2^{n-1}.$$

Because of Proposition 4.2, it is actually commutative.

Corollary 4.2 *If* $w,z\in\Lambda_{\mathrm{even}}(E)$*, then* $w\wedge z=z\wedge w\in\Lambda_{\mathrm{even}}(E)$.

4.3 Tensorization of Linear Maps

4.3.1 Tensor Product of Linear Maps

Let E_0,E_1,F_0,F_1 be vector spaces over K. If $u_j\in\mathscr{L}(E_j;F_j)$, Theorem 4.1 allows us to define a linear map $u_0\otimes u_1\in\mathscr{L}(E_0\otimes E_1;F_0\otimes F_1)$, satisfying

$$(u_0\otimes u_1)(x\otimes y)=u_0(x)\otimes u_1(y).$$

A similar construction is available with an arbitrary number of linear maps $u_j:E_j\to F_j$.

Let us choose bases $\{\mathbf{e}^{01},\dots,\mathbf{e}^{0m}\}$ of E_0, $\{\mathbf{e}^{11},\dots,\mathbf{e}^{1q}\}$ of E_1, $\{\mathbf{f}^{01},\dots,\mathbf{f}^{0n}\}$ of F_0, $\{\mathbf{f}^{11},\dots,\mathbf{f}^{1p}\}$ of F_1. Let A and B be the respective matrices of u_0 and u_1 in these bases. Then

$$\begin{aligned}(u_0\otimes u_1)(\mathbf{e}^{0i}\otimes\mathbf{e}^{1j}) &= \left(\sum_k a_{ki}\mathbf{f}^{0k}\right)\otimes\left(\sum_\ell b_{\ell j}\mathbf{f}^{1\ell}\right)\\ &=\sum_{k,\ell}a_{ki}b_{\ell j}\mathbf{f}^{0k}\otimes\mathbf{f}^{1\ell}.\end{aligned}$$

This shows that the $((k,l),(i,j))$-entry of the matrix of $u_0\otimes u_1$ in the tensor bases is the product $a_{ki}b_{\ell j}$. If we arrange the bases $(\mathbf{e}^{0i}\otimes\mathbf{e}^{1j})_{i,j}$ and $(\mathbf{f}^{0k}\otimes\mathbf{f}^{1\ell})_{k,\ell}$ in lexicographic order, then this matrix reads blockwise

$$\begin{pmatrix}a_{11}B & a_{12}B & \dots & a_{1m}B\\ a_{21}B & \ddots & & \vdots\\ \vdots & & & \\ a_{n1}B & \dots & & a_{nm}B\end{pmatrix}.$$

This matrix is called the *tensor product* of A and B, denoted $A\otimes B$.

4.3.2 Exterior Power of an Endomorphism

If $u \in \mathbf{End}(E)$, then $u \otimes u \otimes \cdots \otimes u =: u^{\otimes k}$ has the property that $u(L_k) \subset L_k$. Therefore there exists a unique linear map $\Lambda^k(u)$ such that

$$u(x^1 \wedge \cdots \wedge x^k) = u(x^1) \wedge \cdots \wedge u(x^k).$$

Proposition 4.3 *If A is the matrix of $u \in \mathbf{End}(E)$ in a basis $\{\mathbf{e}^1, \dots, \mathbf{e}^n\}$, then the entries of the matrix $A^{(k)}$ of $\Lambda^k(u)$ in the basis of vectors $\mathbf{e}^{j_1} \wedge \cdots \wedge \mathbf{e}^{j_n}$ are the $k \times k$ minors of A.*

Proof. This is essentially the same line as in the proof of the Cauchy–Binet formula (Proposition 3.4). □

Corollary 4.3 *If* $\dim E = n$ *and $u \in \mathbf{End}(E)$, then $\Lambda^n(u)$ is multiplication by* $\det u$.

4.4 A Polynomial Identity in $\mathbf{M_n}(\mathbf{K})$

We already know a polynomial identity in $\mathbf{M}_n(K)$, namely the Cayley–Hamilton theorem: $P_A(A) = 0_n$. However, it is a bit complicated because the matrix is involved both as the argument of the polynomial and in its coefficients. We prove here a remarkable result, where a multilinear application vanishes identically when the arguments are arbitrary $n \times n$ matrices. To begin with, we introduce some special polynomials in *noncommutative* indeterminates.

4.4.1 The Standard Noncommutative Polynomial

Noncommutative polynomials in indeterminates $X_1, \dots, X_\ell$ are linear combinations of *words* written in the alphabet $\{X_1, \dots, X_\ell\}$. The important rule is that in a word, you are not allowed to permute two distinct letters: $X_i X_j \neq X_j X_i$ if $j \neq i$, contrary to what occurs in ordinary polynomials.

The *standard polynomial* $\mathscr{S}_\ell$ in noncommutative indeterminates $X_1, \dots, X_\ell$ is defined by

$$\mathscr{S}_\ell(X_1, \dots, X_\ell) := \sum \varepsilon(\sigma) X_{\sigma(1)} \cdots X_{\sigma(\ell)}.$$

Hereabove, the sum runs over the permutations of $\{1, \dots, \ell\}$, and $\varepsilon(\sigma)$ denotes the signature of σ. For instance, $\mathscr{S}_2(X,Y) = XY - YX = [X,Y]$. The standard polynomial is thus a tool for measuring the defect of commutativity in a ring or an algebra: a ring R is Abelian if $\mathscr{S}_2$ vanishes identically over $R \times R$.

The following formula is obvious.

Lemma 4. *Let $A_1, \dots, A_r \in \mathbf{M}_n(K)$ be given. We form the matrix $A \in \mathbf{M}_n(\Lambda(K^r)) \sim \mathbf{M}_n(K) \otimes_K \Lambda(K^r)$ by*

$$A = A_1\mathbf{e}^1 + \cdots + A_r\mathbf{e}^r.$$

We emphasize that A has entries in a noncommutative ring.

Then we have

$$A^\ell = \sum_{i_1<\cdots<i_\ell} \mathscr{S}_\ell(A_{i_1},\dots,A_{i_\ell})\,\mathbf{e}^{i_1}\wedge\cdots\wedge\mathbf{e}^{i_\ell}.$$

In particular, we have

$$A^r = \mathscr{S}_r(A_1,\dots,A_r)\,\mathbf{e}^1\wedge\cdots\wedge\mathbf{e}^r. \tag{4.4}$$

The other important formula generalizes the well-known identity $\mathrm{Tr}[A,B]=0$. To begin with, we easily have

$$\mathscr{S}_\ell(X_{\sigma(1)},\dots,X_{\sigma(\ell)}) = \varepsilon(\sigma)\mathscr{S}_\ell(X_1,\dots,X_\ell), \qquad \forall\sigma\in\mathbf{S}_\ell.$$

Applying this to a cycle, we deduce

$$\mathscr{S}_\ell(X_2,\dots,X_\ell,X_1) = (-1)^{\ell-1}\mathscr{S}_\ell(X_1,\dots,X_\ell).$$

Because $\mathrm{Tr}(A_2\cdots A_\ell A_1) = \mathrm{Tr}(A_1\cdots A_\ell)$, we infer the following.

Lemma 5. *If ℓ is even and $A_1,\dots,A_\ell\in\mathbf{M}_n(R)$ (R a commutative ring), then*

$$\mathrm{Tr}\,\mathscr{S}_\ell(A_1,\dots,A_\ell) = 0.$$

Proof. If ℓ is even, we have

$$\mathrm{Tr}\,\mathscr{S}_\ell(A_1,\dots,A_\ell) = -\mathrm{Tr}\,\mathscr{S}_\ell(A_2,\dots,A_\ell,A_1) = -\mathrm{Tr}\,\mathscr{S}_\ell(A_1,\dots,A_\ell),$$

the first equality because this is true even before taking the trace, and the last equality because of $\mathrm{Tr}(AB)=\mathrm{Tr}(BA)$. If $2x=0$ implies $x=0$ in R, we deduce

$$\mathrm{Tr}\,\mathscr{S}_\ell(A_1,\dots,A_\ell) = 0.$$

For instance, this is true if $R=\mathbb{C}$. Because $\mathrm{Tr}\,\mathscr{S}_\ell(\cdots)$ belongs to $\mathbb{Z}[Y_1,\dots,Y_{\ell n^2}]$, it must vanish as a polynomial. Thus the identity is valid in every commutative ring R. □

4.4.2 The Theorem of Amitsur and Levitzki

A beautiful as well as surprising fact is that $\mathbf{M}_n(K)$ does have some amount of commutativity, although it seems at first glance to be a paradigm for noncommutative algebras.

Theorem 4.4 (A. Amitsur and J. Levitzki) *The standard polynomial in* $2n$ *non-commutative indeterminates vanishes over* $\mathbf{M}_n(K)$*: for every* $A_1,\dots,A_{2n}\in\mathbf{M}_n(K)$*, we have*

$$\mathscr{S}_{2n}(A_1,\dots,A_{2n})=0_n. \tag{4.5}$$

Comments

- This result is accurate in the sense that $\mathscr{S}_k$ does not vanish identically over $n\times n$ matrices. For instance, $\mathscr{S}_k$ does not vanish over the $(2n-1)$-uplet of matrices $\mathbf{e}^n\otimes\mathbf{e}^1$, $\mathbf{e}^{n-1}\otimes\mathbf{e}^1$, ..., $\mathbf{e}^1\otimes\mathbf{e}^1$, $\mathbf{e}^1\otimes\mathbf{e}^2$, ..., $\mathbf{e}^1\otimes\mathbf{e}^n$.
- Actually, it is known that $\mathscr{S}_{2n}\equiv 0_n$ is the only polynomial identity of degree less than or equal to $2n$ over $\mathbf{M}_n(\mathbb{C})$.
- Theorem 4.4 has long remained mysterious, with a complicated proof, until Rosset published a simple proof in 1976. This is the presentation that we give below.
- We notice that inasmuch as $\mathscr{S}_{2n}$ may be viewed as a list of n^2 elements of $\mathbb{Z}[Y_1,\dots,Y_{2n^3}]$, the identity (4.5) must be valid for matrices whose entries are independent *commuting* indeterminates m_{ij}^{α} with $1\le\alpha\le 2n$ and $1\le i,j\le n$. The theorem is thus a list of n^2 identities in $\mathbb{Z}\left[m_{ij}^{\alpha}\,|\,1\le\alpha\le 2n,\,1\le i,j\le n\right]$.

Proof. (Taken from Rosset [32].)

We thus assume that $K=\mathbb{C}$. As above, we form the matrix $A\in\mathbf{M}_n(\Lambda(\mathbb{C}^{2n}))$ by

$$A=A_1\mathbf{e}^1+\cdots A_{2n}\mathbf{e}^{2n}.$$

Because of (4.4), what we have to prove is that $A^{2n}=0_n$.

The matrix A has the flaw of having noncommutative entries. However, Corollary 4.2 tells us that the entries of A^2 belong to the abelian ring $\Lambda_{\text{even}}(\mathbb{C}^{2n})$. We thus may apply Proposition 3.24 (recall that it is valid for matrices with entries in an Abelian ring R in which $mx=0$ implies $x=0$ whenever m is a positive integer): in order to prove that $A^{2n}=(A^2)^n=0_n$, it is sufficient to prove the identities

$$\mathrm{Tr}(A^{2k})=0,\qquad 1\le k\le n.$$

The latter follow immediately from Lemmas 4 and 5. □

Exercises

1. Show that the rank of an element of $E\otimes F$ is bounded by $\min\{\dim E,\dim F\}$.
2. Let $u\in\mathbf{End}(E)$ be a diagonalizable map, with eigenvalues $\lambda_1,\dots,\lambda_n$, counting multiplicities. Show that $u^{\otimes k}$ and $\Lambda^k(u)$ are diagonalizable and identify their eigenvalues.
3. Let $u\in\mathbf{End}(E)$ be given. Show that the following formula defines an endomorphism over $T^k(E)$, which we denote $u^{\oplus k}$,

$$x^1 \otimes \cdots \otimes x^k \mapsto (u(x^1) \otimes \cdots \otimes x^k) + (x^1 \otimes u(x^2) \otimes \cdots \otimes x^k) \\ + \cdots + (x^1 \otimes \cdots \otimes u(x^k)).$$

a. If u is diagonalizable, show that $u^{\oplus k}$ is so and compute its eigenvalues.
b. Show that $u^{\oplus k}$ passes to the quotient, yielding an endomorphism over $\Lambda^k(E)$, which we denote $u^{\wedge k}$:

$$x^1 \wedge \cdots \wedge x^k \mapsto (u(x^1) \wedge \cdots \wedge x^k) + (x^1 \wedge u(x^2) \wedge \cdots \wedge x^k) \\ + \cdots + (x^1 \wedge \cdots \wedge u(x^k)).$$

c. Again, show that $u^{\wedge k}$ is diagonalizable if u is so, and compute its eigenvalues.

4. Complete the proof of Proposition 4.3.
5. Prove that $\mathscr{S}_{2n-1}$, when applied to the matrices E^{ij} for either $i = 1$ or $j = 1$ (this makes a list of $2n - 1$ matrices), where

$$(E^{ij})_{k\ell} = \delta_i^k \delta_j^\ell, \qquad 1 \le i, j, k, \ell \le n,$$

gives a nonzero matrix.
6. If $A \in \mathbf{M}_n(k)$ is alternate, we define $\mathbb{A} \in \mathbf{M}_n(\Lambda(k^n))$ by

$$\mathbb{A} := \sum_{i<j} a_{ij} \mathbf{e}^i \wedge \mathbf{e}^j.$$

We assume that $n = 2m$. Prove that

$$\mathbb{A}^m = \mathrm{Pf}(A) \mathbf{e}^1 \wedge \cdots \wedge \mathbf{e}^n.$$

Hint: Use the expansion formula established in Exercise 25 of Chapter 3.